Global Bifurcation of Shilnikov Type in a Double-Gyre Ocean Model

By BALASUBRAMANYA T. NADIGA[1] and BENJAMIN P. LUCE[2]

[1]Earth and Environmental Sciences, MS-B296
Los Alamos National Laboratory, Los Alamos NM 87545
email:balu@lanl.gov

[2] Theoretical Division, MS-B284
Los Alamos National Laboratory, Los Alamos, NM 87545
email:luceb@lanl.gov

The dynamics of an idealized wind-driven double-gyre circulation in an ocean basin are studied from a dynamical systems point of view in an effort to better understand its variability. While previous analyses of this circulation have mostly dealt with local bifurcations of steady states and limit cycles, this study demonstrates the importance of considering global bifurcations as well. In one case, a coherent picture of the global dynamics spanning a range of parameters from where there are only stable steady-state solutions to where there is chaotic eddy-shedding is presented. A simple but novel use of power spectra along with dynamical projections of the dynamics suggests that just beyond the regime in which there are only stable steady-states, the system exhibits a complicated global bifurcation known as "Shilnikov Phenomenon".

1. Introduction

The world oceans are an important constituent of our climate system. They account for a significant fraction of the poleward heat transport (e.g., see Gill (1982)) and are the main sources of long time scales in climate dynamics. The dynamics of ocean circulation are, therefore, expected to play a major role in climate variability. However, even the response of the oceans themselves to specified mechanical and thermodynamic forcings—wind-driven circulation and thermohaline circulation respectively—is, in general, very complicated. While simple models are useful in understanding various individual features of this complicated response, realistic descriptions invariably require the use of ocean general circulation models (OGCMs). The downside of OGCMs, however, is their immense complexity resulting in part from the very large range of spatial and temporal scales and the large number of different physical processes involved.

The feature of ocean circulation of interest to us in this article is its variability. To reproduce the observed variability of ocean circulation in models requires, besides using realistic OGCMs, an extensive tuning of parameters and parameterizations of the models. On the other hand, even simple models are capable of significant variability (see e.g., chapter 2 in Pedlosky (1996) and references therein, and McCalpin & Haidvogel (1996)). While it would be almost hopeless to carry out a systematic analysis of the origin of variability in the OGCMs, we think that such an attempt can now be made at least for some of the simple models. Although it is not clear to what extent the results of such analysis will carry over to OGCMs, study of the simple models is a logical place to start, and may shed light on some of the fundamental dynamical mechanisms involved.

Simple models that have been particularly useful in understanding various aspects of the dynamics of midlatitude wind-driven ocean circulation are those that assume the hydrostatic and possibly also the quasi-geostrophic (QG) approximation, simplify the vertical representation to a few layers, idealize the domain into a closed rectangular basin, and idealize the wind forcing (see e.g., chapters 2 & 3 in Pedlosky (1996) and references therein). In this article, we consider a commonly studied QG model of this type wherein the upper ocean is represented by a single active layer overlying a deeper, quiescent, "half" layer. The most serious shortcoming of such a model is that it does not support baroclinic instabilty; and our justification for studying it is that the variability even for this case is not well understood.

Considering a barotropic model of the sub-tropical ocean gyre circulation forced by steady wind stress, computations performed by Bryan (1963), for the first time, demonstrated a transformation of flow from a classical Stommel solution to one with an inertial recirculation to unsteady solutions as the degree of nonlinearity was increased. Veronis (1966), using a different dissipation operator, then pointed out a different transformation of steady states as the degree of nonlinearity increased, with more Fofonoff-like solutions emerging at large nonlinearities. These works were followed by various other investigations which noted the spontaneous generation of mesoscale variability in ocean circulation driven by steady winds at moderate values of nonlinearity. It was then pointed out by Holland & Haidvogel (1981), in the context of a baroclinic QG model, that a lower frequency variability (than that of the mesoscales) appears at still higher values of nonlinearity. Le Provost & Verron (1987), by considering unsteady computations of the barotropic double-gyre circulation, identified two types of destabilization—meandering of the midlatitude eastward jet and Rossby wave radiation from the westward return flow, and Moro (1988) considering computations of steady states, but with asymmetric double-gyre wind stress demonstrated the possiblity of stable standing meanders in the midlatitude jet

However, detailed studies of variability in simple models of wind-driven ocean circulation began only more recently with the work of Jiang, Jin, & Ghil (1995), when they showed the presence of multiple equilibria, and periodic and aperiodic solutions in the context of a double-gyre circulation driven by symmetric wind stress and using a reduced-gravity shallow water model. (Henceforth, when we refer to the double-gyre circulation, the reader should assume that we are actually referring to an idealized realization of it in a simple model rather than its more realistic counterparts, and if we do not explicitly specify it, that it is driven by symmetric wind stresses.) At about the same time, Ierley & Sheremet (1995) demonstrated the presence of multiple equilibria in a barotropic single-gyre QG model, Cessi & Ierley (1995) mapped the (anti-symmetric and symmetry-breaking) steady states of a barotropic QG double-gyre model as a function of frictional and forcing parameters, and studied their stability, and Speich, Dijkstra, & Ghil (1995) did the same for the model used in Jiang, Jin, & Ghil (1995). (Note that while the steady states of the barotropic QG case with no quiescent deep layer are the same as those for its reduced-gravity counterpart, the time-dependent dynamics of the two are different.)

McCalpin & Haidvogel (1996), on the other hand, integrated a reduced-gravity QG model of the double-gyre circulation to study its low-frequency variability and suggested that the low-frequency variability that they observed was associated with irregular transitions among three preferred states—a high energy state with the mid-latitude jet penetrating deep into the basin, a low energy state with a weakly penetrating jet, and an intermediate energy state. (These preferred states have been linked to the unstable steady states of that model by various authors (e.g., Primeau (1998) and Scott & Straub (1998).)) While Dijkstra & Katsman (1997) examined bifurcations of steady states and formation of limit cycles for the reduced-gravity and two layer QG models of double-gyre circulation, Primeau (1998) mapped out the bifurcation of steady states for a barotropic QG model but with a different dissipation operator than used in Cessi & Ierley (1995). This was followed by the Berloff-Meacham program, which considered, rather extensively, the low-dimensional dynamics of the single gyre circulation using both barotropic and baroclinic QG models (Berloff & Meacham (1997), Meacham & Berloff (1997), Berloff & Meacham (1998)). In this context, they identified a sequence of Hopf bifurcations giving rise to an unstable toroidal attractor that, then, gives way to a strange attractor. Berloff & McWilliams (1999) consider the reduced-gravity and two layer QG models of double-gyre circulation with both symmetric and asymmetric wind forcing. In each of the cases, they find a primary (intermonthly) and a secondary (interannual) Hopf bifurcation and then go on to characterize the differences in the low-frequency variability of their different cases at higher nonlinearities using empirical orthogonal functions. Various aspects of the double-gyre circulation are considered in a recent study by Ghil, Feliks, & Sushama (2000). Finally, in a recent article (published since this work was completed), Meacham (2000) in trying to explain large amplitude, low-frequency variability of aperiodic solutions in a barotropic double-gyre circulation, associated it with homoclinic bifurcations.

Thus, with the exception of Meacham (2000), all the low-dimensional dynamics that have been identified in the context of wind-driven circulation are local bifurcations of steady states and periodic solutions, and a comprehensive study of the role of more complicated solution types is still lacking. We note that a possible special solution in the problem under consideration is (what is called in the nonlinear dynamics literature as) a "homoclinic orbit"—which in the present context would mean a time-dependent double-gyre solution which tends asymptotically to a steady-state solution (discussed previously) in both forward and backward time. (An orbit would be called "heteroclinic" if it connected two different asymptotic solutions.) Such an orbit provides the fundamental recurrence mechanism which makes possible the repeated "stretching and folding" of phase space that is characteristic of dynamical chaos. Furthermore, it is now well recognized that bifurcations involving homoclinic orbits can act as "organizing centers" for the dynamics of low-dimensional dynamical systems. That is, these solutions often play a fundamental role in the bifurcation of new solutions as parameters are varied, and a crucial role in the onset of chaotic behavior. It is also easy to imagine how the recurrent structure of a homoclinic orbit can provide a setting wherein new periodic solutions can be created—bifurcating from the homoclinic orbit as parameters are varied. For examples of homoclinic and hetroclinic orbits, and their importance in the dynamics of highly truncated (ordinary differential equation) models of geophysical flows, see Ghil & Childress (1987).

In this paper we report evidence which suggests that the dynamics of the double-gyre circulation (in a reduced-gravity QG model) just beyond the regime in which there are only stable equilibria (steady-states) exhibit a complicated global bifurcation which is known in the literature as "Shilnikov Phenomenon" (SP). We deduce this by comparing power spectra and dynamical projections of the double-gyre circulation with known characteristics of the SP. SP, which was first partially described by Shilnikov (1965), and later studied by Glendinning & Sparrow (1984), Gaspard, Kapral, & Nicolis (1984), Tresser (1984), and others, involves the collision of limit cycles with saddle-foci equilibria in phase space creating homoclinic orbits and other dynamical structures. Thus, while previous bifurcation analysis of the double-gyre circulation have all involved only local bifurcations, to the best of our knowledge our evidence for SP seems to be first instance of considering a global bifurcation in this context.

As has been shown in the previous studies, when nonlinearity is initially increased in the symmetrically forced double-gyre circulation, antisymmetric steady states give way to asymmetric steady states and then to periodic solutions, through a series of local bifurcations. Moreover, in the context of the beta-plane approximation, it is physically easy to see that the dynamics of the linearization about any one of the steady state solutions would involve Rossby waves. In particular, considering the antisymmetric steady state solution branch beyond the first (codimension one, pitchfork) bifurcation when it loses stability, the simplest picture of the dynamics of the linearization about that solution would be that of a saddle-focus, with Rossby waves comprising the two-dimensional stable oscillatory mode. It is then easy to imagine the periodic solutions (asymmetric limit cycles, after they are possibly involved in other local bifurctions) colliding with this saddle-focus, providing the basic framework necessary for SP.

Much of the work on Shilnikov chaos has been in the three-dimensional (phase space) setting, and it is this three-dimensional theory that we consider here. The relevance of such a theory to the double-gyre arises from the fact that although the system of partial differential equations (pde) describing such a system (e.g., the reduced-gravity QG model) is capable, *a priori*, of exhibiting an arbitrarily large number of degrees of freedom in its dynamics, dissipative mechanisms suppress instabilities at wavenumbers above a threshold. This threshold increases with increasing wind stress, allowing finer scale structure with more degrees of freedom to emerge as forcing increases. In the regime considered in this paper, the wind stress is small enough that only a few instabilities are present in the time-asymptotic dynamics and these involve only a small number of degrees of freedom. For many pde systems with driving and damping, such as the forced two-dimensional Navier-Stokes equations in a bounded domain, the finite dimensionality of the asymptotic dynamics can be made rigorous via the concept of inertial manifolds (Constantin, Foias & Temam (1985)). While such a demonstration is likely to be possible for the double-gyre circulation, we remark that the question of the *exact* effective dynamical dimension for a pde in a given parameter regime is often a difficult question, and in practice, dynamicists tend to rely on numerical measurements such as Lyapunov dimension instead of rigorous theory. Although we do not fully address this question in this article, there do appear to be some interesting and subtle aspects to the question of dynamical dimension for the double-gyre circulation, which we discuss further in section 6.

Finally, in the context of using a double-gyre wind forcing, we note that although the details of bifurcation diagrams are bound to be dependent on the exact form of the dissipation operator and the associated boundary conditions, the possibility of "internal compensation" is expected to reduce such a dependence. Internal compensation is the process wherein a) the potential vorticity of a fluid parcel is changed by input of wind stress curl in say the sub-tropical gyre, b) the fluid parcel is advected across the zero wind stress curl line to the sub-polar gyre, and c) the potential vorticity of that fluid parcel is reset in part by input of wind stress curl of the opposite sign by virtue of being in the sub-polar gyre (see e.g., Marshall (1984), Pedlosky (1996), Primeau (1998), Greatbatch & Nadiga (1999)). The inviscid nature of this transport-mechanism allows a weak dependence of the dynamics on dissipative processes in a weakly dissipative regime. It is, however, clear that once symmetry is broken the importance of dissipative processes has to reemerge. For example, it is perhaps because of internal compensation that low-frequency variability of the kind found by McCalpin & Haidvogel (1996) using a QG model persists in the primitive equation model studies of Spall (1996), the equilibria found by Jiang, Jin, & Ghil (1995) in the shallow water model are perturbed versions of those found by Cessi & Ierley (1995) in the corresponding QG model, and the time-mean circulation of turbulent double-gyre simulations in Greatbatch & Nadiga (1999) are practically the same (their Figs. 4 and 5) when either freeslip or super-slip boundary conditions are used. It is also in the same vein that Scott & Straub (1998) point out that certain aspects of their nonsymmetric solutions appear insensitive to (poorly known) frictional parameters, when these parameters are smaller than certain critical values. In such cases, they point out that the flow organizes itself in such a way as to reduce the input of energy from the wind forcing when friction is reduced (also see Greatbatch & Nadiga (1999)).

An outline of the rest of the paper is as follows. In section 2, we describe a simple model of the double-gyre circulation that we use, including details of the setup and the numerical integration scheme. In section 3, we briefly introduce the dynamical concepts used in this article and give a brief sketch of the Shilnikov phenomenon. After explaining the data types that we use in the analysis of the double-gyre circulation in section 4, we present evidence for the Shilnikov Phenomenon in the double-gyre circulation in section 5. Section 5 also considers two other better understood examples to validate our methodology. In section 6, we briefly discuss the issue of dimension of the double-gyre circulation in relation to exact mass conservation (Kelvin wave parameterization) and conclude in section 7.

2. The Reduced-Gravity Quasi-Geostrophic Model

We consider the reduced-gravity QG model in vorticity form to study the nonlinear dynamics of an idealized sub-tropical/sub-polar gyre system. In this description of large scale circulation, a single active layer of water of constant density ρ_0 and variable thickness h, representative of the upper-ocean, overlies an infinitely deep and quiescent bottom layer of density $\rho_0 + \Delta\rho$. The interface between these two layers, across which there is no mass transport, models the behavior of the permanent thermocline. We choose the equilibrium depth of the upper layer to be 1000 meters, its density ρ_0 to be 1022 kg m^{-3} and the density difference between the upper and lower layers, $\Delta\rho$, to be about 3.3 kg m^{-3}. We consider a rectangular model domain extending 1000 kilometers in the East-West (x) direction and 2000 kilometers in the North-South (y) direction, centered at 30 degrees North. The setup of the domain is shown in Fig. 1. The Coriolis parameter is assumed to vary as $f = f_0 + \beta y$ (β-plane approximation), where $f_0 = 2\Omega \sin\theta = 7.3 \times 10^{-5}$ s^{-1} corresponds to a latitude, θ, of 30 degrees and β the latitudinal variation of rotation is assumed to be 2×10^{-11} m^{-1} s^{-1}. The upper layer flow is then described by

$$\frac{\partial \zeta}{\partial t} - \frac{f_0^2}{g'H}\frac{\partial \psi}{\partial t} + J[\psi, \zeta] + \beta\frac{\partial \psi}{\partial x} = \frac{1}{\rho_0 H}\nabla \times \tau + D, \tag{2.1}$$

where ψ is the streamfunction,

$$\zeta = \nabla^2\psi$$

is the relative vorticity, $J[\ ,\ ]$ is the Jacobian operator given by

$$J[\psi, \zeta] = -\frac{\partial \psi}{\partial y}\frac{\partial \zeta}{\partial x} + \frac{\partial \psi}{\partial x}\frac{\partial \zeta}{\partial y},$$

τ is the wind stress forcing, and D represents dissipation. The quantity g', the reduced gravity, is related to the gravitational acceleration g by

$$g' = \frac{\Delta\rho}{\rho_0} g.$$

The streamfunction ψ is constant along the boundary to ensure that there is no flow normal to the boundary, but this constant can have a time-dependence:

$$\psi_B = c(t). \tag{2.2}$$

The wind stress τ is chosen to have an idealized double gyre structure (see Fig. 1) given by

$$\tau(x,y) = \left(\tau_0 \cos(\pi \frac{y}{L}), 0\right),$$

where τ_0 is the magnitude of the horizontal wind stress (in N m^{-2}). Note that with this wind stress, which is symmetric about $y = 0$, the dynamics of the double-gyre circulation is invariant under the transformation

$$\psi \to -\psi, y \to -y. \tag{2.3}$$

The dissipation operator D is chosen to be a combination of biharmonic mixing and Rayleigh damping of momentum, resulting in

$$D = -A_0\zeta - A_4\nabla^4\zeta,$$

following Marshall (1984), to emphasize the importance of the eddy field in transporting potential vorticity across the zero wind stress curl line. While the highly scale selective biharmonic mixing allows for an effective removal of small-scale enstrophy which results from nonlinear cascade processes, the scale insensitive Rayleigh damping allows for an equilibration of the energy of the system. For biharmonic friction, we choose freeslip boundary conditions:

$$\zeta = 0,$$

and

$$\frac{\partial^2\zeta}{\partial n^2} = 0,$$

where n is normal to the boundary.

A nondimensionalization of (2.1) using a length scale corresponding to the horizontal extent of the domain L, a velocity scale U given by the Sverdrup relation

$$U = \frac{\pi\tau_0}{\rho_0 H\beta L},$$

and a time scale corresponding to the Sverdrup gyre-turnaround time L/U allows (2.1) to be written as

$$\frac{\partial q}{\partial t} + J[\psi, q] = \sin(\pi y) - \frac{A_0}{\beta L}\zeta - \frac{A_4}{\beta L^5}\nabla^4\zeta, \tag{2.4}$$

where q is the potential vorticity:

$$q = Ro\zeta - \frac{L^2}{L_R^2}\psi + y,$$

Ro is the Rossby number:

$$Ro = \frac{U}{\beta L^2},$$

and L_R is the Rossby radius of deformation:

$$L_R = \sqrt{\frac{g'H}{f_0^2}}.$$

Finally, for convenience, we represent the inertial scale, $\sqrt{\frac{U}{\beta}}$, as a percentage of the horizontal domain size, L, by δ_i:

$$\delta_i = 100\sqrt{Ro}. \tag{2.5}$$

The midlatitude of the basin is now at $y = 0$, the northern and southern boundaries are at $y = \pm 1$, respectively, and the western and eastern boundaries at $x = 0$ and $x = 1$ respectively (see Fig. 1). We choose to vary the Rossby number, or in effect the magnitude of the wind stress, in our experiments for chosen values of dissipation coefficients. While we have experimented with various values for the dissipation coefficients and correspondingly different ranges of wind stress magnitude, for the case we consider in detail here, the Rayleigh damping time scale was about 193 days, the value of A_4 was 2×10^{10} m^4 s^{-1} and the corresponding maximum windstress magnitude was about 0.03 N m^{-2}.

Equation (2.4) is finite differenced using the Arakawa discretization (Arakawa (1966)) for the Jacobian and centered differencing for the spatial discretization of the other terms. The time stepping is carried out using a (nominally) fifth order embedded Runge-Kutta Cash-Karp scheme (Press et al. (1992)). In addition to providing adaptive time step control and better time accuracy compared to the usual leapfrog time discretization, this forward-time discretization obviates the need for time-filtering used to remove the computational mode in the leapfrog discretization. In all our computations, we fix the resolution at 20 kilometers or $0.02L$.

3. Sketch of the Shilnikov Bifurcation

In this section, we briefly sketch the Shilnikov Phenomenon (SP), assuming only an elementary knowledge of dynamical systems theory on the part of the reader. (General introductions to dynamical systems may be found in Wiggins (1990) and Guckenheimer & Holmes (1983)). We will describe the SP for three-dimensional dissipative dynamical systems which possess complete inversion symmetry. Such a symmetry corresponds to the symmetry of the two gyres (2.3) (which our model possesses *if and only if* the domain and wind stress forcing are both chosen to be symmetric). Our summary is largely based on work by Glendinning & Sparrow (1984) and a second paper by Glendinning (1984).

To illustrate the SP, consider a one-parameter dynamical system of the form

$$\begin{aligned} \dot{x} &= -\lambda x + \omega y + P(x, y, z; R), \\ \dot{y} &= -\omega x - \lambda y + Q(x, y, z; R), \\ \dot{z} &= \rho z + W(x, y, z; R), \end{aligned} \tag{3.1}$$

which has real parameters λ, $\rho > 0$, and R (note that the phase variables x, y, and z are not directly related to the spatial coordinates of our double-gyre model). In addition, P, Q, and W are assumed to be real-analytic functions of x, y, and z which vanish along with their first derivatives when $x = y = z = 0$. The latter, along with the absence of constant terms, implies that the origin is a stationary point, i.e. a solution trajectory with the initial condition $x(0) = y(0) = z(0) = 0$ which simply remains fixed. The form of the linear terms (involving parameters ρ, λ, and ω) are such that the fixed point at the origin is of "saddle-focus type", with one positive real eigenvalue, ρ, and a pair of complex eigenvalues, $-\lambda \pm i\omega$. The positive eigenvalue ρ implies the existence of a pair of one-dimensional *unstable* manifolds on which phase trajectories depart from the origin at an exponential rate ρ (for example, the *initial* part of the homoclinic trajectory in Figure 2a). Likewise, the complex pair of eigenvalues with negative real part $(-\lambda \pm i\omega)$ implies that the origin has associated with it a two-dimensional *stable* invariant manifold on which solutions tend asymptotically to the origin in a spiral with an exponential radial contraction rate λ and oscillation frequency ω (for example, the latter part of the homoclinic trajectory sketched in Figure 2a).

A "homoclinic" orbit occurs if there is an intersection of the stable and unstable manifolds. Such an intersection immediately implies the existence of a trajectory which tends asymptotically to the fixed point as $t \to \pm\infty$. We now assume that the functions P, Q, and W are chosen such that a homoclinic of system (3.1) exists at $R = 0$. Symbolically, we represent this homoclinic solution by $(x, y, z) = (x^{\mathrm{h}}(t; \vec{r}_0), y^{\mathrm{h}}(t; \vec{r}_0), z^{\mathrm{h}}(t; \vec{r}_0))$, where $\vec{r}_0$ is any initial condition on this orbit at $t = 0$. An example of such a homoclinic to a saddle-focus is illustrated in Fig. 2a. Assume now also that P, Q, and W are invariant under the inversion

$$(x, y, z) \to -(x, y, z), \tag{3.2}$$

so that the system (3.1) possesses inversion symmetry. This implies that any solution of (3.1) $(x, y, z) = (x(t; \vec{r}_0), y(t; \vec{r}_0), z(t; \vec{r}_0))$ has associated with it another solution $(x, y, z) = (-x(t; -\vec{r}_0), -y(t; -\vec{r}_0), -z(t; -\vec{r}_0))$. There is, therefore, another homoclinic obtained by inverting the first, and such a pair, which interleave one another, is illustrated in Fig. 2b.

The dynamical consequences of such a pair of homoclinics depend crucially on the ratio of the real parts of the eigenvalues which we denote by δ:

$$\delta \equiv -\frac{\lambda}{\rho}. \tag{3.3}$$

Shilnikov (1965) first showed that when δ lies in the interval $(0, 1)$, such homoclinic solutions are accompanied at $R = 0$ by the existence of an countable set of periodic and an uncountable set of aperiodic solutions (i.e. chaotic structure) in the proximity of the homoclinic. Shilnikov also showed that in this case there is an infinite, countable set of "multiple pulse" homoclinic solutions in any R-interval containing the origin, each having its own sets of periodic and aperiodic solutions. Later work by Glendinning & Sparrow (1984), Gaspard, Kapral, & Nicolis (1984), Tresser (1984), and Glendinning (1984) has partially explained how this complicated state of affairs comes about as the bifurcation parameter R is varied, and it is these results that are most relevant presently.

Before describing these results, we first digress to recap a few relevant details about bifurcations and periodic orbits:

• In the presence of inversion symmetry, we note that a periodic orbit can always be classified as either asymmetric or symmetric, with asymmetric orbits always occurring in pairs in which each member transforms into the other under the inversion operation, while symmetric orbits are unique. Both types of solutions are illustrated in Fig. 3.

• Secondly, recall that there are four types of local bifurcations common to dynamical systems—the Hopf, saddle-node, period-doubling, and symmetry-breaking bifurcations. Figure 4 illustrates the bifurcation diagrams associated with these four. A brief description of these follow. (For detailed explanations, see e.g., Wiggins (1990), