

DYNAMO ACTION AND THE SUN

Michael Proctor¹

Abstract. Solar magnetic fields are produced by dynamo action in or just below the solar convection zone. After an account of observations of solar magnetic activity, the fundamentals of dynamo theory are reviewed, and the ideas applied to a discussion of different models of dynamo action in the Sun.

1 Introduction

The Sun's magnetic field appears in many different manifestations. At the largest scales there are sunspots and pores, and their associated active regions, extending over a significant fraction of the Sun's surface. These features form part of a coherent cyclical pattern, and must be explained by some global process. At intermediate scales there are "plages" of significant though intermittent field associated with active regions. Away from active regions there are the network fields that appear near the edges of the supergranulation, the deeper seated and larger scale convection system that underlies the granules. On these smaller scales there is evidence of a dynamo process taking place, which is disorganised rather than global and which works in an entirely different manner. In this review I shall discuss solar observations that indicate the existence of dynamo action, and then describe some theoretical ideas that go some way to explaining the observations. The review is intended to be pedagogical in nature and emphasises basic ideas. For more advanced reviews of the solar dynamo see Weiss(1994), Ossendrijver(2003) and Tobias(2005).

2 Indicators of Solar Magnetic Activity

2.1 Sunspots

The most visible sign that the Sun has magnetic activity is the presence of sunspots on the surface. These have been seen since ancient times, but were first systemati-

¹ DAMTP, University of Cambridge
Wilberforce Road, Cambridge CB3 0WA, UK

cally observed by Galileo, who noted that individual spots emerged in mid latitudes and rotated at a rate dependent upon latitude. He also recognised that they had a finite lifetime of the order of one month. (Figure 1(a)). At this stage of course there was no link to any magnetic properties. After Galileo's time, sunspot observations became more systematic, and it was observed that the locations of sunspot appearance had a cyclical component, with the median latitude decreasing, in time over an 11y period, with a new cycle then beginning at higher latitudes. This leads to the famous "butterfly diagram" (Figure 2).

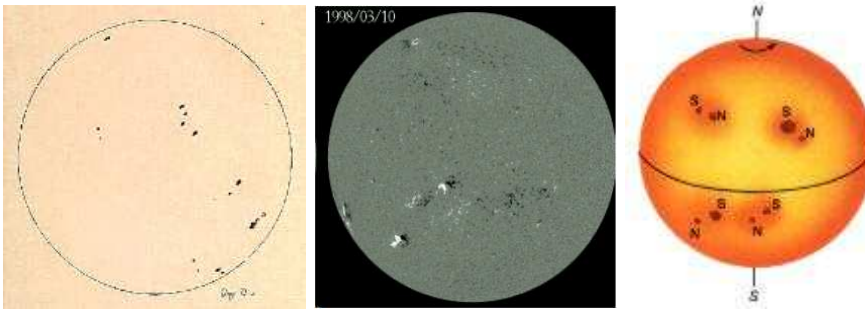


Fig. 1. (a) An early drawing by Galileo of sunspots. The solar rotation axis points up and to the right. (b) A magnetogram of the Sun (Kitt Peak National Observatory). Black and white colorations show positive and negative polarity of the line of sight field. (c) Schematic showing the tilt of active regions.

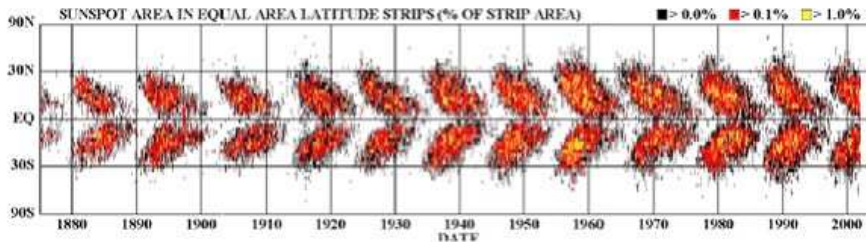


Fig. 2. The butterfly diagram. Shown are incidences of sunspots with latitude and time (Courtesy of D.Hathaway)

This diagram is very revealing. While it shows clearly that there is some sort of cyclical effect, with near symmetry between the two hemispheres, the symmetry is not exact, and furthermore there are obvious modulations between cycles in the

range of altitudes at which sunspots appear. Continuation of the record back in time shows that there have been periods of very low sunspot activity, such as the so-called Maunder Minimum in the 17th century. Modulations of the cycle are discussed below.

Much more recently, it has become possible to measure magnetic fields on the Sun, and it was quickly discovered that sunspots are associated with large line-of-sight magnetic fields, of $O(3000G)$. Hale's pioneering observations showed that sunspots are associated with (magnetically) active regions, which typically appear in pairs of opposite polarity, with the leading spot/region (measuring in the direction of rotation) having different polarities in the N. and S. hemispheres at any particular point in the cycle. This polarity is reversed every other cycle, so that the approximate period of cyclic activity is really 22y. A further interesting observation is that the two parts of each region or sunspot are tilted by $\sim 5 - 10^\circ$ with respect to the equator so that the leading spot is closer to the equator (Figure 1(c)). This systematic effect has its origin in the dynamics of the magnetic fields, as discussed below. A discussion of the fine structure of sunspots is beyond the scope of this article: for a comprehensive review see Thomas & Weiss(2004). It is important, however, to note that not all active regions have a fully developed sunspot with a central dark umbra and a filamentary outer penumbra. Smaller umbral type structures without penumbrae (pores) can also appear, typically with smaller fluxes, and sometimes there are no spots at all. We conclude that the sunspot number can be a sensitive function of the flux emerging in active regions.

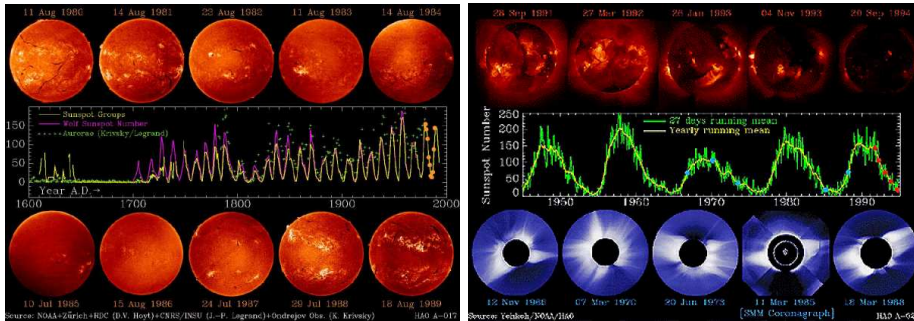


Fig. 3. Various measures of the solar cycle, compiled by NASA. Shown are several measures of sunspot activity, together with observations of the corona, and two different types of X-ray emission, at time points marked by circles on the sunspot curves. It can be seen that the X-ray activity is greatly reduced at sunspot minimum.

2.2 Other indicators of the cycle

The cycle can be detected in other ways than the sunspot record. The NASA compilation pictures shown in Figure 3 illustrate a number of different measures

of magnetic activity. Among these are coronal observations, and X-ray emissions, which is a diagnostic for magnetic activity higher up in the solar atmosphere than the photosphere. All these measures show the 11y cycle and the progression towards the equator.

The (weak) poloidal field of the Sun can also be measured. It is generally dipolar, though by no means dominantly so, with magnitude of only a few Gauss. The dipole moment reverses approximately every 11y, but with a different phase than the sunspot cycle. This, together with Hale's law indicates a global magnetic field reversal every 11y.

2.3 Modulation of the cycle

As previously mentioned, the sunspot cycle, and its other indicators, are far from regular and show modulation in amplitude, and to a lesser extent frequency, on a variety of timescales. The most obvious period of modulation has seen one almost total cessation of sunspot activity between about 1650 and 1700 (the Maunder Minimum). This appears to be part of a succession of modulations with a timescale of the order of 200+y, known as grand minima. Other frequencies can be detected; there is a prominent contribution from a period of order 88y (the Gleissberg cycle). The Maunder Minimum was associated with cold weather in Northern Europe (the Little Ice Age) (Figure 4), suggesting a relation between sunspot activity and solar activity in general. A historical account can be found in Eddy(1976).



Fig. 4. The Little Ice Age: painting by Hondius of a Frost Fair on the Thames in the 17th Century.

Before about 1600 there were no direct observations of sunspots. However it is possible to use secondary data, believed to be well correlated with magnetic activity, to continue observations into the distant past. The majority of information is provided by the incidence of ^{14}C in tree rings, and ^{10}Be in Arctic ice cores. These isotopes are radioactive, and decay at a predictable rate, so their incidence when the rings or ice were formed can be estimated. Any anomalies can be attributed to the level of cosmic ray activity, which is lower when the Sun is at its most active since the radiation belts round the Earth are more substantial and let fewer cosmic rays through. Thus the isotope abundances are inversely correlated with magnetic activity. These measures can be continued back many thousands of years, and give strong evidence of modulations of activity on the scales mentioned above, as shown in Figure 5. Interestingly, detailed examination of the Maunder Minimum period indicates that some sort of cyclic activity continues throughout the period when no sunspots were seen. See for example Beer et al.(1998).

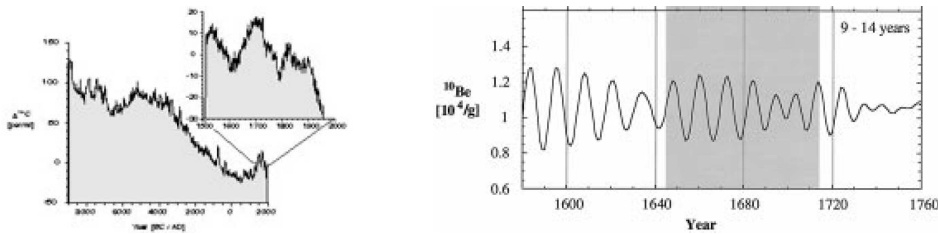


Fig. 5. (a) ^{14}C anomalies for the last 8000y, showing several “grand minima”, the last well correlated (inversely) with the sunspot record.(from Beer(2000) (b) Filtered ^{10}Be data showing a clear 11y period throughout the Maunder Minimum. From Beer et al.(1998).

2.4 Magnetic activity on other stars

The Sun is not the only body to manifest magnetic activity. By observing Ca IIR emission profiles, closely correlated with magnetic activity, the Mount Wilson Survey has revealed cyclical behaviour in a large number of late type stars (Figure 6).

While there is large variation in the behaviour, the general trend seems to be that cyclical behaviour is less periodic and more disordered for higher stellar rotation rates. It would seem therefore that if these cycles are caused by the same mechanism as in the Sun then rotation plays an important role (as I shall confirm below).

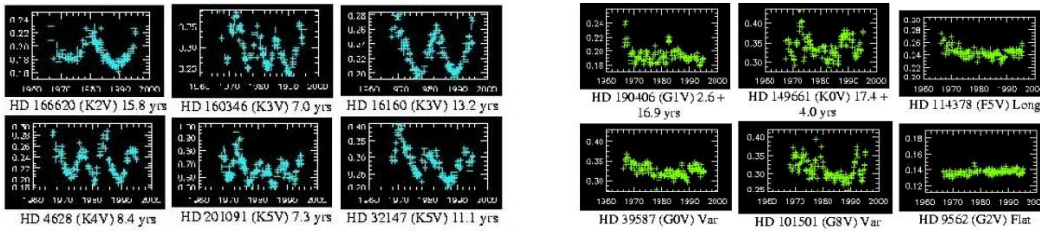


Fig. 6. Stellar cycles from the Mount Wilson Survey. Higher rotation rates correspond to less ordered cycles

3 The Solar convection zone

If the solar cycle is generated within the Sun, then it is important to understand its structure. The inner radiative core is stably stratified, and while it may be the seat of strong magnetic fields, these are unlikely to play a direct role in the solar cycle. The outer 30% of the solar atmosphere, the convection zone or CZ, is superadiabatic, and is in vigorous turbulent motion. The observed surface manifestations of this convection are the solar granulation and supergranulation, but in the interior the motion is probably dominated by sinking plumes of cool plasma, with broad warm upflows. The interaction of rotation with this convection produces the observed solar differential rotation, which can now be detected at depth due to helioseismology. The profile of the differential rotation with latitude and depth is shown in Figure 7.

It can be seen that there is equatorial acceleration, which depends only weakly on radius until the base of the convection zone is reached, where it changes rapidly to an almost solid body rotation. The profile is essentially symmetric about the equator. A further diagnostic of cyclical behaviour can be found in the small modulations of the differential rotation. Figure 8 shows the behaviour of these modulations (“torsional waves”), which take the form of waves propagating towards the equator at lower latitudes and polewards at higher values. Their period is about 11y. It will be seen that these waves are a natural consequence of the mechanism responsible for the cycle.

4 The dynamo process

4.1 The induction equation and the Lorentz force

It is generally agreed that the magnetic activity observed on the Sun is due to *dynamo action* – the process by which kinetic energy is converted into magnetic energy by Faraday induction. We use Faraday’s equation $\dot{\mathbf{B}} = -\nabla \times \mathbf{E}$ to relate

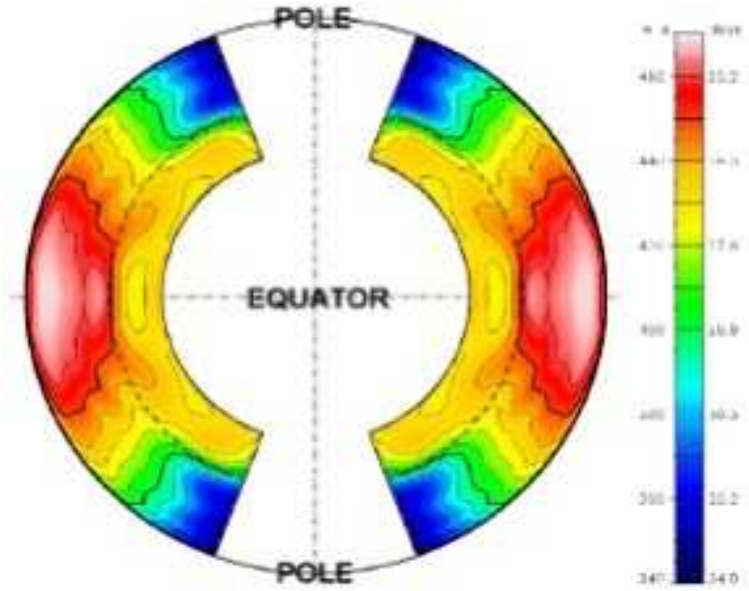


Fig. 7. A profile of the solar differential rotation, deduced from helioseismology

electric field to rate of change of flux, and use Ohm's Law in a moving medium $\mathbf{j} = \sigma(\mathbf{E} + \mathbf{U} \times \mathbf{B})$ to relate electric field to current. Finally we use Ampère's law $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$, neglecting displacement currents since we are considering times much longer than the electromagnetic crossing time. Ohm's law, unlike the others, is phenomenological, and certainly not correct near the top of the solar atmosphere, but it is sufficient for modelling purposes. Eliminating \mathbf{E} , we obtain the *induction equation*

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{U} \times \mathbf{B}) (\text{Induction}) - \nabla \times (\eta \nabla \times \mathbf{B}) (\text{Diffusion}); \quad \nabla \cdot \mathbf{B} = 0; \quad \eta = (\mu_0 \sigma)^{-1} \quad (4.1)$$

The quantity η is the magnetic diffusivity. If \mathcal{U}, \mathcal{L} are typical velocity and length scales, then we can define two time scales $\tau_A = \mathcal{L}/\mathcal{U}$ (advective time) and $\tau_D = \mathcal{L}^2/\eta$ (diffusion time). Their ratio is given by the magnetic Reynolds number $R_m = \tau_D/\tau_A = \mathcal{U}\mathcal{L}/\eta$.

The induction equation gives the *kinematics* of the field. The *dynamics* of the field is given by the Lorentz force

$$\mathbf{j} \times \mathbf{B} = \mu_0^{-1} \left(-\nabla \left(\frac{1}{2} |\mathbf{B}|^2 \right) + (\mathbf{B} \cdot \nabla) \mathbf{B} \right) \quad (4.2)$$

This body force is quadratic in the field strength. The second form in Equation (4.2) shows that the force can be divided into magnetic pressure, leading to a reduction in gas pressure in strong field regions, and a curvature force, which

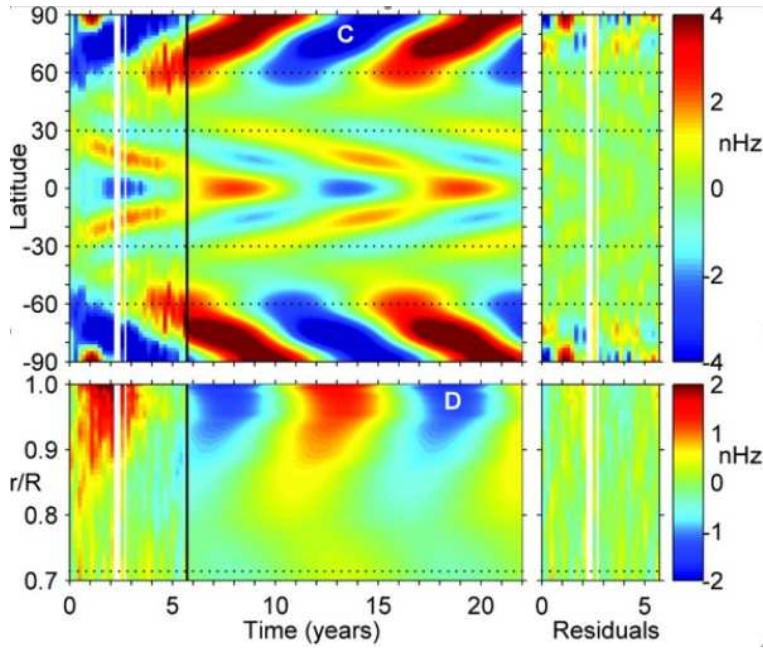


Fig. 8. Torsional waves deduced from helioseismology. Upper figure, surface anomaly as a function of latitude and time. Lower figure: behaviour with depth at fixed latitude. The right hand parts of the figures are extrapolated. Figure taken from Vorontsov et al.(2002)

acts to straighten magnetic field lines. The magnetic pressure is important in evacuating regions of strong field, which can lead to significant buoyancy effects, an important component in some scenarios of the cycle as discussed below.

Returning to the induction equation, we note that if R_m is sufficiently large, diffusion can be ignored except when the field has very small length scales. It can then be shown that field lines move with the fluid (Alfvén's theorem). In a finite domain, field lines will be stretched and folded. This will increase the magnetic energy. Think of a cylinder of magnetic field lines, and deform it so that it is twice its length. The flux $\int \mathbf{B} \cdot d\mathbf{S}$ through the cylinder is unchanged, so the field strength doubles. Thus the magnetic energy is quadrupled. In order actually to increase the flux in a closed system, the field lines must be folded. The Stretch-Twist-Fold dynamo of Vainshtein and Zeldovich, shown in Figure 9, in which a circular tube of flux is stretched and then twisted and folded so as to produce a tube with the same shape but twice the original flux. Repeated indefinitely, this will lead to an exponential increase in magnetic energy. The role of diffusion is subtle here, since flux can only increase through a finite cross-section if diffusion is non-zero. In a finite domain, stretching is always accompanied by folding, and this can lead to very short scales of variation of the field, enhancing the effects of diffusion. It is not obvious that even at large values of R_m the effects of folding outweigh those of

stretching. In fact if all fields and flows are confined to a plane it can be shown that the folding and subsequent diffusion always win, and all fields ultimately decay, even for very large R_m . Figure 9 also shows the effects of persistent folding in a model flow.

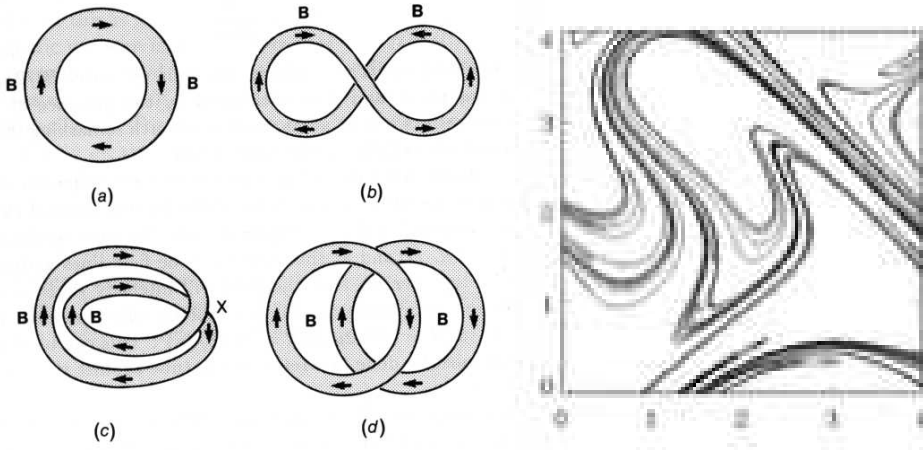


Fig. 9. (a) Diagrammatic illustration of the STF mechanism, from Roberts(1994) (b) An example of highly folded fields due to stretching and twisting at large R_m

4.2 Necessary conditions

The problem of finding dynamo action is harder than it might seem since various simplifications render the dynamo ineffective. Here I briefly mention two different kinds of result, Cowling's theorem and Backus' necessary condition. These considerations are covered in more detail in Moffatt(1978).

