FLUCTUATIONS IN BABCOCK-LEIGHTON DYNAMOS. I. PERIOD DOUBLING AND TRANSITION TO CHAOS

PAUL CHARBONNEAU,[1] CÉDRIC ST-JEAN, AND PIA ZACHARIAS[2]

Département de Physique, Université de Montréal, C. P. 6128, Succursale Centre-Ville, Montréal, QC H3C 3J7, Canada

ABSTRACT

We present a large series of numerical simulations of the solar magnetic activity cycle based on the Babcock-Leighton mechanism for the regeneration of the solar poloidal magnetic field. While the primary cycle period changes very little as the dynamo number is increased, the model shows a well-defined transition to chaos through a sequence of period-doubling bifurcations, i.e., the sequential appearance of modulations of the primary cycle's amplitude, with associated periods equal to twice the periods characterizing the amplitude variations prior to a given bifurcation. This behavior arises through the unavoidable time delay built into this type of solar dynamo model, rather than through the effects of complex, nonlinear magnetic back-reaction on the fluid motions driving the dynamo process. It is noteworthy that a chaotic regime exists in this numerical model, given that the only non-linearity present is a simple algebraic amplitude-quenching factor in one of the governing partial differential equations. The results also represent a rare instance in which the complex dynamical behavior of a spatially extended, diffusive solar dynamo model can be reproduced in detail on the basis of the simplest of low-order dynamical systems, namely a one-dimensional iterative map. The numerical results also demonstrate the central role of meridional circulation in setting the primary cycle period in this class of dynamo models; despite variations by many orders of magnitude in the dynamo number and concomitant large and sometimes even chaotic variations in amplitude, the cycle period remains tightly locked to the meridional circulation turnover time.

Subject headings: Sun: activity — Sun: magnetic fields

1. THE SOLAR CYCLE

The solar activity cycle originates in the periodic polarity reversal of the Sun's internal large-scale magnetic field. Compared to the Earth's magnetic field, this temporal variation of the Sun's field is fairly regular and takes place on decadal timescales, which is some 9 orders of magnitude shorter than the expected ohmic dissipation time. The solar cycle thus requires a vigorous, active regeneration process operating on a comparable timescale.

It is now generally agreed that the physical origin of the cycle itself is in a dynamo process powered by the inductive action of fluid motions in the Sun's interior (for recent reviews, see Petrovay 2000; Stix 2002; Ossendrijver 2003; Rüdiger & Arlt 2003). Beginning in the late 1950s, magnetographic observations have revealed that the Sun's poloidal magnetic field component also undergoes cyclic variations along with its toroidal component (as mapped by sunspots), with the large-scale poloidal field peaking in strength at times of sunspot minima and flipping polarity around times of sunspot cycle maxima.

Helioseismology, the art and science of mapping the solar interior from acoustic waves detected spectroscopically via the Doppler shift they induce at the surface, has now pinned down with unprecedented accuracy the thermodynamic structure of and large-scale flows through most of the Sun's interior (see Christensen-Dalsgaard 2002). In particular, it has revealed that the observed surface latitudinal (pole-to-equator) differential rotation persists through the outer 30% of the Sun's interior and then vanishes across a thin layer, dubbed the "tachocline," coinciding approximately with the interface between the stable internal core and the unstable, convecting outer envelope.

Such an internal differential rotation, symmetric about the equator and acting upon a poloidal field antisymmetric about the equator, readily produces a sunspot-compatible toroidal field on decadal timescales. The second half of the dynamo process, turning this newly generated toroidal field into a new poloidal field, has proved much harder to pin down. Various toroidal-to-poloidal production mechanisms have been put forth, including turbulent inverse MHD cascade (the so-called α-effect of mean field theory; see Moffatt 1978), the surface decay of sunspot pairs (also known as the Babcock-Leighton mechanism; more on this shortly), and a regenerative process based on various instabilities of the differential rotation and/or toroidal magnetic field (Ferriz-Mas et al. 1994; Ossendrijver 2000; Thelen 2000; Dikpati & Gilman 2001).

Such issues notwithstanding, the physical origin of the cycle's amplitude variations remains a controversial topic in itself. Explanations typically fall into one of two broad categories: stochastic forcing or dynamical nonlinearities. Vigorous turbulence is known to pervade the Sun's outer convective layers, where the dynamo is believed to operate, in part or in toto. Thus, it would be very surprising if any dynamo cycle operating therein were nicely regular (e.g., Hoyng 1988, 1993; Choudhuri 1992; Moss et al. 1992; Hoyng et al. 1994; Ossendrijver et al. 1996). Stochastic forcing is, therefore, a perfectly viable working hypothesis. The observed strength of the dynamo-generated solar magnetic field is also such that dynamical (nonlinear) feedback on the driving fluid motions is not only possible but, in fact, expected, and this has been shown to lead to marked amplitude fluctuations in a variety of solar cycle models (see, e.g., Gilman 1983; Brandenburg et al. 1991; Tobias 1997; Küker et al. 1999; Moss & Brooke 2000 and references therein).

In this paper we focus on a third class of nonlinear feedback mechanism, based on the notion of time delay in the dynamo mechanism. We do so in the context of solar cycle models based

---

[1] Also at Centre Observatoire du Mont Mégantic.
[2] Also at Universität Freiburg, Germany.

on the Babcock-Leighton mechanism, which we describe in § 2, but in fact the general idea is applicable to any dynamo model in which the source regions for the two field components, poloidal and toroidal, are spatially segregated. In such circumstances, there exists a physically plausible procedure whereby the MHD dynamo equations can be reduced to a one-dimensional iterative map, as outlined in § 3. The dynamical behavior predicted from such maps finds a striking counterpart in the dynamical behavior of a more realistic model based on numerical solutions of the MHD induction equations in two spatial dimensions (§ 4). We conclude (§ 5) by positioning our results in the wider context of solar cycle fluctuation studies.

2. BABCOCK-LEIGHTON SOLAR CYCLE MODELS

2.1. *The Surface Decay of Sunspots*

Sunspots are the photospheric manifestation of undulating magnetic flux ropes rising through the Sun's convective envelope to emerge at the photosphere. Observational support for this notion is well established (see Schüssler 1994 and references therein). As a consequence of disconnection from the underlying flux system and/or buffeting from surrounding turbulent convective fluid motions, sunspots do eventually decay; in doing so, they release a fraction of their magnetic flux into the photosphere. Under the joint action of turbulent diffusion, differential rotation, and poleward meridional circulation, this flux accumulates in the polar regions until it overwhelms the "old" poloidal flux and causes a reversal of the poloidal field. This global flux release and transport process has now been observed in great detail by space-borne solar monitoring instruments for more than two sunspot cycles. It has also been successfully modeled as a two-dimensional advection-diffusion transport problem on a sphere, with model results comparing favorably to observations (see Wang et al. 1989, 2002; Wang & Sheeley 1991; Schrijver et al. 2002). A crucial question in the dynamo context, and one that has yet to be resolved, is whether this observed phenomenon is a mere side effect or an active component of the underlying magnetic cycle.

Decaying sunspots end up contributing a net dipole moment to the photospheric field because the lines joining the centers of sunspot pairs of opposite polarities show a statistically significant, systematic pattern of tilt with respect to the east-west (longitudinal) direction. This is perhaps most easily understood by considering the spherical harmonic decomposition of a bipolar structure on a sphere. If the line segment joining each pole is aligned with the ϕ-direction, all $m = 0$ terms in the modal decomposition have zero amplitude. If, on the other hand, this line segment is inclined with respect to the ϕ-direction, then nonvanishing $m = 0$ contributions arise in the decomposition. These collectively add up to an axisymmetric poloidal field. The joint action of dissipation and differential rotation rapidly destroys all $m \neq 0$ modes, but the $m = 0$ modes, in contrast, decay on their much longer purely dissipative timescales, leaving meridional circulation enough time to concentrate a portion of the corresponding axisymmetric poloidal magnetic field toward the solar poles. Since sunspot pairs originate with the Sun's deep-seated toroidal magnetic field, the net effect is to convert a fraction of that toroidal field to a large-scale poloidal field. The second half of the dynamo loop, producing a toroidal field from a preexisting poloidal component, is readily achieved by the shearing associated with differential rotation.

