Article:
Weiss, NO and Tobias, SM (2016) Super-modulation of the Sun's magnetic activity: the effects of symmetry changes. Monthly Notices of the Royal Astronomical Society, 456 (3). pp. 2654-2661. ISSN 0035-8711

https://doi.org/10.1093/mnras/stv2769

Mon. Not. R. Astron. Soc. **000**, 1–8 (2015) Printed 20 November 2015 (MN LATEX style file v2.2)

Super-modulation of the Sun's magnetic activity: the effects of symmetry changes

N. O. Weiss[1] and S. M. Tobias[2]
[1] *Department of Applied Mathematics & Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, UK*
[2] *Department of Applied Mathematics, University of Leeds, Leeds LS2 9JT, UK*

Accepted 2015 December 31. Received 2015 June 30; in original form 2015 March 1

ABSTRACT
In this paper we argue that the solar activity record, as revealed by telescopic observations and proxy data from the abundances of cosmogenic isotopes, is consistent with the action of a deterministic nonlinear chaotic dynamo. In particular we claim that the long timescale 'super-modulation' apparent in the isotopic record can be ascribed to switching of the dynamo between two different modulational patterns. The first (which is currently in operation) involves deep grand minima and occasional changes in symmetry triggered by these minima. The second, which exhibits only weak modulation and no grand minima, arises as a consequence of symmetry breaking. These processes are demonstrated for highly idealized simple models of the nonlinear dynamo equations.

Key words: dynamo – sun: sunspots – activity – magnetic fields.

1 INTRODUCTION

Although the eruption of individual active regions at the solar surface can only be modelled as a stochastic process, the overall pattern of cyclic magnetic activity shows systematic reversals of the toroidal field that demand a different approach. As first proposed by Zeldovich, Ruzmaikin and Sokoloff (1983), the smoothed cycles have to be treated as nonlinear oscillations – and their aperiodic modulation should be represented within the framework of nonlinear dynamics, rather than as a consequence of random perturbations (e.g. Weiss 1985, 2010). The earliest telescopic observations of sunspots in the seventeenth century revealed considerable activity, which was interrupted by the Maunder Minimum (Eddy 1976; Thomas & Weiss 2008), as recently confirmed by the detailed study of Usoskin et al. (2015). Proxy measurements of the abundances of the cosmogenic isotopes ^{14}C (in tree rings) and ^{10}Be (in polar ice cores) confirm that similar grand minima have recurred aperiodically over the past 10,000 years (e.g. Usoskin 2013). Simplified nonlinear models of the solar cycle, governed by ordinary differential equations, imply that this modulation is an example of deterministic chaos with sensitive dependence on initial conditions (e.g. Tobias, Weiss & Kirk 1995; Knobloch, Tobias & Weiss 1998; Tobias 2002; Spiegel 2008; Weiss & Thompson 2008), while similar patterns of behaviour also appear for models based on nonlinear partial differential equations (e.g. Tobias 1997; Küker, Arlt & Rüdiger 1999; Pipin 1999; Weiss 2005, 2010; Bushby 2006; Tobias & Weiss 2007a,b).

Solutions that are governed by partial differential equations follow chaotic trajectories in the phase space of the system and stochastic perturbations (provided that they are not too large) merely shift a solution to the neighbourhood of a nearby chaotic trajectory, representing a chaotic solution that *shadows* the actual behaviour of the system (Ott 1993). Within the phase space there are several invariant subspaces: one is the (nonmagnetic) hydrodynamic subspace, which becomes unstable to magnetohydrodynamic (MHD) perturbations if dynamo action sets in. In the simplest case, where the large-scale fields are rendered axially symmetric, there are two distinct invariant MHD subspaces, one with *dipole* symmetry and toroidal magnetic fields that are antisymmetrical about the equator, and the other with *quadrupole* symmetry and toroidal magnetic fields that are symmetrical. Both of these subspaces are contained in a general MHD subspace that includes large-scale fields for which these symmetries are broken (Tobias & Weiss 2007b). Dipole or quadrupole symmetry may then give way to more general mixed-mode solutions that lack any form of symmetry about the equator.

The Sun's magnetic activity is typical of that in a slowly rotating late-type star with a deep convective zone (Baliunas et al. 1995; Wright 2004; Santos et al. 2010). The 11-year Schwabe activity cycle corresponds to distributions of sunspots that are approximately symmetrical about the solar equator (e.g. Arlt et al. 2013). Strict symmetry is violated owing both to the quasi-random emergence of sunspots and to the chaotic behaviour of the overall cycle. Similarly, although the Sun's poloidal magnetic field is, on average, an-

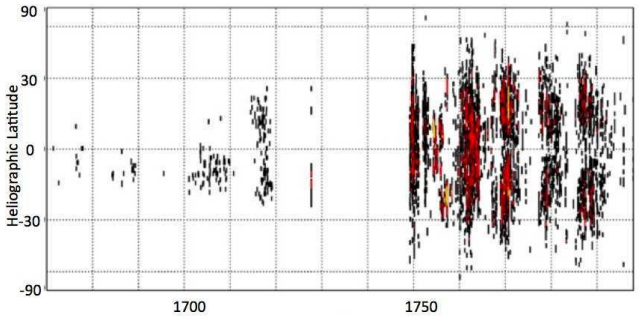

Figure 1. Observed incidence of sunspots from 1670 to 1800, shown as an incomplete butterfly diagram. Results up to 1719 cover the latter part of the Maunder Minimum, when there is a hemispheric pattern, and the first weak normal cycle (Ribes & Nesme-Ribes 1993). The observations by Staudacher (Arlt 2009a) combined with those of Zucconi (Cristo et al. 2011) and Hamilton (Arlt 2009b) show the return to cyclic behaviour up to the Dalton Minimum at the end of the 18th century. Note, however, the non-standard pattern of Cycle 1, around 1760: like Cycle 0, spots are most frequent at the equator; they appear more or less simultaneously at all latitudes and there is no evidence of the characteristic butterfly pattern. Arlt suggests that this results from a brief excursion from dipolar to quadrupolar symmetry of the Sun's magnetic field. (After Arlt & Weiss 2015.)

tisymmetrical about the equator, this antisymmetry breaks down around the times of sunspot maximum, when the dipole field reverses. As is only to be expected, the fields near the two poles become decoupled and one reverses before the other, leaving a poloidal field that is temporarily quadrupolar – while the toroidal field remains very definitely dipolar (DeRosa, Brun & Hoeksema 2012). Nevertheless, the long-term dipole symmetry is fairly soon restored.