Cowling's theorem. No axisymmetric magnetic field can be sustained by dynamo action. This can only be possible anyway when the velocity field is axisymmetric. We can see the problem by expressing axisymmetric fields and flows in terms of their zonal and poloidal parts. We write $\mathbf{B} = \nabla \times (A\mathbf{e}_\phi) [\equiv \mathbf{B}_p] + B\mathbf{e}_\phi$, $\mathbf{U} = \mathbf{u}_p + s\Omega\mathbf{e}_\phi$, and assume η constant; then

$$\frac{\partial A}{\partial t} + \frac{1}{s}\mathbf{u}_p \cdot \nabla (sA) = (\eta + \beta) \left(\nabla^2 - \frac{1}{s^2} \right) A \quad (4.3)$$

$$\frac{\partial B}{\partial t} + s\mathbf{u}_p \cdot \nabla \left(\frac{B}{s} \right) = s\mathbf{B}_p \cdot \nabla \Omega + \eta \left(\nabla^2 - \frac{1}{s^2} \right) B \quad (4.4)$$

It can be seen that the equation for A is of advection-diffusion type, with no term in B . It can be shown by energy arguments that A decays exponentially in

mean-square; when A is negligibly small there is nothing to sustain B , which can then be shown to decay also by similar arguments. It should be noted, however, from Equation (4.4) that axisymmetric differential rotation does provide a means of producing strong zonal fields from poloidal fields, by stretching out field lines. Even mildly non-axisymmetric fields can be excluded, provided that the essential topology of the field is preserved.

Thus any working dynamo has to have a fully three-dimensional magnetic field (though not, note, a three-dimensional velocity field).

Backus' necessary condition In a finite stationary conductor, magnetic fields can be shown to decay at a finite rate, with timescale $\mathcal{O}(\tau_D)$. It follows that fields cannot grow for arbitrarily small velocities. Backus' result uses the energy equation

$$\frac{1}{2} \frac{\partial}{\partial t} \int |\mathbf{B}|^2 d\mathbf{x} = \int \mathbf{B} \cdot (\mathbf{B} \cdot \nabla \mathbf{u}) d\mathbf{x} - \eta \int |\nabla \mathbf{B}|^2 d\mathbf{x}, \quad (4.5)$$

and calculus of variations techniques to show that in a sphere of radius a surrounded by insulator, $R_m \equiv \max(|\nabla \mathbf{u}| a^2 / \eta) > \pi^2$ for dynamo action. This precludes any attempt to get at dynamos by expanding in powers of R_m . These and other results are described in Moffatt(1978).

4.3 Small and large scale dynamos

It has recently been recognised that there can be two very different types of dynamo action. In the first, *large-scale* or *mean-field* dynamos (to be described later), the magnetic field is significant on a scale much larger than that of the flow. However it is quite possible to produce dynamo action on scales comparable to those of the flow. Well away from solar active regions, magnetic fields are emerging as small bipolar pairs. Although there is no net flux produced, the amount of unsigned flux is much greater than that associated with solar active regions. This flux is broken up by the convection and ends up in intergranular lanes. There seems to be very little connexion between this small scale field and the large scale cycle. Observations are described in Title(2000). The subject has been reviewed by Cattaneo & Hughes(2001).

The most likely explanation for the “magnetic carpet” is that the fields are being produced by small scale dynamo action, on scales larger than granular, and so presumably at depths corresponding to the super- or mesogranulation. Emerging flux is eventually recycled into deeper layers, perhaps by the pumping effect due to anisotropy of the convection, as discussed below. In Figure 10 shows observations of magnetic elements in the photosphere and their relationship to the line of sight flows.

While as we shall see mean-field dynamos rely on rotation to be effective, small scale dynamo action does not seem to need rotation, relying on stretching. It can be modelled successfully by investigating Boussinesq (i.e. essentially incompressible) convection in a fluid layer with large temperature difference). The equations

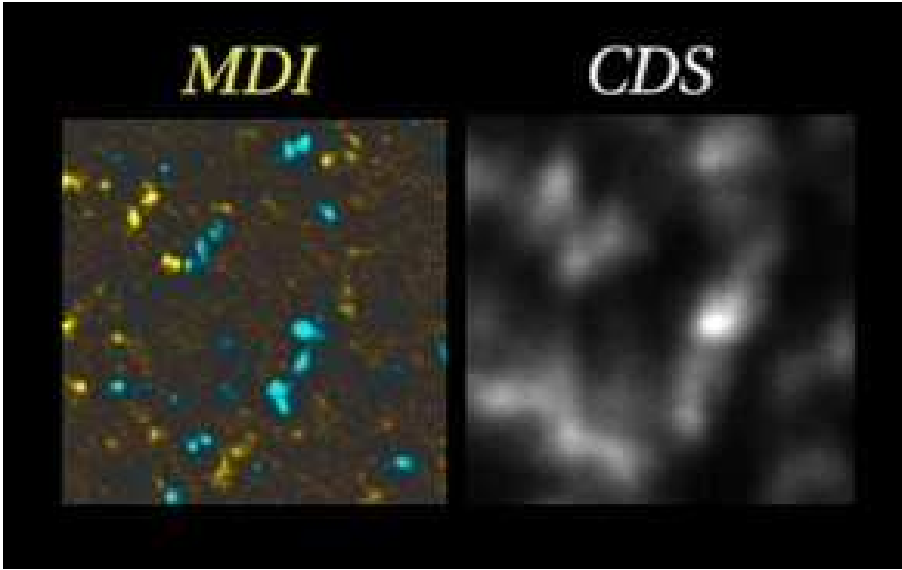


Fig. 10. Observations of the magnetic carpet (courtesy A.Title). Left panel shows small scale magnetic elements; right panel shows line of sight flow velocities.

of motion and heat conduction are solved together with the induction equation, as in Cattaneo(1999). The important parameter here is the magnetic Prandtl number $P_m = \nu/\eta$. It is relatively easy to obtain dynamo action if P_m is not too small. Unfortunately near the solar surface P_m may be as small as 10^{-6} , which makes dynamo action harder to excite at values of R_m available to calculation; nonetheless it is believed that even for this small value of P_m dynamo action can be found for sufficiently vigorous motion. Figure 11 shows the results of such a calculation; the similarity with the previous figure is striking. The magnetic field energy grows as a result of dynamo action until the Lorentz forces act on the flow to stop the growth. The magnetic energy is less than, but of a similar order to, the kinetic energy of the flow in the equilibrated state.

While the small-scale dynamo is particularly observable at the photosphere, since R_m is very large everywhere it probably operates elsewhere as well. The consequences of this are simply not known at present, and the effect has largely been ignored in the study of the solar cycle. I shall return to this point later on.

4.4 Fast and slow dynamos

The Sun operates at very large values of R_m . Thus the two timescales, τ_A and τ_D are very different, and one can ask which of them is relevant for the growth of dynamo fields. On the one hand, Faraday's law states that at $R_m \rightarrow \infty$ the amount

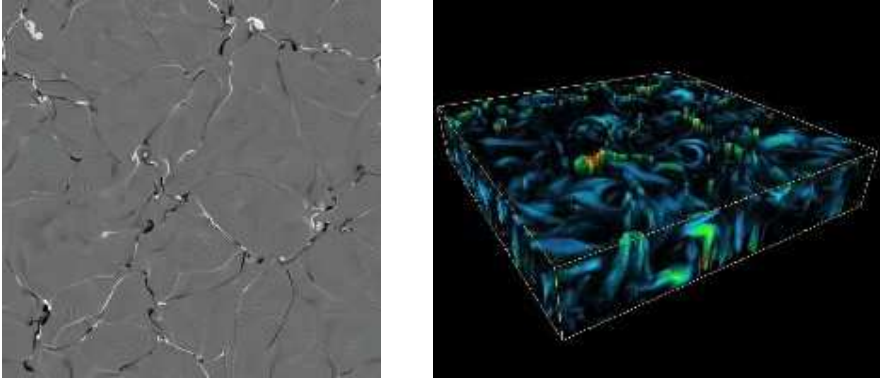


Fig. 11. A small scale dynamo. Left panel, vertical fields at the top of the convective layer. Note the intermittency and the concentration at the edges of convective cells. This should be compared with Figure 10. Right panel, view of the entire convection cell, showing regions of high magnetic energy. Note the large filling factor. Courtesy of F. Cattaneo

of flux through any conductor cannot change. This suggests that growth rates will tend to zero at large R_m . On the other hand if fields are shredded to produce small length scales then diffusion will always be important, and we might expect the scale τ_A to be relevant. Dynamos of the former type are called *slow*; of the latter type, *fast*. It has been shown that for simple velocity fields, that do not have chaotic particle paths, any resulting dynamo will always be slow, while most turbulent flows will lead to fast dynamos. The evolved magnetic fields in fast dynamos will always be highly structured, with a fractal or multifractal structure in the large R_m limit. It should be emphasised that the fast growth rate is (in order of magnitude) the fastest available, in that the growth rate is bounded above by the maximum rate of strain in the flow, typically $\sim \tau_D^{-1}$. Of course if a fast dynamo reaches equilibrium due to dynamical interactions, it is neither fast nor slow, but neutral since the growth rate is zero! In Figure 12 the relation between stretching properties of a flow and the distribution of field is shown for a simple velocity field depending on only two space coordinates. The whole subject of fast dynamos is treated in the monograph by Gilbert & Childress(1995).

4.5 The large scale solar dynamo

The contrast between the dynamo action discussed above and the large scale phenomenon of the solar cycle could not be more dramatic. The cycle is global in nature: it has cyclical variation: and stellar observations seem to indicate a crucial dependence on the rotation. Nonetheless the consensus is that the solar cycle is maintained by a dynamo process, though one of a very different kind. It should be emphasised that, in contrast to the geomagnetic field, which would decay on the

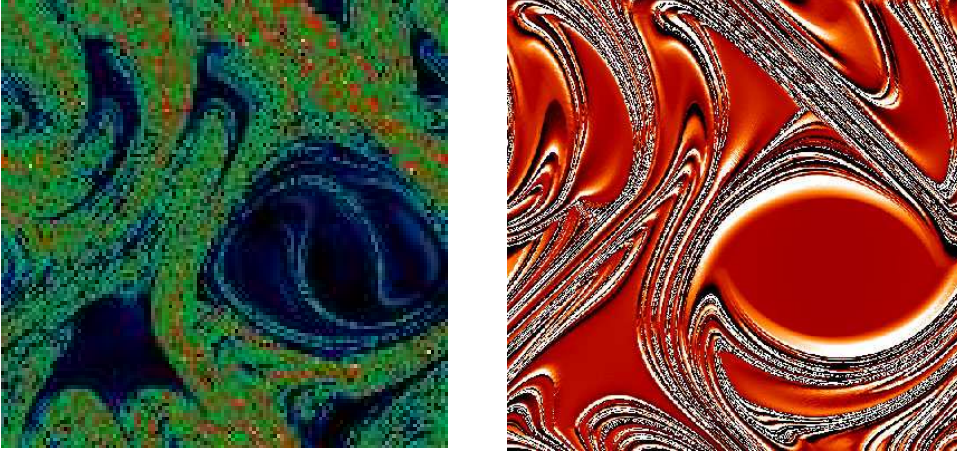


Fig. 12. The relation between stretching and magnetic field distribution in a simple flow (the “CP flow” of Galloway & Proctor(1992)). Left panel, Liapunov exponents in the (x, y) -plane showing regions of exponential stretching (Courtesy F.Cattaneo and D.W.Hughes) ; right panel, regions of strong field in the same flow, from Brummell et al.(1995)

geophysically short timescale of c. 15,000y unless maintained by dynamo action, the resistive time τ_D for the solar field is very long ($\sim 10^{10}$ y), so that dynamo action is not needed to account for its continued existence. But the 11y cycle, involving a reversal of all field components, cannot occur through diffusion alone. It has been suggested, though, that the solar field is essentially a fossil field, existing from the time of its formation, acted upon by oscillatory zonal flows which lead to an oscillating toroidal field by the mechanism discussed in subsection 4.2. If this were the case, however, the torsional oscillations would have the same period as the complete cycle, namely 22y, since they would reverse with the field. In fact, as we have seen, the torsional oscillations shown in Figure 8 have a period of 11y, because they are a dynamic response (through the quadratic Lorentz force) to the oscillating field. This shows definitively that the oscillator theory is untenable and that the dynamo mechanism must provide the explanation. In the following I will critically discuss various models based on the dynamo concept.

5 Mean field dynamos

5.1 The α -effect

The most popular method of producing a tractable model of the solar cycle is that introduced in its essentials by Parker(1955b), and developed and formalised by Krause & Rädler(1980) and Moffatt(1978) He supposed that the velocity and magnetic fields exist on two disparate scales ℓ and $L \gg \ell$. We write $\mathbf{B} = \langle \mathbf{B} \rangle + \mathbf{b}$,

$\mathbf{U} = \langle \mathbf{U} \rangle + \mathbf{u}$ where the brackets denote an average over the short length scale and by definition $\langle \mathbf{b} \rangle = \langle \mathbf{u} \rangle = 0$. Taking the average of the induction equation Equation (4.1), we obtain

$$\frac{\partial \langle \mathbf{B} \rangle}{\partial t} = \nabla \times (\langle \mathbf{U} \rangle \times \langle \mathbf{B} \rangle) + \nabla \times \mathcal{E} - \nabla \times (\eta \nabla \times \langle \mathbf{B} \rangle); \quad \mathcal{E} = \langle \mathbf{u} \times \mathbf{b} \rangle \quad (5.1)$$

$$\frac{\partial \mathbf{b}}{\partial t} = \nabla \times (\langle \mathbf{U} \rangle \times \mathbf{b} + \mathbf{u} \times \langle \mathbf{B} \rangle) - \nabla \times (\eta \nabla \times \mathbf{b}) + \nabla \times (\mathbf{u} \times \mathbf{b} - \mathcal{E}) \quad (5.2)$$

It can be seen that the averaged equation Equation (5.1) now has a new term involving \mathcal{E} formed from the fluctuating quantities. This “mean emf” then has to be determined. If it is supposed that the small-scale field \mathbf{b} owes its existence to $\langle \mathbf{B} \rangle$, i.e. there is no small-scale dynamo, then we can regard \mathbf{b} as a linear functional of $\langle \mathbf{B} \rangle$ and thus hope to find an expression for \mathcal{E} in terms of $\langle \mathbf{B} \rangle$ and the velocity field. If this is local in space and time, we can make the ansatz

$$\mathcal{E}_i = \alpha_{ij} \langle \mathbf{B} \rangle_j + \beta_{ijk} \frac{\partial \langle \mathbf{B} \rangle_j}{\partial x_k} = \alpha \langle \mathbf{B} \rangle_i - \beta (\nabla \times \langle \mathbf{B} \rangle)_i \text{ if isotropic} \quad (5.3)$$

the tensor α_{ij} is actually a pseudo-tensor, which changes sign under reflection, while β_{ijk} is a true tensor. This shows that in flows possessing mirror-symmetry, the symmetric part of α_{ij} must vanish, leaving only an antisymmetric part which acts like a mean flow. Since this will not add any extra physics, we conclude that lack of mirror symmetry is an essential ingredient of the large scale dynamo.

The effective mean flow introduced due to lack of isotropy or inhomogeneity is called magnetic pumping. It has an important effect in the solar convection zone, where the downflows are in the form of narrow plumes, distinct from the more diffuse rising regions. The consequence is that magnetic fields are pushed downwards. This effect is discussed more fully in the treatment of the deep layer dynamo scenario below. For a review of mean flow effects see Moffatt(1983).

In the isotropic case, the mean emf has two parts. The term in β is just an additional diffusivity (the “turbulent diffusivity”) while the term in α represents an emf parallel to the mean field (the α -effect). There is an attractive, though over simplified, physical explanation of the α -effect, originally due to Parker (Parker(1955b)). Consider a helical gyre acting on a uniform field. The field line is pulled up and twisted. The action of diffusion leads to the detachment of a loop of flux, which corresponds to a component of emf parallel to the original field. If the twist of the loop is small, this component has the opposite sign to the *helicity* $\langle \mathbf{u} \cdot \nabla \times \mathbf{u} \rangle$ of the small-scale flow. The situation is illustrated in Figure 13

The α -effect allows one to get round Cowling’s theorem, and thus to develop axisymmetric models of the cycle. Making a the same decomposition as Equation (4.3).Equation (4.4), and without assuming isotropy, we obtain (with the

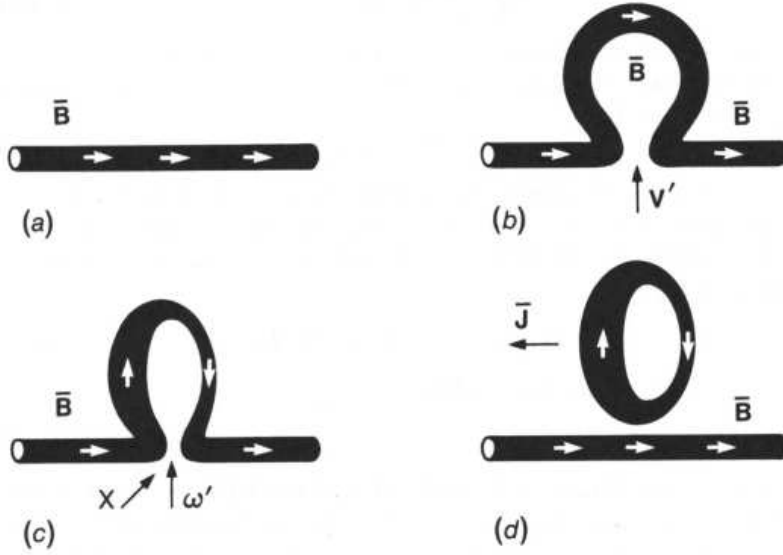


Fig. 13. The mean field mechanism of Parker, after Roberts(1994)

decomposition $\langle \mathbf{B} \rangle = \nabla \times (A \mathbf{e}_\phi) [\equiv \mathbf{B}_p] + B \mathbf{e}_\phi$, $\langle \mathbf{u} \rangle = \mathbf{u}_p + s \Omega \mathbf{e}_\phi$

$$\frac{\partial A}{\partial t} + \frac{1}{s} \mathbf{u}_p \cdot \nabla (sA) = \alpha_1 B + (\eta + \beta) \left(\nabla^2 - \frac{1}{s^2} \right) A \quad (5.4)$$

$$\frac{\partial B}{\partial t} + s \mathbf{u}_p \cdot \nabla \left(\frac{B}{s} \right) = \nabla \times (\alpha_2 \nabla \times (A \mathbf{e}_\phi)) + s \mathbf{B}_p \cdot \nabla \Omega + (\eta + \beta) \left(\nabla^2 - \frac{1}{s^2} \right) B \quad (5.5)$$

The values of α may be different in the two equations. From Equation (5.5) it may be seen that there are two ways in which toroidal field may be produced from poloidal. When the differential rotation is large, the term in Ω will dominate, and so the term in α_2 is normally neglected. The resulting system is known as an $\alpha\Omega$ dynamo. In some circumstances the α_2 term is retained, leading to an “ $\alpha^2\Omega$ dynamo”. From now on we ignore α_2 when Ω is present.

5.2 Parker dynamo waves

It is easy to show that the $\alpha\Omega$ dynamo can lead to growing magnetic fields. Consider a cartesian geometry, modelling a thin spherical shell near the tachocline. α is taken constant, while $\omega \sim \omega' r$, with $\mathbf{u}_p = 0$. We set $A, B = \mathcal{R}(a(x, t)e^{iKr}, b(x, t)e^{iKr})$, where x is colatitude. Thus we have the model system

$$\frac{\partial a}{\partial t} = \alpha b + \eta \left(\frac{\partial^2 a}{\partial x^2} - K^2 a \right), \quad \frac{\partial b}{\partial t} = \omega' \frac{\partial a}{\partial x} + \eta \left(\frac{\partial^2 b}{\partial x^2} - K^2 b \right) \quad (5.6)$$

We can find solutions $\propto e^{i(\sigma t + kx)}$, $k > 0$, where

$$\sigma = \eta K^2 \left[- \left(\frac{k^2}{K^2} + 1 \right) + \sqrt{\frac{k}{K}} \cdot \sqrt{|\mathcal{D}|} (1 + i \operatorname{sign} \mathcal{D}) \right]; \quad \mathcal{D} \equiv \frac{\omega' \alpha}{2\eta^2 K^3} \quad (5.7)$$

\mathcal{D} , the *dynamo number*, depends on the product of α and Ω . It can be seen that $\Re\sigma > 0$, giving growing solutions, if \mathcal{D} satisfies

$$|\mathcal{D}| > \sqrt{\frac{K}{k}} \left(1 + \frac{k^2}{K^2} \right)^2; \quad |\mathcal{D}| > \mathcal{D}_c = \frac{16}{3\sqrt{3}} \quad (5.8)$$

The imaginary part of σ is never zero, so all solutions take the form of travelling waves. The direction of travel is given by the sign of \mathcal{D} . If $\mathcal{D} > 0$ waves move towards the poles, if $\mathcal{D} < 0$ towards the equator. It is known from the differential rotation profile in Figure 7 that the sign of the radial angular velocity gradient changes at mid latitudes, and together with plausible arguments about the sign of α derived from dynamical considerations (the twist in the Parker cyclonic picture is due to Coriolis forces, and is odd about the equator and has the sign given by the twist in surface bipolar pairs as in Figure 1(c)), yields the observed propagation of magnetic activity. Thus the butterfly diagram is in some sense explained!

5.3 Linear spherical models

The Parker wave model can be improved by setting the mean field equations in a spherical geometry. Since the theory is kinematic at this stage, the distribution of α and Ω can be chosen freely. Physical considerations, however, dictate that α should be odd, and Ω even, about the equator. We also expect that the meridional flow \mathbf{u}_P has even radial and odd polar components. Then consideration of Equation (5.4), Equation (5.5) (with α_2 also odd) shows that solutions may be divided into two types:

Dipole Parity: B odd, A even;

Quadrupole Parity: B even, A odd.

(It is possible for solutions to occur without either of these symmetries, but in that case the dynamical consequence of the Lorentz force is that the velocity fields also lose their symmetry.)

If either of the symmetries hold, then the dynamo number \mathcal{D} is odd about the equator. If we choose \mathcal{D} to be negative in the northern hemisphere, then we can find solutions (of either parity), in which wave-like structures travel towards the equator. If we identify the occurrence of active regions with large values of the toroidal field, then this produces a butterfly diagram in accordance with observations. In Figure 14 we show examples of marginal solutions of dipole and quadrupole parity.