Solar cycle models based on this mechanism were first proposed by Babcock (1961) and Leighton (1964, 1969) and are now usually referred to as "Babcock-Leighton models" (here-

after B-L models). Following ever-growing difficulties in reconciling helioseismic results on internal flows with the long-favored dynamo models based on mean field electrodynamics, B-L models have been enjoying a vigorous revival in the ongoing helioseismic era of solar cycle modeling (see Wang et al. 1991; Durney 1995; Dikpati & Charbonneau 1999, hereafter DC99; Nandy & Choudhuri 2001). As of this writing, they remain solid contenders as explanatory models of the solar cycle, comparing quite favorably with any other class of current solar dynamo models at similar levels of complexity.

The tilt of sunspot pairs with respect to the zonal direction is a crucial component of the B-L poloidal field generation mechanism. Simulations of magnetic flux ropes buoyantly rising through the Sun's convective envelope have shown that this tilt builds up under the action of the Coriolis force acting on the internal, axis-aligned flow that develops in the course of the flux tube's rise through the convective envelope (Fan et al. 1993; D'Silva & Choudhuri 1993; Caligari et al. 1995). An important aspect of this mechanism is that it is subject to both a lower and an upper operating threshold. Flux ropes of strength in excess of about 100 kG rise too quickly for the Coriolis force to have significant effects; thus, the resulting sunspot pairs are east-west aligned and have no axisymmetric poloidal component to release during their subsequent decay. Flux ropes of strength inferior to about 10 kG show no systematic pattern of tilt for a number of reasons, including deflection to high latitudes and randomization of their spatial orientations by turbulent convective fluid motions (Longcope & Fisher 1996). In addition, it appears that such weak flux ropes cannot become unstable on sufficiently short timescales; even if they could, they could not retain their structural integrity in the course of their buoyant rise (Schüssler 1994; Moreno-Insertis et al. 1995).

2.2. *Axisymmetric Formulation of Solar Cycle Models*

All partial differential equation (PDE)-based numerical solutions discussed further below make use of the general B-L modeling formalism of DC99, to which we refer the interested reader for all details regarding the specification of the differential rotation, meridional circulation, and magnetic diffusivity profiles. This model uses solar-like profiles for these quantities and yields cyclic solutions that compare very well to a number of observed features of the solar cycle.

The DC99 model describes the time evolution of an axisymmetric, large-scale magnetic field in spherical geometry and in the kinematic regime. Working in the usual spherical polar coordinates (r, θ, ϕ), an axisymmetric magnetic field can be written as

$$\boldsymbol{B}(r, \theta, t) = \nabla \times [A(r, \theta, t)\hat{\boldsymbol{e}}_\phi] + B(r, \theta, t)\hat{\boldsymbol{e}}_\phi, \quad (1)$$

where A describes the poloidal component of the magnetic field and B its toroidal component. Note that this formulation ensures that $\nabla \cdot \boldsymbol{B} = 0$, as required by Maxwell's equations. The MHD induction equation can then be separated into two PDEs for A and B, which take the form

$$\frac{\partial A}{\partial t} + \frac{1}{\varpi}\boldsymbol{u}_p \cdot \nabla(\varpi A) = \eta\left(\nabla^2 - \frac{1}{\varpi^2}\right)A + S(B), \quad (2)$$

$$\frac{\partial B}{\partial t} + \varpi\nabla \cdot \left(\frac{\boldsymbol{u}_p B}{\varpi}\right) = \eta\left(\nabla^2 - \frac{1}{\varpi^2}\right)B + \frac{1}{r}\frac{\partial n}{\partial r}\frac{\partial(rB)}{\partial r}$$
$$+ \varpi(\nabla \times A\hat{\boldsymbol{e}}_\phi)\cdot(\nabla\Omega),$$

$$(3)$$

where $\varpi = r \sin \theta$, $\Omega(r, \theta)$ is the angular velocity profile, $\boldsymbol{u}_p(r, \theta)$ is the meridional flow, and $\eta(r)$ is the net magnetic diffusivity. The two PDEs are linearly coupled via the shearing term ($\propto \nabla \Omega$) and nonlinearly coupled through the nonlocal source term $S(B)$ appearing in equation (2), which represents the production of a poloidal magnetic field by the B-L mechanism (see § 2.3 of DC99). With the flow fields (differential rotation and meridional circulation) specified as steady, this source term is the only nonlinearity left in the kinematic regime considered here. The following functional form is adopted in all computations reported on below:

$$S(r, \theta, B) = \frac{s_0}{4} \left[1 + \mathrm{erf} \left(\frac{B - B_1}{w_1} \right) \right] \left[1 - \mathrm{erf} \left(\frac{B - B_2}{w_2} \right) \right] f(r, \theta) B,$$

$$(4)$$

where the toroidal field B is to be evaluated at the corresponding colatitude at the core-envelope interface, and $f(r, \theta)$ is a function of spatial coordinates that concentrates the source term in the surface layers while enforcing the minimal latitudinal dependency compatible with results from thin flux tube simulations (see DC99 for further details; also § 2.1 of Charbonneau & Dikpati 2000). Note that the source term so defined is nonlocal, which is meant to account for the fact that sunspot pairs ultimately originate with the rapid rise to the surface of toroidal flux tubes forming in the vicinity of the core-envelope interface from the dynamo-generated large-scale magnetic field. The numerical coefficient s_0 then sets the absolute strength of the source term.

Unless noted otherwise, in all numerical calculations reported on below, the numerical parameters defining the form of the nonlinearity take the values $B_1 = 60$, $w_1 = 20$, $B_2 = 100$, and $w_2 = 80$ kG. The source term then "turns on" at a toroidal field strength of $\sim B_1 - w_1 = 40$ kG and turns off at $\sim B_2 + w_2 = 180$ kG, in basic agreement with the aforementioned simulations of rising thin flux tubes. This implies that unlike in the original model of DC99, the numerical model defined herein is now characterized by both upper and lower operating thresholds on the magnetic field strength.

Three dimensionless quantities arising in the scaling of the model's governing equations are of interest in what follows. The first is a dynamo number (C_S), which will serve as the control parameter in the numerical calculations to be reported on presently. It essentially measures the efficiency of poloidal magnetic field production via the B-L mechanism (a magnetic energy source) relative to ohmic dissipation (a magnetic energy sink):

$$C_S = \frac{s_0 R}{\eta_T},$$

$$(5)$$

where R is the Sun's radius. The quantity η_T is the (turbulent) magnetic diffusivity in the Sun's convective envelope, which is also used to define a diffusion time τ:

$$\tau = \frac{R^2}{\eta_T}.$$

$$(6)$$

Physically, this diffusion time represents the adjustment time of the model to externally imposed forcing and is also a typical timescale for the damping of transients associated with initial conditions (for example). In the simulations discussed here, we adopt $\eta_T = 1.7 \times 10^{11}$ cm^2 s^{-1}, which leads to $\tau = 920$ yr with the Sun's radius used as a typical length scale (for some purposes the thickness of the envelope would perhaps be more

appropriate; this would reduce the diffusion time to ~ 100 yr for our adopted value of η_T). The third important dimensionless quantity is the magnetic Reynolds number, which measures the relative efficiency of meridional circulation and magnetic diffusion in transporting magnetic fields across the domain,

$$\mathrm{Rm} = \frac{u_0 R}{\eta_T},$$

$$(7)$$

where u_0 is a characteristic meridional flow speed. Note that the solar radius is again used as a typical length scale.

