

# An Elementary Introduction to Solar Dynamo Theory

Arnab Rai Choudhuri  
Department of Physics  
Indian Institute of Science  
Bangalore — 560012. India

March 7, 2007

## Abstract

The cyclically varying magnetic field of the Sun is believed to be produced by the hydromagnetic dynamo process. We first summarize the relevant observational data pertaining to sunspots and solar cycle. Then we review the basic principles of MHD needed to develop the dynamo theory. This is followed by a discussion how bipolar sunspots form due to magnetic buoyancy of flux tubes formed at the base of the solar convection zone. Following this, we come to the heart of dynamo theory. After summarizing the basic ideas of a turbulent dynamo and the basic principles of its mean field formulation, we present the famous dynamo wave solution, which was supposed to provide a model for the solar cycle. Finally we point out how a flux transport dynamo can circumvent some of the difficulties associated with the older dynamo models.

## 1 Introduction

Several other lecturers in this Winter School must have already convinced you that magnetic fields play a very important role in the Sun. From sunspots to coronal heating, from solar flares to coronal mass ejections—all these apparently diverse phenomena have magnetic fields as their ultimate cause. It is, therefore, no wonder that one of the most fundamental questions of solar physics is: how is the magnetic field of the Sun generated? It is this question which I shall address in my lectures.

We believe that the magnetic fields of not only the Sun, but of all astronomical bodies, are produced by a process called the *hydromagnetic dynamo process*. Let me begin by making a comment about this name, which may seem a little bit peculiar at the first sight. All physics students have to learn the basic principles of an ordinary electromagnetic dynamo. In an ordinary dynamo, we have a conducting coil which is made to rotate in a magnetic field, cutting magnetic flux lines and thereby producing an e.m.f. by Faraday's law of electromagnetic

induction. We do not have discrete coils inside an astronomical body, but sometimes we can have rotating blobs of plasma which are also good conductors of electricity. If a magnetic field is present in the region where plasma blobs are rotating, then Faraday's law generalized to a continuum situation suggests that an e.m.f. can be induced within a rotating plasma blob. Under favourable circumstances, this e.m.f. can reinforce the magnetic field present in the astronomical body. In other words, this hydromagnetic dynamo process can build up a magnetic field within the astronomical body starting from a seed field, but we do need a seed field to begin with. Scientists working on galactic magnetic fields have to bother about the origin of the seed field on which a galactic dynamo can work. However, when studying solar or stellar dynamos, we usually do not bother about the seed field question. Since all stars form from interstellar gas clouds which seem to have magnetic fields, we tacitly assume that there must have been some initial seed field on which the stellar dynamo could have started operating. In solar dynamo theory, we study how the dynamo process sustains the solar magnetic field in the form in which we find it today.

It should be apparent from this brief sketch of the dynamo process that we need to consider the dynamics of the plasma inside the Sun to understand how the solar dynamo operates. Much of my presentation will be devoted to developing the theoretical framework for this purpose. But before we get into theory, I would first like to present some relevant observational data in § 2.

## 2 Basic observational data

Although the solar magnetic fields give rise to many different solar phenomena, it is the sunspots which provide the best detailed information about the behaviour of the solar magnetic fields. So, while discussing observational data, we shall first focus our attention on sunspots, which are dark regions of concentrated magnetic field on the solar surface, and then discuss about the Sun's weak magnetic field outside sunspots. The crucial test for a solar dynamo model is that it has to explain the observational data we are going to summarize in this section.

Hale (1908) was the first to discover the evidence of Zeeman effect in sunspot spectra and made the momentous announcement that sunspots are regions of strong magnetic fields. This is the first time that somebody found conclusive evidence of large-scale magnetic fields outside the Earth's environment. The typical magnetic field of a large sunspot is about 3000 G (i.e. of order  $10^4$  times stronger than the magnetic field on the surface of the Earth).

Even before it was realized that sunspots are seats of solar magnetism, several persons have been studying the statistics of sunspots. Schwabe (1844) noted that the number of sunspots seen on the solar surface increases and decreases with a period of about 11 years. Now we believe that the Sun has a cycle with twice that period, i.e. 22 years. Since the Sun's magnetic field changes its direction after 11 years, it takes 22 years for the magnetic field to come back to its initial configuration. Carrington (1858) found that sunspots seemed to appear

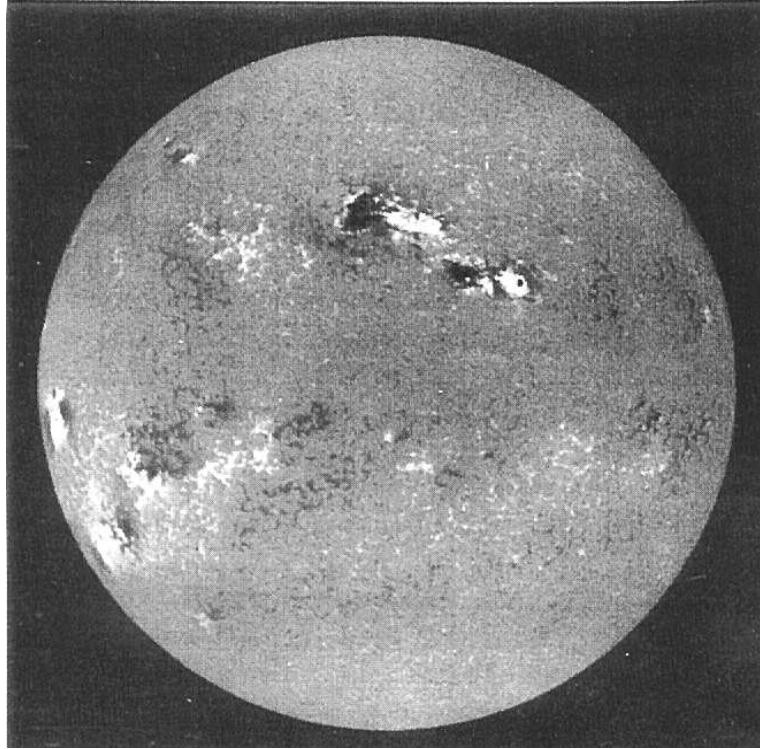


Figure 1: A magnetogram image of the full solar disk. The regions with positive and negative magnetic polarities are respectively shown in white and black, with grey indicating regions where the magnetic field is weak.

at lower and lower latitudes with the progress of the solar cycle. In other words, most of the sunspots in the early phase of a solar cycle are seen between  $30^\circ$  and  $40^\circ$ . As the cycle advances, new sunspots are found at increasingly lower latitudes. Then a fresh half-cycle begins with sunspots appearing again at high latitudes. Individual sunspots live from a few days to a few weeks.

After finding magnetic fields in sunspots, Hale and his coworkers made another significant discovery (Hale *et al.*, 1919). They found that often two large sunspots are seen side by side and they invariably have opposite polarities. The line joining the centres of such a bipolar sunspot pair is, on an average, nearly parallel to the solar equator. Hale's coworker Joy, however, noted that there is a systematic tilt of this line with respect to the equator and that this tilt increases with latitude (Hale *et al.*, 1919). This result is usually known as *Joy's law*. The tilts, however, show a considerable amount of scatter around the mean given by Joy's law. The sunspot in the forward direction of rotation is called the leading spot and the other the following spot. The tilt is such that the leading spot is usually found nearer the equator than the following spot. It was also noted

that the sunspot pairs have opposite polarities in the two hemispheres. In other words, if the leading sunspot in the northern hemisphere has positive polarity, then the leading sunspot in the southern hemisphere has negative polarity. This is clearly seen in Figure 1, which is a magnetic map of the Sun's disk obtained with a magnetogram. The regions of positive and negative polarities are shown in white and black respectively. The polarities of the bipolar sunspots in both hemispheres get reversed from one half-cycle of 11 years to the next half-cycle.

After the development of the magnetogram by Babcock and Babcock (1955), it became possible to study the much weaker magnetic field near the poles of the Sun. This magnetic field is of the order of 10 G and reverses its direction at the time of solar maximum (i.e. when the number of sunspots seen on the solar surface is maximum) (Babcock, 1959). This shows that this weak, diffuse field of the Sun is somehow coupled to the much stronger magnetic field of the sunspots and is a part of the same solar cycle. Low-resolution magnetograms show the evidence of weak magnetic field even in lower latitudes. The true nature of this field is not very clear. It has been found (Stenflo, 1973) that the magnetic field on the solar surface outside sunspots often exists in the form of fibril flux tubes of diameter of the order of 300 km with field strength of about 2000 G (large sunspots have sizes larger than 10,000 km). One is not completely sure if the field found in the low-resolution magnetograms is truly a diffuse field or a smearing out of the contributions made by fibril flux tubes. Keeping this caveat in mind, we should refer to the field outside sunspots as seen in magnetograms as the 'diffuse' field. It was found that there were large unipolar matches of this diffuse field on the solar surface which migrated poleward (Bumba and Howard, 1965). Even when averaged over longitude, one finds predominantly one polarity in a belt of latitude which drifts poleward (Howard and LaBonte, 1981; Wang, Nash and Sheeley, 1989). The reversal of polar field presumably takes place when sufficient field of opposite polarity has been brought near the poles.

