Space Sci Rev (2009) 144: 53–66
DOI 10.1007/s11214-008-9435-z

The Solar Dynamo

N.O. Weiss · M.J. Thompson

Received: 21 July 2008 / Accepted: 2 September 2008 / Published online: 1 October 2008
© Springer Science+Business Media B.V. 2008

Abstract It is generally accepted that the strong toroidal magnetic fields that emerge through the solar surface in sunspots and active regions are formed by the action of differential rotation on a poloidal field, and then stored in or near the tachocline at the base of the Sun's convection zone. The problem is how to explain the generation of a reversed poloidal field from this toroidal flux—a process that can be parametrised in terms of an α-effect related to some form of turbulent helicity. Here we first outline the principal patterns that have to be explained: the 11-year activity cycle, the 22-year magnetic cycle and the longer term modulation of cyclic activity, associated with grand maxima and minima. Then we summarise what has been learnt from helioseismology about the Sun's internal structure and rotation that may be relevant to our subject. The ingredients of mean-field dynamo models are differential rotation, meridional circulation, turbulent diffusion, flux pumping and the α-effect: in various combinations they can reproduce the principal features that are observed. To proceed further, it is necessary to rely on large-scale computation and we summarise the current state of play.

Keywords Sun · Sunspots · Magnetic fields · Dynamos

1 Introduction

The principal task of dynamo theory, as applied to the Sun, is to explain the systematic global behaviour of solar magnetic fields. In this broad-brush review, we begin by summarising the key properties of magnetic fields that are observed at and above the surface of the Sun, including not only the cyclic variation of solar activity but also its modulation on much longer

N.O. Weiss
DAMTP, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WA, UK
e-mail: now@damtp.cam.ac.uk

M.J. Thompson (✉)
School of Mathematics & Statistics, University of Sheffield, Sheffield S3 7RH, UK
e-mail: Michael.Thompson@sheffield.ac.uk

time-scales. It is clear that these fields are generated below the photosphere, at levels where they have not as yet been directly measured, but helioseismology has succeeded in revealing the possibly associated patterns of motion in the solar interior, including both differential rotation and meridional flows. In Sect. 3 we describe these results, with some emphasis on the slender solar tachocline, which is generally agreed to play an essential part in the dynamo process (Tobias and Weiss 2007a). In the next section we briefly review the current state of mean field dynamo theory, before going on, in Sect. 5, to discuss the various models that have been proposed in order to describe the solar dynamo. Then, in Sect. 6, we consider progress beyond mean field models and towards direct numerical simulation of stellar dynamos. In the final section we comment briefly on attempts to predict the future amplitude of solar activity and try to estimate the expected lifetime of the current grand maximum. Solar and stellar dynamos have already been the subject of several recent reviews (e.g. Tobias 2002; Ossendrijver 2003; Choudhuri 2003; Rüdiger and Hollerbach 2004; Charbonneau 2005; Solanki et al. 2006).

2 Magnetic Activity on the Sun

Sunspots are the most striking manifestations of solar activity (Thomas and Weiss 2008) and their incidence, as measured by area occupied or the traditional sunspot number R, follows the irregular 11-year activity cycle that is best demonstrated by the butterfly diagram in Fig. 1. The true nature of solar activity was only revealed when Hale discovered that dark spots are the sites of strong magnetic fields. Sunspots typically occur as a pair, aligned approximately with a parallel of latitude but with the leading spot (in the sense of the solar rotation) usually somewhat closer to the equator. The leading and following spots have

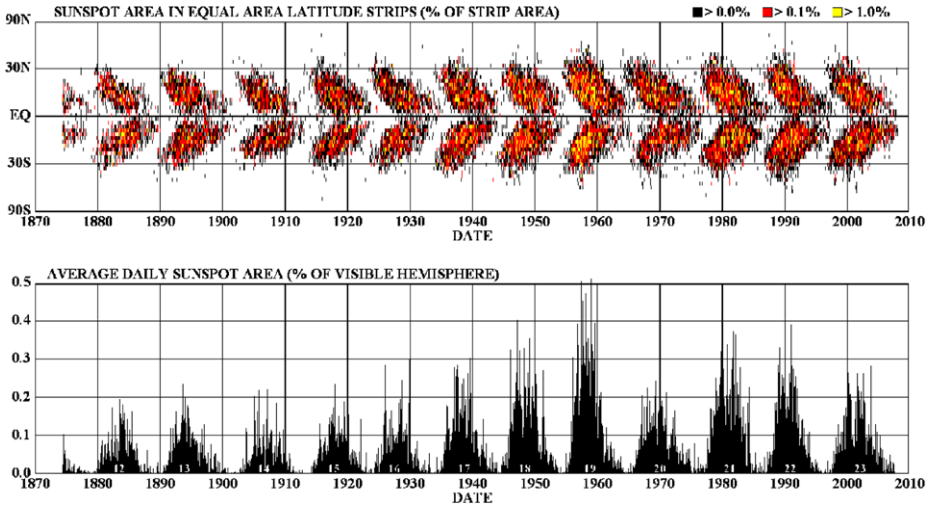

Fig. 1 Cyclic activity on the Sun since 1874. The *lower panel* shows the daily sunspot area, as a percentage of the visible hemisphere that is covered by sunspots and averaged over individual solar rotations, which varies quasiperiodically with an 11-year period. The butterfly diagram in the *upper panel* shows the corresponding incidence of sunspots as a function of latitude and time. At the beginning of a new cycle, spots appear around latitudes of $\pm 30°$. The activity zones spread until they extend to the equator, and then gradually die away, disappearing at the equator as the first spots of the next cycle appear at higher latitudes. (Courtesy of D.H. Hathaway)

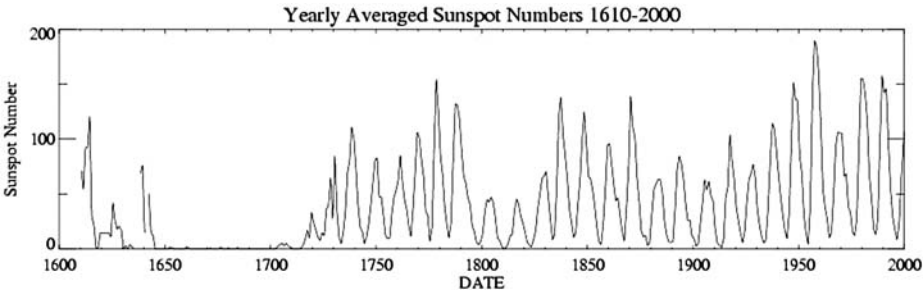

Fig. 2 Annual values of the sunspot number from the first telescopic observations in 1610 to 2000. The amplitude of cyclic activity has varied irregularly, with a prominent interval of inactivity in the seventeenth century—the Maunder Minimum. (Courtesy of D.H. Hathaway)

opposite magnetic polarities, and the appearance of sunspot pairs at the photosphere is best understood as caused by the emergence of a toroidal flux tube from deep in the underlying convection zone. Hale found that the polarities of the spots are consistent in each hemisphere but antisymmetric about the equator. Moreover, these polarities reverse at the end of each 11-year activity cycle, so that there is a magnetic cycle with a period of 22 years. The sunspots are associated with active regions, and there are smaller scale fields on a wide range of scales all over the solar surface. Near the poles there are weaker unipolar fields that are most prominent at sunspot minimum, and reverse at sunspot maximum.