3. REDUCTION TO A ONE-DIMENSIONAL ITERATIVE MAP

3.1. *Time Delays in Babcock-Leighton Models*

A crucial aspect of solar cycle models based on the B-L mechanism is the fact that the two source terms in the governing equations (2) and (3) are segregated in two distinct regions of the spatial domain. The B-L mechanism operates in the surface layers, while production and storage of the sunspot-generating toroidal field takes place much deeper in the interior, at or immediately beneath the core-envelope interface. For the dynamo to operate, these two source regions must somehow "communicate."

In the most recent solar cycle models based on the B-L mechanism, this communication is assumed to take place via advection by a quadrupolar meridional flow pervading the convective envelope. Such a flow has been detected helioseismically down to $r/R_\odot \simeq 0.85$ (see Braun & Fan 1998; Schou & Bogart 1998). The basic idea is illustrated in cartoon form in Figure 1. Let P_n represent the amplitude of the high-latitude, surface (s_1) poloidal magnetic field in the late phases of cycle n, i.e., after the polar field has reversed. The poloidal field P_n is advected downward by meridional circulation ($s_1 \to s_2$), where it then starts to be sheared by the differential rotation while also being advected equatorward ($s_2 \to s_3$). This leads to the growth of a new low-latitude (s_3) toroidal flux system, T_{n+1}, which becomes buoyantly unstable ($s_3 \to s_4$) and starts producing sunspots (s_4), which subsequently decay and start releasing the poloidal flux P_{n+1} associated with the new cycle $n + 1$. Poleward advection and accumulation of this new flux at high latitudes ($s_4 \to s_1$) then obliterates the old poloidal flux P_n, and the above sequence of steps begins anew.

Meridional circulation plays a key role in this "conveyor belt" model of the solar cycle. First and foremost, it effectively sets—and even regulates—the cycle period (DC99; Charbonneau & Dikpati 2000; Hathaway et al. 2003). In doing so, it also introduces a long time delay in the dynamo mechanism ("long" in the sense of being comparable to the cycle period). In contrast, the production of the poloidal field from the deep-seated toroidal field is a "fast" process, growth rates and buoyant rise times for sunspot-forming toroidal flux ropes being of the order of a few months (see Moreno-Insertis 1986; Fan et al. 1993; Caligari et al. 1995 and references therein). The time delay associated with advection by meridional circulation turns out to have rich dynamical consequences.

3.2. *From PDEs to Iterative Maps*

The long time delay inherent in B-L models of the solar cycle makes it possible to express the cycle-to-cycle amplitude variations as a simple one-dimensional iterative map (Durney 2000; Charbonneau 2001). The procedure runs as follows (for details see Charbonneau 2001). In view of the preceding conveyor-belt argument, in the kinematic regime the toroidal field strength
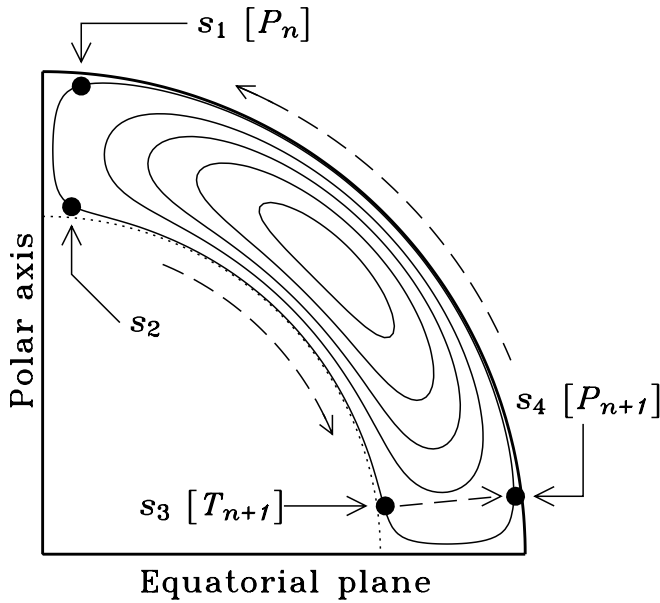
Fig. 1.—Operation of a solar cycle model based on the B-L mechanism. The diagram is drawn in a meridional quadrant of the Sun. Streamlines of meridional circulation are plotted as solid lines. This flow is poleward in the outer layers, sinks at the pole, and is equatorward along the core-envelope interface at $r/R_\odot \simeq 0.7$ (*dotted line*). Poloidal fields accumulated in the surface polar regions (s_1) at cycle n must first be advected down to the core-envelope interface before production of the toroidal field for cycle $n + 1$ may commence. This introduces a long time delay in the dynamo process (*see text*).

(T_{n+1}) at cycle $n + 1$ is linearly proportional to the poloidal field strength (P_n) of cycle n, i.e.,

$$T_{n+1} = aP_n. \qquad (8)$$

In writing down equation (8), we are neglecting any ohmic dissipation of either poloidal or toroidal components during the magnetic field's advection from the surface polar regions to the low-latitude regions of the tachocline ($s_1 \rightarrow s_2 \rightarrow s_3$ in Fig. 1). In contrast, the strength of the poloidal field at cycle $n + 1$ is *non-linearly* proportional to the toroidal field strength of that cycle,

$$P_{n+1} = f(T_{n+1})T_{n+1}, \qquad (9)$$

where the function $f(T_{n+1})$ measures the efficiency of surface poloidal field production from the deep-seated toroidal field by the B-L mechanism. After normalizing the field amplitudes, substitution of equation (8) into equation (9) leads immediately to a one-dimensional iterative map,

$$p_{n+1} = \alpha f(p_n)p_n, \qquad (10)$$

where the p_n values are normalized amplitudes, and the normalization constants and the constant a in equation (8) have been absorbed into the definition of the map's parameter α. Equation (10) is called a "map" because it allows the calculation of a cycle amplitude in terms of the amplitudes of preceding cycles (here only the previous cycle) and is said to be "one-dimensional" because it involves only a single dynamical variable, namely the amplitude iterate p_n.

3.3. *Dynamical Behavior*

The nonlinear function $f(p_n)$ in equation (10) must be chosen so as to catch the essential features of the B-L mechanism, most importantly the fact that it can only operate in a finite range of
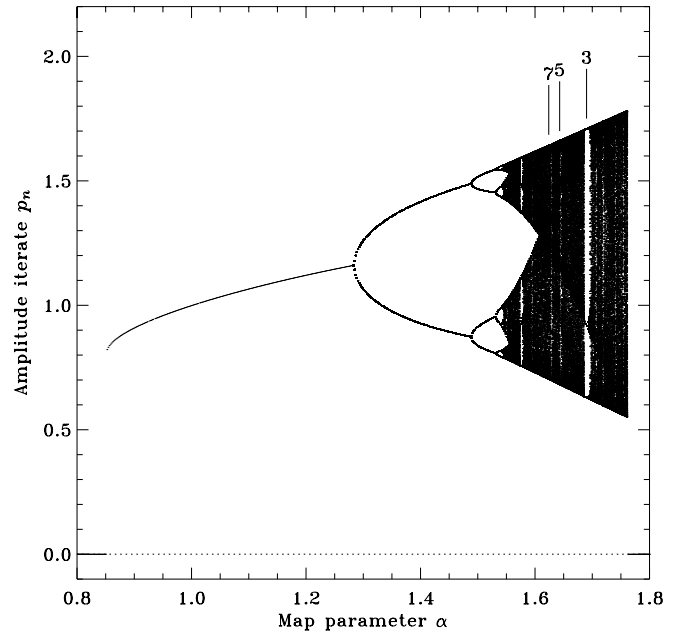

Fig. 2.—Bifurcation diagram for the one-dimensional iterative map defined by eqs. (10) and (11). It shows the usual features of one-dimensional single-hump maps, most notably a transition to chaos via a sequence of period-doubling bifurcations. Multiperiodic windows are interspersed throughout the chaotic regime, the first three widest being indicated by vertical lines and labeled (n) in terms of their n-fold periodicity.