Figure 2 shows the distribution of both sunspots and the weak, diffuse field in a plot of latitude vs. time. The various shades of grey indicate values of longitude-averaged diffuse field, whereas the latitudes where sunspots were seen at a particular time are marked in black. The sunspot distribution in a time-latitude plot is often referred to as a *butterfly diagram*, since the pattern (the regions marked in black in Figure 2) reminds one of butterflies. Such butterfly diagrams were first plotted by Maunder (1904). Historically, most of the early dynamo models concentrated on explaining the distribution of sunspots and ignored the diffuse field. Only during the last few years, it has been realized that the diffuse fields give us important clues about the dynamo process and they should be included in a full self-consistent theory. The aim of such a theory should be to explain diagrams like Figure 2 (i.e. not just the butterfly diagram).

Here I may mention an important study made from the photographs taken in the Kodaikanal Observatory. There were no direct measurements of the Sun's weak, diffuse magnetic field in the early decades of the 20th century. However, one can still obtain information about this field by indirect means. Dark filaments are found on the solar surface at the boundaries between the positive and negative regions of the weak, diffuse field. One can draw some conclusions about

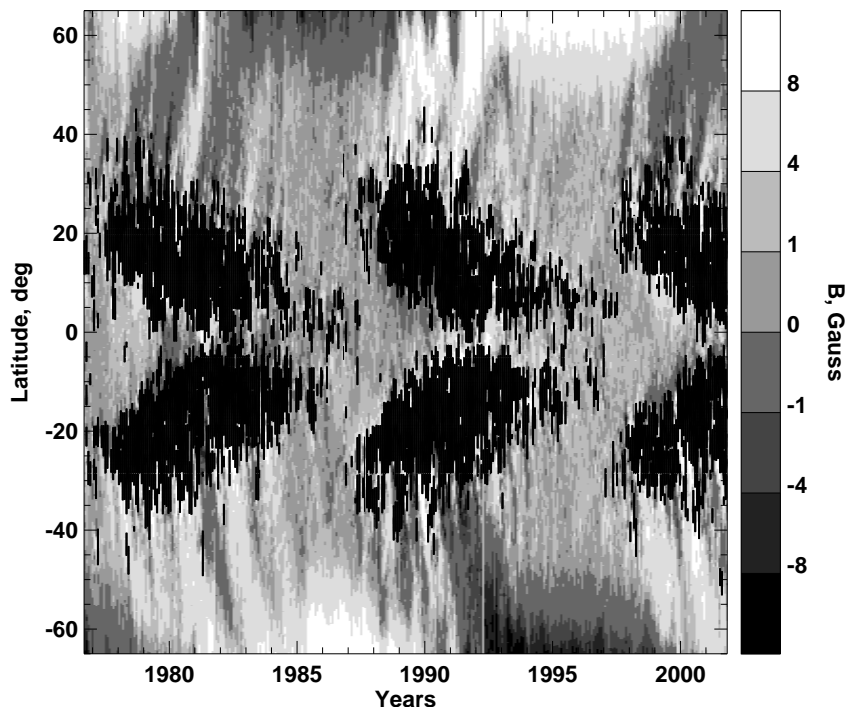


Figure 2: Shades of grey showing the latitude-time distribution of longitudinally averaged weak, diffuse magnetic field ( $B$  is in Gauss) with a ‘butterfly diagram’ of sunspots superimposed on it.

the evolution of the weak, diffuse field by studying the positions of these dark filaments. By analyzing the photographs of the Sun taken regularly from the Kodaikanal Observatory for nearly a century, Makarov and Sivaraman (1989) could show the poleward migration of the weak, diffuse field even when there were still no measurements of this magnetic field.

We have provided above a summary of the various regular features in the Sun’s activity cycle. One finds lots of irregularities and fluctuations superposed on the underlying regular behaviour, as can be seen in Figure 2. These irregularities are more clearly visible in Figure 3, where the number of sunspots seen on the solar surface is plotted against time. Galileo was one of the first persons in Europe to study sunspots at the beginning of the 17th century. After Galileo’s work, sunspots were almost not seen for nearly a century! Such a grand minimum has not occurred again in the last 300 years.

It may be noted that all the observations discussed above pertain to the Sun’s surface. We have no direct information about the magnetic field underneath the Sun’s surface. The new science of helioseismology, however, has provided us

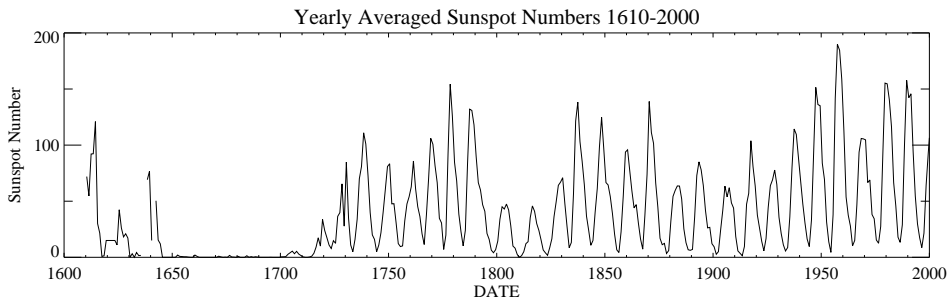


Figure 3: The monthly number of sunspots plotted against time for the period 1610–2000.

lots of information about the velocity field underneath the solar surface. Other lecturers in the Winter School have given an introduction to helioseismology. We shall have occasions to refer to some of the helioseismic findings in our discussion later. It is to be remembered that heat is transported by convection in the outer layers of the Sun from about  $0.7R_{\odot}$  to  $R_{\odot}$  (where  $R_{\odot}$  is the solar radius). This region is called the convection zone, within which the plasma is in a turbulent state. The job of a theorist now is to construct a detailed model of the physical processes in this turbulent plasma such that all the surface observations of magnetic fields are properly explained—a fairly daunting problem, of which the full solution is still a distant dream.

### 3 Basics of Magnetohydrodynamics

After summarizing the relevant observations in the previous section, we now look at the basics of theory. Although air is a collection of molecules, we can treat air as a continuum fluid while analyzing many problems of air flow. Exactly similarly, even though a plasma is a collection of positively charged ions and negatively charged electrons, we can regard the plasma also as a continuum fluid while discussing many problems in plasma dynamics. The branch of plasma physics in which the plasma is treated as a continuum fluid is called *magnetohydrodynamics*, abbreviated as MHD. It should be kept in mind that there are many problems in plasma physics which require a more microscopic treatment of plasma and one has to go beyond MHD. For example, the propagation of electromagnetic waves through a plasma is an important problem in plasma physics which cannot be handled with MHD. However, when we study non-relativistic ( $|\mathbf{v}| \ll c$ ) bulk motions of the plasma under the influence of gravitational and magnetic forces, MHD is usually adequate. For almost all problems related to the dynamics of plasma underneath the Sun's surface, MHD provides the appropriate theoretical framework.

Due to the electrical attraction between positive and negative charges in a plasma, these opposite charges usually remain very well mixed. In other words,

if we take a volume element of plasma which has sufficiently large number of particles, we would find that the positive and negative charges almost exactly balance each other. Does that mean that a volume element of a plasma is exactly like a volume element of an ordinary non-ionized gas? The main difference is that an ordinary non-ionized gas is a bad conductor of electricity. On the other hand, it is possible for the electrons and ions in a plasma to move systematically with respect to each other, thereby giving rise to a current, even though volume elements of the plasma remain charge-neutral. In other words, in MHD we regard a plasma as a special kind of fluid which is a very good conductor of electricity. Since volume elements of the plasma remain charge-neutral, we do not expect large-scale electric fields in the plasma. But currents flowing through the plasma can give rise to large-scale magnetic fields.

One combines Maxwell's equations with the equations of fluid mechanics to derive the basic equations of MHD. It is beyond the scope of my lectures to give a systematic derivation of the basic equations. I refer to my book (Choudhuri, 1998) for a systematic introduction to MHD. Here I shall just write down some of the basic equations and discuss their significance. The velocity field  $\mathbf{v}$  and the magnetic field  $\mathbf{B}$  in a plasma interact with each other according to the following MHD equations:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla) \mathbf{v} = -\frac{1}{\rho} \nabla \left( p + \frac{B^2}{2\mu} \right) + \frac{(\mathbf{B} \cdot \nabla) \mathbf{B}}{\mu\rho} + \mathbf{g}, \quad (1)$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \lambda \nabla^2 \mathbf{B}. \quad (2)$$

Here  $\rho$  is density,  $p$  is pressure,  $\mathbf{g}$  is gravitational field and

$$\lambda = \frac{1}{\mu\sigma} \quad (3)$$

is called magnetic diffusivity ( $\sigma$  is electrical conductivity). Equation (1) is essentially the Euler equation of fluid mechanics, to which magnetic forces have been added. It is clear from (1) that the magnetic field has two effects: (i) it gives rise to an additional pressure  $B^2/2\mu$ ; and (ii) the other magnetic term  $(\mathbf{B} \cdot \nabla) \mathbf{B}/\mu\rho$  is of the nature of a tension along magnetic field lines.