The cycles in Fig. 1 are asymmetric, with the sunspot area, and the corresponding sunspot number R, rising more rapidly than it falls; it is also apparent that stronger cycles rise faster and are shorter. The longer record of R in Fig. 2 shows that the cycle's amplitude has varied widely during the past 300 years: the peak value of R approached 200 in 1958, but in the early nineteenth century the maxima reached no more than 50. Far more striking is the interval from 1645 to 1715—the Maunder Minimum (Eddy 1976; Ribes and Nesme-Ribes 1993)—when sunspots almost completely disappeared.

To investigate the long-term modulation of solar activity we must turn to proxy data—and fortunately such datasets exist. The magnetic fields that are carried out into the heliosphere by the solar wind deflect galactic cosmic rays. When these energetic particles impinge on the earth's atmosphere they give rise to the production of radioactive isotopes such as ^{14}C (which is preserved in trees) and ^{10}Be (which is preserved in polar icecaps) whose abundances therefore vary in antiphase with solar magnetic activity. Tree rings can be dated and seasonal variations are apparent in ice cores; thus the 11-year activity cycle can be followed back for hundreds of years, while its envelope has been established for many millennia. Figure 3 shows how this envelope of solar magnetic activity has varied over the past 9000 years. The apparently chaotic pattern of modulation, giving rise to grand maxima and grand minima, persists throughout this period. Power spectra nevertheless reveal persistent periodicities in both records (Damon and Sonett 1991; Stuiver and Braziunas 1993; Beer 2000; Wagner et al. 2001): in addition to the basic 11 year (Schwabe) cycle, there are well-defined peaks corresponding to periods of 88 years (the Gleissberg cycle), 210 years (the de Vries cycle) and 2300 years (the Hallstatt cycle). As Figs. 1 and 2 show, we happen to be living during an episode of exceptionally high activity—but this episode is not unique, and there have been several comparable intervals within the past millennium.

It is also worth pointing out that similar patterns of cyclic activity can be detected in other slowly rotating late-type stars with deep convection zones, like the Sun. There are also

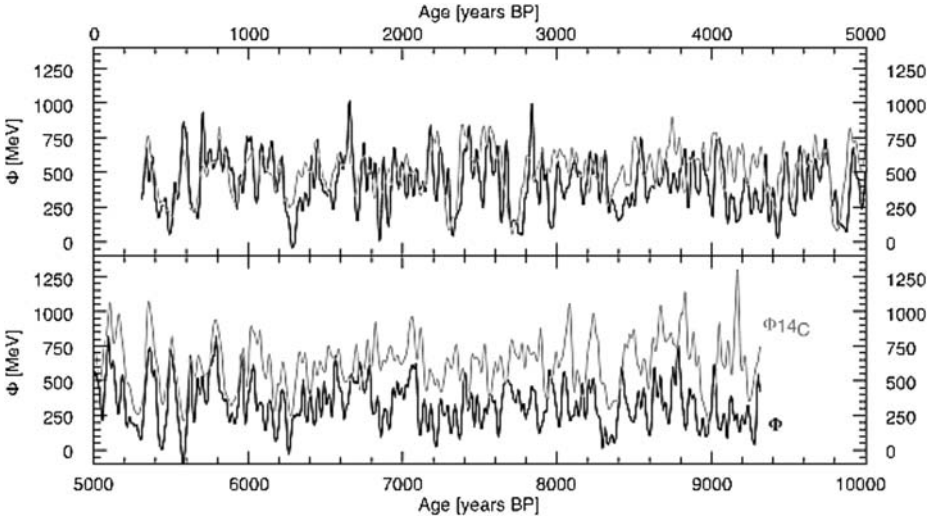

Fig. 3 Modulation of solar activity over a 9000 yr interval from 304 to 9315 BP (1646 AD to 7365 BC), as shown by the solar modulation potential Φ derived from ^{10}Be abundances in the GRIP ice core (*black line*) and that derived from the ^{14}C production rate in tree rings (*grey line*). Both records are high-pass filtered to reduce the effects of changes in the geomagnetic field and other long-term variations. The two records are initially in close agreement, though they gradually drift apart. (From Vonmoos et al. (2006))

examples of similar stars that are magnetically quiescent, and perhaps undergoing grand minima.

3 Helioseismology and Internal Properties of the Sun

Helioseismology uniquely is capable of imaging the solar interior. Waves that are essentially acoustic (modified by the stratification and by magnetic fields) are generated by turbulent motions in the upper part of the convection zone. These waves set up resonant global modes of the whole Sun, and the frequencies of these modes are used to infer properties of the solar interior. The local properties of the wave fields can also be used to infer subsurface conditions under localised features such as sunspots. In this section we summarise the helioseismic findings relevant to the solar dynamo.

Solar models indicate that in the outer envelope of the Sun the temperature gradient required to transport the observed luminosity by radiation is too steep to be stable, thus requiring this region to be convectively unstable. The transport of heat by convection is efficient and so the resulting stratification throughout the bulk of the convection zone in such models is very close to being adiabatic. These model predictions are confirmed by helioseismology. (A good review of the helioseismological results on the solar structure and on internal abundances is that of Basu and Antia (2007).) By measuring the gradient of sound speed, which changes rather abruptly where the heat transport changes between radiation and convection, Christensen-Dalsgaard et al. (1991) located the base of the convective envelope at a fractional radius $r/R_\odot = 0.713 \pm 0.003$. More recent work has confirmed this result with even higher precision (Basu 1998). Moreover, the location of the base of the convection zone appears to be independent of latitude (Basu and Antia 2001).