These early numerical results could not benefit from modern data concerning the differential rotation profile or any mechanical constraints on α , and so cannot give any detailed information as to the morphology of the field. Nonetheless the

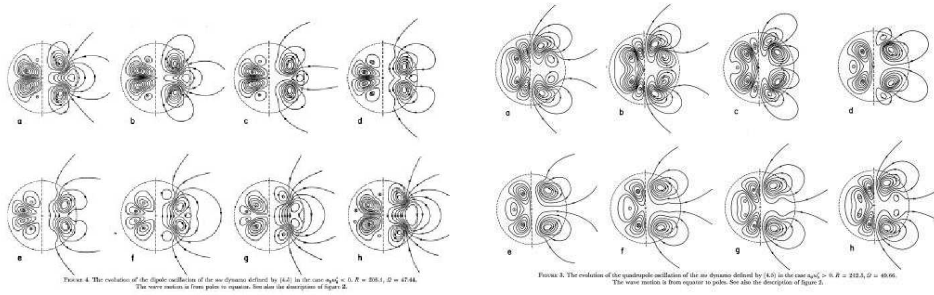


Fig. 14. Marginal solutions of a linearised $\alpha - \Omega$ dynamo model. Left panel: Dipole parity. Right panel: Quadrupole parity. Both pictures show half a complete cycle. From Roberts(1972)

results are highly suggestive and the mean field *ansatz* has proved very durable in a wide variety of attempts to model the field. Perhaps the final word for models of this kind was given by Bushby(2003). Figure 15, from that paper, shows an equilibrated nonlinear solution (of which more later) with the correct differential rotation profile. We see again propagation towards the equator at low latitudes, but weaker poleward propagation at high latitudes, in line with the suggestions from the torsional oscillation profile shown in Figure 8.

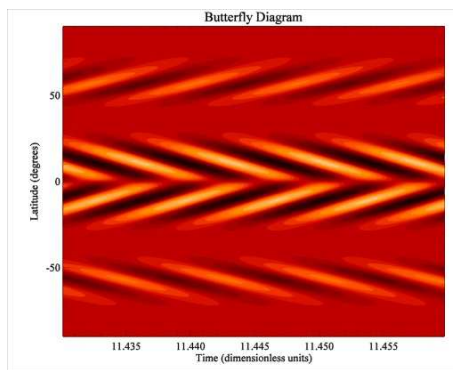


Fig. 15. Latitude/time plot for a nonlinear dynamo model, from Bushby(2003). Propagating structures can be seen at both high and low latitudes

5.4 Problems with the mean-field ansatz

While the above scenario is superficially very attractive, closer examination reveals problems with the actual evaluation of α and β . The equation Equation (5.2) for the fluctuating field is problematical to solve unless the last term (called the 'pain in the neck' term by Moffatt), involving products of fluctuating quantities, can be neglected. There are two circumstances in which this might be done. If R_m , based on the short length scale ℓ , is very small, then $|\mathbf{b}| = \mathcal{O}(R_m)|\langle\mathbf{B}\rangle| \ll |\langle\mathbf{B}\rangle|$. In that case α can be calculated with some degree of rigour: the result is (in the isotropic case with $\langle\mathbf{U}\rangle = 0$)

$$\alpha \sim -\frac{\ell^2}{3\eta}\langle\mathbf{u} \cdot \nabla \times \mathbf{u}\rangle \quad (5.9)$$

This shows the clear relation between the mean emf and the mean *helicity* $\langle\mathbf{u} \cdot \nabla \times \mathbf{u}\rangle$ of the small scale flow. This bears out the Parker picture discussed above. Unfortunately, though, the Sun has very large values of R_m , even at the smallest scales of turbulent convection!

The alternative, proposed by Parker and others, is that R_m is *large*, so that diffusion can be ignored. In this case we might expect $|\langle\mathbf{B}\rangle| \ll |\mathbf{b}|$ due to field line stretching, but it is assumed that the velocity field becomes decorrelated in time, so that the correlated part of \mathbf{b} is small. Then for a postulated correlation time τ_c we can again ignore the awkward terms and obtain the ansatz (when $\langle\mathbf{u}\rangle = 0$)

$$\alpha = -\frac{\tau_c}{3}\langle\mathbf{u} \cdot \nabla \times \mathbf{u}\rangle \quad (5.10)$$

where once again α has the opposite sign to the helicity.

This approximation, which formalizes the Parker picture, is very appealing. But there is little evidence that it works well. Calculations by Courvoisier using the so-called CP-flow of Galloway and Proctor (Galloway & Proctor(1992)) namely

$$\mathbf{u} = (\psi_y, -\psi_x, \psi), \quad \psi(x, y) = \sqrt{\frac{3}{2}}(\cos(x + \cos t) + \sin(y + \sin t)), \quad (5.11)$$

which has chaotic streamlines and good mixing properties, shows a very different picture for large R_m , as illustrated in Figure 16.

Admittedly, this flow is not typical of turbulence. But similar problems arise for more realistic flows. Recent work (Cattaneo & Hughes(2006))examines mean field dynamo action in a rotating convective layer of Boussinesq fluid. When the convection is not too vigorous, the flow does not act as a small-scale dynamo but it has significant helicity induced by the rotation, antisymmetric about the midpoint of the layer. One might expect therefore that if the α -effect is calculated by measuring the emf induced by an imposed uniform magnetic field then this too might be significant. In fact not only is the induced emf very small, but it does not appear to converge for large R_m as the "short-sharp" theory demands.

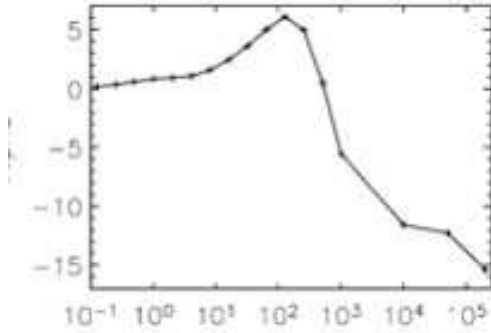


Fig. 16. Plot of α as a function of R_m for the Galloway-Proctor flow (from Courvoisier et al (2006, to appear)). Note that α changes sign as R_m increases and does not appear to converge to a value independent of R_m as $R_m \rightarrow \infty$.

Figure 17 shows the form of the convection, and typical plots of the emfs in one run as a function of time. Interestingly, convection in a much smaller domain, for which the flow is much more coherent due to boundary constraints, gives much higher values of α . The small values in the large box must therefore be due to some kind of decoherence phenomenon. One might expect similar difficulties in the solar context too. In the Hughes-Cattaneo calculation, even when a dynamo does get going for more vigorous convection, there is no sign of a significant large-scale component. This difficulty may be remedied by introducing the effects of large scale shear, which may improve coherence.

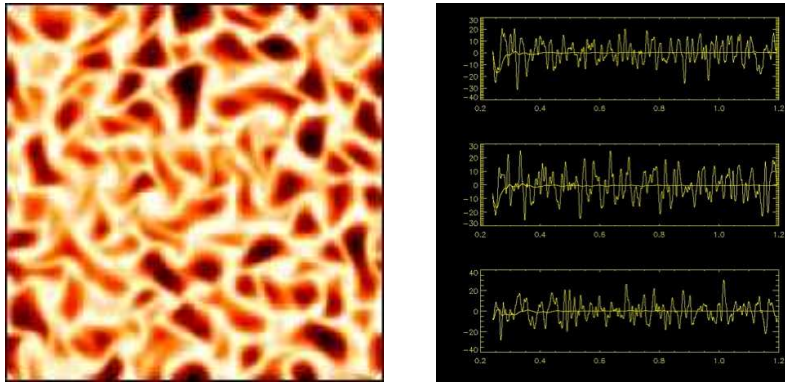


Fig. 17. Decoherence in dynamo action due to convection in a rotating Boussinesq layer (Cattaneo & Hughes(2006)). Left panel, plot of the temperature field near the top of the layer, showing the scale for convection. Right panel; time traces of the emf produced in response to an imposed uniform field.

5.5 Dynamics of mean field dynamos

If there is a dynamo process driven by the mean field effect, possibly in conjunction with large scale shear, the fields will grow until they exert a dynamical effect on the flow. The growth of the magnetic energy can be halted either due to the large scale Lorentz forces affecting the large scale flow (the ‘‘Malkus-Proctor mechanism’’), Malkus & Proctor(1975), or by the small scale forces affecting the mean field coefficients (α - and β -quenching). In the solar context, where R_m is large, the small scale fields are much larger than the large scale ones, so it may be supposed that the quenching process is the more important. We expect the small-scale flow to be altered when $|\mathbf{b}|^2/\mu_0 \sim \rho|\mathbf{u}|^2 \equiv B_{eq}^2/\mu_0$, that is when the small-scale field reaches energy equipartition with the flow. This means that the α -effect in particular will be reduced when the mean field is still well below equipartition values., This is called *catastrophic quenching*, and is hard to reconcile with the observed significant size of large scale solar fields. Using these ideas in a variety of models leads to a quenching law of the form

$$\alpha(\langle \mathbf{B} \rangle) = \frac{\alpha_0}{1 + R_m^\gamma \langle \mathbf{B} \rangle^2 / B_{eq}^2} \quad (5.12)$$

for some $\gamma > 0$ (and often found to be unity). There is much current controversy as to the relevance of such quenching laws, with some workers arguing that the law given by Equation (5.12) can only hold for very special sets of boundary conditions. The question remains open.

However a more important criticism of quenching laws of the type above is that it is assumed that there is no small-scale dynamo. In fact, when the small scale R_m is very large, it is likely that a small scale dynamo can exist, so that the underlying basic state is MHD turbulence, with self-generated small-scale fields. One might expect then that the emf due to the imposition of a large scale field would take a different form.

There is a variant of the formula in Equation (5.10), pertinent to this case. It was originally derived in a closure model of MHD turbulence, but has never been properly validated in a general context. The revised formula takes the form

$$\alpha = -\frac{\tau_c}{3} (\langle \mathbf{u} \cdot \nabla \times \mathbf{u} \rangle - \langle \mathbf{b} \cdot \nabla \times \mathbf{b} \rangle) \quad (5.13)$$

There is some confusion in the literature as to whether the quantities appearing in the averages are the actual small scale fields, or the values they would take if no mean field were imposed. Only in the case of small imposed field can the question be answered unambiguously (given the many approximations involved!). Suppose that there is a preexisting state of MHD turbulence with small scale fields \mathbf{u}, \mathbf{b} , and no mean field or velocity. If a small mean field $\langle \mathbf{B} \rangle$ is imposed, then there are

small perturbations \mathbf{u}' , \mathbf{b}' that depend linearly on $\langle \mathbf{B} \rangle$; they obey the equations

$$\left. \begin{aligned} \frac{\partial \mathbf{u}'}{\partial t} &= -\nabla p + \frac{1}{\mu_0 \rho} \langle \mathbf{B} \rangle \cdot \nabla \mathbf{b} \\ \frac{\partial \mathbf{b}'}{\partial t} &= \langle \mathbf{B} \rangle \cdot \nabla \mathbf{u} \end{aligned} \right\} + \text{diffusion terms} \quad (5.14)$$

$$\begin{aligned} \mathcal{E} &= \alpha \langle \mathbf{B} \rangle = \langle \mathbf{u} \times \mathbf{b}' \rangle + \langle \mathbf{u}' \times \mathbf{b} \rangle \\ &\approx \int_0^{\tau_c} \left(\langle \mathbf{u} \times \mathbf{b}' \rangle + \langle \mathbf{u}' \times \mathbf{b} \rangle \right) dt \end{aligned} \quad (5.15)$$

and substituting from Equation (5.14) we have, assuming isotropy (itself only appropriate for small $\langle \mathbf{B} \rangle$) the relation

$$\alpha \approx -\frac{\tau_c}{3} \left(\langle \mathbf{u} \cdot \nabla \times \mathbf{u} \rangle - \frac{1}{\mu_0 \rho} \langle \mathbf{b} \cdot \nabla \times \mathbf{b} \rangle \right). \quad (5.16)$$

This shows that the existence of a previous small-scale field alters the α -effect by an amount called the ‘‘magnetic torsality’’, clearly related to the magnetic helicity. However it is far from clear that this result makes any sense for larger imposed fields. Indeed, it is always possible for the α -effect to be calculated entirely from the induction equation, if the velocity field is known. There are few results available that do not depend on smoothing approximations as described above. One such, assuming a homogeneous statistically steady state, yields the result

$$\alpha |\langle \mathbf{B} \rangle|^2 = -\eta \langle \mathbf{b} \cdot \nabla \times \mathbf{b} \rangle \quad (5.17)$$

which brings out the essential influence of diffusion in dynamo action, even for very large R_m (for more discussion see Proctor(2003)).