Equation (2) is the key equation in MHD and is called the *induction equation*, since Faraday's law of electromagnetic induction is incorporated within it. If  $V$ ,  $B$  and  $L$  are the typical values of velocity, magnetic field and length scale, then the two terms on the r.h.s. of (2) are of order  $VB/L$  and  $\lambda B/L^2$ . The ratio of these two terms is a dimensionless number, known as the magnetic Reynolds number, given by

$$R_m = \frac{VB/L}{\lambda B/L^2} = \frac{VL}{\lambda}. \quad (4)$$

Since  $R_m$  goes as  $L$ , it is expected to be much larger in astrophysical situations than it is in the laboratory. In fact, usually one finds that  $R_m \gg 1$  in astrophysical systems and  $R_m \ll 1$  in laboratory-size objects. Hence the behaviours of

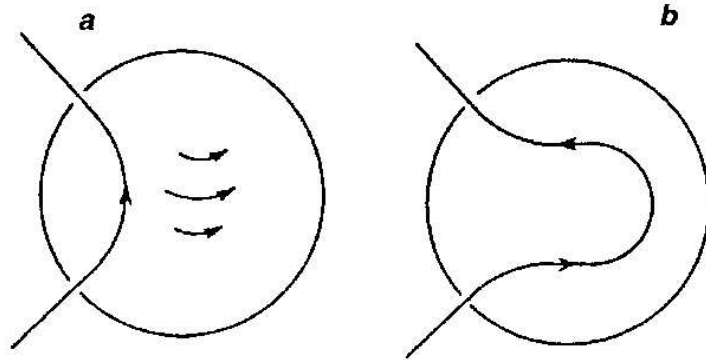


Figure 4: The production of a strong toroidal magnetic field underneath the Sun's surface. **a.** An initial poloidal field line. **b.** A sketch of the field line after it has been stretched by the faster rotation near the equatorial region.

magnetic fields are very different in laboratory plasmas and astrophysical plasmas. For example, it is not possible to have a laboratory analogue of the Sun's self-sustaining magnetic field. Only recently it has been possible to demonstrate the dynamo mechanism in the laboratory (Gailitis *et al.*, 2000). If  $R_m \gg 1$  in an astrophysical system, then the diffusion term in (2) is negligible compared to the term preceding it. In such a situation, it can be shown that the magnetic field is frozen in the plasma and moves with it. This result was first recognized by Alfvén (1942) and is often referred to as *Alfvén's theorem of flux-freezing*. A proof of this fundamental theorem can be found in Choudhuri (1998, §4.6). Several other lecturers in this Winter School also must be using this theorem in their theoretical discussions.

It is known that the Sun does not rotate like a solid body. The angular velocity at the equator is about 20% faster than that at the poles. Because of the flux freezing, this differential rotation would stretch out any magnetic field line in the toroidal direction (i.e. the  $\phi$  direction with respect to the Sun's rotation axis). This is indicated in Figure 4. We, therefore, expect that the magnetic field inside the Sun may be predominantly in the toroidal direction.

We already mentioned in §2 that energy is transported by convection in the layers underneath the Sun's surface. To understand why the magnetic field remains concentrated in structures like sunspots instead of spreading out more evenly, we need to study the interaction of the magnetic field with the convection in the plasma. This subject is known as *magnetoconvection*. The linear theory of convection in the presence of a vertical magnetic field was studied by Chandrasekhar (1952). The nonlinear evolution of the system, however, can only be found from numerical simulations pioneered by Weiss (1981). It was found that space gets separated into two kinds of regions. In certain regions, magnetic field is excluded and vigorous convection takes place. In other regions, magnetic field gets concentrated, and the tension of magnetic field lines suppresses convection



in those regions. Sunspots are presumably such regions where magnetic field is piled up by surrounding convection. Since heat transport is inhibited there due to the suppression of convection, sunspots look darker than the surrounding regions.

Although we have no direct information about the state of the magnetic field under the Sun's surface, it is expected that the interactions with convection would keep the magnetic field concentrated in bundles of field lines throughout the solar convection zone. Such a concentrated bundle of magnetic field lines is called a *flux tube*. In the regions of strong differential rotation, therefore, we may have the magnetic field in the form of flux tubes aligned in the toroidal direction. If a part of such a flux tube rises up and pierces the solar surface as shown in Figure 5b, then we expect to have two sunspots with opposite polarities at the same latitude. But how can a configuration like Figure 5b arise? The answer to this question was provided by Parker (1955a) through his idea of magnetic buoyancy. We have seen in (1) that a pressure  $B^2/2\mu$  is associated with a magnetic field. If  $p_{\text{in}}$  and  $p_{\text{out}}$  are the gas pressures inside and outside a flux tube, then we need to have

$$p_{\text{out}} = p_{\text{in}} + \frac{B^2}{2\mu} \quad (5)$$

to maintain pressure balance across the surface of a flux tube. Hence

$$p_{\text{in}} \leq p_{\text{out}}, \quad (6)$$

which often, though not always, implies that the density inside the flux tube is less than the surrounding density. If this happens in a part of the flux tube, then that part becomes buoyant and rises against the gravitational field to produce the configuration of Figure 5b starting from Figure 5a. We shall discuss in the next section how the idea of magnetic buoyancy can be elaborated to model different aspects of bipolar sunspots.

## 4 Formation of bipolar sunspots

A look at Figure 4 now ought to convince the reader that the sub-surface toroidal field in the two hemispheres should have opposite polarity, if it is produced from a poloidal field of dipolar nature as sketched in Figure 4. If this toroidal field rises due to magnetic buoyancy to produce the bipolar sunspot pairs, then we expect the bipolar sunspots to have opposite polarities in the two hemispheres as seen in Figure 1. We thus see that combining the ideas of flux freezing, magnetoconvection and magnetic buoyancy, we can understand many aspects of the bipolar sunspot pairs. Flux freezing tells us that the magnetic field in the interior of the Sun should be stretched in the toroidal direction by differential rotation. Magnetoconvection calculations suggest that the magnetic field should exist in the form of flux tubes within the convection zone. Finally, we use the idea of magnetic buoyancy to explain how bipolar sunspots arise. To understand the

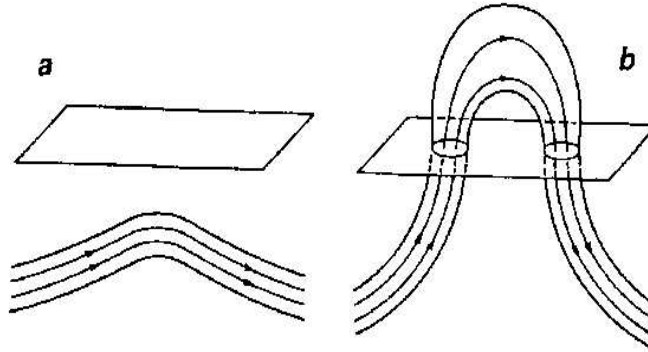


Figure 5: Magnetic buoyancy of a flux tube. **a.** A nearly horizontal flux tube under the solar surface. **b.** The flux tube after its upper part has risen through the solar surface.

formation of bipolar sunspots more fully, we have to look at some of the details like where exactly inside the Sun the toroidal field is formed and why parts of it become buoyant.

Since the toroidal field is generated by the stretching of the poloidal field by differential rotation, we expect the strongest toroidal fields to be generated in the regions where differential rotation (i.e. the gradient of angular velocity) is strongest. One of the major achievements of helioseismology in the last two decades was to map the distribution of angular velocity in the interior of the Sun. You must have learned about this in the lectures of Antia, and you must be knowing that there is thin layer called *tachocline* at the bottom of the solar convection zone where strong differential rotation is concentrated. It is within this tachocline that we expect the strong toroidal field to be generated. If interaction with convection keeps the magnetic field in the form of flux tube, then the toroidal field generated in the tachocline may be in the form of flux rings going around the rotation axis of the Sun. Now, it can be shown that magnetic buoyancy is particularly destabilizing in the interior of the convection zone, where convective instability and magnetic buoyancy reinforce each other. On the other hand, if a region is stable against convection, then magnetic buoyancy can be partially suppressed there (see, for example, Parker, 1979, §8.8). Since the toroidal flux tube is produced at the bottom of the convection zone, we may expect some parts of it to come into the convection zone and become buoyant, whereas other parts may remain underneath the bottom of the convection zone and stay anchored there due to the suppression of magnetic buoyancy. A part of the flux tube coming within the convection zone is expected to rise and eventually reach the solar surface to form sunspots, as sketched in Figure 5. In order to model the formation of bipolar sunspots, we have to study the dynamics of flux tubes rising through the convection zone due to magnetic buoyancy.

The best way to study this problem is to treat it as an initial-value problem.

First, an initial configuration with a magnetic flux ring at the bottom of the convection zone, having a part coming inside the convection zone, is specified, and then its subsequent evolution is studied numerically. The evolution depends on the strength of magnetic buoyancy, which is in turn determined by the value of the magnetic field. We shall give arguments in §6 why most of the dynamo theorists till the early 1990s believed that the magnetic energy density should be in equipartition with the kinetic energy density of convection, i.e.

$$\frac{B^2}{2\mu} \approx \frac{1}{2}\rho v^2. \quad (7)$$

This suggests  $B \approx 10^4$  G on the basis of standard models of the convection zone. If we use full MHD equations to study the evolution of the flux tube, then the calculations become extremely complicated. However, if the radius of cross-section of the flux tube is smaller than the various scale heights, then it is possible to derive an equation for flux tube dynamics from the MHD equations (Spruit, 1981; Choudhuri, 1990). Even this flux tube equation is a sufficiently complicated nonlinear equation and has to be solved numerically. The evolution of such magnetic flux tubes due to magnetic buoyancy (starting from the bottom of the convection zone) was studied by Choudhuri and Gilman (1987) and Choudhuri (1989). It was found that the Coriolis force due to the Sun's rotation plays a much more important role in this problem than what anybody suspected before. If the initial magnetic field is taken to have a strength around  $10^4$  G, then the flux tubes move parallel to the rotation axis and emerge at very high latitudes rather than at latitudes where sunspots are seen. Only if the initial magnetic field is taken as strong as  $10^5$  G, then magnetic buoyancy is strong enough to overpower the Coriolis force and the magnetic flux tubes can rise radially to emerge at low latitudes.