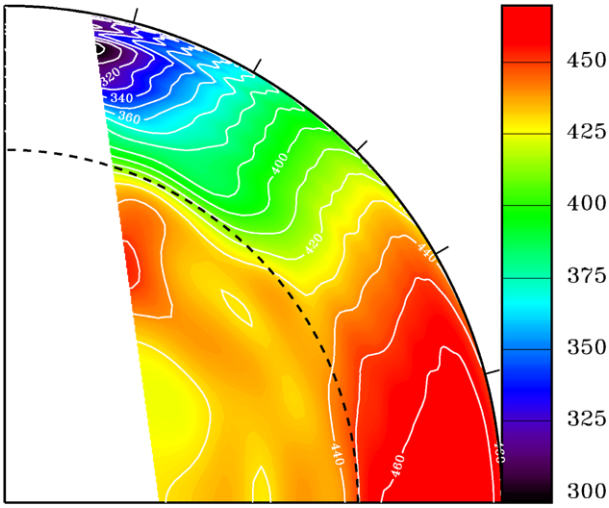

Fig. 4 The internal rotation of the Sun, as determined from observations by the MDI instrument on board the SOHO satellite. The solar equator is along the *horizontal axis*, the pole along the *vertical axis*. Values of $\Omega/2\pi$ are shown, in nHz. The *dashed line* indicates the base of the convection zone, and *tick marks* are at 15° intervals in latitude. (From Thompson et al. (2003))

This measurement actually refers to the extent of the essentially adiabatically stratified region: this may include a region of convective overshooting insofar as the overshoot region is also adiabatically stratified. Simplistic models of convective overshooting at the base of the solar convection zone would indicate that beneath the adiabatically stratified region there would then be a rather sharp adjustment to the subadiabatic gradient required to transport flux in the radiative interior. Such an abrupt transition would give rise to a characteristic signature in the p-mode frequencies, which can be used to measure the location and sharpness of the transition. Analyses of the observed frequencies have indicated that the extent of overshooting of this nature, if any, is small, no more than about one tenth of a pressure scale-height (Basu and Antia 1994; Monteiro et al. 1994; Christensen-Dalsgaard et al. 1995). One possibility is that the overshoot region is corrugated, which would appear smoother when spherically averaged as it is in this seismic analysis. Another possibility is that the overshoot produces a gentler subadiabatic transition: such models have been realised by Rempel (2004).

A major achievement of helioseismology has been to determine the internal rotation of the Sun over much of the solar interior (Fig. 4). It has long been known that at the surface the Sun rotates faster at the equator than at high latitudes: helioseismology has shown that this latitudinal differential rotation persists through the convection zone, with contours of isorotation mostly aligned nearly radially. There are two regions of prominent radial shear: a near-surface shear layer and a layer now known as the tachocline between the convection zone and the radiative interior beneath. The results for the radiative interior are consistent with solid-body rotation, though the rotation of the inner, energy-generating core is still uncertain. Hence the tachocline is not only a region of radial shear but also a transition layer from latitudinal differential rotation to latitudinally independent rotation.

The helioseismic results for the tachocline region are summarised in the review by Christensen-Dalsgaard and Thompson (2007). Kosovichev (1996) obtained the first quantitative results on the location and thickness of the tachocline. He adopted a functional dependence $\Phi(r)$ for the transition in depth of the latitudinal differential rotation of the form

$$\Phi(r) = \frac{1}{2}\left[1 + \mathrm{erf}\,(2(r - r_c)/w)\right] , \tag{1}$$

where erf is the error function. This provides for a continuous step function varying monotonically from zero to one, with a characteristic width w and centred on radial location $r = r_c$. Kosovichev found that the tachocline thus defined was centred at $r_c/R_\odot = 0.692 \pm 0.05$ and had width $w/R_\odot = 0.09 \pm 0.04$. More recent determinations are summarised by Christensen-Dalsgaard and Thompson (2007). The modern result of Basu and Antia (2003), converted to the same functional form as used by Kosovichev, is that the tachocline location varies from $r_c/R_\odot = 0.692 \pm 0.002$ at low latitudes to $r_c/R_\odot = 0.710 \pm 0.002$ at 60° latitude. At the same latitudes the corresponding tachocline widths are $w/R_\odot = 0.033 \pm 0.007$ and $w/R_\odot = 0.076 \pm 0.010$, respectively. Thus the tachocline is prolate and its width is greater at high latitudes: at low latitudes the tachocline is essentially wholly beneath the base of the convection zone, whereas at high latitudes it is roughly centred upon it.

Since aspects of the Sun vary with the solar cycle, it may be expected that the frequencies of the solar oscillations may vary also, at some level. Indeed the frequencies do vary with the solar cycle, and the variability has a very strong correlation with surface magnetic activity. Moreover, analyses of where the frequency variations originate indicate that the predominant causes of variability are located in the very superficial outer layers of the Sun (Antia et al. 2001). There has been no accepted direct seismological detection of magnetic fields in the tachocline or deeper interior. There are only upper bounds from helioseismology on the possible field strength in and near the tachocline, and these are of order several hundred kilogauss since that is the strength of field required at those depths to get any sensible frequency variation (Roberts and Campbell 1986). The presence of a magnetic field may be detectable through its effect on the thermal structure or on the bulk fluid motions. There have been hints of temporal variations in the wave-speed in the region (Chou and Serebryanskiy 2002; Baldner and Basu 2008). There have also been detections of modulations in the rotation rate just above and beneath the base of the tachocline, with an apparent quasi-periodic 1.3-year oscillation (Howe et al. 2000a). More certain are the relatively weak banded zonal flows, sometimes called torsional oscillations, that were first detected in the surface rotation rate and have since been revealed by helioseismology to penetrate at least one third of the way down through the convection zone, and possibly even to its base (Howe et al. 2000b; Vorontsov et al. 2002; Antia et al. 2008). At mid- and lower latitudes the banded flows migrate towards the equator over the solar cycle, in step with the active latitudes. At higher latitudes there is a poleward branch whose strength waxes and wanes with the cycle.

There is a greater possibility to detect magnetic fields directly in the upper convection zone than in the tachocline, since their dynamical importance is likely to be greater there. There is a persistent suggestion of a wave-speed anomaly at about 60° latitude and at a fractional radius of about $r/R_\odot = 0.9$, which could be caused by a fractional thermal perturbation of about 10^{-4} or a magnetic field strength of about 50 kG (Antia et al. 2000; Dziembowski et al. 2000). However, one might expect such a feature, if magnetic, to vary with the solar cycle, and little variation is discernable (Antia et al. 2003). Local helioseismic techniques such as ring analysis and time-distance helioseismology have detected flows and apparent thermal or magnetic anomalies under sunspots and active regions (e.g. Zhao and Kosovichev 2003; Haber et al. 2004).

Local techniques have also succeeded in revealing the subsurface meridional flow in the upper convection zone. Longitudinally and temporally averaged over a couple of months, the flows are poleward and largely steady from one epoch to another. Over the outer 15 Mm or so the flow speed is also largely independent of depth, at about 20–30 m s^{-1} (Haber et al. 2002). There are indications that in the approach to the last solar maximum the meridional flow in the northern hemisphere developed a counter-cell at mid-latitudes, with poleward flow persisting at those latitudes only in a superficial surface layer. Unfortunately the evidence from helioseismology for what the meridional flow is at greater depths is as yet inconclusive.