The temporal behaviour of sunspots follows Maunder's characteristic butterfly pattern, with fields whose dipole symmetry reverses from one cycle to the next, giving rise to the 22-year Hale magnetic cycle. This pattern was interrupted during the Maunder Minimum in the seventeenth century: scarcely any spots were seen between 1660 and 1671, and almost all those detected by regular observations at the Paris Observatory between 1670 and 1715 lay in the southern hemisphere of the Sun (Ribes and Nesme-Ribes 1993), as shown in Fig. 1. Such *hemispheric* behaviour is characteristic of a mixed-mode pattern (Sokoloff & Nesme-Ribes 1994). Since then, the averaged sunspot distributions have normally been symmetric about the solar equator, with activity starting at intermediate latitudes and spreading poleward as well as towards the equator. The dipole symmetry implies that there are scarcely any spots at the equator. Between 1750 and 1775, however, the spot patterns are anomalous: spots appear more or less simultaneously at all latitudes and are actually most frequent in the vicinity of the equator, as shown by Staudacher's observations in Fig. 1. This suggests a brief episode with quadrupole symmetry and sunspot fields of the same sign in both hemispheres (Arlt 2009a). Thus there is evidence of symmetry changes in the actual sunspot record. Field reversals and associated symmetry changes arise also in several related systems, notably the geodynamo (e.g. Roberts and Glatzmaier 2000; Pétrélis et al. 2009; Pétrélis & Fauve 2010; Gallet et al. 2012) and the VKS experiment, where dynamo action is driven by ferromagnetic rotors in a tank of molten sodium (Ravelet et al. 2008; Monchaux et al. 2009; Berhanu et al. 2010).

These symmetry changes can conveniently be explored by studying simple illustrative models, based on appropriate low-order systems of ordinary differential equations. Such models will be discussed in Section 3, where we exhibit examples of cyclic behaviour with dipole, quadrupole and mixed-mode symmetry and demonstrate not only chaotic modulation but also transitions from one symmetry to another. These model systems also display two different types of behaviour (Knobloch, Tobias & Weiss 1998). In the first, the basic cycle is modulated in amplitude without the appearance of grand minima, while records of activity in the second type are punctuated by grand minima, when activity becomes extremely small. Similar effects can also be found not only in mean-field models based on partial differential equations but also in direct numerical simulations, albeit for parameters far removed from those applicable to the solar interior (Raynaud & Tobias 2015).

Our aim in this paper is to relate features of the long-term behaviour of the Sun's magnetic field over the past 10,000 years to such symmetry changes. With this in mind, we first summarize, in the next section, the history of solar activity as deduced from the concentrations of the cosmogenic isotopes ^{14}C, in tree trunks, and ^{10}Be, in polar ice cores (Steinhilber et al. 2012). These concentrations can be measured with great precision using mass spectrometry or, in the case of ^{10}Be, the much more accurate technique of accelerator mass spectrometry. Here we emphasize the contrast between episodes that show recurrent Maunder-like grand minima and intervals when the modulation is much weaker and grand minima are absent (McCracken et al. 2013a,b). In the final section, we argue that this long-term pattern of *super-modulation* is a consequence of symmetry changes, most probably from a predominantly dipolar to a mixed-mode cycle. Correspondingly, there are switches from the second type of modulation to the first and back again. This interpretation of the record is the principal result of our paper, some of whose contents were announced in a preliminary report by Arlt & Weiss (2014).

2 THE PALAEOMAGNETIC HISTORY OF THE SOLAR CYCLE

In this section we review the long-term records of the abundances of cosmogenic isotopes, specifically of ^{14}C (in tree trunks) and of ^{10}Be (in polar ice cores, both from Greenland and from the Antarctic), over the Holocene period from 9400 BP to the present (Steinhilber et al. 2010, 2012). These records are summarized in Fig. 1 of McCracken et al. (2013b). They are based on measurements of the normalized cosmogenic production rate of ^{10}Be from the GRIP icecore in Greenland (Yiou et al. 1997; Muscheler et al. 2004), combined with measurements of the Greenland NGRIP, Milcent and Dye-3 icecores, together with the EDML, South Pole and Dome Fuji ^{10}Be records from Antarctica, and the production rate Q^{14}C of ^{14}C derived from the INTCAL09 ^{14}C record (Reimer et al. 2009). These datasets are combined using a Principal Components Analysis (Steinhilber at al. 2012; Abreu et al. 2013) to eliminate the effects of local cli-

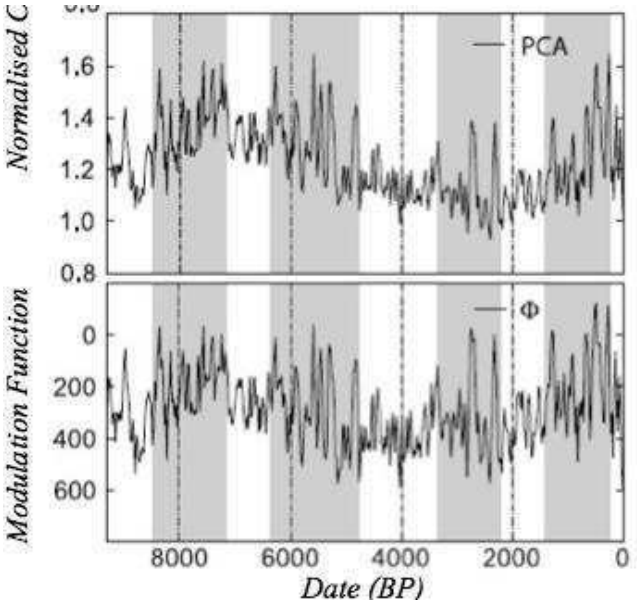

Figure 2. The cosmogenic isotope record since 9400 BP (where 0 BP is defined as 1950 CE) expressed relative to the interval 1944–1988. The upper panel shows the combined normalized cosmogenic production rate of ^{10}Be and ^{14}C as estimated after a Principal Components Analysis. The lower panel shows the corresponding variation of the Modulation Function Φ, after correction for estimated changes in the geomagnetic field; note that the scale is inverted so as to match the variation in radionuclide production. Grand minima appear as clusters of upward spikes, lying within the shaded regions, which are separated by episodes of diminished modulation. These clusters are confined to the intervals that are shaded in the figure. Although the grand minima are not strictly periodic, they do recur with a well-defined mean period of around 200 years (the de Vries cycle). (After McCracken et al. 2013b.)

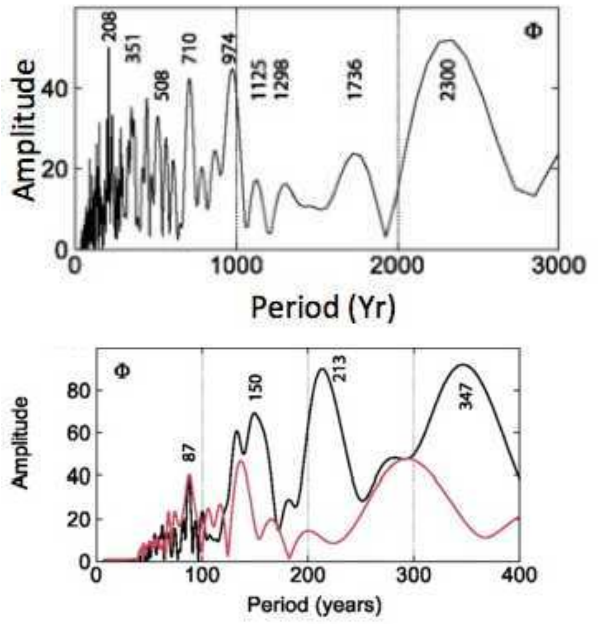

Figure 3. Fourier analysis of the modulation function Φ in Fig. 2. Upper panel: Fourier amplitudes as a function of their periods for the entire interval. Note the sharp line at the de Vries period of 208 years, and the broad, inadequately resolved peak around the Hallstatt period of 2300 years. Lower panel: the contrast between the shaded intervals (with strong modulation) and the weakly modulated intervening intervals, without grand minima in Fig. 2. The upper (black) curve shows the spectrum of Φ over the interval from 6300 BP to 4300 BP. This interval is shaded in Fig. 3 and it includes between four and six grand minima. The lower (red in the online version) curve represents the spectrum for the interval from 4700 BP to 3500 BP, which is devoid of any grand minima. In both panels the various peaks are labelled by their periods in years. (After McCracken et al. 2013b)