D'Silva and Choudhuri (1993) extended these calculations to look at the tilts of emerging bipolar regions at the surface. These tilts are also produced by the action of the Coriolis force on the rising flux tube. Figure 6 taken from D'Silva and Choudhuri (1993) shows the observational tilt vs. latitude plot of bipolar sunspots (i.e. Joy's law) along with the theoretical plots obtained by assuming different values of the initial magnetic field. It is clearly seen that theory fits observations only if the initial magnetic field is about  $10^5$  G. Apart from providing the first quantitative explanation of Joy's law nearly three-quarters of a century after its discovery, these calculations put the first stringent limit on the value of the toroidal magnetic field at the bottom of the convection zone. Several other authors (Fan, Fisher and DeLuca, 1993; Caligari *et al.*, 1995) soon performed similar calculations and confirmed the result. The evidence is now mounting that the magnetic field at the bottom of the convection zone is indeed much stronger than the equipartition value given by (7). We already mentioned that the tilts of active regions have a large amount of scatter around the mean given by Joy's law. In fact, it is found that active regions often emerge with initial tilts inconsistent with Joy's law and then the tilts change in the next few days to come closer to values given by Joy's law (Howard, 1996). Longcope and Choudhuri (2001) have argued that the vigorous convective turbulence in

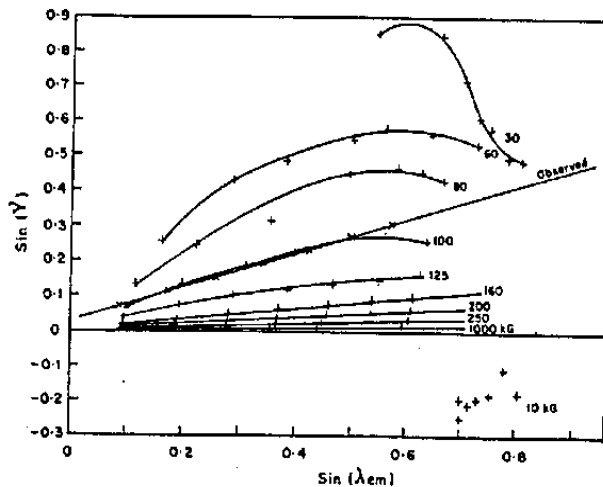


Figure 6: Plots of  $\sin(\text{tilt})$  against  $\sin(\text{latitude})$  theoretically obtained for different initial values of magnetic field indicated in kG. The observational data indicated by the straight line fits the theoretical curve for initial magnetic field 100 kG (i.e.  $10^5$  G). Reproduced from D'Silva and Choudhuri (1993).

the top layers of the convection zone exerts a random force on the tops of the rising flux loops, causing the scatter around the Joy's law, and then the tilt of the flux tube relaxes to the appropriate value after emergence when it is no longer kicked by convective turbulence.

We thus see that flux tube dynamics simulations can explain many aspects of bipolar sunspot formation and also put a tight constraint on the value of the toroidal magnetic field at the bottom of the convection zone. We shall find that this constraint is going to play a very important role in dynamo theory. After understanding how to explain various properties of individual bipolar sunspots, we now turn our attention to providing theoretical explanations for the statistical properties of many sunspots, such as the 22-year cycle and the latitude drift. We now have to get into the heart of dynamo theory to explain these things.

## 5 The turbulent dynamo and mean field MHD

We begin our discussion of dynamo theory by raising the question whether it is possible for motions inside the plasma to sustain a magnetic field. Ideally, one would like to solve (1) and (2) to understand how velocity and magnetic fields interact with each other. Solving these two equations simultaneously in any non-trivial situation is an extremely challenging job. In the early years of dynamo research, one would typically assume a velocity field to be given and

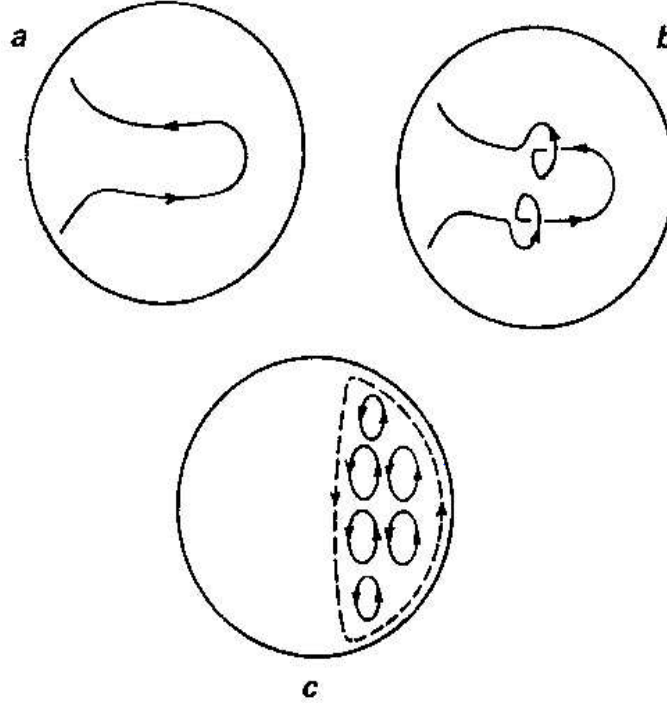


Figure 7: Different stages of the dynamo process. See text for explanation.

then solve (2) to find if this velocity field would sustain a magnetic field. This problem is known as the *kinematic dynamo problem*. One of the first important steps was a negative theorem due to Cowling (1934), which established that an axisymmetric solution is not possible for the kinematic dynamo problem. One is, therefore, forced to look for more complicated, non-axisymmetric solutions.

A major breakthrough occurred when Parker (1955b) realized that turbulent motions inside the solar convection zone (which are by nature non-axisymmetric) may be able to sustain the magnetic field. We have indicated in Figure 4 how a magnetic field line in the poloidal plane may be stretched by the differential rotation to produce a toroidal component. Parker (1955b) pointed out that the uprising hot plasma blobs in the convection zone would rotate, as they rise, because of the Coriolis force of solar rotation (just like cyclones in the Earth's atmosphere) and such helically moving plasma blobs would twist the toroidal field shown in Figure 7a to produce magnetic loops in the poloidal plane as shown in Figure 7b. Keeping in mind that the toroidal field has opposite directions in the two hemispheres and helical motions of convective turbulence should also have opposite helicities in the two hemispheres, we conclude that the poloidal loops in both hemispheres should have the same sense as indicated in Figure 7c.

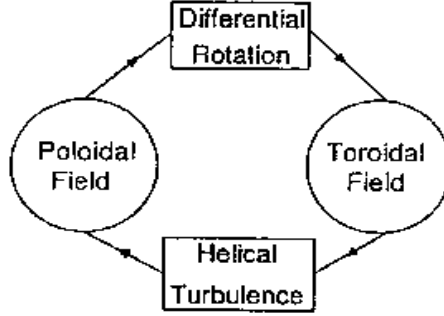


Figure 8: Schematic representation of Parker's idea of the turbulent dynamo.

Although we are in a high magnetic Reynolds number situation and the magnetic field is nearly frozen in the plasma (i.e. molecular resistivity is negligible), we expect the turbulent mixing to give rise to an effective diffusion and the poloidal loops in Figure 7c should eventually coalesce to give the large-scale poloidal field as sketched by the broken line in Figure 7c.

Figure 8 captures the basic idea of Parker's turbulent dynamo. The poloidal and toroidal components of the magnetic field feed each other through a closed loop. The poloidal component is stretched by differential rotation to produce the toroidal component. On the other hand, the helical turbulence acting on the toroidal component gives back the poloidal component. Parker (1955b) developed a heuristic mathematical formalism based on these ideas and showed by mathematical analysis that these ideas worked. However, a more systemic mathematical formulation of these ideas had to await a few years, when Steenbeck, Krause and Rädler (1966) developed what is known as mean field MHD. Some of the basic ideas of mean field MHD are summarized below.

Since we have to deal with a turbulent situation, let us split both the velocity field and the magnetic field into average and fluctuating parts, i.e.

$$\mathbf{v} = \overline{\mathbf{v}} + \mathbf{v}', \quad \mathbf{B} = \overline{\mathbf{B}} + \mathbf{B}'. \quad (8)$$

Here the overline indicates the average and the prime indicates the departure from the average. On substituting (8) in the induction equation (2) and averaging term by term, we obtain

$$\frac{\partial \overline{\mathbf{B}}}{\partial t} = \nabla \times (\overline{\mathbf{v}} \times \overline{\mathbf{B}}) + \nabla \times \mathcal{E} + \lambda \nabla^2 \overline{\mathbf{B}} \quad (9)$$

on remembering that  $\overline{\mathbf{v}'} = \overline{\mathbf{B}'} = 0$ . Here

$$\mathcal{E} = \overline{\mathbf{v}' \times \mathbf{B}'} \quad (10)$$

is known as the *mean e.m.f.* and is the crucial term for dynamo action. This term can be perturbatively evaluated by a scheme known as the *first order*

*smoothing approximation.* Here we shall present only final results without going through derivation and ask readers wishing to learn about first order smoothing approximation to consult §16.5 of Choudhuri (1998), where full derivations can be found. If the turbulence is isotropic, then this approximation scheme leads to

$$\mathcal{E} = \alpha \overline{\mathbf{B}} - \beta \nabla \times \overline{\mathbf{B}}, \quad (11)$$

where

$$\alpha = -\frac{1}{3} \overline{\mathbf{v}' \cdot (\nabla \times \mathbf{v}')} \tau \quad (12)$$

and

$$\beta = \frac{1}{3} \overline{\mathbf{v}' \cdot \mathbf{v}'} \tau. \quad (13)$$