toroidal field strength. The product of error functions, as introduced previously, can again be used here:

$$f(p_n) = \frac{\alpha}{4}\left[1 + \mathrm{erf}\left(\frac{p_n - B_1}{w_1}\right)\right]\left[1 - \mathrm{erf}\left(\frac{p_n - B_2}{w_2}\right)\right], \qquad (11)$$

now with $B_1 = 0.6$, $w_1 = 0.2$, $B_2 = 1.0$, and $w_2 = 0.8$, so as to obtain the same shape for the nonlinearity as in the two-dimensional numerical model (cf. § 2.2). The bifurcation diagram for the resulting iterative map is presented in Figure 2. The diagram is constructed by plotting as small dots successive values of the amplitude iterate p_n for increasing values of the map parameter α. For low values of α, one finds $p_{n+1} = p_n$, and p_n ends up looking like a single-valued function of α. For α exceeding 1.283, however, the amplitude iterate starts to alternate between "high" and "low" values, bracketing the former single amplitude value. This is called "period doubling" in dynamical system jargon; however, it must be emphasized that the primary cycle period remains unaltered, with "doubling" referring instead to the number of amplitude modulation periods superimposed on the primary cycle. The same pattern eventually repeats at higher α values, leading to multiply periodic as well as aperiodic solutions. This bifurcation diagram shows the usual generic characteristics of one-dimensional single-hump iterative maps (see, e.g., Holton & May 1993), namely:

1. A first bifurcation from the trivial solution, here supercritical and occurring at $\alpha_1 = 0.851$.
2. A sequence of period-doubling bifurcations, beginning at $\alpha_2 = 1.283$ ($1 \rightarrow 2$), then $\alpha_4 = 1.488$ ($2 \rightarrow 4$), $\alpha_8 = 1.531$, etc.
3. A chaotic regime starting at the end of the period-doubling cascade.
4. A set of periodic windows interspersed across the chaotic domain. The widest is a 3 period window, spanning the range $1.6855 \leq \alpha \leq 1.6958$ and preceded by a narrower 5 period

window at $1.6422 \leq \alpha \leq 1.6453$, itself preceded by an even narrower 7 period window at $1.6226 \leq \alpha \leq 1.6233$.

This rich dynamical behavior ultimately originates with the time delay in the dynamo process, which is what makes possible the reduction of the governing equations to a one-dimensional iterative map in the first place. Not surprisingly, the introduction of ad hoc time delays in dynamo models has long been known to lead to cycle amplitude fluctuations (see, e.g., Yoshimura 1978). However, in the model considered here, the time delay is not at all ad hoc but is, rather, unavoidable, as in any other dynamo model in which the source regions for the poloidal and toroidal magnetic field components are spatially segregated. The B-L model being a particularly clear-cut example of such a situation, one is then led to anticipate that the map's rich dynamical behavior should find its counterpart in the original, more realistic, spatially extended, diffusive axisymmetric model that inspired the map formulation. This is the quest to which we now turn.

4. COMPARISON WITH A SPATIALLY EXTENDED NUMERICAL MODEL

4.1. *Numerical Solutions and Their Analysis*

We now seek the counterpart of the iterative map's behavior in the two-dimensional axisymmetric dynamo model of DC99, modified to include a nonlinearity with a lower threshold, as described in § 2.2. Starting from an arbitrary but finite-amplitude initial condition, the axisymmetric dynamo equations (2) and (3) are integrated forward in time using a finite element-based numerical scheme with implicit time-stepping. Solutions are followed over up to 40,000 time steps spanning 10 diffusion times to ensure that all transients have died out. We compute solutions for dynamo numbers spanning more than 3 orders of magnitude to fully sample the dynamical behavior of the model. Our reference model below has parameter values $C_\Omega = 8.4 \times 10^4$, $\mathrm{Rm} = 420$, $\Delta\eta = 500$, $r_b = 0.65$, $d_1 = 0.025$, and $r_c = 0.7$ (see § 2 of DC99 for descriptions and definitions of these model parameters).

The first step in the subsequent analysis is to compute the magnetic energy at every time step,

$$E_B(t) = \frac{1}{8\pi} \int \boldsymbol{B}^2(t)\, dV, \tag{12}$$

where the integral is carried out over the full spatial volume of the simulation. An autocorrelation analysis is then performed to identify periodicities in the time series. We first compute a standard autocorrelation coefficient,

$$P(L) = \frac{\sum_{k=1}^{N-L} \left[E_B(t_k) - \bar{E}_B \right] \left[E_B(t_{k+L}) - \bar{E}_B \right]}{\sum_{k=1}^{N} \left[E_B(t_k) - \bar{E}_B \right]^2}, \tag{13}$$

where $E_B(t_k)$ is the magnetic energy at the kth time step, $\bar{E}_B$ is the mean of the time series, and L is the lag measured in time steps. Whenever an autocorrelation coefficient larger than 0.5 is found, the solution is tentatively flagged as "periodic," and the last full period in the time series is scanned to identify the number of peaks in $E_B(t)$ (excluding endpoints) having distinct values. When autocorrelation fails, a solution is deemed "nonperiodic," and all peaks in the time series are retained for further analysis.

4.2. *Bifurcation Diagram*

To construct a bifurcation diagram conceptually equivalent to that depicted in Figure 2, we now plot, as a function of the
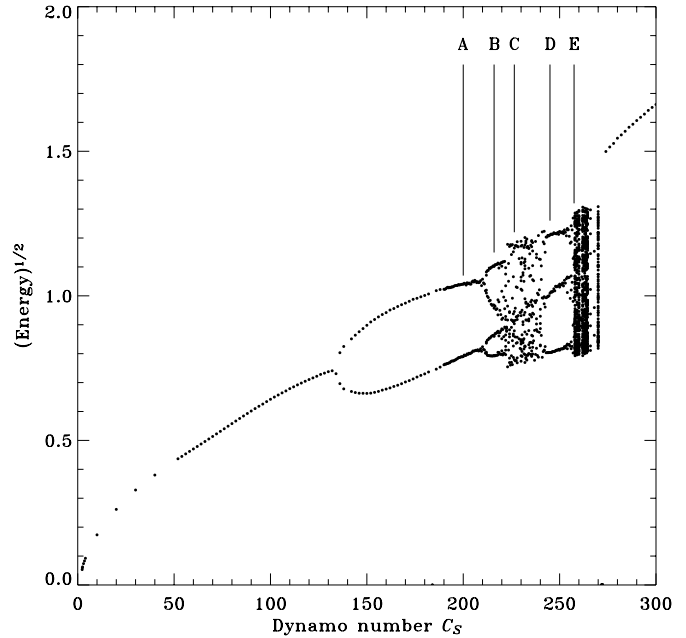

Fig. 3.—Bifurcation diagram reconstructed from a sequence of numerical dynamo solutions with increasing dynamo numbers C_S. Amplitudes are measured in time series of magnetic energy. Time series for a few representative individual solutions, flagged here by vertical lines, are depicted in Fig. 4. In this and the following figures, magnetic energies have been arbitrarily but consistently scaled to values of order unity.

dynamo number C_S (equivalent to the map parameter α in eq. [10]), the various distinct peaks in $E_B(t)$ identified in the autocorrelation analysis.[3] The result is shown in Figure 3, and a few representative magnetic energy time series are shown in Figure 4 (*left column*).