matic variations, and the results are summarized in Fig. 2. The upper panel shows the result of the Principal Components Analysis, which provides the best available measure of the normalized cosmogenic production of radionuclides (McCracken et al. 2013b). Most of the long-term variation is probably caused by changes in the Earth's magnetic field. The lower panel shows (on an inverted scale) the Modulation Function Φ, a measure of the strength of the interplanetary field that represents the energy lost by a cosmic ray proton that reaches Earth at 1 AU (Vonmoos, Beer & Muscheler 2006). This record is corrected for estimated changes in the geomagnetic field (Yang et al. 2000). Both panels display 22-yr averages and therefore show only longer-term modulation of the underlying Schwabe and Hale cycles.

The long-term records of cosmogenic abundances that are compared in Fig. 2 reveal recurrent Maunder-like grand minima, with a typical spacing of around 200 years (the de Vries cycle) but these grand minima are themselves confined to clusters (Damon & Sonett 1991; Stuiver & Braziunas 1993). The most recent of these clusters includes the familiar Maunder, Spörer, Wolf and Oort episodes, while clusters typically recur after around 2300 years (the Hallstatt cycle). Between these episodes the modulation is much weaker and grand minima are absent (McCracken et al. 2013a,b). We shall claim later that this long-term super-modulation results from symmetry changes, most likely from dipolar to mixed-mode cycles.

Here we focus on the information contained in the records of cosmogenic abundances in Fig. 2. First of all, we show in Fig. 3 a periodic analysis of the entire record, presented as Fourier amplitudes. This spectrum shows some robust periodicities, most notably the de Vries period of 208 years (which corresponds to the spacing between grand minima). This is a very striking feature of the entire record (McCracken et al. 2013b). Given its finite duration, the broad feature in the Fourier spectrum that is centred on a period of 2300 years is clearly not adequately resolved. Nevertheless, it does match the Hallstatt period that is characteristic of super-modulation.

2.1 Modulation of cyclic activity through most of the Holocene

A clearer impression of changes in behaviour can be gained by separating out the episodes that contain grand minima, which are shaded in Fig. 2, and contrasting them with the unshaded intervals where grand minima are lacking. McCracken et al. (2013b) focus on the interval 6300–4700 BP, which includes at least four very prominent grand minima, and the subsequent interval 4700–3500 BP, during which

there are no recognizable grand minima. The lower panel in Fig. 3 contrasts the Fourier amplitude spectra for these two intervals. (The peaks are naturally less well resolved than those for the full record.) The spectrum for the first interval shows clear peaks at periods of approximately 140–150 yr, 208 yr and 350 yr; by contrast, the interval without grand minima has weaker peaks at 140 yr and 290 yr – but nothing at all at 208 yr or 350 yr. It is apparent from Fig. 3 that the two spectra are qualitatively different. (Interestingly, however, the Gleissberg period of 88 yr shows up in both of them.) Taking the whole record in Fig. 2, McCracken et al. (2013b) find that the 208 yr de Vries period is prominent during the intervals that contain grand minima but absent during those where grand minima are missing. In other words, the Sun's cyclic activity shows not only modulation with the occurrence of grand minima and associated grand maxima but also *supermodulation*, with episodes during which behaviour is qualitatively different and the grand minima disappear. The challenge for theoreticians is to provide an explanation for this unexpected behaviour.

3 SYMMETRY PROPERTIES OF DETERMINISTIC MODELS GOVERNED BY DIFFERENTIAL EQUATIONS

In this section we explore the properties of some simple deterministic dynamo models that shed light on the long-term modulation of the solar cycle, and emphasize the role of symmetry changes (Tobias and Weiss 2007a). The simplest models of stellar dynamos share a similar bifurcation structure, depending on a dynamo number, D, proportional to the star's rotation rate. As $|D|$ is increased from zero successive Hopf and pitchfork bifurcations give rise to branches of nonlinear oscillatory solutions, involving dipole, quadrupole, multipole and mixed-mode solutions, most of which are unstable (Jennings & Weiss 1991). These branches form a complicated web and stability is transferred from one branch to another at symmetry-breaking bifurcations. Generalizing from these results, it is apparent that oscillatory dynamos must involve a multiplicity of patterns, only a few of which are stable. Relevant behaviour has been demonstrated in an idealized nonlinear Cartesian model, governed by partial differential equations (Beer, Tobias & Weiss 1998; Knobloch, Tobias & Weiss 1998; Arlt & Weiss 2014). The complicated nonlinear oscillations arise owing to the quadratic back reaction of the Lorentz force on differential rotation in the model — a deterministic process. These results demonstrate how cyclic behaviour may be interrupted by grand minima, during which symmetry properties are broken. In particular, the toroidal field may be limited to one hemisphere when the solution is entering or leaving a grand minimum. From time to time, there may also be a transition from cycles with dipole symmetry to cycles with quadrupole symmetry — where toroidal fields appear at the equator as in Fig. 1 – and such transitions recur throughout the record. When the PDE solutions are projected onto a three-dimensional space, whose co-ordinates are the dipolar and the quadrupolar magnetic energies, together with the fluctuating kinetic energy, most of the behaviour is confined to the dipolar subspace but there are recurrent excursions into the quadrupolar subspace. In this representation there are rapid transverse oscillations