The question of the nature of α -quenching is the subject of ongoing controversy. On the one hand it is argued that the existence of large scale fields near equipartition values must imply that any quenching only becomes effective when the mean field reaches equipartition, contrary to the results described above. For a discussion see Diamond et al.(2005), and for a rather different perspective see Brandenburg & Subramanian(2004). It would then follow that the calculations leading to R_m -dependent (catastrophic) quenching must be irrelevant, for example because the boundary conditions are not properly taken into account (for example, in Equation (5.17) the helicity is conserved, while if there are open boundaries this will not be true). On the other hand, it is maintained that the quenching process is essentially local in nature, relying on locally generated small-scale fields, and so boundary conditions should be irrelevant. An agreed answer will take some time to appear.

6 Nonlinear mean field dynamo models

In spite of the clear failings of the mean field mechanism to deliver a quantitative description of the mean field dynamo process, it is the best model available

at present. A detailed discussion of how the model can best be applied can be found in Brandenburg(2005). Until recently (allowing as it does the discussion of axisymmetric mean fields with a large saving in computer time) was the only way of treating fully dynamic dynamos. There have been many models of this type, with the nonlinearity provided by a model of α -quenching or with the effects of the Lorentz forces on the mean fields. Early studies used the Malkus-Proctor effect; more recently in the solar context the mean effects of the small-scale fields on the Reynolds stresses (“ Λ -quenching”). Whatever the origin of the back reaction the primary effect on an $\alpha\Omega$ dynamo is on the zonal flow field. This mechanism not only will lead, via Lenz’s Law, to a reduction in the growth rate of an initially unstable dynamo, but will also result in zonal flow anomalies with a period half that of the cycle, as observed.

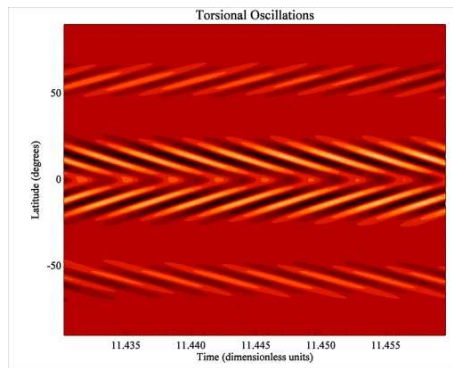


Fig. 18. Latitude/time plot of zonal flow velocity anomalies for a nonlinear dynamo model, from Bushby(2003) (Compare Figure 15). As observed (see e.g. Figure 8) the period of the fluctuations is half that of the field shown in the other picture.

An important effect of nonlinearity is the selection between solution branches with different parity and time dependence. While the linear problem admits in general solutions of dipole or quadrupole parity, which can be either (in a confined geometry) steady or oscillatory. Nonlinear interactions will in general select a small number (often just one) type of solution which is stable in a given parameter range, and also give rise to solutions of mixed parity, with no symmetry about the equator. The details of such selections are highly model dependent; an example is shown in Figure 19.

Even the simplest dynamically realistic models of the effects of the Lorentz force involve a separate evolution equation for the zonal flow. This will introduce a separate timescale, namely the diffusive decay time of the flow. There are many

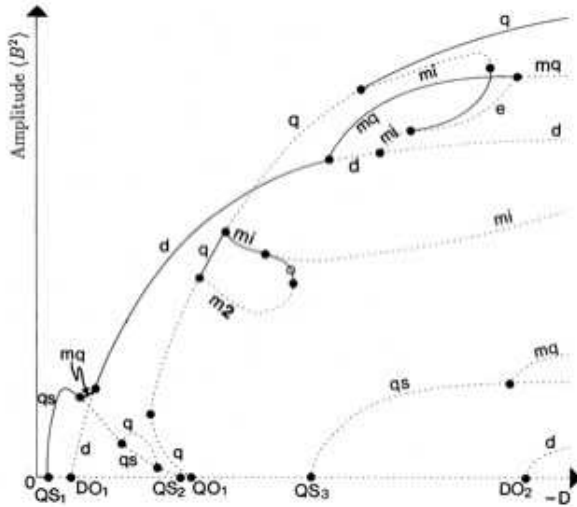


Fig. 19. Interaction of solutions of different parities in a simple nonlinear dynamo model (Jennings & Weiss (1991)). d, q, m denote dipole, quadrupole and mixed mode branches, while capital letters and O, S show the types of bifurcation from the zero state. Heavy lines show stable branches. The dynamo number D is shown as negative to obtain waves of activity propagating towards the equator.

models extant, involving this new timescale, that show more complex dynamical behaviour than simple regular cycles. The principal new effects are modulation of the amplitude and phase of the cycle, and changes in the parity of the solution. In fact changes in the parity of the solution and low values of the wave amplitude often go together, as shown in the simulation of Pipin (1999) (Figure 20). Here we see regular modulations of the amplitude; when the latter is small the parity of the solution changes from -1 , corresponding to pure dipole, to near zero, denoting a mixed solution. Here is a possible explanation of grand minima, and of the interesting observation of Ribes & Nesme-Ribes (1993) that the sunspot record was highly asymmetric with respect to the equator at the close of the Maunder Minimum (Figure 22). For more elaborate models the modulation behaviour can be more complex, even effectively chaotic, bringing the results even closer to what we know about the actual irregularities in the cycle. But before taking this robustness as evidence of the correctness of the mean field model, it must be emphasised that the behaviour is principally due to the symmetries of the system. Any other model respecting the same symmetries might be expected to show a similar variety of behaviour. It is encouraging, however, that the results of very simple low order dynamical systems, incorporating the bare bones of the physics, show very similar behaviour to much larger PDE models, with full resolution in two dimensions, as shown in Knobloch et al. (1998) (Figure 21).

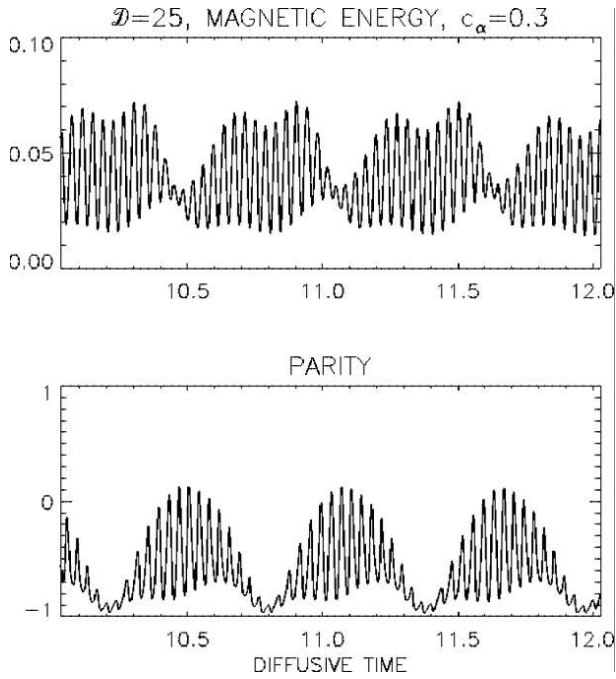


Fig. 20. Periodic modulation in a simple nonlinear dynamo model. Top graph shows time evolution of amplitude, bottom graph the evolution of parity (normalised value of quadrupole energy - dipole energy). From Pipin(1999) Note the change of parity near amplitude minima.

7 Other dynamo scenarios

7.1 The Parker Interface Model

While the distributed α -effect model has had successes, as we have seen, there are numerous shortcomings associated with its use. Various authors have tried to improve the model to take account of the dynamical conditions occurring in the deep convective zone. The simplest of these enhancements was given by Parker(1993) , who noted that the rapid zonal differential rotation associated with the tachocline was in a region where convection was not vigorous, so that any mean field effects

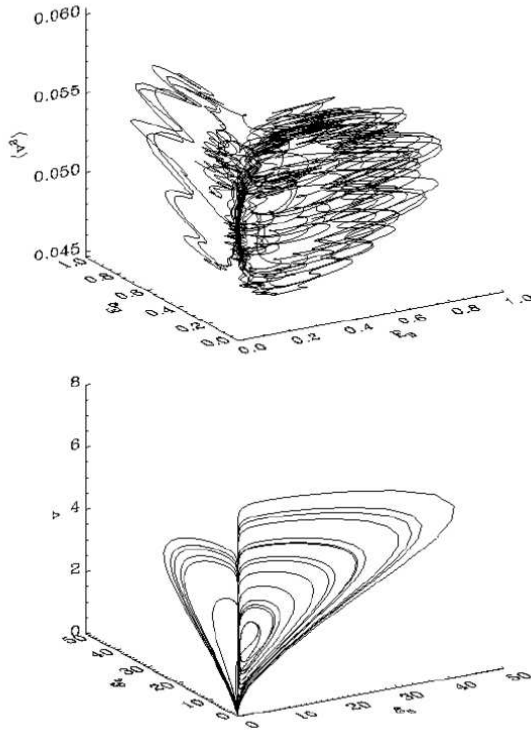


Fig. 21. Comparison of behaviour of a low order dynamo model (bottom graph) with that in a fully resolved PDE model. Each graph shows the behaviour of the dipole and quadrupole energies as functions of time. It can be seen that the system spends most of its time close to one parity state, with excursions to the other parity or to mixed parity induced by episodes of low total amplitude. From Knobloch et al.(1998)

would be small. Thus his model has an upper region where convection is vigorous (so that the α and β -effects are large), but differential rotation can be neglected, and a lower region where α and β are small but there is significant shear. Equations for dynamo waves can be set up; effectively the system is of $\alpha\Omega$ type but with the two elements disjoint in space. Consequently the dynamo waves are confined to the neighbourhood of the interface. In Figure 24 we see a typical solution. Note the different length scales in the radial direction as a consequence of the different effective diffusion rates.

The Parker analysis is purely kinematic, in that the diffusion rates etc are prescribed. However it has been shown that if the diffusion rates are linked non-linearly to the field strength via a quenching model, then a self-consistent model can be achieved.

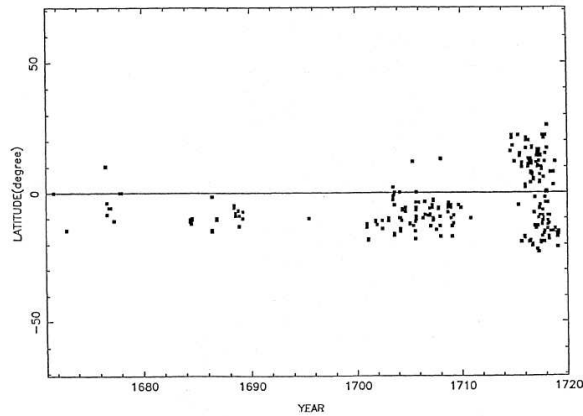


Fig. 22. Observations of sunspot latitudes towards the end of the Maunder Minimum. It can be seen that there is a marked preponderance of sunspots in the southern hemisphere (corresponding to a mixed parity solution) as the sunspot number rises, with symmetry only reestablished during the last cycle shown. From Ribes & Nesme-Ribes(1993).

7.2 Conveyor belt models

The traditional mean field models regard the whole convection zone as a potential source of poloidal field. However there may be problems in some of the models in reconciling the phase of the cycle as evidenced by the sunspot record with the phase of the poloidal field at the poles, which appears to be correlated with the sunspot field at earlier times. The original models of Babcock and Leighton regarded the tilt of sunspot pairs with respect to the equator as providing the sole source of poloidal flux. More recent work by Charbonneau, Dikpati and others, (e.g Dikpati & Charbonneau(1999)) updates this idea, by using the observed meridional flow from pole to equator at the surface as the "conveyor belt" that brings poloidal field to the poles, and then draws it down to the base of the convection zone, to produce toroidal field via the Ω -effect. In practical terms the model resembles the mean field models described earlier, but the crucial α term is only non-zero near the surface, and depends not on the local toroidal field but on its value at the base of the convection zone. Convincing agreement is claimed with time scale and phase relationships observed in the Sun.

There is no role for the convection zone in early versions of the model. This has the good feature that there is no problem with α -quenching, as the local mean zonal field does not appear in the model, and the field is in any case less dynamically active at the base of the convection zone. However there are some difficulties with the model, quite apart from its overparametrization of the flux emergence process. The first objection is partially philosophical: the model assigns great importance

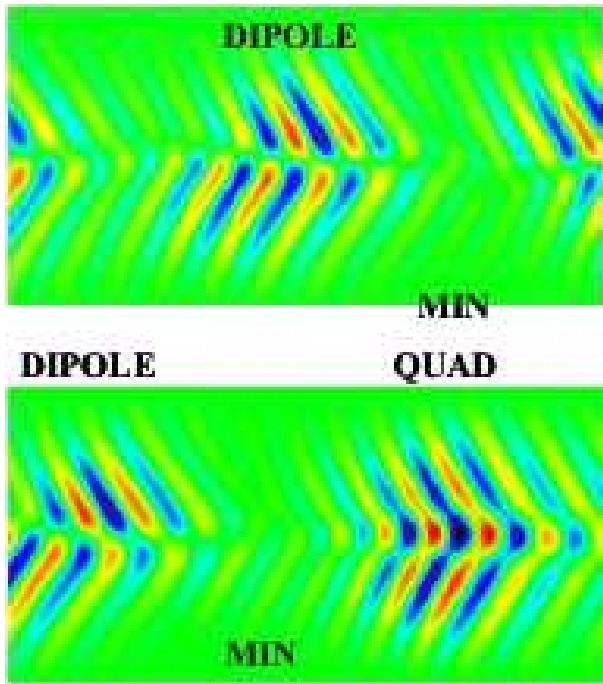


Fig. 23. Grand minima and parity changes in a 2D dynamo model. Shown is the zonal field strength: red and blue indicate opposite polarities. During the first minimum the dipole parity is preserved, but after the second one the parity is reversed. In both cases there are significant asymmetries in the field structure when the amplitude is small. Adapted from Beer et al.(1998)

to the sunspot field, which is highly visible, but is not necessarily representative of structures below the surface. It seems more likely that the sunspots are the symptoms of the process that drives the cycle, rather than its cause. In addition, the model does not account for the apparent continuation of the cyclical process during the grand minima, when we know (at least in the case of the Maunder Minimum) that no sunspots were present to provide the necessary flux.

More recent versions of the model, e.g Dikpati & Gilman(2001) and Dikpati et al.(2004), incorporate an alternative production mechanism for poloidal field, due to MHD instabilities of the tachocline. This new mechanism avoids some of the problems referred to above, but the MHD instability mechanism has still not been shown unambiguously to give poloidal field of the correct structure. Nonetheless, many qualitative features of the cycle, at least as far as the zonally averaged fields are concerned, have been successfully simulated.

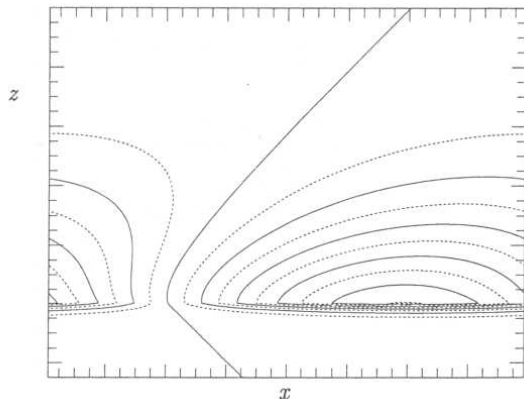


Fig. 24. Marginal travelling wave solution to the Parker interface dynamo, from Parker(1993). The wave travels to the left.

7.3 Direct numerical simulation

It is now becoming possible to contemplate the direct simulation of the solar dynamo. A pioneering start has been made by the JILA group (Brun et al(2004)) who have solved the anelastic equations of motion and heat conduction in a rotating spherical shell. The anelastic approximation, in which the deviation of the density and entropy fields from spherically symmetric basic states is small, filters out sound waves and allows faster integration.

The results of the simulations are highly promising so far as the global velocity field is concerned (Figure 25). There is vigorous convection, which is dominated by concentrated downflows. The Reynolds stresses produced by these flows drive a differential rotation which is not unlike that of the Sun, though the angular velocity tends to be constant more on cylinders than on radial lines. There is no physics leading to a tachocline, so although the differential rotation decreases with depth it does not approach solid body rotation at the base of the shell.

The velocity field obtained readily leads to dynamo action, and the magnetic energy equilibrates at about one tenth of the KE. However the dynamo bears little resemblance to the solar field, being dominated by the scales of the convection, with no significant small scale part (Figure 26). In fact, the results bear a strong resemblance to the simulations of Cattaneo and Hughes described earlier for a rotating convective layer. There is no evidence of cyclic behaviour. The dynamo may be characterised as "small scale" even though it is of global extent! It is possible that such disordered dynamos may be relevant to stars without radiative cores, where there is no tachocline to organise the cycle. For the Sun it seems that more careful attention to conditions at the base of the convection zone will be needed if a model showing cyclic behaviour is to be found.

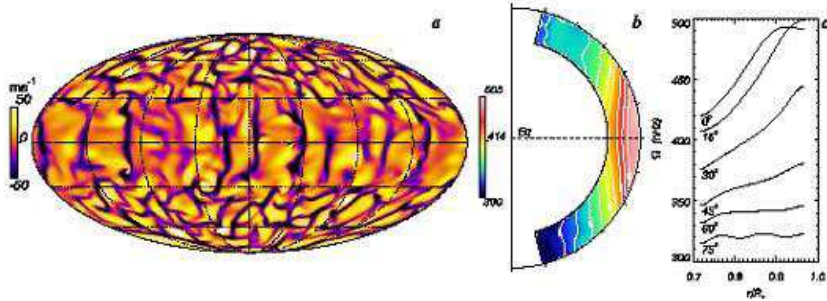


Fig. 25. Velocity field from the simulations of Brun et al(2004). Left panel: radial velocity at the surface; centre panel: contours of angular velocity (compare Figure 8); right panel: differential rotation as a function of depth at several latitudes.

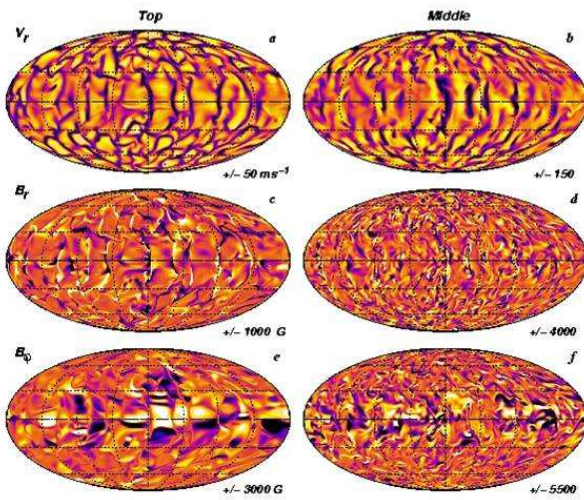


Fig. 26. Dynamo action in a sphere. Radial velocity, radial field and zonal field at the top of the convection zone and at mid-depth

7.4 Deep-seated Scenario

Finally in this section we discuss what would appear to be the developing consensus concerning the operation of the dynamo. This may be called the 'deep-seated scenario' as the controlling part of the mechanism operates at or below the base

of the convection zone. The operation is best described by means of a diagram, shown here courtesy of N.Brummell (Figure 27). It is supposed that there is poloidal field in the lower part of the convection zone. This is pumped down by the anisotropic convection, due to the mean field effect discussed earlier (see e.g. Tobias et al.(2001)), and finds itself below the convection zone, A in the region of strong radial shear. This produces zonal field due to the Ω -effect. Eventually this becomes unstable due to the operation of magnetic buoyancy, and discrete loops of buoyant flux rise up into the convection zone. Some of these loops remain coherent and may rise to the surface giving rise to active regions. A proportion of the flux, however, remains in the convection zone due to shredding by the convection and the pumping effect(Figure 28). The operation of the Coriolis force ensures that poloidal field is produced by twisting of the rising tubes. This allows the cycle to begin again.