4 Dynamo Theory

Transient amplification of magnetic fields by sheared transverse flows (for example, by differential rotation) is distinct from dynamo action (Mestel 1999). There is also an important distinction between small-scale dynamo action, which allows the generation of disordered magnetic fields, averaging to zero, by turbulent convection, for instance in the solar photosphere (Cattaneo 1999; Vögler and Schüssler 2007), and the large-scale global dynamos with which we are concerned. In a planet like the Earth, with a resistive decay time $\tau_\eta \approx 7000$ yr in its core, a dynamo is needed to maintain a magnetic field that has been present for at least 3.5×10^9 yr and reverses on a time scale much longer than τ_η. In the Sun, by contrast, $\tau_\eta \approx 10^{10}$ yr, and yet there is a rapidly oscillating field. A straightforward hydromagnetic oscillator can be ruled out, for it would require variations in velocity with a 22 yr period, and there is no sign of them. The challenge to dynamo theory is then to explain both the cyclic variations of solar magnetic activity and their longer term modulation.

Cowling's theorem prohibits the maintenance of a purely axisymmetric magnetic field. Most dynamo models have treated azimuthally averaged fields, which can be split into a poloidal (meridional) component $\mathbf{B}_\mathrm{P} = \nabla \times (A\mathbf{e}_\phi)$ and a toroidal (azimuthal or zonal) component $\mathbf{B}_\mathrm{T} = B_\phi \mathbf{e}_\phi$. Then we expect the overall field to exhibit dipole symmetry, with B_ϕ antisymmetric about the equator. The principal mechanisms responsible for maintaining these fields are differential rotation, which generates $\mathbf{B}_\mathrm{T}$ from $\mathbf{B}_\mathrm{P}$ (as first realised by Ferraro, Walén and Cowling) and cyclonic eddies, which give rise to the α-effect (first introduced by Parker) and can generate a reversed $\mathbf{B}_\mathrm{P}$ from $\mathbf{B}_\mathrm{T}$. These processes have to compete with enhanced turbulent diffusion, the β-effect.

These models rely on mean field dynamo theory (Moffatt 1978; Parker 1979; Krause and Rädler 1980; Roberts 1994; Mestel 1999) in its simplest form. The idea here is to separate the magnetic field $\mathbf{B}$ and the velocity $\mathbf{U}$ into mean and fluctuating parts and to define a suitable averaging procedure, so that

$$\mathbf{B} = \overline{\mathbf{B}} + \mathbf{b}, \qquad \mathbf{U} = \overline{\mathbf{U}} + \mathbf{u}, \tag{2}$$

where $\langle \mathbf{b} \rangle = \langle \mathbf{u} \rangle = 0$. Then the averaged induction equation takes the form

$$\partial \overline{\mathbf{B}}/\partial t = \nabla \times (\overline{\mathbf{U}} \times \overline{\mathbf{B}}) + \nabla \times \boldsymbol{\mathcal{E}} + \eta \nabla^2 \overline{\mathbf{B}}\,, \tag{3}$$

where $\boldsymbol{\mathcal{E}} = \langle \mathbf{u} \times \mathbf{b} \rangle$ and the magnetic diffusivity η is assumed to be uniform. The procedure then is to assume a separation of length scales so that we can write

$$\mathcal{E}_i = \alpha_{ij} B_j + \beta_{ijk} \frac{\partial B_j}{\partial x_k} + \cdots. \tag{4}$$

If we separate α_{ij} into a symmetric part $\overline{\alpha}_{ij}$ and an antisymmetric part $\gamma_k \epsilon_{ijk}$, and consider pseudo-isotropic (non-mirror symmetric) turbulence we may then set $\overline{\alpha}_{ij} = \alpha \delta_{ij}$ and $\beta_{ijk} = \beta \epsilon_{ijk}$, with the pseudo-scalar $\alpha \neq 0$, whence (3) becomes

$$\partial \overline{\mathbf{B}}/\partial t = \nabla \times [(\overline{\mathbf{U}} + \boldsymbol{\gamma}) \times \overline{\mathbf{B}}] + \nabla \times (\alpha \overline{\mathbf{B}}) + (\beta + \eta) \nabla^2 \overline{\mathbf{B}}. \tag{5}$$

We expect that $\beta \gg \eta$, though molecular diffusion is none the less an essential part of the dynamo process. The α-effect typically depends on the presence of 'gyrotropic' turbulence with a net kinetic helicity $H = \langle \mathbf{u} \cdot \nabla \times \mathbf{u} \rangle$ in the small-scale motion. By further making a quasilinear approximation (first order smoothing), and assuming that the magnetic Reynolds

number $R_m \gg 1$ but the correlation time $\tau_c \ll \ell/u$, where u, ℓ are typical values of the velocity and length scale, it can be shown that $\alpha \approx -\frac{1}{3}\tau_c H$, while $\beta \approx \frac{1}{3}u^2\tau_c$ (e.g. Mestel 1999). Analogous results hold if $R_m \ll 1$, with τ_c replaced by the Ohmic decay time ℓ^2/η. If neither of these conditions is satisfied, there is no straightforward relationship between H and α.

The mean field induction (5) is widely used, though it should be borne in mind that this set of approximations can only be justified if there is indeed a separation of scales and either the magnetic Reynolds number $R_m \ll 1$ or $\tau_c \ll \ell/u$. Neither of the latter conditions is valid in the Sun. It follows then that we can only regard the mean field coefficients as physically plausible parametrisations of turbulent processes in the convection zone of a star.

5 Solar Dynamo Models

Mean field dynamo models rely on a number of ingredients, only two of which are constrained by observations. The most obvious contribution is from differential rotation, especially in the tachocline, where the angular velocity Ω is known from helioseismology, as explained in Sect. 3, and the shear gives rise to an ω-effect. Next is meridional flow, revealed by both helioseismic and surface measurements, which show an average poleward flow near the surface; continuity requires that this flow must reverse at some depth—and it may reverse more than once (Mitra-Kraev and Thompson 2007)—but its speed at the base of the convection zone is not determined. Then there is pumping of magnetic flux down the gradient of turbulent intensity and into the tachocline (the γ-effect: Tobias et al. 2001; Dorch and Nordlund 2001). (The actual roles of rotation, shear and penetration are discussed by Tobias elsewhere in these proceedings.) Finally come turbulent diffusion (the β-effect) and, most importantly, the crucial α-effect. Traditionally, this last has been ascribed to helicity produced by the effect of the Coriolis force acting on turbulent convection. Unfortunately, recent numerical studies by Cattaneo and Hughes (2006, 2008) have shown that α is negligible for convection in a rotating Boussinesq layer (though small-scale dynamo action is readily found). A further complication is that α is liable to catastrophic quenching in the nonlinear regime, so that its effective value becomes

$$\alpha_{\text{eff}} = \frac{\alpha}{1 + R_m^q B^2/B_0^2}\,, \tag{6}$$

with $0 < q \le 2$, where B_0 is the equipartition field strength (Vainshtein and Cattaneo 1992; Diamond et al. 2005; Hughes 2007a, 2007b). In the Sun, where $R_m \gg 1$, this would imply that α is quenched when the mean field B is less than 1 G. Numerical experiments on turbulence driven by helical forcing (Cattaneo and Hughes 1996) and on rotating compressible magnetoconvection (Ossendrijver et al. 2001) provide support for such catastrophic quenching, with $q = 1$. As an alternative to a distributed α-effect, it has therefore been proposed that the Coriolis force, acting on instabilities driven by magnetic buoyancy at or near the tachocline, may generate helicity and so produce a net α-effect (Ferriz-Mas et al. 1994; Schmitt et al. 1996; Brandenburg and Schmitt 1998; Thelen 2000a, 2000b).