Here  $\tau$  is the correlation time of turbulence. On substituting (11) in (9), we get

$$\frac{\partial \overline{\mathbf{B}}}{\partial t} = \nabla \times (\overline{\mathbf{v}} \times \overline{\mathbf{B}}) + \nabla \times (\alpha \overline{\mathbf{B}}) + (\lambda + \beta) \nabla^2 \overline{\mathbf{B}} \quad (14)$$

It follows from (12) that  $\alpha$  is a measure of average helical motion in the fluid. It is this coefficient which describes the production of the poloidal component from the toroidal component by helical turbulence. This term would go to zero if turbulence has no net average helicity (which would happen in a non-rotating frame). The other coefficient  $\beta$  is of the nature of a diffusion coefficient, as can be seen in (14). Since this is basically diffusion induced by turbulence, it is called *turbulent diffusion*. This is usually much larger than the molecular diffusion  $\lambda$  so that  $\lambda$  can be neglected in (14). I may mention that turbulent diffusion is a phenomenon which we very often encounter in our daily lives. Suppose we put sugar in a cup of coffee and do not stir it. Then sugar will mix with the coffee by molecular diffusion, which is an inefficient process and will take a very long time (the coffee will get cold by that time). On the other hand, if we stir the coffee with a spoon, then we produce some turbulence in the coffee cup, which mixes the sugar much more quickly by turbulent diffusion.

Equation (14) is known as the *dynamo equation* and has to be solved to understand the generation of magnetic field by the dynamo process. A variant of this equation was first derived by rather intuitive arguments in the classic paper of Parker (1955b). The mean field MHD developed by Steenbeck, Krause and Rädler (1966) put this equation on a firmer footing. In the kinematic dynamo approach, one has to specify a velocity field  $\overline{\mathbf{v}}$  and then solve (14). After developing the basic dynamo theory, Parker (1955b) showed that the dynamo equation (14) admits of a periodic wave solution when there is differential rotation. In the next section, we shall present this famous dynamo wave solution. For many years, it was believed that the solar cycle is caused by such a periodic dynamo wave propagating towards the equator inside the convection zone of the Sun. Sunspots were supposed to be produced in the regions where the wave had its crest. The equatorward propagation of the crest provided an explanation why sunspots appear at lower and lower latitudes with the progress of the solar cycle. The period of the solar cycle was supposed to be just the

period of the dynamo wave. The original derivation of dynamo wave by Parker (1955b) reproduced in the next section used Cartesian geometry and could be done analytically. Later Steenbeck and Krause (1969) were the first to solve the dynamo equation (14) numerically in a spherical geometry with appropriate boundary conditions and produced a theoretical butterfly diagram of sunspots. Although for many years this seemed to provide a very elegant explanation of the solar cycle, we now think that the solar dynamo does not operate exactly this way for the following reason. One of the main assumptions in the theoretical model outlined above is that the toroidal field is twisted by the helical turbulence to produce the poloidal field, as sketched in Figure 7b. Since the magnetic tension of the toroidal field would resist any twisting, such a twisting can be possible only if the toroidal field is not too strong. The maximum possible value of the toroidal field was supposed to be the equipartition value given by (7). We already pointed out in §4 that the flux tube rise simulations of D’Silva and Choudhuri (1993), quickly followed by similar works by other groups, established that the toroidal field at the bottom of the solar convection zone has to be of order  $10^5$  G, whereas the equipartition value of magnetic field there is only about  $10^4$  G. The convective turbulence will certainly not be able to twist a toroidal field as strong as  $10^5$  G. We, therefore, think that the model of the dynamo presented in this section is not the appropriate model for the solar dynamo, although this model is probably applicable to dynamo processes in many other astrophysical systems. In §7 we shall come to the question how, according to our current belief, the solar dynamo works.

## 6 Parker’s dynamo wave solution

We now present what is the perhaps the most famous solution of the dynamo equation (14) obtained by Parker in his classic paper (1955b). This solution was originally proposed as a model for the solar cycle. As we already pointed out, we now no longer think that this is the appropriate model for the solar dynamo. However, this solution was historically very influential and still gives us a deep insight into the behaviour of the dynamo.

We saw in §5 that the main idea of Parker’s turbulent dynamo (as summarized in Figure 8) is that the poloidal and toroidal magnetic fields can sustain each other—the toroidal field arises due to the action of differential rotation on the poloidal field and the poloidal field is then generated back from the toroidal field by helical turbulence. We now want to discuss the mathematical representation of these ideas on the basis of the dynamo equation (14). Since we shall be concerned only with mean fields in this section, let us drop the overline sign to denote averages. So  $\mathbf{v}$  and  $\mathbf{B}$  will stand for the mean velocity field and mean magnetic field respectively. Denoting the diffusion coefficient by  $\lambda_T$  with the subscript  $T$  to remind us that it is primarily turbulent diffusion, we write the dynamo equation (14) in the form

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}) + \nabla \times (\alpha \mathbf{B}) + \lambda_T \nabla^2 \mathbf{B}. \quad (15)$$



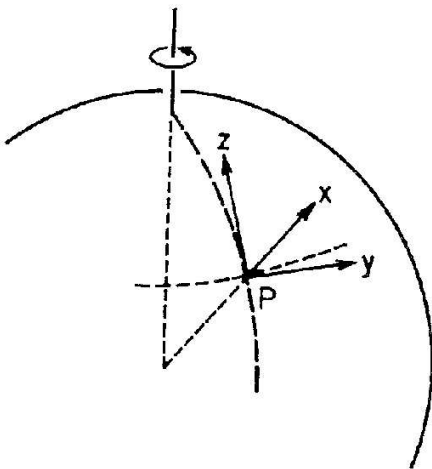


Figure 9: Local Cartesian coordinates at a point in the northern hemisphere of a spherical body.

For applications to the Sun or to the Earth, one has to write down (15) in spherical coordinates. Some of the basic features, however, can be understood quite well by solving (15) in a local Cartesian coordinate system at a point in the spherical object. Figure 9 shows a local Cartesian system at a point in the northern hemisphere of a spherical body like the Sun. The  $x$  axis corresponds to the radially outward direction, the  $y$  axis to the toroidal (i.e.  $\phi$ ) direction and the  $z$  axis to the direction of increasing latitude. Since the statistics of sunspots show a tendency for equatorward migration, we want to find a wave-like solution of (15) propagating in the negative  $z$  direction. As Cowling's theorem is not applicable for mean fields, we expect to find solutions symmetric around the rotation axis. With respect to the local Cartesian system, axisymmetry means symmetry in  $y$ , i.e. we look for solutions for which  $\partial/\partial y = 0$ .

The toroidal magnetic field in our representation is simply the component  $B_y \mathbf{e}_y$ . The poloidal magnetic field is a two-dimensional solenoidal (i.e. zero divergence) vector field with field lines lying in the  $xz$  plane. The most convenient way of representing the poloidal field is to write it in the form  $\nabla \times [A(x, z) \mathbf{e}_y]$ , where  $A(x, z)$  is a function which can be easily shown to be constant on the poloidal field lines in the  $xz$  plane. Hence the full magnetic field can be written as

$$\mathbf{B} = B_y(x, z) \mathbf{e}_y + \nabla \times [A(x, z) \mathbf{e}_y], \quad (16)$$

where  $B_y$  and  $A$  respectively represent the toroidal and the poloidal components. The mean velocity field is due to the differential rotation and hence has a component in the  $y$  direction alone. Although this velocity field  $v_y$  can be a function of both  $x$  and  $z$ , let us take it to be a function of  $x$  alone, since this

results in simpler solutions. So we write

$$\mathbf{v} = v_y(x)\mathbf{e}_y \quad (17)$$

such that the velocity shear is given by

$$G = \frac{\partial v_y}{\partial x}.$$

We shall now show that it is possible to have plane wave solutions propagating in the  $z$  direction, if the velocity shear  $G$  and the coefficients  $\alpha$ ,  $\lambda_T$  are taken as constants.

Substituting (16) and (17) in (15), it is straightforward to show that the  $y$  component of (15) becomes

$$\frac{\partial B_y}{\partial t} = GB_x - \alpha \nabla^2 A + \lambda_T \nabla^2 B_y. \quad (18)$$

Apart from the  $y$  component, the other components of (15) can be put in the form

$$\nabla \times \left( \frac{\partial A}{\partial t} \mathbf{e}_y - \alpha B_y \mathbf{e}_y - \lambda_T \nabla^2 A \mathbf{e}_y \right) = 0.$$

The easiest way to satisfy this equation is to take

$$\frac{\partial A}{\partial t} = \alpha B_y + \lambda_T \nabla^2 A. \quad (19)$$

It is straightforward to see that if the magnetic field is taken to be of the form (16) with  $B_y$  and  $A$  satisfying (18) and (19), then the basic dynamo equation (15) is satisfied. Hence we now have to solve the two simultaneous equations (18) and (19).

Before solving these two equations (18) and (19), let us try to understand their physical significance. The simpler equation (19) gives the evolution of the poloidal field. If the term  $\alpha B_y$  were not there, then it would have been a simple diffusion equation and would imply that any poloidal field diffuses away. The additional term  $\alpha B_y$  acts as the source term which generates the poloidal field. We have seen that  $\alpha$  is a measure of helical motion in the turbulence. Hence this source term corresponds to the production of the poloidal field as a result of helical motions twisting the toroidal field  $B_y$ . The other equation (18) gives the evolution of the toroidal field and has two source terms (the first two terms on the right hand side). The term  $GB_x$  corresponds to the velocity shear of differential rotation stretching out the  $B_x$  component of the poloidal field to produce the toroidal field. Just as the helical motion can twist the toroidal field to produce the poloidal field, it can also twist the poloidal field to produce the toroidal field as well. This is the origin of the other source term  $-\alpha \nabla^2 A$  in (18). If the astrophysical system has strong differential rotation, then the differential rotation term  $GB_x$  is much larger than the other source term  $-\alpha \nabla^2 A$ , and this other term can be neglected. Then (18) becomes

$$\frac{\partial B_y}{\partial t} = -G \frac{\partial A}{\partial z} + \lambda_T \nabla^2 B_y. \quad (20)$$

Equations (19) and (20) together constitute the mathematical representation of Figure 8.

If the term  $-\alpha\nabla^2 A$  in (18) is neglected (as we are doing here), then the dynamo is known as an  $\alpha\Omega$  dynamo. When one uses spherical coordinates, the rotation is usually denoted by  $\Omega$  as we shall see in §7. Hence, for the  $\alpha\Omega$  dynamo equations written in spherical geometry, the source term for the poloidal field involves  $\alpha$  and the source term for the toroidal field involves  $\Omega$ , giving rise to the name  $\alpha\Omega$  dynamo. On the other hand, if the astrophysical system does not have much differential rotation, then the term  $GB_x$  in (18) can be neglected and the other source term  $-\alpha\nabla^2 A$  retained. Such a dynamo is called an  $\alpha^2$  dynamo, because the source terms for both the components involve the  $\alpha$ -coefficient. The Earth's magnetic field is produced by an  $\alpha^2$  dynamo operating in the Earth's core which does not have much differential rotation.

We now wish to show that (19) and (20) admit of propagating wave-like solutions of the type  $\exp[i(\omega t + kz)]$  for constant  $\alpha$ ,  $\lambda_T$  and  $G$ . Let us begin by trying solutions of the form

$$A = \hat{A} \exp(\sigma t + ikz), \quad B_y = \hat{B} \exp(\sigma t + ikz). \quad (21)$$

Substituting in (19) and (20), we have

$$(\sigma + \lambda_T k^2) \hat{A} = \alpha \hat{B}$$

and

$$(\sigma + \lambda_T k^2) \hat{B} = -ikG \hat{A}.$$

Combining these two equations, we get

$$(\sigma + \lambda_T k^2)^2 = -ik\alpha G, \quad (22)$$

from which

$$\sigma = -\lambda_T k^2 \pm \left( \frac{i-1}{\sqrt{2}} \right) \sqrt{k\alpha G}. \quad (23)$$

Let us take  $k$  to be positive. We now separately discuss the two cases of  $\alpha G > 0$  and  $\alpha G < 0$ .

First consider  $\alpha G > 0$ . For the dynamo maintenance of magnetic fields, we must have  $\text{Re } \sigma \geq 0$ . This is possible only if we choose the negative sign in (23) so that

$$\sigma = -\lambda_T k^2 + \left( \frac{k\alpha G}{2} \right)^{1/2} - i \left( \frac{k\alpha G}{2} \right)^{1/2}. \quad (24)$$

This expression makes it clear that the dynamo problem has the character of a stability calculation. If  $\alpha G$ , which gives the combined effect of helical turbulence and differential rotation, is larger than a critical value, then only is it possible for magnetic fields to grow. Otherwise magnetic fields decay away. We introduce a dimensionless parameter called the dynamo number defined as

$$N_d = \frac{|\alpha G|}{\lambda_T^2 k^3}. \quad (25)$$

It is easy to see from (24) that the condition for the dynamo growth of the magnetic field (i.e. for  $\text{Re } \sigma \geq 0$ ) is

$$N_d \geq 2. \quad (26)$$

For the marginally sustained magnetic field (i.e. for  $N_d = 2$ ), the eigenmodes are of the form

$$A, B_y \sim \exp \left[ -i \left( \frac{k\alpha G}{2} \right)^{1/2} t + ikz \right]. \quad (27)$$

This corresponds to a wave propagating in the positive  $z$  direction, i.e. in the poleward direction (see Figure 9).

We now consider the case  $\alpha G < 0$  for which again the negative sign in (23) has to be chosen such that

$$\sigma = -\lambda_T k^2 + \left( \frac{k|\alpha G|}{2} \right)^{1/2} + i \left( \frac{k|\alpha G|}{2} \right)^{1/2}. \quad (28)$$

It is straightforward to see that the condition for dynamo growth (26) remains the same, with  $N_d$  still defined by (25). The marginally sustained eigenmodes, however, now become

$$A, B_y \sim \exp \left[ i \left( \frac{k|\alpha G|}{2} \right)^{1/2} t + ikz \right]. \quad (29)$$

This gives an equatorward propagating wave as desired in the solar context. On this ground, it used to be assumed in the early years of dynamo research that  $\alpha G$  in the northern hemisphere of the Sun must be negative.

Let us recapitulate what determines the direction of the propagation vector. We took  $\mathbf{v}$  along the  $y$  direction with its variation in the  $x$  direction (i.e.  $\partial v_y / \partial x \neq 0$ ). Then the dynamo wave propagates in the third  $z$  direction. It is to be noted that the wavefronts for our solution correspond to infinite planes perpendicular to the  $z$  axis. This is because we took  $\alpha$ ,  $\lambda_T$  and  $G$  as constants. In a realistic situation, one has to solve the dynamo equation in a finite region with suitable boundary conditions on the boundaries. We shall not get into a discussion of boundary conditions here, because they make the problem considerably more complicated. We hope that the plane wave solution for constant coefficients gives some idea on how to approach dynamo problems.

## 7 The Babcock–Leighton mechanism and flux transport dynamo

Although the very elegant dynamo wave solution presented in the previous section seemed to provide a natural explanation for the periodicity and equatorward propagation associated with solar cycle, we have already pointed out at the end of §5 that eventually it was realized that this is probably not the correct model

for the solar dynamo. If the toroidal field at the bottom of the convection zone is of order  $10^5$  as concluded by D'Silva and Choudhuri (1993) and others, then helical turbulence would not be able to twist the toroidal field as required in the dynamo model presented in §5–6. If it is not possible for helical turbulence to produce the poloidal field as indicated in Figure 7, then we need an alternative mechanism for the production of the poloidal field. In fact, an alternative mechanism for poloidal field generation was proposed by Babcock (1961) and Leighton (1969) several decades ago. Within the last few years, several dynamo theorists have invoked this Babcock–Leighton mechanism for generating the poloidal field in view of the difficulty in twisting a very strong toroidal field by helical turbulence.

Let us now explain what this Babcock–Leighton mechanism is. We pointed out in §2 that bipolar sunspots have tilts increasing with latitude, in accordance with Joy's law. Then we discussed in §4 how this law was explained by D'Silva and Choudhuri (1993) by considering the action of the Coriolis force on rising flux tubes. Now, a typical sunspot lives for a few days and the magnetic field of the sunspot diffuses in the surrounding region by turbulent diffusion after its decay. When a tilted bipolar sunspot pair with the leading spot nearer the equator and the following spot at a higher latitude decays, the polarity of the leading sunspot gets more diffused in the lower latitudes and the polarity of the following sunspot gets more diffused in the higher latitudes. This essentially gives rise to a poloidal field at the solar surface. Since sunspots form from the toroidal field due to magnetic buoyancy, a tilted bipolar sunspot pair can be viewed as a conduit through which a part of the toroidal field ultimately gets transformed into the poloidal field. This is the basic idea of poloidal field generation proposed by Babcock (1961) and Leighton (1969). If the toroidal field is so strong that helical turbulence cannot twist it as required in the dynamo model discussed in §5–6, then the Babcock–Leighton mechanism seems like the natural option for the generation of the poloidal field.

We mentioned in §2 that weak, diffuse magnetic field of the Sun outside sunspots migrates poleward with progress of the solar cycle. It is known for about two decades that the plasma near the Sun's surface continuously flows from the equator to the pole, with a maximum flow speed of about  $20 \text{ m s}^{-1}$  in the mid-latitudes. This is known as the *meridional circulation*. By assuming that the weak, diffuse magnetic field arises from the decay of tilted bipolar sunspots by the Babcock–Leighton mechanism and is then carried by the meridional circulation, Wang, Nash and Sheeley (1989) could model the evolution of the weak, diffuse field on the Sun's surface. Now the material brought from the equatorial region to the polar region by the meridional circulation somehow has to be returned back, if we do not want a gradual pile-up of additional material near the poles. We believe that there is a counter-flowing part of the meridional circulation underneath the Sun's surface bringing back the material from the polar region to the equatorial region. Although there is so far no direct observational evidence that such a counter-flow really exists underneath the Sun's surface, it has to be there to avoid the pile-up of material near the poles. It is usually believed that the meridional circulation is driven by the turbulent

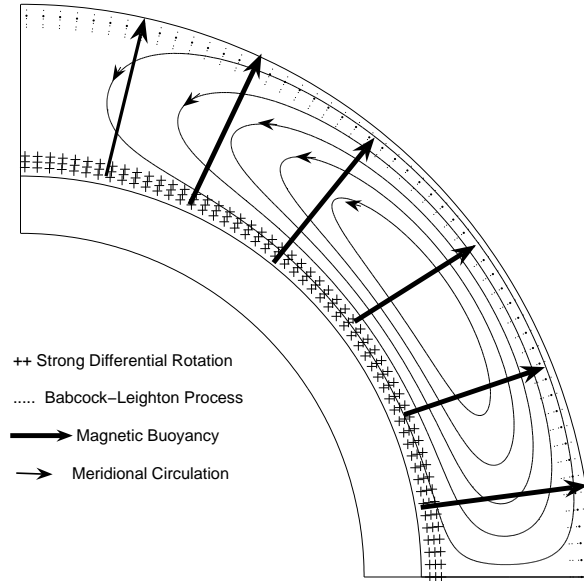


Figure 10: A cartoon explaining how the solar dynamo works.

stresses in the convection zone, although we do not yet have a very satisfactory theory of the meridional circulation. It is then natural to assume that the counter-flow will be in the bottom layers of the convection zone. Figure 10 indicates how the streamlines of meridional circulation are expected to look like.