Striking similarities are immediately seen between Figure 3 and the bifurcation diagram arising from the one-dimensional iterative map (Fig. 2). As with the latter, nontrivial behavior materializes along the sequence of increasing α as a supercritical bifurcation at $C_{S,1} \simeq 2.4$, followed by a first period-doubling bifurcation at $C_{S,2} \simeq 132$ and a second at $C_{S,4} \simeq 210$. [No attempt was made to identify with very high accuracy the numerical values of the dynamo numbers corresponding to the bifurcation points; this would have consumed much CPU time in view of the very slow relaxation to the stable branch(es) in the vicinity of a bifurcation point.] This is followed ($225 \lesssim C_S \lesssim 242$) by a regime in which the vast majority of solutions are nonperiodic and after which well-defined multiperiodic behavior emerges once again. These are now triply periodic solutions, just as in the first wide multiperiodic window encountered in the chaotic regime for the attractor of Figure 2 ($\alpha \simeq 1.69$). As in Figure 2, chaotic nonperiodicity resumes at $C_S \simeq 256$, again through a sequence of period doubling (a hint of which is visible in Fig. 3 [$C_S \simeq 254$]), up to $C_S \simeq 270$, where a drastic transition takes place in dynamo behavior (more on this shortly) and once again yields singly periodic solutions as C_S is further increased. Although it is not immediately apparent in the figure, the numerical solutions include a narrow 5 period window at $C_S \simeq 226.5$, which is the second-widest periodic window "predicted" by the one-dimensional iterative map, at $\alpha \simeq 1.644$ (see Fig. 2).

---

[3] In Fig. 3 we actually plot the square root of the magnetic energy peaks as a measure of the cycle amplitude in order to facilitate the comparison with Fig. 2, itself constructed directly from the amplitude iterates p_n.
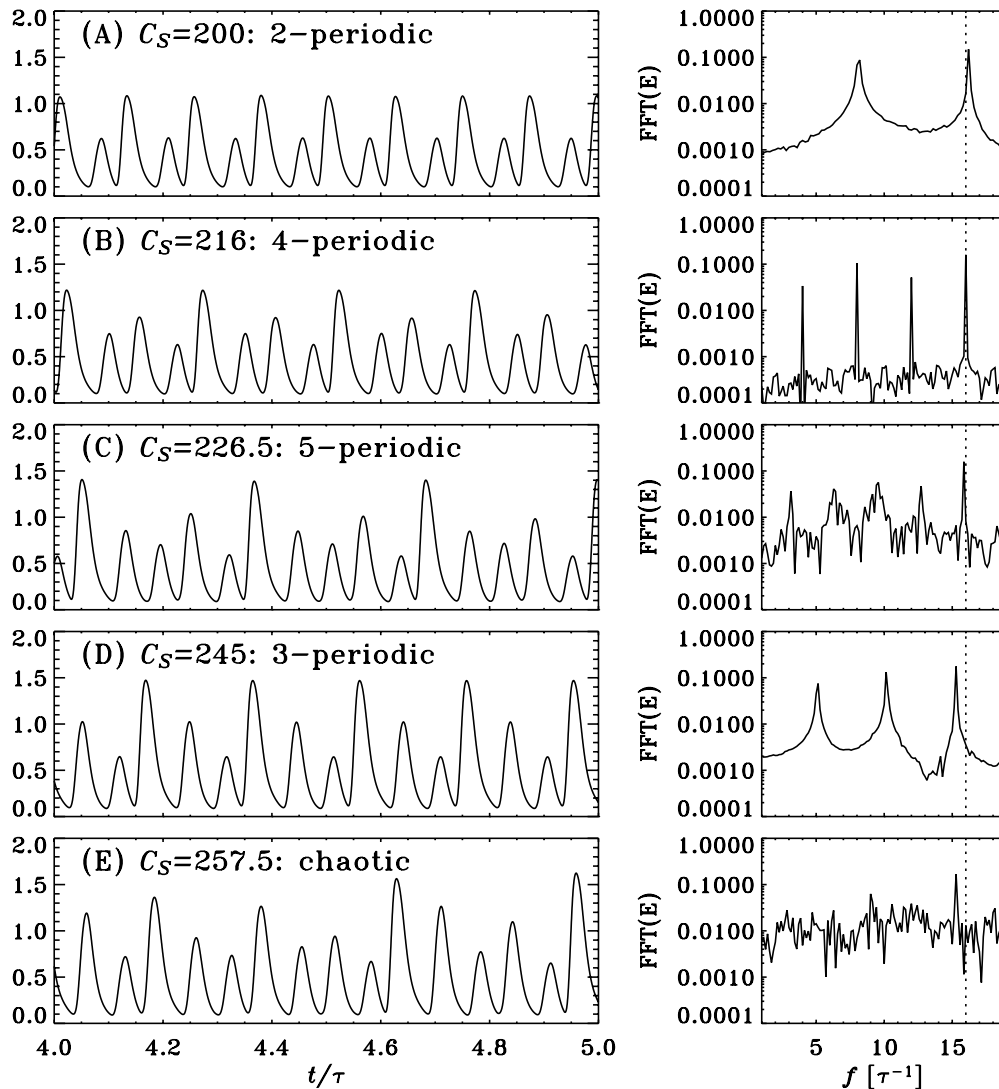
Fig. 4.—Time series (*left column*) and power spectra (*right column*) for a sample of numerical solutions at values of C_S indicated by vertical lines in Fig. 3. Interestingly, the magnetic energy curves resemble observed sunspot cycles in that their rising phases are of significantly shorter duration than their more leisurely decaying phases.

The distinction between multiperiodic and nonperiodic solutions can be quite striking upon computing power spectra of the corresponding time series, as depicted in the right column of Figure 4 for time series plotted in the left column. Multiperiodic solutions show narrow peaks in frequency standing clearly above a noisy background. The chaotic solution (*bottom frame*), on the other hand, shows a much flatter spectrum. The 5 period solution has the least well defined peaks, yet these remain clearly visible; examination of the full time series reveals that this solution occasionally "lapses" into nonperiodic behavior for some half-dozen cycles at a time before 5-periodic behavior resumes again, thus accounting for the poor quality of the power spectrum.

Closer examination of the power spectra in Figure 4 soon reveals that the frequency peak at $f \simeq 16\tau^{-1}$ is always present in all spectra (although slightly shifted in the bottom two) and, moreover, acts as the base frequency from which period doubling develops. This is readily visualized upon stacking power spectra for all solutions making up Figure 3 for increasing values of the dynamo number C_S. This results in a form of bifurcation diagram in frequency, rather than amplitude. One such diagram is shown in Figure 5 and should be carefully compared

to Figure 3. The primary frequency stands out quite clearly in this diagram and persists in a well-defined manner even through nonperiodic intervals.

This primary frequency is associated with the turnover time of meridional circulation, which is known to be the primary determinant of cycle period in this class of dynamo models (see, e.g., § 4 of DC99). Charbonneau & Dikpati (2000) also showed that meridional circulation can also *regulate* the cycle period in the presence of high-amplitude stochastic fluctuation, effectively acting as a "clock" ensuring sustained phase coherence over many cycles. The results presented here reinforce this conclusion. It is quite noteworthy that the primary cycle frequency shows up so well throughout the two chaotic intervals in Figure 5, given that the cycle amplitude is fluctuating chaotically from one cycle to the next (cf. Fig. 4*e*). Here the primary cycle period is roughly solar, varying from ∼25 yr at $C_S = 100$ to 57 yr at $C_S = 200$. The discrepancy is not worrisome, since a twofold increase in u_0 (from 10 to 20 m s^{-1}, well within observational bounds) together with relatively minor variations in the form of the circulation profile could easily bring the dynamo period down to ∼10 yr.
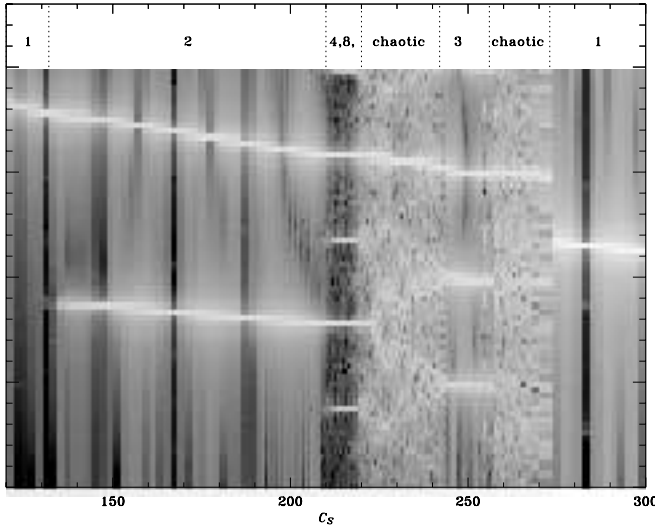
Fig. 5.—Variations of power spectra as a function of dynamo number C_S. The dotted tick marks along the top of the figure indicate the boundaries between various types of cyclic behavior, labeled by their n-fold periodicity. As one moves from left to right, period-doubling bifurcations correspond to the sudden appearance of well-defined power ridges at half the frequency of existing ridges. A frequency of 10 τ^{-1} corresponds here to a cycle period of 92 yr.

4.3. *Is It Really Chaos?*

Obviously, care must be taken before declaring this or that solution "chaotic" simply on the basis of having failed to detect periodicity in a time series that samples 100 oscillations or so. Since sensitivity to initial conditions is the hallmark of deterministic chaos, the allegedly chaotic nature of our nonperiodic solutions is best assessed by introducing a small perturbation in such a numerical solution and tracking the solution's further evolution in comparison to an unperturbed counterpart.

The result of such an exercise is shown in Figure 6 for a solution in the nonperiodic window following the first sequence of period doubling (see Fig. 3). A perturbation of relative amplitude 10^{-4} was introduced at time t^* along a thin spherical shell centered on the core-envelope interface in the model, and the evolution followed for 5000 time steps thereafter (amounting to 1.25 magnetic diffusion times). The expectation is that after a transient phase during which the spatially localized perturbation spreads across the spatial domain and lasting presumably some significant fraction of the diffusion time, the two solutions will start to diverge exponentially. This is indeed what happens in the numerical solutions, as shown in Figure 6. The solid line in Figure 6a is the magnetic energy time series $E_B(t)$ for the reference (unperturbed) solution, starting at the time t^* when the perturbation was introduced. The dashed line is the corresponding time series $E_B^*(t)$ for the perturbed solution. The two solutions are indistinguishable on this plot up to about 0.8 diffusion time past the introduction of the perturbation, but then they diverge markedly.

Figure 6b shows a time series of the normalized unsigned difference $\Delta(t)$ between the two magnetic energy time series, defined as

$$\Delta(t) = \left| \frac{E_B^*(t) - E_B(t)}{E_B^*(t) + E_B(t)} \right|. \tag{14}$$

This time series is very noisy, so we smoothed it with a 1000 iteration wide boxcar averaging filter, resulting in the thick solid line plotted in Figure 6b. The quantity $\Delta(t)$ varies ir-
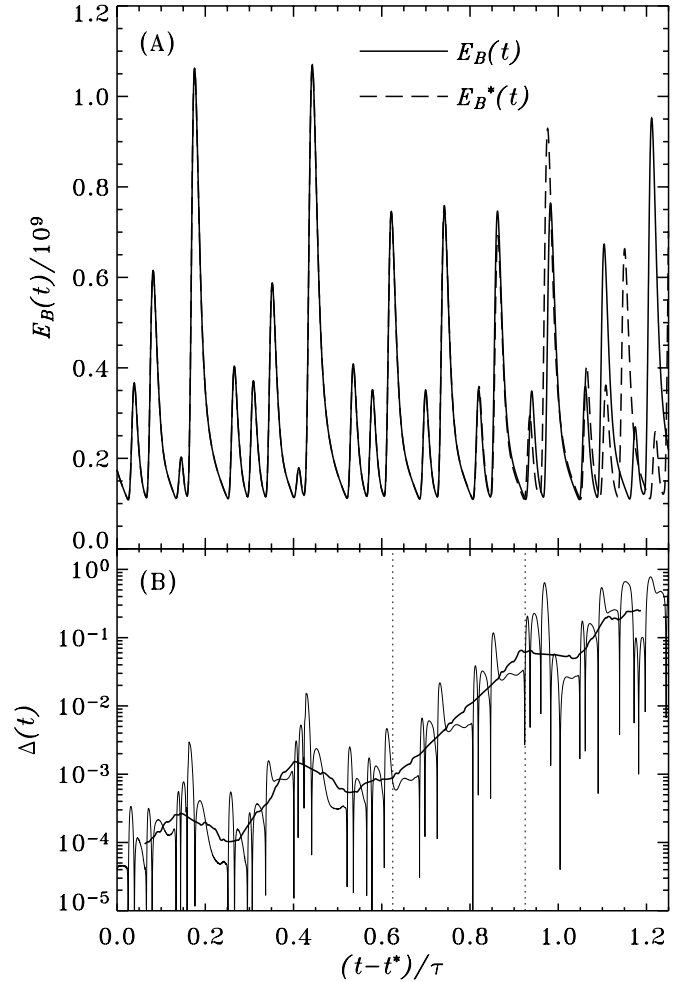

Fig. 6.—Sensitivity to initial conditions in a nonperiodic solution. (*a*) The solid line shows a reference time series of magnetic energy $E_B(t)$, and the dashed line shows the corresponding time series $E_B^*(t)$ for a solution in which a spatially localized perturbation of normalized amplitude 10^{-4} was introduced at time t^*. (*b*) The normalized unsigned difference between the two time series (*thin solid line*), on which is superimposed a smoothed version (via a 1000 iteration running boxcar filter). The two vertical dotted lines bracket a phase of exponential increase in $\Delta(t)$. The exponential divergence of the two solutions indicates sensitivity to initial conditions, the hallmark of deterministic chaos.

regularly at first, corresponding to the transient phase associated with the spatial spread of the perturbation, but then enters a phase of exponential increase during which $\Delta(t)$ increases by more than 2 orders of magnitude. After about 1 diffusion time, the two solutions have become entirely decorrelated, and thereafter $\Delta(t)$ remains of order unity. This is exactly the behavior expected from deterministic chaos and confirms that our "chaotic" solutions are indeed chaotic in the strict sense of the word.

4.4. *Behavior at High C_S*

The sudden drop in frequency so striking at $C_S = 272$ in Figure 5 heralds a switch to a qualitatively distinct dynamo behavior. At all smaller dynamo numbers, the primary cycle frequency is that associated with the turnover time of meridional circulation, with only weak dependence on the dynamo number. However, upon reaching $C_S = 272$, the primary frequency drops abruptly by some 25%, while chaotic solutions give way equally suddenly to a singly periodic cycle with amplitude now increasing much more rapidly with C_S than prior to the transition

(see Fig. 3). Note also that this dynamo branch has no counterpart in the bifurcation diagram arising from the iterative map (Fig. 2).

This high-C_S branch is one in which the kinematic time delay associated with meridional circulation no longer plays a leading role in the dynamical behavior. Instead, at the onset of the cycle the magnetic field grows extremely rapidly, leading to a quenching on the source term so swift that the cycle cannot really get underway; i.e., circulation no longer has time to carry the poloidal field equatorward. Instead, the toroidal magnetic field grows until the poloidal field begins to decay resistively, which in turn leads to the resistive decay of the toroidal field. When the latter has fallen to amplitudes low enough for the source term to turn on again, a new cycle begins.

4.5. Robustness of Results

We have obtained results qualitatively similar to those presented above for other choices of amplitude-limiting nonlinearities and over substantial portions of the parameter space defined by the DC99 dynamo model, as modified to include a lower threshold on the B-L source term. Not surprisingly, the solutions are most sensitive near the bifurcation points to the size of the spatial mesh used, the time-step size, and other purely numerical details.

The bifurcation structure is most maplike in models with a high Reynolds number, a high magnetic diffusivity contrast between core and envelope, and/or circulation penetrating below the core-envelope interface. The high Reynolds number requirement (meaning low envelope diffusivity η_T for a fixed circulation speed u_0; cf. eq. [7]) is important here; if the envelope magnetic diffusivity is large, the diffusion time across the envelope can become comparable to or even shorter than the advection time by the meridional flow, which leads to a "short-circuiting" of the dynamo conveyor belt and a loss of memory of past cycle amplitudes. Such solutions typically do not show a transition to chaos through period-doubling bifurcations and instead often remain singly periodic over a very wide range of dynamo numbers. The dynamical behavior discussed here thus characterizes B-L models operating in the advection-dominated regime; for $u_0 \sim 20$ m s^{-1}, this requires $\eta_T \lesssim 4 \times 10^{11}$ cm^2 s^{-1} or, equivalently, Rm $\gtrsim 300$.

While transition to chaos via a sequence of period-doubling bifurcations characterizes the model's behavior in wide regions of parameter space, the degree of resemblance to the bifurcation structure calculated using the iterative map is often not as good as that between Figures 2 and 3. The distinction hinges on the degree of invariance of the solution's spatial planforms to the overall amplitude of the cycle. This effect is illustrated in Figure 7, which shows time-latitude ("butterfly") diagrams of the toroidal magnetic field constructed at the core-envelope interface $r/R = 0.7$. The top solution is the $C_S = 228$ solution from the sequence whose bifurcation diagram is shown in Figure 3, while the bottom solution has a much higher meridional flow speed (parameter values $\eta_T = 1.67 \times 10^{11}$ cm^2 s^{-1}, $C_\Omega = 84,000$, $C_S = 40$, Rm $= 2335$, $\Delta\eta = 500$, and $r_b = 0.68$). In the first case, successive cycles have more or less the same spatiotemporal structure in this diagram, while this clearly is not the case for the second solution. In the former case, the conveyor belt always operates in the same manner independently of cycle amplitude, with toroidal magnetic fields peaking around $60°$ latitude; in the latter case, however, the cycles can look very different from one cycle to the next, with the field peaking at polar latitudes for some cycles and showing a mid- to low-latitude branch of comparable strength for others. In such a
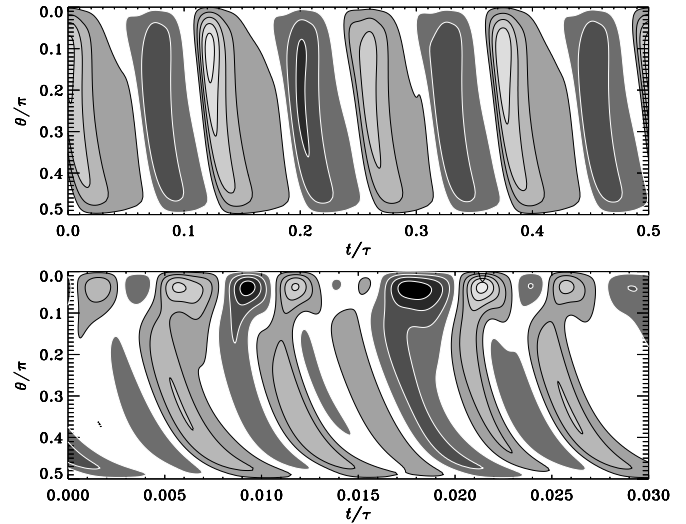

Fig. 7.—Two butterfly diagrams for chaotic solutions in two different parameter regimes. The gray scale indicates the toroidal magnetic field at the core-envelope interface in the model ($r/R = 0.7$), in which sunspot-producing toroidal flux ropes presumably form prior to their buoyant destabilization and rise through the overlying convective envelope. In the top diagram, the solution's planform remains more or less the same independent of the overall amplitude; this is not true of the bottom diagram. Because of this, the bifurcation diagram for the first solution is well reproduced by the one-dimensional iterative map, while the second is not (see text).

situation the simple conveyor-belt argument of Figure 1 cannot be expected to hold very well.

A numerical bifurcation diagram for the sequence of solutions to which that plotted in the bottom panel of Figure 7 belongs shows similarities as well as important differences with that produced by the iterative map (Fig. 2). As the dynamo number is increased, cyclic activity appears at a supercritical bifurcation, but the singly periodic regime rapidly undergoes a first period doubling, eventually transiting to chaos via further period doublings, in the usual manner. However, as the dynamo number is further increased, chaos almost immediately gives way to multiperiodic behavior with multiple pairs of branches gradually merging as C_S is further increased until stable, doubly periodic behavior is recovered. Chaos then reappears abruptly and persists to much higher dynamo numbers, within which the widest imbedded periodic window is of even periodicity. None of these latter features are characteristic of classical, single-hump iterative maps.

With toroidal fields often peaking at high latitudes, neither one of the time-latitude diagrams plotted in Figure 7 compares well to the sunspot butterfly diagram. However, stability analyses of toroidal flux ropes stored immediately beneath the core-envelope interface indicate that flux ropes stored at high latitudes are much harder to destabilize because of the strong magnetic tension associated with the greater curvature about the polar axis (see, e.g., Fig. 1 in Ferriz-Mas et al. 1994). The high-latitude toroidal fields in Figure 7 may well remain invisible on a sunspot butterfly diagram. Concentration of toroidal fields closer to the equatorial regions can also be achieved by allowing for deeper penetration of the meridional flow into the tachocline (e.g., Nandy & Choudhuri 2002; also DC99), although the recent boundary layer analysis of Gilman & Miesch (2004) indicates that this is probably not a viable alternative. At any rate, the time-delayed nonlinear behavior is largely insensitive to the detailed form of the butterfly diagram as long as the "conveyor belt" operates in an amplitude-independent manner, as discussed above.

5. CONCLUDING REMARKS

In this paper, we have presented a series of numerical simulations of the solar magnetic activity cycle showing a well-defined transition to chaos via a sequence of period-doubling bifurcations as the numerical value of the model's source parameter is increased. Multiperiodic and chaotic behavior originates with the long time delay built into the dynamo process, as confirmed by the very good agreement of the numerical results with the behavior predicted using a simple one-dimensional iterative map, at all but very high dynamo number values.

In the solar cycle context, we are aware of no other example in which a PDE-based two-dimensional model incorporating only a simple amplitude-limiting quenching nonlinearity can be shown to undergo transition to chaos through a sequence of period-doubling bifurcations. One related case in torus geometry has been well documented by Brooke & Moss (1994), but the transition to chaos could not be resolved beyond the first period-doubling bifurcation.

Numerous procedures have been proposed whereby the dynamo-governing equations are truncated to yield a dynamical system described by a set of nonlinearly coupled ordinary differential equations resembling the famed Lorenz system (e.g., Ruzmaikin 1981; Weiss et al. 1984). Alternately, the search for chaotic behavior has been made in various more elaborate and spatially extended dynamo models, as in the present paper, but including various forms of nonlinear feedback mechanisms. Here also a variety of irregular behavior has been observed, with some deemed chaotic (see, e.g., Belvedere et al. 1990; Brandenburg et al. 1991; Schmalz & Stix 1991; Roald & Thomas 1997). Attempts have also been made to connect the two approaches by looking for similarity between the bifurcation structure of low-order dynamical systems and that reconstructed from numerical solutions to spatially extended dynamo models, including dynamical nonlinearities. In some cases the comparison is, in fact, quite favorable (Knobloch et al. 1998), suggesting that suitably constructed low-order dynamical systems do catch some essential features of the full problem and that the nonlinear behaviors observed therein are generic. On the other hand, some important aspects of the nonlinear behavior are also found to depend, sometimes quite sensitively, on what one would have hoped would be relatively minor details of the model's formulation (see, e.g., Phillips et al. 2002 and discussion therein). At any rate, in all these cases, the nonlinearities built into the dynamical systems and introduced in the numerical models relate directly to the back-reaction of the dynamo-generated magnetic field on the fluid motions powering the dynamo.