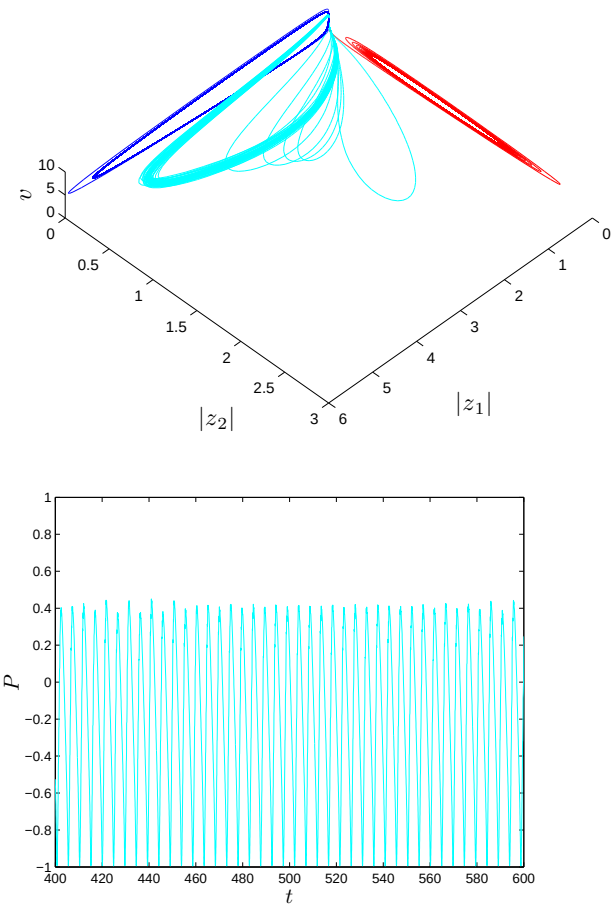

Figure 4. ODE model with three independent stable solutions. (a) Three-dimensional projection showing a dipolar trajectory (in blue), a quadrupolar trajectory (in red) and a chaotic mixed-mode trajectory (cyan). The projection is onto the three-dimensional space spanned by the rms energies z_1, z_2 of the dipolar and quadrupolar fields, and the value v of the symmetric fluctuating rms velocity. The basic ("Schwabe") cycles are filtered out, revealing deep grand minima, but the three solutions coexist and remain distinct. The co-ordinates denote the rms energies $|z_1|$ and $|z_2|$ in the dipole and quadrupole fields as well as the symmetric velocity v. (b) Variation with time of the parity P, for the mixed-mode solution only, over a representative interval. Note that the solution reverts to dipole symmetry as it approaches the origin in the phase space formed by $|z_1|$, $|z_2|$ and v.

that correspond to a "Schwabe" cycle, while the large-scale modulation represents a "de Vries" cycle.

This structure becomes more apparent when the problem is reduced to a sixth-order system of ordinary differential equations. For details of this system, see the Appendix; from a mathematical point of view it is a simplified version of the normal form equations that describe a double saddle-node/Hopf bifurcation (Knobloch, Tobias & Weiss 1998); the behaviour exhibited by solutions is therefore expected to be robust. The dipole and quadrupole magnetic fields are then represented by the two complex quantities z_1 and z_2, while the symmetric and antisymmetric velocity components are represented by the real quantities v and w, respectively

(Knobloch, Tobias & Weiss 1998). Then the "Schwabe" and "Hale" cycles can be filtered out by forming $|z_1|$ and $|z_2|$, so that the "de Vries" modulation is more clearly revealed.

In Fig. 4(a) trajectories representing three separate stable solutions are projected onto the three-dimensional space whose co-ordinates are $|z_1|$ and $|z_2|$ (representing the rms magnetic energies in the dipole and quadrupole fields), together with the symmetric velocity perturbation v. The dipole and quadrupole solutions are confined to the planes $—z_2| = 0$ and $|z_1| = 0$, respectively, while both components contribute to the noisier mixed-mode solution. Since the basic "Hale" and "Schwabe" periodicities (of order 0.5 dimensionless time units) have been filtered out, only the long-term modulation, associated with grand minima and with characteristic time scales of order 5 dimensionless units, is revealed. Such behaviour can be summarized by introducing the parity

$$P = (E_Q - E_D)/(E_Q + E_D), \qquad (1)$$

where $E_Q = |z_2|^2$, $E_D = |z_1|^2$ are the magnetic energies of the quadrupole and dipole fields, following Brandenburg et al. (1989), Knobloch & Landsberg (1996) and Knobloch, Tobias & Weiss (1998). Thus $|P| \leqslant 1$ and pure quadrupole and pure dipole fields have parities +1 and -1, respectively. Dynamo models allow not only pure dipole or quadrupole fields but also mixed-mode behaviour, with positive values for both E_D and E_Q, as indicated in Fig. 4. If we allow ourselves to move away from highly symmetric dynamo models to more general systems we can easily find examples that possess multiple solutions, with different symmetries present for the same parameter values. Then we can distinguish between two different types of modulation. For modulation of Type 1, the parity varies considerably while the cycles undergo fluctuations that may involve episodes of moderately reduced activity but without the appearance of deep grand minima (Knobloch & Landsberg 1996). For modulation of Type 2 the parity need not vary significantly (except when the total energy $E_Q + E_D$ is very small) but grand minima are frequent and flipping of symmetry can occur (as will be illustrated in Figs. 5 and 7 below). It follows from this discussion that grand minima can be expected to appear in nonlinear dynamos; they may also be associated with changes of parity. When the total energy $E_D + E_Q$ is very small, a slight perturbation that favours one or other symmetry can easily result in such a parity change.

The lower panel of Fig. 4 shows how the parity P varies for nearly periodic oscillations of the mixed mode solution, alternating between vigorous asymmetric behaviour and almost dipolar fields as the trajectory approaches the origin. Fig. 5 illustrates a very different pattern of behaviour, with a single stable mixed-mode solution that flips between quadrupole and (almost) dipole symmetry, as demonstrated by the plot of the parity P in Fig. 5(b). Note that the bursts of cyclic activity are interspersed with deep grand minima, as trajectories approach the origin in the 3D phase space formed by $|z_1|, |z_2|$ and v. This is clearly an example of Type 2 behaviour, associated with a global bifurcation at the origin that leads to chaotic modulation.

Fig. 6 illustrates a very different pattern of behaviour, with a single stable mixed-mode solution that keeps clear of the origin but covers a significant range of kinetic energy, with relatively mild fluctuations in both dipolar and

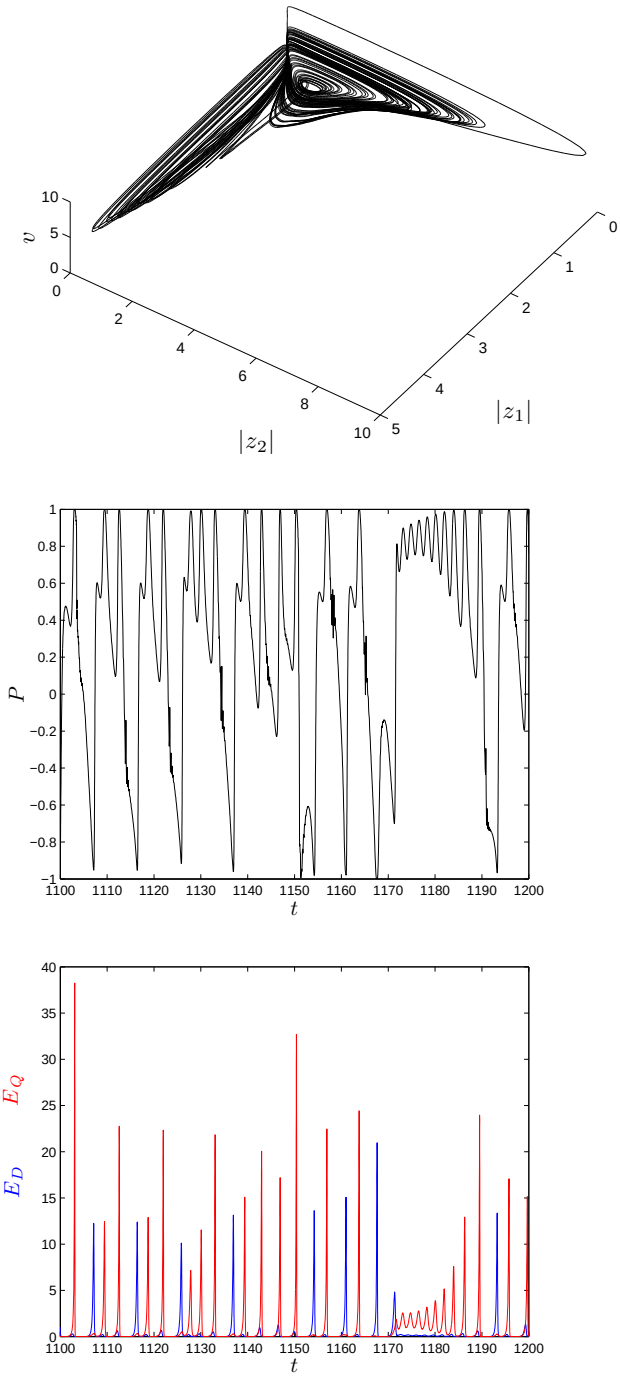

Figure 5. As Fig. 4 but for an ODE model with a single stable solution, in the form of a mixed mode that flips from almost dipolar to quadrupolar symmetry. (a) A projection showing the chaotic mixed-mode trajectory, with episodes that lie close to the dipolar and quadrupolar planes. (b) A section of the record, showing variation of the polarity P, which flips between quadrupole ($P = 1$) and dipole ($P = -1$) symmetries with intervals of mixed-mode behaviour. (c) The corresponding energies E_D and E_Q in the dipole and quadrupole fields, shown in blue and red, respectively, in the online version. Note the presence of recurrent grand minima as trajectories approach the heteroclinic bifurcation at the origin in the 3D phase space.

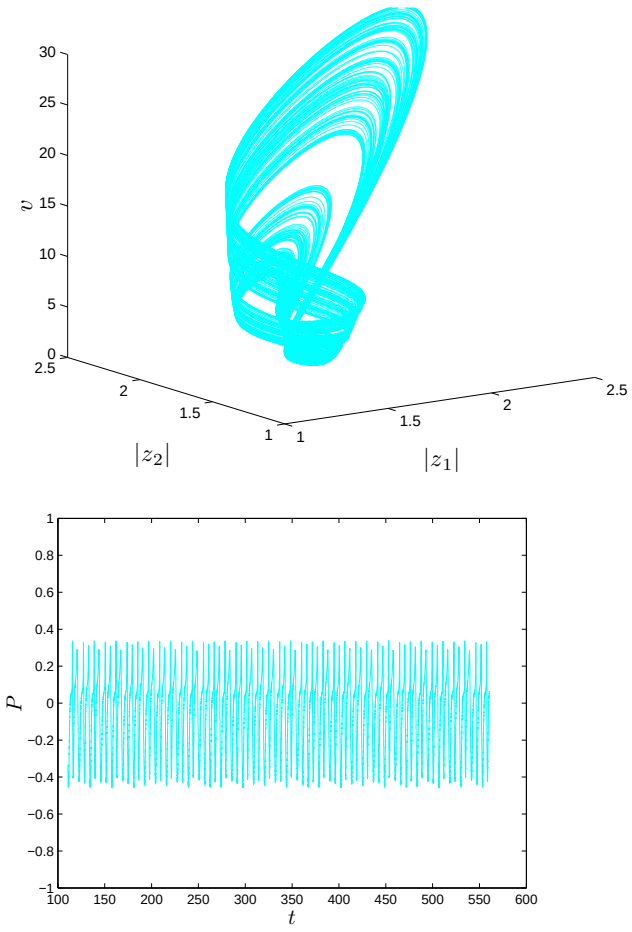

Figure 6. An example of aperiodic Type 1 modulation for ODEs. The trajectory corresponds to a chaotic mixed-mode solution that never approaches close enough to the origin in the 3D phase space for grand minima to occur. (a) The mixed-mode attractor in the three-dimensional phase space of the system. The energies vary significantly and the velocity varies considerably, while the parity of the mixed mode reverses sign without deviating far from zero, as shown in panel (b). This example serves as a model for the Sun's magnetic behaviour during the unshaded intervals in Fig. 2, when grand minima are lacking.

quadrupolar magnetic energies. The trajectory is certainly chaotic but it does not approach the origin: the solution is therefore of Type 1.

More immediately significant are the model solutions illustrated in Fig. 7, which display an example of flipping between a strongly modulated dipole (in blue) and a mixed-mode solution (in cyan). The energy plots show modulation, associated with grand minima, but the dipole fields are significantly stronger than the mixed-mode fields, although the latter can swing from pure dipoles to almost pure quadrupoles, as shown by the parity plot. It is also interesting to note that the minima have a longer recurrence period in the pure dipole phase than in the mixed phase for this solution. This example is relevant to the long-term behaviour of the solar dynamo, as we explain below.

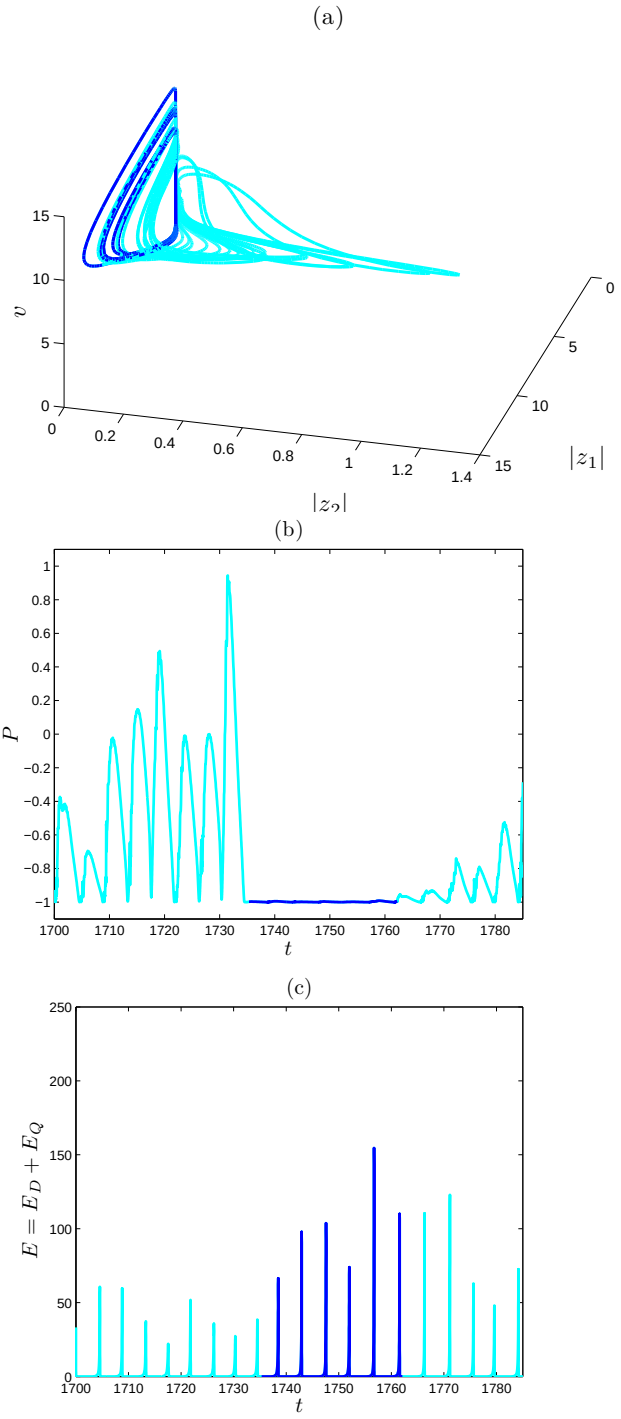

Figure 7. Type 2 modulation in an ODE model with alternating, strongly modulated, pure dipole (in blue) and mixed mode (cyan) solutions. (a) The attractor, showing flipping between the strongly modulated solutions. (b) The parity $\mathcal{P}$ showing an interval of pure dipole behaviour enclosed between episodes of mixed-mode modulation, which can even approach pure quadrupole behaviour. (c) The corresponding magnetic energies in the dipolar and mixed mode solutions over a representative time interval. Note the presence of recurrent grand minima, interrupted by large scale episodic bursts, which are more prominent for dipole fields than for mixed-mode fields.

4 SUPER-MODULATION AND SYMMETRY CHANGES IN THE SOLAR DYNAMO

The 400-year record of telescopic observations provides direct evidence of sporadic symmetry changes in the distribution of sunspots – from symmetric, apparently dipolar, patterns to asymmetric, hemispheric patterns during the Maunder Minimum, as well as a brief excursion from dipole to quadrupole behaviour in the eighteenth century, as shown in Fig. 1. The much longer proxy record derived from cosmogenic isotope abundances demonstrates that the Maunder Minimum is just the most recent of a series of grand minima that span the past 10,000 years. Within this time span, grand minima recur aperiodically (the de Vries cycle). Closer inspection reveals that this pattern is itself modulated on a longer timescale of around 2300 years (the Hallstatt cycle): episodes with recurrent grand minima alternate with intervals of weaker modulation, during which grand minima are absent – as is apparent from Fig. 2. It is with the origin of this super-modulation (possibly associated with hemispheric behaviour) that our paper is concerned.

We interpret this long-term modulation as a consequence of symmetry changes in the nonlinear dynamo that drives the solar cycle. The simple dynamo models described in the previous section reveal not only modulation of a basic cycle due to the appearance of grand minima but also examples of super-modulation, associated with transitions from dipole or quadrupole symmetry to mixed-mode solutions with significantly weaker modulation. We associate the grand minima with trajectories that approach a homoclinic or heteroclinic bifurcation in the phase space of the system; the associated tangle provides not only a route to chaotic behaviour but also an opportunity for switching from dipolar to quadrupolar symmetry or vice versa (as explained by Knobloch, Tobias & Weiss 1998). We propose that super-modulation, on the other hand, is associated with transitions from dipole or quadrupole symmetry to *mixed-mode* solutions that exhibit much weaker modulation. Such transitions have already been exhibited for simple nonlinear dynamo models in Section 3; our calculations indicate that they should be regarded as a generic feature of nonlinear dynamos, and therefore that super-modulation is to be expected not just in the Sun but also in other stars as well.

Normal solar activity over the last two centuries has, on average, shown symmetry about the equator, with minor transient deviations. Those occur most notably when polar fields reverse at different times around sunspot maximum, giving rise to short intervals with similar polarities around each pole (DeRosa, Brun & Hoeksema 2012). More striking was the behaviour during the last four sunspot cycles at the end of the Maunder Minimum (cf. Fig. 1) when such spots as did appear were all confined to the southern hemisphere. Such hemispheric behaviour requires a combination of dipolar and quadrupolar fields, which reinforce each other in one hemisphere but cancel out in the other. This is illustrated in the example of an α^2 mean-field dynamo discussed by Gallet and Pétrélis (2009), where the hemispherical field oscillates from one hemisphere to the other as the field reverses.

So far we have only considered generic patterns revealed by simple model calculations, involving low-order systems of ordinary differential equations. We expect that similar behaviour occurs also for partial differential equations, including much more elaborate systems that describe the nonlinear solar dynamo. Theoretical representations of the solar cycle have hitherto fallen into two classes. The first (the "Babcock-Leighton" model) relied on a poleward meridional flow that transported poloidal fields generated near the photosphere down to the base of the convection zone, where the flow reversed and strong toroidal fields could be formed by differential rotation. Others have proposed that the meridional flow should have a multicellular structure (Jouve & Brun 2007). The simple Babcock-Leighton picture has meanwhile been demolished by recent observations. Detailed measurements, first of the proper motions of supergranules (Hathaway 2012) and then from helioseismology (Zhao et al. 2013), show conclusively that the "conveyor belt" reverses its direction at a depth of only 65 Mm, leading to an equatorward flow between radii of 0.91 and 0.82 $R_\odot$; at the latter level there is a further reversal, giving rise to a poleward flow – and there may be yet another reversal giving an equatorward flow near the base of the convection zone at a radius of $0.71R_\odot$ (Hazra, Karak & Choudhuri 2014).

In order to model the solar dynamo successfully, theoreticians must therefore have recourse to direct numerical simulation of the interactions between convection, rotation and magnetic fields, both in the Sun's convection zone and in the tachocline that lies below it. The simplest and most straightforward approach is to assume the Boussinesq approximation, with a solenoidal velocity field, in a rotating spherical shell (Grote & Busse 2000; Busse & Simitev 2005). In some of these models, the oscillating fields remain in one hemisphere as they reverse, as apparently happened during the Maunder Minimum. Of course, we have no direct evidence of the spatial properties of the Sun's magnetic field during the intervals without grand minima in Fig. 2 – but we may speculate that they correspond to the appearance of mixed-mode patterns with weaker hemispheric fields.

Simulations of three-dimensional convection in spherical shells adopt the anelastic approximation and follow two different approaches. In the first, simple models are constructed to analyse the nonlinear interactions that may be important for dynamo action and modulation. Typically these models, although turbulent, analyse the changes in the morphology of solutions relatively close to the onset of dynamo action. Even in this restricted parameter regime the dynamo may exhibit a wide range of nonlinear behaviour — with solutions undergoing both Type 1 and Type 2 modulation, and indeed transiting between them (Raynaud & Tobias 2015). Furthermore, bistability between complicated nonlinear states (as demonstrated here for the simple set of ODEs) is often observed. The second approach is to attempt to mimic solar convection and dynamo action (choosing for example basic state entropy profiles derived from solar models, adopting the solar rotation rate and attempting to achieve a high degree of turbulence). The most ambitious of these studies that have so far appeared do indeed succeed in reproducing cyclic magnetic activity, together with meridional velocities that have a complicated, time-dependent structure. At high latitudes there is an equatorward flow at the tachocline, with poleward motion in most of the convection zone, giving way to equatorward flow above $0.9R_\odot$ and rapid poleward velocities nearer to the surface (Passos, Charbonneau & Miesch 2015; Hanasoge et al. 2015). The corresponding

azimuthal fields accumulate near the tachocline and display dipole symmetry; they reverse cyclically and have been followed for 40 activity cycles, with an average period of 40 years (Passos & Charbonneau 2014; Passos, Charbonneau & Miesch 2015). The activity cycles vary in a manner consistent with weak chaotic modulation but so far without any indication of grand minima.

Most recently, Augustson et al. (2015) have described an ambitious and very relevant model calculation for a solar-mass star rotating at three times the angular velocity of the Sun. As in similar investigations by Brown et al. (2011), the predominantly toroidal field in the star's interior takes the form of magnetic 'wreaths'. This model reproduces 24 activity cycles and does show one brief and comparatively shallow grand minimum. We stress here that these solutions are still a long way from the relevant parameter regime for dynamo action — indeed they are significantly below the asymptotic regime of high magnetic Reynolds number necessary to demonstrate some dynamo effects (Tobias & Cattaneo 2013). However, we anticipate that yet more ambitious computations will in due course be able to represent not only the occurrence of grand minima but also the super-modulation that we have described above.

ACKNOWLEDGMENTS

We thank Rainer Arlt, Jürg Beer, Paul Bushby, Basile Gallet, Edgar Knobloch, François Pétrélis, Mark Miesch, and Ed Spiegel for helpful and constructive discussions. We are also grateful to the referee, Sacha Brun, for many constructive suggestions.

REFERENCES

Abreu, J.A, Beer, J., Steinhilber, F., et al., 2013, Space. Sci. Rev., 176, 343
Arlt, R., 2009a, Solar Phys., 255, 143
Arlt, R., 2009b, Astron. Nachr., 330, 311
Arlt, R., Lessu, R., Giese, N., Mursula, I.G., Usoskin, I.G., 2013, MNRAS, 433, 3165
Arlt, R., Weiss, N.O., 2014, Space Sci. Rev., 186, 525
Augustson, K., Brun, A.S., Miesch, M., Toomre, J., 2015, ApJ, 809, 149
Baliunas, S.L., Donahoe, R.A., Soon, W.H., et al., 1998, ApJ, 438, 269
Beer J., Tobias S.M., Weiss, N.O., 1998, Solar Phys., 181, 237
Berhanu, M., Gallet, B., Monchaux, R., Bourgoin, M., Odier, P., et al., 2009, J. Fluid Mech., 641, 217
Brandenburg A., Krause F., Meinel R., Moss D., Tuominen I., 1989, A&A, 213, 411
Brown, B.P., Miesch, M.S., Browning, M.K., Brun, A.S., Toomre, J., 2011, ApJ 731, 69.
Bushby, P.J., 2006, MNRAS, 371,
Busse, F.H., Simitev, R., 2005, Astron. Nachr., 326, 231
Cristo, A., Vaquero, J.M., Sánchez-Bajo, F., 2011, J. Atm. Solar-Terr. Phys., 73, 187
Damon, P.E., Sonett, C.P., 1991, in The Sun in Time, ed. C.P. Sonett, M.S. Giampapa, M.S. Matthews, (Tucson, University of Arizona Press), 360
DeRosa, M.L., Brun, A.S., Hoeksema, J.T., 2012, ApJ, 757, 96
Eddy, J.A., 1976, Science, 192, 1189
Gallet, B., Herault, J., Laroche, C., Pétrélis, F., Fauve, S., 2012, GAFD, 106, 468
Gallet, B., Pétrélis, F., 2009, Phys. Rev. E, 80, 035302
Grote, E., Busse, F.H., Phys. Rev. E, 62, 4457, 2000
Hanasoge, S., Miesch, M.S., Roth, M., Schou, J., Schüssler, M., & Thompson, M.J., 2015, Space Sci. Rev., in press
Hazra, G., Karak, B.B., & Choudhuri, A.R., 2015, ApJ, 782, 93
Hathaway, D.H., 2012, ApJ, 760, 84
Jouve, L., & Brun, A.S., 2007, A&A, 474, 239
Knobloch, E., & Landsberg, A.S., 1996, MNRAS, 278, 294
Knobloch, E., Tobias S.M., Weiss N.O., 1998, MN, 297, 1123
Küker, M., Arlt, R., Rüdiger, G., 1999, A&A, 343, 977
McCracken, K., Beer, J., Steinhilber, F., Abreu, J., 2013a, Space Sci. Rev., 176, 59
McCracken, K.G., Beer, J., Steinhilber, F., Abreu, J., 2013b, Solar Phys., 286, 609
Monchaux, R., Berhanu, M., Aumaître, S., Chiffaudel, A., Daviaud, F., et al., 2009, Phys. Fl., 21, 035108
Muscheler, R., Beer, J., Wagner, G., et al., 2004, Earth. Planet. Sci. Lett., 219, 325
Ott, E., 1993, Chaos in Dynamical Systems, Cambridge University Press, Cambridge.
Passos, D., Charbonneau, P., 2014, A&A, 568, A113
Passos, D., Charbonneau, P., Miesch, M., 2015, ApJ, 800, L18
Pétrélis, F., Fauve, S., 2010, Phil. Trans. Roy. Soc. A, 368, 1595
Pétrélis, F., Fauve, S., Dormy, E., Valet, J.-P., 2009, PRL, 102, 144503
Pipin, V..V. 1999, A&A, 346, 295
Ravelet, F., Berhanu, M., Monchaux, R., Aumaître, S., Chiffault, A., et al. 2008, PRL, 101, 074502
Ribes, J.C., Nesme-Ribes E., 1993, A&A, 276, 549
Roberts, P.H. & Glatzmaier, G.A., 2000, Rev. Mod. Phys., 72, 1081
Santos, N.C., Gomes da Silva, J., Lovis, C., Melo, C., 2010, A&A, 511, A54
Sokoloff, D.D., Nesme-Ribes E., 1994, A&A, 288, 293
Spiegel, E.A., 2008, Space Sci. Rev., 144, 25
Steinhilber, F., Abreu, J.A., Beer, J., McCracken, K.G., 2010, J. Geophys. Rev., 115, A01104
Steinhilber, F., Abreu, J.A., Beer, J., et al., 2012, PNAS, 109, 5967
Stuiver, M., Braziunas, T.F., 1993, Holocene, 3, 289
Thomas, J.H., Weiss, N.O., 2008, Sunspots and Starspots. Cambridge University Press, Cambridge.
Tobias S.M., 1997, A&A, 322, 1007
Tobias S.M., 2002, Astron. Nachr., 323, 417
Tobias S.M., Weiss, N.O., Kirk V., 1995, MN, 273, 1150
Tobias S.M., Cattaneo, F., 2013, Nature, 497, 463
Tobias S.M., Weiss N.O., 2007a in Mathematical Aspects of Natural Dynamos, ed. E. Dormy & A.M. Soward, 281. CRC Press, Boca Raton.
Tobias S.M., Weiss N.O., 2007b in The Solar Tachocline, ed. D.W. Hughes, R.Rosner & N.O.Weiss, 320. Cambridge University Press, Cambridge.
Usoskin, I.G., 2013, Living Rev Solar Phys., 10, 1