Since the meridional circulation sinks underneath the Sun's surface near the pole, the poloidal magnetic field, after being brought to the pole, must be carried downward and brought to the bottom of the convection zone by the meridional circulation. We have repeatedly referred to the famous helioseismology result that the differential rotation is concentrated in the tachocline at the bottom of the convection zone. Once the poloidal field is brought to the tachocline, the differential rotation can act on it and produce the strong toroidal field. We have discussed in §4 that tubes of toroidal field produced at the bottom of the convection zone can rise in parts due to magnetic buoyancy to produce the sunspots. All these ideas are put together in a cartoon form in Figure 10, explaining how the solar dynamo operates. The toroidal field is produced in the tachocline by the differential rotation stretching out the poloidal field. Then this toroidal field rises due to magnetic buoyancy to produce bipolar sunspots at the solar surface, where the poloidal field is generated by the Babcock–Leighton mechanism from these bipolar sunspots. The poloidal field so generated is carried by the meridional circulation first to the polar region and then underneath the surface to the tachocline to be stretched by the differential rotation—thus completing the cycle. This type of dynamo is called a *flux transport dynamo*. Most of the dynamo theorists at the present time believe that the solar dynamo operates in

this way. Wang, Sheeley and Nash (1991) proposed the idea of the flux transport dynamo. Choudhuri, Schüssler and Dikpati (1995) and Durney (1995) were the first to construct two-dimensional models of the flux transport dynamo to demonstrate that such a dynamo really does work.

How do we treat the flux transport dynamo mathematically? We can still begin with the basic dynamo equation (15). In the original derivation for this equation sketched in §5, the coefficient  $\alpha$  was supposed to be a measure of helical turbulence in the system as evident from (12). It follows from (19) that  $\alpha$  finally appears in the theory as a measure of poloidal field generation from the toroidal field. Even the Babcock–Leighton process can be phenomenologically described with the help of the coefficient  $\alpha$  concentrated near the solar surface. Since this  $\alpha$  in the Babcock–Leighton process does not arise from helical turbulence, it will no longer be given by (12), although it appears in the dynamo equation in exactly the same form. Using spherical coordinates, we write the magnetic field as

$$\mathbf{B} = B(r, \theta)\mathbf{e}_\phi + \nabla \times [A(r, \theta)\mathbf{e}_\phi], \quad (30)$$

which has a form very similar to (16) for Cartesian geometry. Here again  $B(r, \theta)$  is the toroidal component and  $A(r, \theta)$  represents the poloidal component. We can write the velocity field as  $\mathbf{v} + r \sin \theta \Omega(r, \theta)\mathbf{e}_\phi$ , where  $\Omega(r, \theta)$  is the angular velocity in the interior of the Sun and  $\mathbf{v}$  is the velocity of meridional circulation having components in  $r$  and  $\theta$  directions. On substituting these forms of magnetic and velocity fields in (15), it can be shown that the poloidal and toroidal components of an  $\alpha\Omega$  dynamo evolve according to the following equations

$$\frac{\partial A}{\partial t} + \frac{1}{s}(\mathbf{v} \cdot \nabla)(sA) = \lambda_T \left( \nabla^2 - \frac{1}{s^2} \right) A + \alpha B, \quad (31)$$

$$\frac{\partial B}{\partial t} + \frac{1}{r} \left[ \frac{\partial}{\partial r}(rv_r B) + \frac{\partial}{\partial \theta}(v_\theta B) \right] = \lambda_T \left( \nabla^2 - \frac{1}{s^2} \right) B + s(\mathbf{B}_p \cdot \nabla)\Omega + \frac{1}{r} \frac{d\lambda_T}{dr} \frac{\partial}{\partial r}(rB), \quad (32)$$

where  $s = r \sin \theta$ . The dynamo model indicated in the cartoon form in Figure 10 is achieved by using the helioseismically determined  $\Omega$  which has a strong gradient at the bottom of the convection zone, by choosing a meridional circulation  $\mathbf{v}$  as indicated in the Figure 10 and by making  $\alpha$  concentrated near the solar surface to represent the Babcock–Leighton process. Additionally, we have to specify magnetic buoyancy in some suitable fashion. We have to solve the coupled partial differential equations (31) and (32) numerically, since no non-trivial calculations are possible analytically. Our research group in Bangalore has developed a numerical code *Surya* for solving the flux transport dynamo problem. I can send the code *Surya* and a detailed guide for using it (Choudhuri, 2005) to anybody who sends a request to my e-mail address [arnab@physics.iisc.ernet.in](mailto:arnab@physics.iisc.ernet.in).

As we pointed out in §5–6, it was believed in the early years of dynamo research that Parker’s dynamo wave provides the explanation for the solar cycle. We saw in §6 that the condition for equatorward propagation of such a wave is  $\alpha G < 0$  in the northern hemisphere. Not much was known about the nature of differential rotation  $G$  before the development of helioseismology. In the case of

$\alpha$  which is supposed to be given by (12), not only its value, but even its sign was uncertain! Dynamo theorists of earlier times used to make assumptions about  $G$  and  $\alpha$  which seemed reasonable at that time (but does not appear reasonable now), making sure that their product was negative in the norther hemisphere. One of the attractive aspects of current flux transport dynamo models is that both the toroidal and poloidal field generation mechanisms are directly based on observations. The angular velocity  $\Omega$  is now completely mapped by helioseismology as discussed by Antia. So there is now no uncertainty in the toroidal field generation term in the dynamo equation (the middle term in the right hand side of (32)). On the other hand, we directly observe the poloidal field generation by the Babcock–Leighton process on the Sun’s surface. When we directly observe the basic processes, we do not have the freedom of choosing our parameters arbitrarily. It appears that both  $\alpha$  and  $G$  are positive in the regions of northern hemisphere where sunspots are seen. That is why it was initially not clear whether a flux transport model would be able to explain even such a basic observation as the equatorward drift of sunspots with the solar cycle. Choudhuri, Schüssler and Dikpati (1995) succeeded in settling this issue. Since the toroidal field from which sunspots form is produced in the tachocline at the bottom of the convection zone where the meridional circulation is equatorward (as seen in Figure 10), Choudhuri, Schüssler and Dikpati (1995) showed that a sufficiently strong meridional circulation can force an equatorward propagation of the toroidal field even if  $\alpha G$  is positive. Thus the condition that  $\alpha G$  has to be negative in the northern hemisphere, which was regarded almost sacrosanct for several decades, is no longer applicable in the presence of meridional circulation. With this demonstration by Choudhuri, Schüssler and Dikpati (1995), it became clear that the flux transport dynamo provides an attractive model for the solar cycle.

Dikpati and Charbonneau (1999), Küker, Rüdiger and Schultz (2001) and Nandy and Choudhuri (2002) were amongst the first to construct models of flux transport dynamo using the angular velocity distribution mapped by helioseismology. Many of these initial models were of exploratory nature and did not always fit different aspects of observational data. However, as dynamo theorists kept gaining more experience of building models, it has eventually been possible to construct models with such combinations of parameters that most aspects of observational data are reproduced reasonably well. Figure 11 taken from Chatterjee, Nandy and Choudhuri (2004) shows a theoretically calculated butterfly diagram of sunspots superposed on contours of constant  $B_r$  at the surface in a time-latitude plot. This figure has to be compared with the observational Figure 2. Given the fact that this was one of the first serious attempts of reproducing this observational figure, perhaps most of you will agree that it is not a bad first theoretical fit.



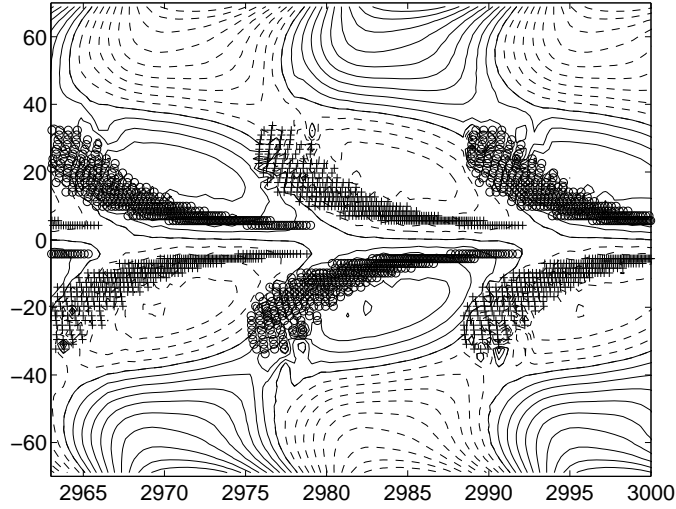


Figure 11: A theoretical butterfly diagram of sunspots superposed on contours of constant  $B_r$  at the solar surface in a time-latitude plot. This figure is taken from Chatterjee, Nandy and Choudhuri (2004).