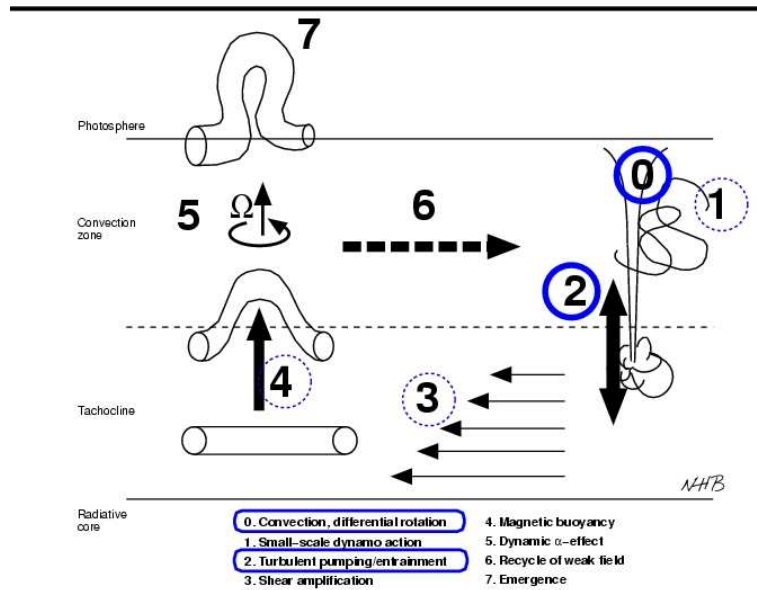


Fig. 27. Schematic of the operation of the “deep-seated” scenario. For explanation see text

In this model the α -effect is crucial in producing the poloidal field. However the energetics differ fundamentally from the mean-field or conveyor-belt scenario. There the field arises as an instability of the turbulent sheared flow, and draws its

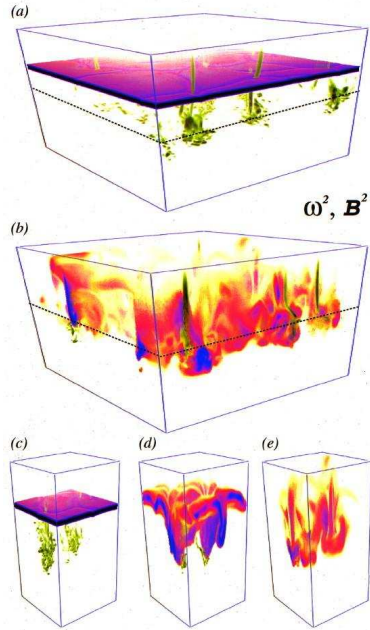


Fig. 28. Combined action of magnetic buoyancy and pumping on an initially uniform layer of strong field. In spite of the instability of the upper surface of the layer, almost all the field is eventually transported into the stable region at the base of the layer (adapted from Tobias et al.(2001)). The lower figures show an analogous calculation in a deep layer without a lower stably stratified part: the effect is the same.

energy from the flow in general. Here the process is driven primarily by the shear, with the mean field effects an essential but very minor player. In this scenario the active regions, far from being the primary source for the poloidal field, are in fact just the symptoms of the cycle, which continues at a deeper level even when no sunspots are present. This is of course consistent with the occurrence of grand minima, when cyclical behaviour apparently continued in the absence of sunspots.

The operation of this mechanism depends crucially on the effects of magnetic buoyancy. Because of magnetic pressure, concentrations of magnetic field are at a lower gas pressure, and therefore presumably at lower density, than their surroundings. The resulting loss of equilibrium (or instability for special field configurations) was first discussed by Newcomb(1961) and Parker(1955a), and has since been extensively investigated (see e.g. Tobias(2005)). The great majority of the investigations concern a pre-existing field, and its evolution with or without an ambient shear. An early example is given by Wissink et al.(2000), which shows that three-dimensional flux structures can occur spontaneously from a simple initial configuration (Figure 29). It is not clear that these calculations shed light directly on the dynamic process whereby the zonal field is continuously evolving,

rather than existing as a time independent basic state. Gilman and co-workers have shown that the tachocline shear-flow could itself be unstable to dynamical instabilities, modified and perhaps enhanced by the ambient field, even when buoyancy is not important. In actual fact both the buoyancy and the shear-flow mechanisms will probably be operating.

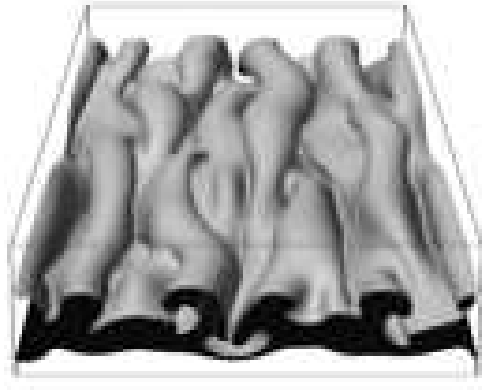


Fig. 29. Three dimensional development of the buoyancy instability (Wissink et al.(2000)). The original configuration was a sheet of uniform horizontal field. The initial instability is two-dimensional and the three-dimensional arching occurs as a secondary effect.

There have been several attempts to bring magnetic buoyancy into the form of a dynamo process (e.g Brandenburg & Schmitt(1998)). Thelen(2000a) and Thelen(2000b) solved the stability problem for a zonal field approximating to the conjectured solar field; the eigenfunctions of the instability were then used to construct an α -effect whose magnitude depended on position and on the local strength of the field, and a nonlinear mean field dynamo thus constructed. In fact magnetic buoyancy can lead to dynamo action even without turbulence and rotation! Cline et al.(2003) have found a simple dynamo in a periodic box, entirely driven by shear in a stratified system, in which the strong but unstable field is produced dynamically by a given two-dimensional shear. Instabilities of the induced flux tubes lead to toroidal currents and to a self-sustaining field. Further development of this scenario to provide for cyclical behaviour, and proper treatment of the role of the convection zone will undoubtedly be the focus of much interest over the next few years.

8 Conclusion

In this review I have attempted to survey the current level of understanding of the fundamental theory of the solar cycle. The study of the subject is at an exciting stage, as the long reign of the mean field dynamo models reach their end, and computing power is now such that it is possible to conduct detailed simulations of the complex dynamical processes that contribute to the cycle, though an integrated description of the global process is still some way off. In addition, more detailed observations of distant stars will lead to a better understanding of the role of the tachocline in producing regular behaviour, as suggested by the deep seated scenario.

Acknowledgements

I thank my friends and colleagues Nigel Weiss, David Hughes, Fausto Cattaneo and Steve Tobias for many enlightening discussions and for help in finding images, Michel Rieutord for his hospitality at Toulouse and Aussois, and François Rincon for assistance with the figures. s

References

- Beer, J. 2000, *Space Sci. Rev.* 94,53
- Beer, J., Tobias, S.M. & Weiss, N.O. 1998, *Solar Phys.*, 181, 237.
- Brandenburg, A. 2005, *Astrophys.J.*, 625, 539.
- Brandenburg, A. & Schmitt, D. 1998, *Astron. Astrophys.*, 338, L55
- Brandenburg, A. & Subramanian, K. 2004 *Phys. Rep.*, 141, 1502.
- Brummell, N. et al. 1995, *Science*, 269,1370.
- Brun, A.S., Miesch, M.S. & Toomre, J. 2004, *Astrophys J.*, 614, 1073.
- Bushby, P. J. 2003, *MNRAS*, 342, L15
- Cattaneo, F. 1999, *Astrophys.J*, 515, L39.
- Cattaneo, F. & Hughes, D.W. 1996, *Phys. Rev. E* 54, R4532.
- Cattaneo, F. and Hughes, D.W. 2001, *Astron, Geophys.*, 42(3), 18.
- Cattaneo, F. & Hughes, D.W. (2006). *J.Fluid Mech.*, in press.
- Cline, K.S., Brummell, N.H. & Cattaneo, F. 2003, *Astrophys. J.*, 588, 630.
- Diamond, P.H., Hughes, D.W. & Kim, E. 2005. In "Fluid Dynamics and Dynamos in Astrophysics and Geophysics" ed. A.M. Soward, C.A. Jones, D.W. Hughes & N.O. Weiss (CRC Press, London), p. 145.
- Dikpati, M. & Charbonneau, P. 1999, *Astrophys J.*, 518, 508.
- Dikpati M., Gilman P. A., 2001, *Astrophys. J.*, 559, 428.
- Dikpati, M., de Toma, G., Gilman, P.A. Arge, C.N. & White, O.R. 2004, *Astrophys. J.*, 601, 1136.
- Eddy, J.A. 1976, *Science*, 192, 1189.
- Galloway, D.J. & Proctor, M.R.E. 1992, *Nature*, 356,691.

- Gilbert, A.D. and Childress, S. 1995, "Stretch, Twist, Fold: The Fast Dynamo". Springer-Verlag.
- Jennings, R.L. and Weiss, N.O. 1991, MNRAS, 252,249.
- Knobloch, E., Tobias, S.M. & Weiss, N.O. 1998, MNRAS, 297, 1150.
- Krause, F. & Rädler, K.-H. 1980, "Mean-field Magnetohydrodynamics and Dynamo Theory", Akademie-Verlag, Berlin.
- Mason, J., Hughes, D.W. & Tobias, S.M. 2002, Astrophys J., 580, L89.
- Malkus, W.V.R. & Proctor, M.R.E. 1975. J.Fluid Mech., 67, 417.
- Moffatt, H.K. 1978, "Magnetic Field Generation in Electrically Conducting Fluids", Cambridge University Press.
- Moffatt, H.K. 1983, Rep. Prog. Phys. 46, 621.
- Newcomb, W.A. 1961, Phys. Fluids A,4, 391.
- Ossendrijver, M.A.J.H. 2003, Astr. and Space Sci. Rev., 11, 287.
- Parker, E.N. 1955a. Astrophys. J.,121, 491.
- Parker, E.N. 1955b, Astrophys. J.122, 293.
- Parker, E.N. 1993, Astrophys. J., 408, 707.
- Pipin, V.V. 1999, Astr. Astrophys., 346, 295.
- Proctor, M.R.E. 2003, in "Stellar and Astrophysical Fluid Dynamics", Thompson & Christensen-Dalsgaard, eds., Cambridge University Press.
- Ribes, J.C. & Nesme-Ribes E. 1993, Astr. Astrophys., 276, 549.
- Roberts, P.H. 1972, Phil. Trans Roy Soc. Lond. A272, 663
- Roberts, P.H. (1994). In *Lectures on Solar and Planetary Dynamos*, ed. M.R.E. Proctor & A.D. Gilbert (Cambridge University Press, Cambridge), p. 1.
- Thelen, J.-C. 2000a, MNRAS, 315, 155.
- Thelen, J.-C. 2000b MNRAS, 315, 165.
- Thomas, J.H. & Weiss, N.O. 2004, Ann. Rev. Astron. Astrophys., 42, 517.
- Title, Alan 2000, Phil. Trans Roy Soc. Lond. A358, 657
- Tobias, S.M. 2002, Phil. Trans. R. Soc. A360, 2741.
- Tobias, S.M. 2005, In "Fluid Dynamics and Dynamos in Astrophysics and Geophysics", ed. A.M. Soward, C.A. Jones, D.W. Hughes & N.O. Weiss (CRC Press, London), p. 193.
- Tobias S. M., Brummell N. H., Clune T. L., Toomre J., 2001, ApJ, 549, 1183.
- Vorontsov S.V., Christensen-Dalsgaard, J., Schou J., Strakhov G.N. & Thompson M.J. 2002, Science, 296, 101.
- Wissink, J.G., Hughes, D.W., Matthews, P.C. & Proctor, M.R.E. 2000, MNRAS, 318, 501.
- Weiss, N.O. 1994, in "Lectures on Solar and Planetary Dynamos", Gilbert & Proctor, eds, Cambridge University Press.