Usoskin, I.G., Arlt, R., Asvestari, A., Hawkins, E., Käpylä, M., et al., 2015, A&A, in press
Vonmoos, M., Beer, J., & Muscheler, R., 2008, J. Geophys. Rev., 111, A10105
Weiss, N.O., 1985, J. Stat. Phys., 39, 477
Weiss, N.O., 2005, Astron. Nachr., 326, 157
Weiss, N.O., 2010, A&G, 51, 3.9
Weiss, N.O., Thompson, M.J., 2008, Space Sci. Rev., 144, 53
Wright, J.T., ApJ, 128, 1273, 2004
Yang, S., Odah, H., & Shaw, J., 2000, Geophys. J. Int., 140, 158
Yiou, F., Raisbeck, G.M., Baumgartner, S., et al., 1997, J. Geophys. Rev., 102, 26783
Zeldovich, Y.B., Ruzmaikin, A.A., Sokoloff, D.D., 1983, Magnetic Fields in Astrophysics. Gordon & Breach, New York.
Zhao, J., Bogart, R.S., Kosovichev, A.G., Duvall, T.L., Jr, Hartlep, T., 2013, ApJ, 774, L29

Table A1. Parameter values for the various cases considered here.

Fig.	μ	σ	a	a'	b	b'	c	c'	τ_1	τ_2
4	0.500509	0.486	0.5-0.5 i	0.4-0.4 i	-0.1	-0.2	0.2	0.0	1.0	0.9
5	1.39095	0.470	0.38-0.38 i	0.5 -0.5 i	-2.0	-1.8	0.0	0.0	1.0	1.1
6	0.08	0.376	-0.9+3.0 i	-0.9+3.0 i	0.5+0.5 i	0.5+0.5 i	0.5+0.5 i	0.5 + 0.5 i	0.101	0.201
7	0.800509	0.571	0.5-0.5 i	0.47-0.47 i	-0.1	-0.2	0.2	0.0	1.0	0.9
Fig.	e_1	e_2	ϵ	ϵ'	δ	δ'	β	β'		
4	1.31	1.81	1.1-1.1 i	1.1-1.1 i	-1.0+ i	-1.0 + i	0.49 -0.49 i	0.43 -0.43 i		
5	1.31	2.6	1.1 -1.1 i	1.1 -1.1 i	-1.0+ i	-0.965 +0.965 i	0.43-0.43 i	1,39 - 1.25 i		
6	0.2	0.002	0.0002 -0.001 i	0.0002 -0.001 i	0.0002 -0.001 i	0.0002 -0.001 i	0.0002 -0.001 i	0.0002 -0.001 i		
7	1.31	1.81	1.1-1.1 i	1.1-1.1 i	-1.0+ i	-1.0 + i	0.49 -0.49 i	0.43 -0.43 i		

APPENDIX A: THE LOW-ORDER MODEL SYSTEM

We supply here some further details of the model systems used to generate the results presented above in Section 3. Following Knobloch et al (1998) we introduce the normal form equations that govern a double-Hopf/pitchfork bifurcation. Such equations are structurally stable; their solutions have generic properties and display behaviour that is therefore likely to be robust. Near onset, the complex amplitudes z_1 and z_2 of the dipole and quadrupole fields satisfy

$$\dot{z}_1 = (\mu + \sigma + \mathrm{i}\omega_1)\, z_1 + a|z_1|^2 z_1, \tag{A1}$$

$$\dot{z}_2 = (\mu + \mathrm{i}\omega_2) z_2 + a'|z_2|^2 z_2. \tag{A2}$$

If the two Hopf bifurcations occur sufficiently close together, so that σ is small, and the frequencies ω_1 and ω_2 are almost equal, resonances become important and so further cubic terms should be added, yielding the equations

$$\dot{z}_1 = (\mu + \sigma + \mathrm{i}\omega_1)\, z_1 + a|z_1|^2 z_1 + b|z_2|^2 z_1 + c|z_2|^2 \bar{z}_1, \tag{A3}$$

$$\dot{z}_2 = (\mu + \mathrm{i}\omega_2) z_2 + a'|z_2|^2 z_2 + b'|z_1|^2 z_2 + c' z_1^2 \bar{z}_2, \tag{A4}$$

where the overbars denote complex conjugates and b, b', c, c' are complex coefficients (Knobloch et al 1998).

We next introduce the velocity components v and w, which are respectively symmetric and antisymmetric about the equator and are driven solely by the magnetic field, against viscous drags. Following Knobloch et al., we represent their evolution by equations of the form

$$\dot{v} = -\tau_1 v + e_1(|z_1|^2 + |z_2|^2),\ \dot{w} = -\tau_2 w + e_2(z_1\bar{z}_2 + z_2\bar{z}_1), \tag{A5}$$

where τ_1, τ_2, e_1 and e_2 are all real parameters. Thus v is driven by either symmetric or antisymmetric fields, while driving of w requires that both z_1 and z_2 should be non-zero. These velocities in turn act on the symmetric and antisymmetric fields and so the equations governing the fields finally take the form

$$\begin{aligned}\dot{z}_1 = {} & (\mu + \sigma + \mathrm{i}\omega_1)\, z_1 + a|z_1|^2 z_1 + b|z_2|^2 z_1 + c z_2^2 \bar{z}_1 \\ & + (\epsilon v + \delta v^2 + \kappa w^2) z_1 + (\beta + \gamma v) w z_2,\end{aligned} \tag{A6}$$

$$\begin{aligned}\dot{z}_2 = {} & (\mu + \mathrm{i}\omega_2) z_2 + a'|z_2|^2 z_2 + b'|z_1|^2 z_2 + c' z_1^2 \bar{z}_2 \\ & + (\epsilon' v + \delta' v^2 + \kappa' w^2) z_2 + (\beta' + \gamma' v) w z_1.\end{aligned} \tag{A7}$$

Nonlinear interactions between the dipole mode z_1 and the quadrupole mode z_2 then suffice to generate chaotic behaviour.

The results quoted in §3 were all obtained using the parameter values $\omega_1 = 12.2108, \omega_2 = 12.51081$, with $\kappa = \kappa' = \gamma = \gamma' = 0$. The remaining parameter values are displayed in the Table.