## 8 Conclusion

I hope that my presentation has given you some idea where solar dynamo theory stands now. We have come a long way from Parker's classic paper published slightly more than half a century ago (Parker, 1955b). We saw that the dynamo wave solution derived in that paper was supposed to provide a model for the solar cycle for a very long time. When flux tube simulations pointed out that the toroidal field at the bottom of the convection zone is probably very strong and cannot be easily twisted by helical turbulence, this model had to be abandoned. Within the last few years, the flux transport dynamo model has emerged as the most promising model for the generation of solar magnetic cycle. I pointed out that many aspects of the flux transport dynamo are based on direct observations. The differential rotation, which generates the toroidal field, is known from helioseismology. On the other hand, we directly see the generation of the poloidal field by the Babcock–Leighton mechanism on the solar surface. However, it should be kept in mind that some aspects of the flux transport dynamo are still rather uncertain. Apart from the reasonable theoretical guess that the streamlines of meridional circulation close at the bottom of the convection zone with a flow towards the equator, we do not know anything more definite about this crucial component of the flux transport dynamo. If a certain combination of parameters gives theoretical results matching observational data, that does not necessarily mean that the combination chosen by us is actually the 'correct' combination. See Choudhuri (2007) for a discussion of the still remaining uncertainties in solar dynamo models.

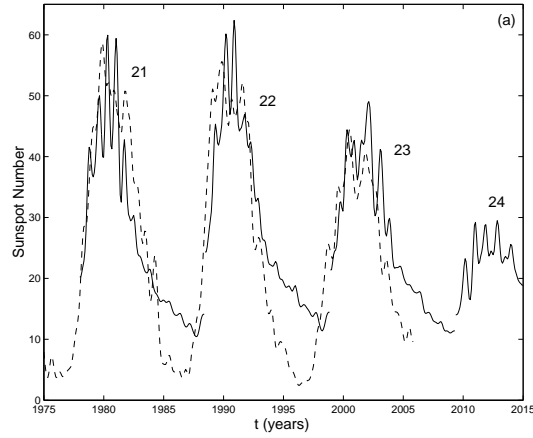


Figure 12: The theoretical monthly sunspot number for the last few years as well as the upcoming next cycle, plotted along with the observational data (dashed line) for the last few years. This figure is taken from Choudhuri, Chatterjee and Jiang (2007).

I should also remind you that there are many important aspects of the solar dynamo problem which I could not even touch upon in my lectures. A look at Figure 3 should convince you that the solar cycle is only roughly periodic. Many irregularities are found superposed on the underlying periodic behaviour. In my lectures, I have mainly focused on the theoretical models of regular periodic behaviour of the solar dynamo and have not discussed the irregularities, since our theoretical understanding of these irregularities is still very primitive. We saw in §5 that we have to average over fluctuations to get the mean field dynamo equation. Fluctuations present around the mean provide an additional random force acting on the periodic dynamo. Choudhuri (1992) suggested this as the main cause behind irregularities of the solar cycle and that seems to be the most satisfactory explanation. One question connected with the study of irregularities is whether we can model individual cycles and can predict the strength of a particular cycle before its onset. For example, the next maximum of solar activity is scheduled around 2010–11. Whether this forthcoming maximum will be strong or weak is being debated. Since theoretical calculations alone cannot tell us about the nature of stochastic fluctuations associated with a particular cycle, we have feed some observational data into the theoretical model in a suitable way when we want to model individual cycles on the basis of dynamo theory. Only recently research has begun on this subject. Dikpati and Gilman (2006) have predicted that the next solar cycle will be one of the strongest in a long time. On the other, the calculations of Choudhuri, Chatterjee and Jiang (2007) based on a different model using a different methodology suggest that the next cycle will be one of the weakest. Figure 12 taken from Choudhuri, Chatterjee and Jiang (2007) shows the theoretically calculated sunspot number

for the last few cycles along with observational data. It remains to be seen whether the theoretical prediction for the next few years indicated in Figure 12 will be borne out or not.

While we now seem to have a broad-brush understanding of how the solar dynamo works, many pieces of the jigsaw puzzle are still missing. It is unlikely that theoretical research alone will supply all the missing pieces. Perhaps some hitherto unforeseen kinds of observational data will come to our help in completing the picture. We certainly expect solar dynamo theory to remain an exciting research field for years to come.

## References

- [1] Alfvén, H. (1942). *Ark. f. Mat. Astr. o. Fysik* **29B**, No. 2.
- [2] Babcock, H. D. (1959). *Astrophys. J.* **130**, 364.
- [3] Babcock, H. W. (1961). *Astrophys. J.* **133**, 572.
- [4] Babcock, H. W. and Babcock, H. D. (1955). *Astrophys. J.* **121**, 349.
- [5] Bumba, V. and Howard, R. (1965). *Astrophys. J.* **141**, 1502.
- [6] Caligari, P., Moreno-Insertis, F. and Schüssler, M. (1995). *Astrophys. J.* **441**, 886.
- [7] Carrington, R. C. (1858). *Mon. Notic. Roy. Astron. Soc.* **19**, 1.
- [8] Chandrasekhar, S. (1952). *Phil. Mag. (7)* **43**, 501.
- [9] Chatterjee, P., Nandy, D. and Choudhuri, A. R. (2004). *Astron. Astrophys.* **427**, 1019.
- [10] Choudhuri, A. R. (1989). *Solar Phys.* **123**, 217.
- [11] Choudhuri, A. R. (1990). *Astron. Astrophys.* **239**, 335.
- [12] Choudhuri, A. R. (1992). *Astron. Astrophys.* **253**, 277.
- [13] Choudhuri, A. R. (1998). *The Physics of Fluids and Plasmas: An Introduction for Astrophysicists* (Cambridge University Press).
- [14] Choudhuri, A. R. (2005) *The User's Guide to the Solar Dynamo Code SURYA* (available upon request).
- [15] Choudhuri, A. R. (2007). *Adv. Space Res.* (in press).
- [16] Choudhuri, A. R., Chatterjee, P. and Jiang, J. (2007). *Phys. Rev. Lett.* (in press).
- [17] Choudhuri, A. R. and Gilman, P. A. (1987). *Astrophys. J.* **316**, 788.

- [18] Choudhuri, A. R., Schüssler, M. and Dikpati, M. (1995). *Astron. Astrophys.* **303**, L29.
- [19] Cowling, T. G. (1934). *Mon. Notic. Roy. Astron. Soc.* **94**, 39.
- [20] Dikpati, M. and Charbonneau, P. (1999). *Astrophys. J.* **518**, 508.
- [21] Dikpati, M. and Gilman, P. A. (2006). *Astrophys. J.* **649**, 498.
- [22] D'Silva, S. and Choudhuri, A. R. (1993). *Astron. Astrophys.* **272**, 621.
- [23] Durney, B. R. (1995). *Solar Phys.* **160**, 213.
- [24] Fan, Y., Fisher, G. H. and DeLuca, E. E. (1993). *Astrophys. J.* **405**, 390.
- [25] Gailitis, A., *et al.* (2000). *Phys. Rev. Lett.* **84**, 4365.
- [26] Hale, G. E. (1908). *Astrophys. J.* **28**, 315.
- [27] Hale, G. E., Ellerman, F., Nicholson, S. B. and Joy, A. H. (1919). *Astrophys. J.* **49**, 153.
- [28] Howard, R. F. (1996). *Solar Phys.* **169**, 293.
- [29] Howard, R. F. and LaBonte, B. J. (1981). *Solar Phys.* **74**, 131.
- [30] Küker, M., Rüdiger, G. and Schultz, M. (2001). *Astron. Astrophys.* **374**, 301.
- [31] Leighton, R. B. (1969). *Astrophys. J.* **156**, 1.
- [32] Longcope, D. and Choudhuri, A. R. (2002). *Solar Phys.* **205**, 63.
- [33] Makarov, V. I. and Sivaraman, K. R. (1989). *Solar Phys.* **119**, 35.
- [34] Maunder, E. W. (1904). *Mon. Notic. Roy. Astron. Soc.* **64**, 747.
- [35] Nandy, D. and Choudhuri, A. R. (2002). *Science* **296**, 1671.
- [36] Parker, E. N. (1955a). *Astrophys. J.* **121**, 491.
- [37] Parker, E. N. (1955b). *Astrophys. J.* **122**, 293.
- [38] Parker, E. N. (1979). *Cosmical Magnetic Fields* (Oxford University Press).
- [39] Schwabe, S. H. (1844). *Astron. Nachr.* **21**, 233.
- [40] Spruit, H. C. (1981). *Astron. Astrophys.* **98**, 155.
- [41] Steenbeck, M. and Krause, F. (1969). *Astron. Nachr.* **291**, 49.
- [42] Steenbeck, M., Krause, F. and Rädler, K.-H. (1966). *Z. Naturforsch.* **21a**, 1285.
- [43] Stenflo, J. O. (1973). *Solar Phys.* **32**, 41.

- [44] Wang, Y.-M., Sheeley, N. R. and Nash, A. G. (1991). *Astrophys. J.* **383**, 431.
- [45] Wang, Y.-M., Nash, A. G. and Sheeley, N. R. (1989). *Astrophys. J.* **347**, 529.
- [46] Weiss, N. O. (1981). *J. Fluid Mech.* **108**, 247.