It is generally accepted that the strong toroidal fields that emerge in active regions must be stored at the base of the convection zone, in or near the stably stratified tachocline, where the effects of turbulent pumping and flux expulsion can hold large-scale fields down against magnetic buoyancy. The tachocline is also the obvious site of the ω-effect, though it should be noted that the radial and latitudinal gradients of Ω are of comparable importance there

if, as seems likely, $|B_\theta| \gg |B_r|$. The near-surface radial gradient of Ω may also be at least locally significant.

Opinions differ as to the site and origin of the α-effect. Some hold that it is distributed throughout the convection zone, though this meets with the difficulties cited above, and Brandenburg (2005) has argued in favour of a near surface dynamo—see his contribution to these proceedings. Others incorporate a surface α-effect into a flux-transport dynamo, developed from Babcock's (1961) phenomenological model, as extended by Leighton (1969). The crucial component is the observed equatorward inclination of sunspot groups and active regions, caused presumably by Coriolis effects as flux rises upward through the convection zone. Combined with surface diffusion, ascribed to supergranular motion, and the observed meridional flow, this tilt can explain the observed evolution of photospheric magnetic fields, including reversals of the fields at the poles. The same meridional flows, acting as a conveyor belt and assisted by turbulent diffusion, then somehow manage to transport the reversed poloidal fields down to the tachocline, to serve as a seed for the next cycle (e.g. Dikpati and Charbonneau 1999; Choudhuri 2003). Dikpati and Gilman describe the most developed model of this process (Dikpati et al. 2004) in their contribution to these proceedings.

The alternative, following Parker (1993), is to adopt an interface model, with all the essential dynamo processes concentrated around the tachocline, where α can arise either from local convective motion or, more likely, from the nonlinear development of magnetic buoyancy instabilities (Tobias and Weiss 2007a). Photospheric fields then serve only as indicators of the action down below.

The solar dynamo is intrinsically nonlinear: in a nonlinear dynamo model, growth of the field must be limited by some dynamical process, whether by non-catastrophic α-quenching or modification of the ω-effect by the Lorentz force. Evidence for the latter process comes from the observed zonal shear flows (torsional oscillations), with a period of 11 years, described in Sect. 3, which seem to extend throughout the convection zone. Some models invoke microdynamic effects (Λ-quenching) of the Lorentz force (e.g. Kitchatinov et al. 1994; Rüdiger and Hollerbach 2004), while others rely on the macrodynamic Malkus-Proctor effect, driven by the mean field (e.g. Covas et al. 2004, Bushby 2005). Figure 5 shows the results of an idealised calculation, where the change in the sign of $\partial\omega/\partial r$ at mid-latitudes leads to a poleward branch of dynamo waves at high latitudes, as well as the stronger equatorward waves that give rise to sunspots; the high- and low-latitude branches of zonal shear flows are present in the observations. Periodic and aperiodic modulation of cyclic activity has been found in several nonlinear models with spherical geometry (e.g. Küker et al. 1999; Pipin 1999; Bushby 2006).

What then is the current status of mean-field dynamo models of the solar cycle? On the one hand, it is clear that there is a plethora of different models that yield plausible results when the arbitrary parameters are carefully tuned; it is reassuring that the various codes are now being benchmarked and compared (Jouve et al. 2008). On the other hand, although these models are certainly *instructive*, they are only *illustrative*.

6 Beyond Mean-Field Dynamos

The obvious route ahead is via direct numerical simulation but the alternative is to adopt a minimal formulation of the problem. Low-order models can exhibit generic patterns of behaviour, although they lack any detailed predictive power. The third-order normal form equations for a saddle-node/Hopf bifurcation are structurally stable and describe the relevant bifurcation sequence (Tobias et al. 1995). As an appropriate control parameter (the dynamo

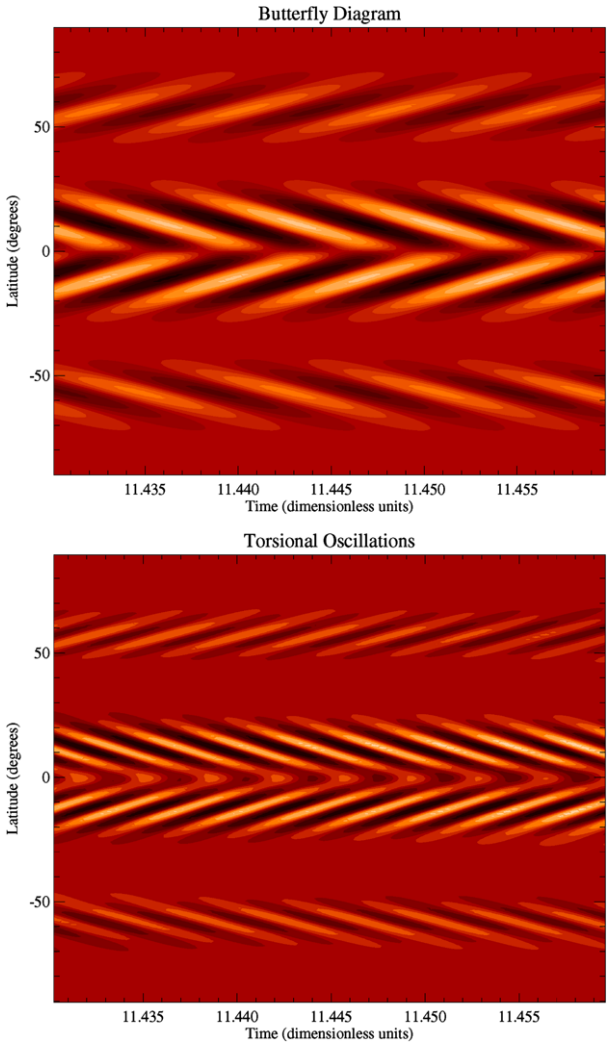

Fig. 5 Nonlinear cyclic behaviour for an idealised model of a spherical interface dynamo, incorporating the macrodynamic Malkus-Proctor effect. *Upper panel*: butterfly diagram showing toroidal fields of opposite signs (with dipole symmetry) at the base of the convection zone. *Lower panel*: the corresponding zonal shear flows (torsional oscillations) with twice the frequency of the magnetic cycle. Note the presence of a polar branch in each panel. (From Bushby (2005))