In contrast, the results presented in this paper indicate that observed fluctuations in the amplitude of the solar cycle may also originate with a simple, kinematic time-delayed feedback effect, rather than with magnetically mediated nonlinear dynamical back-reaction on fluid motions. It must be emphasized that in the context of solar cycle models based on the B-L mechanism of poloidal field regeneration, such a time delay is not at all an ad hoc modeling ingredient but instead represents an unavoidable consequence of the spatial segregation of the two source regions for the toroidal and poloidal magnetic components (see Fig. 1).

The dynamo solutions presented here also confirm the remarkable cycle period regulation achieved by meridional circulation in B-L models. Charbonneau & Dikpati (2000) have demonstrated this property in the context of a singly periodic B-L dynamo model subjected to strong stochastic forcing, and supporting observational evidence has been presented in Hathaway et al. (2003). Here the cycle frequency is found to remain constant within 20% despite period doubling and transition to chaos taking place as the dynamo number is increased by nearly 2 orders of magnitude. Meridional circulation is truly acting as a clock regulating the cycle period in this class of dynamo models. In more general terms, we wish to also emphasize that the inclusion of meridional circulation in the dynamo equations is far from being an ad hoc addition. Meridional flows have now been detected helioseismically down to $r/R \sim 0.8$, and mass conservation requires an equatorward return flow somewhere in the interior. A surface meridional flow has long been known to be an important ingredient for properly reproducing the spatiotemporal evolution characterizing synoptic magnetograms (e.g., Wang et al. 2002 and references therein). Evidence is also accumulating that the speed (and fluctuations) of the surface meridional flow have a determining influence on some observed solar cycle characteristics, such as the relative amplitudes of successive cycles (Hathaway et al. 2003, but see also Schüssler & Schmitt 2004) and the timing of polar field reversals (Dikpati et al. 2004). Including meridional circulation in dynamo models is no longer a mere modeling option; it is there and it plays an important role, and not only in models based on the B-L mechanism of poloidal field regeneration.

Finally, we note that the remarkable agreement between the behavior observed in the spatially extended numerical model and the corresponding one-dimensional iterative map legitimizes the use of such map to investigate phenomena that would be impractical to study using the full numerical model, most notably intermittency. We turn to this interesting issue in the next paper of this series (P. Charbonneau et al. 2005, in preparation).

REFERENCES

Babcock, H. W. 1961, ApJ, 133, 572
Belvedere, G., Pidatella, R. M., & Proctor, M. R. E. 1990, Geophys. Astrophys. Fluid Dyn., 51, 263
Brandenburg, A., Moss, D., Rüdiger, G., & Tuominen, I. 1991, Geophys. Astrophys. Fluid Dyn., 61, 179
Braun, D. C., & Fan, Y. 1998, ApJ, 508, L105
Brooke, J. M., & Moss, D. 1994, MNRAS, 266, 733
Caligari, P., Moreno-Insertis, F., & Schüssler, M. 1995, ApJ, 441, 886
Charbonneau, P. 2001, Sol. Phys., 199, 385
Charbonneau, P., & Dikpati, M. 2000, ApJ, 543, 1027
Choudhuri, A. R. 1992, A&A, 253, 277

Christensen-Dalsgaard, J. 2002, Rev. Mod. Phys., 74, 1073
Dikpati, M., & Charbonneau, P. 1999, ApJ, 518, 508 (DC99)
Dikpati, M., De Toma, G., Gilman, P. A., Arge, C. N., & White, O. R. 2004, ApJ, 601, 1136
Dikpati, M., & Gilman, P. A. 2001, ApJ, 559, 428
D'Silva, S., & Choudhuri, A. R. 1993, A&A, 272, 621
Durney, B. R. 1995, Sol. Phys., 160, 213
———. 2000, Sol. Phys., 196, 1
Fan, Y., Fisher, G. H., & DeLuca, E. E. 1993, ApJ, 405, 390
Ferriz-Mas, A., Schmitt, D., & Schüssler, M. 1994, A&A, 289, 949
Gilman, P. A. 1983, ApJS, 53, 243

Gilman, P. A., & Miesch, M. S. 2004, ApJ, 611, 568
Hathaway, D. H., Nandy, D., Wilson, R. M., & Reichmann, E. J. 2003, ApJ, 589, 665
Holton, D., & May, R. M. 1993, in The Nature of Chaos, ed. T. Mullin (Oxford: Clarendon), 95
Hoyng, P. 1988, ApJ, 332, 857
———. 1993, A&A, 272, 321
Hoyng, P., Schmitt, D., & Teuben, L. J. W. 1994, A&A, 289, 265
Knobloch, E., Tobias, S. M., & Weiss, N. O. 1998, MNRAS, 297, 1123
Küker, M., Arlt, R., & Rüdiger, G. 1999, A&A, 343, 977
Leighton, R. B. 1964, ApJ, 140, 1547
———. 1969, ApJ, 156, 1
Longcope, D. W., & Fisher, G. H. 1996, ApJ, 458, 380
Moffatt, H. K. 1978, Magnetic Field Generation in Electrically Conducting Fluids (Cambridge: Cambridge Univ. Press)
Moreno-Insertis, F. 1986, A&A, 166, 291
Moreno-Insertis, F., Caligari, P., & Schüssler, M. 1995, ApJ, 452, 894
Moss, D., Brandenburg, A., Tavakol, R., & Tuominen, I. 1992, A&A, 265, 843
Moss, D., & Brooke, J. 2000, MNRAS, 315, 521
Nandy, D., & Choudhuri, A. R. 2001, ApJ, 551, 576
———. 2002, Science, 296, 1671
Ossendrijver, M. 2000, A&A, 359, 1205
———. 2003, A&A Rev., 11, 287
Ossendrijver, M., Hoyng, P., & Schmitt, D. 1996, A&A, 313, 938

Petrovay, K. 2000, in The Solar Cycle and Terrestrial Climate, ed. A. Wilson (ESA SP-463; Noordwijk: ESA), 3
Phillips, A., Brooke, J., & Moss, D. 2002, A&A, 392, 713
Roald, C. B., & Thomas, J. H. 1997, MNRAS, 288, 551
Rüdiger, G., & Arlt, R. 2003, in Advances in Nonlinear Dynamics, ed. A. Ferriz-Mas & M. Núñez (London: Taylor and Francis), 147
Ruzmaikin, A. A. 1981, Comments Astrophys., 9, 85
Schmalz, S., & Stix, M. 1991, A&A, 245, 654
Schou, J., & Bogart, R. S. 1998, ApJ, 504, L131
Schrijver, C. J., DeRosa, M. L., & Title, A. M. 2002, ApJ, 577, 1006
Schüssler, M. 1994, in Solar and Astrophysical Magnetohydrodynamic Flows, ed. K. C. Tsinganos (NATO ASI Ser. C, 148; Dordrecht: Kluwer), 17
Schüssler, M., & Schmitt, D. 2004, A&A, 421, 249
Stix, M. 2002, The Sun: An Introduction (2nd ed.; Berlin: Springer)
Thelen, J.-C. 2000, MNRAS, 315, 155
Tobias, S. M. 1997, A&A, 322, 1007
Wang, Y.-M., Lean, J., & Sheeley, N. R., Jr. 2002, ApJ, 577, L53
Wang, Y.-M., Nash, A. G., & Sheeley, N. R., Jr. 1989, Science, 245, 712
Wang, Y.-M., & Sheeley, N. R., Jr. 1991, ApJ, 375, 761
Wang, Y.-M., Sheeley, N. R., Jr., & Nash, A. G. 1991, ApJ, 383, 431
Weiss, N. O., Cattaneo, F., & Jones, C. A. 1984, Geophys. Astrophys. Fluid Dyn., 30, 305
Yoshimura, H. 1978, ApJ, 226, 706