number D) is increased, there is a Hopf bifurcation from a purely hydrodynamic field-free state to a periodic solution, with trajectories that lie on a limit cycle in phase space; this is followed by a second bifurcation, leading to doubly periodic modulated solutions, with trajectories lying on a two-torus; as D is further increased, the torus is destroyed and a complicated series of bifurcations result in the appearance of chaotically modulated oscillations. The same pattern of *in-out* intermittency (Ashwin et al. 1999; Covas et al. 2001) appears in mean-field dynamo models too (Tobias 1996, 1997; Beer et al. 1998). There it is further complicated by transitions between solutions with dipole and quadrupole symmetry (i.e. with B_ϕ symmetric or antisymmetric about the equator), which can be represented in an extended sixth-order system (Knobloch et al. 1998). In this low-order model the cycles are spatially asymmetric as the dynamo emerges from a grand minimum; the same effect appears in mean-field models (Beer et al. 1998)—and this is just what happened at the end of the Maunder Minimum in 1700 (Ribes and Nesme-Ribes 1993). In-out intermittency differs

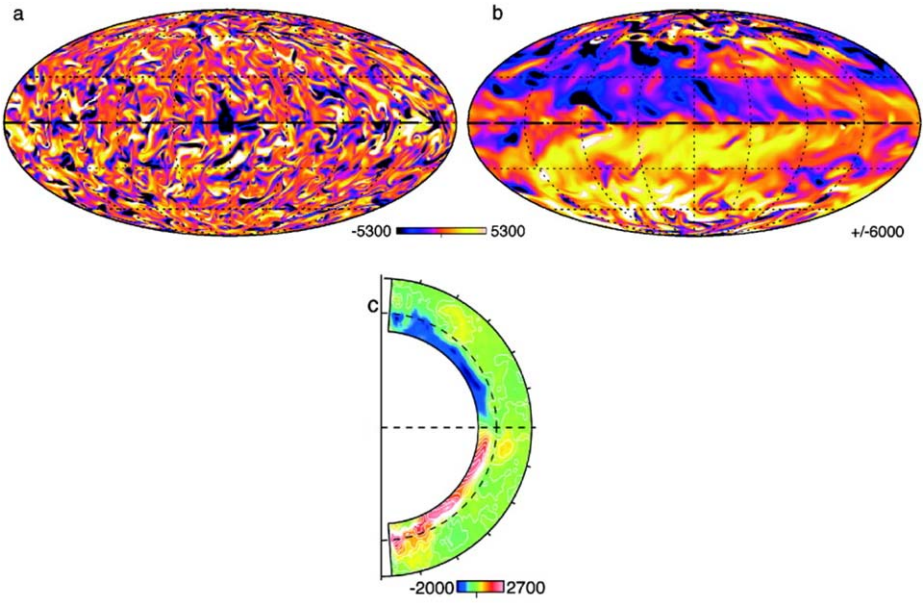

Fig. 6 The toroidal field in a recent computational model of the solar dynamo. Images of the azimuthal field (**a**) in the middle of the convection zone and (**b**) in the region of convective overshoot, both as Mollweide projections, with (**c**) the temporally averaged axisymmetric component of the toroidal field, which has approximate dipole symmetry. (From Browning et al. (2006))

from *on-off* intermittency (Tobias and Weiss 2007b); in the latter case the amplitude of the aperiodic magnetic cycle is controlled by an independent chaotic or stochastic oscillator that is unaffected by magnetic fields (Platt et al. 1993a, 1993b; Brandenburg and Spiegel 2008), as discussed by Spiegel elsewhere in these proceedings.[1]

For more detailed and realistic models one must turn to large-scale computation. This approach has successfully reproduced many key features of the geodynamo (as explained by Christensen et al. in these proceedings). Numerical studies of the solar dynamo were pioneered by Gilman (1983) a quarter-century ago. Although his Boussinesq models could not reproduce the butterflies in Fig. 1, they did establish the generic sensitivity of dynamos to the choice of parameters in the equations: small changes can switch solutions from steady to oscillatory behaviour, from poleward to equatorward travelling waves, or from dipole to quadrupole symmetry. The anelastic approximation was first included by Glatzmaier (1985) and since then anelastic models of the convection zone and of the solar dynamo have grown increasingly elaborate and sophisticated (e.g. Brun et al. 2004; Browning et al. 2006). So far, however, these numerical simulations have all been carried out at parameter values that are far from those that prevail in the solar interior. The current state of the art is illustrated in Fig. 6. Although this model does not exhibit oscillatory behaviour, there is a predominantly azimuthal field, with dipole symmetry, at the base of the convection zone. This is clearly the future: soon we may expect to see more developed models that actually reproduce the known behaviour of the solar cycle.

[1] Mathematically, the latter systems are said to have normal parameters and skew-product structure.

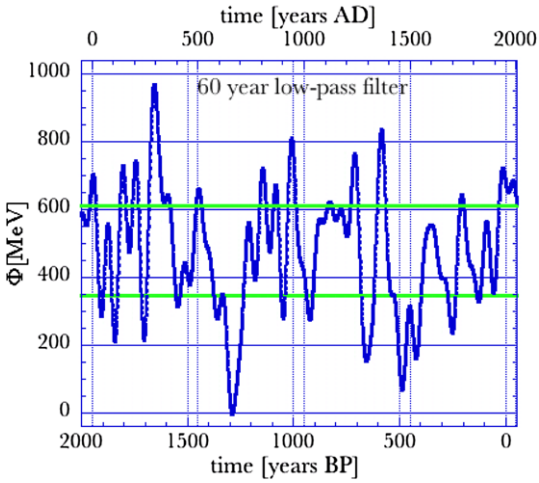

Fig. 7 A composite time series for the solar modulation potential Φ over the past 2000 years. Direct measurements of cosmic rays from neutron monitors (Usoskin et al. 2005) have been merged with ^{10}Be records from the South Pole (McCracken et al. 2004) and Greenland (Vonmoos et al. 2006) and the combined series has been detrended and smoothed (Abreu et al. 2008). *Horizontal lines* define levels of grand maxima and grand minima. The current grand maximum is clearly visible, as is the Maunder Minimum. (Courtesy of J.A. Abreu)

Given such a successful computational model, with both the velocity and the magnetic field known, we might attempt to relate it to mean-field dynamo theory by asking the following three questions. Can one compute a meaningful α from the statistics of the known velocity field $\mathbf{u}$? Can one construct a mean-field dynamo model that mimics the results? Does the α-effect indeed capture the essential physics? We forecast that, although the answer to the first question will be *no*, the other two will be answered in the affirmative.

7 Predicting the Future

Solar activity is apparently chaotic, and thus it is intrinsically difficult to predict its future behaviour. Most efforts have been concerned with the immediate short-term problem of forecasting the next cycle, relying on precursor methods, or on reconstructing the attractor, or a combination of the two. These approaches discussed by Hathaway are in these proceedings.

The long-term problem involves attempts to forecast trends in the envelope of cyclic activity. For the past 80 years the Sun has been experiencing an episode of extreme activity, as shown by the ^{14}C and ^{10}Be proxy records (e.g. Solanki et al. 2004; Steinhilber et al. 2008). The GRIP ice-core provides a ^{10}Be record, already illustrated in Fig. 3, that extends back for more than 9 millennia, and Fig. 7 shows a smoothed composite record for the past 2000 yr. The current grand maximum is unusually high but it is certainly not unprecedented: Steinhilber et al. (2008) have identified 25 comparable events in the GRIP record. It is natural to ask how long this episode can be expected to last. Given its present lifetime of 80 yr, Abreu et al. (2008) estimate a total life expectancy of 95 yr, implying that this grand maximum will terminate within the next few cycles. There is then a possibility that it may be followed by a grand minimum as deep as the Maunder Minimum of the seventeenth century. If so, we may expect a detectable cooling effect on the Earth's climate, but one too small to compensate for global warming caused by anthropogenic greenhouse gases.

References

J.A. Abreu, J. Beer, F. Steinhilber, S.M. Tobias, N.O. Weiss, Geophys. Res. Lett. (2008, in press)

H.M. Antia, S.M. Chitre, M.J. Thompson, Astron. Astrophys. **360**, 335 (2000)
H.M. Antia, S. Basu, F. Hill, R. Howe, R.W. Komm, J. Schou, Mon. Not. R. Astron. Soc. **327**, 1029 (2001)
H.M. Antia, S.M. Chitre, M.J. Thompson, Astron. Astrophys. **399**, 329 (2003)
H.M. Antia, S. Basu, S.M. Chitre, Astrophys. J. **681**, 680 (2008)
P. Ashwin, E. Covas, R. Tavakol, Nonlinearity **12**, 562 (1999)
H.W. Babcock, Astrophys. J. **133**, 572 (1961)
C.S. Baldner, S. Basu, Astrophys. J. (2008, in press)
S. Basu, Mon. Not. R. Astron. Soc. **298**, 719 (1998)
S. Basu, H.M. Antia, Mon. Not. R. Astron. Soc. **269**, 1137 (1994)
S. Basu, H.M. Antia, Mon. Not. R. Astron. Soc. **324**, 498 (2001)
S. Basu, H.M. Antia, Astrophys. J. **585**, 553 (2003)
S. Basu, H.M. Antia, Phys. Rep. **457**, 217 (2007)
J. Beer, Space Sci. Rev. **94**, 53 (2000)
J. Beer, S.M. Tobias, N.O. Weiss, Sol. Phys. **181**, 237 (1998)
A. Brandenburg, Astrophys. J. **625**, 539 (2005)
A. Brandenburg, D. Schmitt, Astron. Astrophys. **338**, L55 (1998)
A. Brandenburg, E.A. Spiegel, Astron. Nachr. **329**, 351 (2008)
M.K. Browning, M.S. Miesch, A.S. Brun, J. Toomre, Astrophys. J. **648**, L157 (2006)
A.S. Brun, M.S. Miesch, J. Toomre, Astrophys. J. **613**, 1253 (2004)
P.J. Bushby, Astron. Nachr. **326**, 218 (2005)
P.J. Bushby, Mon. Not. R. Astron. Soc. **371**, 772 (2006)
F. Cattaneo, Astrophys. J. **515**, L39 (1999)
F. Cattaneo, D.W. Hughes, Phys. Rev. E **54**, 4532 (1996)
F. Cattaneo, D.W. Hughes, J. Fluid Mech. **553**, 401 (2006)
F. Cattaneo, D.W. Hughes, J. Fluid Mech. **594**, 495 (2008)
P. Charbonneau, Living Rev. Sol. Phys. **2**, 2 (2005). www.livingreviews.org/lrsp-2005-2
D.-Y. Chou, A. Serebryanskiy, Astrophys. J. **578**, L157 (2002)
A.R. Choudhuri, in *Dynamic Sun*, ed. by B.N. Dwivedi (Cambridge University Press, Cambridge, 2003), p. 103
J. Christensen-Dalsgaard, M.J. Thompson, in *The Solar Tachocline*, ed. by D.W. Hughes, R. Rosner, N.O. Weiss (Cambridge University Press, Cambridge, 2007), p. 53
J. Christensen-Dalsgaard, D.O. Gough, M.J. Thompson, Astrophys. J. **378**, 413 (1991)
J. Christensen-Dalsgaard, M.J.P.F.G. Monteiro, M.J. Thompson, Mon. Not. R. Astron. Soc. **276**, 283 (1995)
E. Covas, R. Tavakol, P. Ashwin, A. Tworkowski, J. Brooke, Chaos **11**, 404 (2001)
E. Covas, D. Moss, R. Tavakol, Astron. Astrophys. **416**, 775 (2004)
P.E. Damon, C.P. Sonett, in *The Sun in Time*, ed. by C.P. Sonett, M.S. Giampapa, M.S. Matthews (University of Arizona Press, Tucson, 1991), p. 360
P.H. Diamond, D.W. Hughes, E. Kim, in *Fluid Dynamics and Dynamos in Astrophysics and Geophysics*, ed. by A.M. Soward, C.A. Jones, D.W. Hughes, N.O. Weiss (CRC Press, Boca Raton, 2005), p. 145
M. Dikpati, P. Charbonneau, Astrophys. J. **518**, 508 (1999)
M. Dikpati, G. de Toma, P.A. Gilman, C.N. Arge, O.R. White, Astrophys. J. **601**, 1136 (2004)
S.B.F. Dorch, A. Nordlund, Astron. Astrophys. **365**, 562 (2001)
W.A. Dziembowski, P.R. Goode, A.G. Kosovichev, J. Schou, Astrophys. J. **537**, 1026 (2000)
J.A. Eddy, Science **192**, 1189 (1976)
A. Ferriz-Mas, D. Schmitt, M. Schüssler, Astron. Astrophys. **289**, 949 (1994)
P.A. Gilman, Astrophys. J. **53**, 243 (1983)
G.A. Glatzmaier, Astrophys. J. **291**, 300 (1985)
D.A. Haber, B.W. Hindman, J. Toomre, M.J. Thompson, Sol. Phys. **220**, 371 (2004)
D.A. Haber, B.W. Hindman, J. Toomre, R.S. Bogart, R.M. Larsen, F. Hill, Astrophys. J. **570**, 855 (2002)
R. Howe, J. Christensen-Dalsgaard, F. Hill, R.W. Komm, R.M. Larsen, J. Schou, M.J. Thompson, J. Toomre, Science **287**, 2456 (2000a)
R. Howe, J. Christensen-Dalsgaard, F. Hill, R.W. Komm, R.M. Larsen, J. Schou, M.J. Thompson, J. Toomre, Astrophys. J. **533**, L163 (2000b)
D.W. Hughes, in *Solar Tachocline*, ed. by D.W. Hughes, R. Rosner, N.O. Weiss (Cambridge University Press, Cambridge, 2007a), p. 275
D.W. Hughes, in *Mathematical Aspects of Natural Dynamos*, ed. by E. Dormy, A.M. Soward (CRC Press, Boca Raton, 2007b), p. 81
L. Jouve, A.S. Brun, R. Arlt, A. Brandenburg, M. Dikpati, A. Bonanno, et al. Astron. Astrophys. **483**, 949 (2008)
L.L. Kitchatinov, G. Rüdiger, M. Küker, Astron. Astrophys. **292**, 125 (1994)
E. Knobloch, S.M. Tobias, N.O. Weiss, Mon. Not. R. Astron. Soc. **297**, 1123 (1998)

A.G. Kosovichev, Astrophys. J. **469**, L61 (1996)
F. Krause, K.-H. Rädler, *Mean-Field Magnetohydrodynamics and Dynamo Theory* (Akademie, Berlin, 1980)
M. Küker, R. Arlt, G. Rüdiger, Astron. Astrophys. **343**, 977 (1999)
R.B. Leighton, Astrophys. J. **156**, 1 (1969)
K.G. McCracken, F.B. McDonald, J. Beer, G. Raisbeck, F. Yiou, J. Geophys. Res. **109**, A12103 (2004)
L. Mestel, *Stellar Magnetism* (Clarendon Press, Oxford, 1999)
U. Mitra-Kraev, M.J. Thompson, Astron. Nachr. **328**, 1009 (2007)
H.K. Moffatt, *Magnetic Field Generation in Electrically Conducting Fluids* (Cambridge University Press, Cambridge, 1978)
M.J.P.F.G. Monteiro, J. Christensen-Dalsgaard, M.J. Thompson, Astron. Astrophys. **283**, 247 (1994)
M.A.J.H. Ossendrijver, Astron. Astrophys. Rep. **11**, 287 (2003)
M. Ossendrijver, M. Stix, A. Brandenburg, Astron. Astrophys. **376**, 713 (2001)
E.N. Parker, *Cosmical Magnetic Fields: Their Origin and Their Activity* (Clarendon Press, Oxford, 1979)
E.N. Parker, Astrophys. J. **408**, 707 (1993)
N. Platt, E. Spiegel, C. Tresser, Phys. Rev. Lett. **70**, 279 (1993a)
N. Platt, E. Spiegel, C. Tresser, Geophys. Astrophys. Fluid Dyn. **73**, 146 (1993b)
V.V. Pipin, Astron. Astrophys. **346**, 295 (1999)
M. Rempel, Astrophys. J. **607**, 1046 (2004)
J.C. Ribes, E. Nesme-Ribes, Astron. Astrophys. **276**, 549 (1993)
B. Roberts, W.R. Campbell, Nature **323**, 603 (1986)
P.H. Roberts, in *Lectures on Solar and Planetary Dynamos*, ed. by M.R.E. Proctor, A.D. Gilbert (Cambridge University Press, Cambridge, 1994), p. 1
G. Rüdiger, R. Hollerbach, *The Magnetic Universe* (Wiley-VCH, Weinheim, 2004)
D. Schmitt, M. Schüssler, A. Ferriz-Mas, Astron. Astrophys. **311**, L1 (1996)
S.K. Solanki, B. Inhester, M. Schüssler, Rep. Prog. Phys. **69**, 563 (2006)
S.K. Solanki, I.G. Usoskin, M. Kromer, M. Schüssler, J. Beer, Nature **431**, 1084 (2004)
F. Steinhilber, J.A. Abreu, J. Beer, Atmos. Space Sci. Trans. **4**, 1 (2008)
M. Stuiver, T.F. Braziunas, Holocene **3**, 289 (1993)
J.-C. Thelen, Mon. Not. R. Astron. Soc. **315**, 155 (2000a)
J.-C. Thelen, Mon. Not. R. Astron. Soc. **315**, 165 (2000b)
J.H. Thomas, N.O. Weiss, *Sunspots and Starspots* (Cambridge University Press, Cambridge, 2008)
M.J. Thompson, J. Christensen-Dalsgaard, M.S. Miesch, J. Toomre, Annu. Rev. Astron. Astrophys. **41**, 599 (2003)
S.M. Tobias, Astron. Astrophys. **307**, L21 (1996)
S.M. Tobias, Astron. Astrophys. **322**, 1007 (1997)
S.M. Tobias, Philos. Trans. R. Soc. Lond. **360**, 2741 (2002)
S.M. Tobias, N.O. Weiss, in *The Solar Tachocline*, ed. by D.W. Hughes, R. Rosner, N.O. Weiss (Cambridge University Press, Cambridge, 2007a), p. 319
S.M. Tobias, N.O. Weiss, in *Mathematical Aspects of Natural Dynamos*, ed. by E. Dormy, A.M. Soward (CRC Press, Boca Raton, 2007b), p. 281
S.M. Tobias, N.O. Weiss, V. Kirk, Mon. Not. R. Astron. Soc. **273**, 1150 (1995)
S.M. Tobias, N.H. Brummell, T.L. Clune, J. Toomre, Astrophys. J. **549**, 1183 (2001)
I.G. Usoskin, K. Alanko-Huotari, G.A. Kovaltsov, K. Mursula, J. Geophys. Res. **110**, A12108 (2005)
S.I. Vainshtein, F. Cattaneo, Astrophys. J. **393**, 165 (1992)
M. Vonmoos, J. Beer, R. Muscheler, J. Geophys. Res. **111**, A10105 (2006)
S.V. Vorontsov, J. Christensen-Dalsgaard, J. Schou, V.N. Strakhov, M.J. Thompson, Science **296**, 101 (2002)
A. Vögler, M. Schüssler, Astron. Astrophys. **465**, L43 (2007)
G. Wagner, J. Beer, J. Masarik, P.W. Kubik, W. Mende, C. Laj, G.M. Raisbeck, F. Yiou, Geophys. Res. Lett. **28**, 303 (2001)
J. Zhao, A.G. Kosovichev, Astrophys. J. **591**, 446 (2003)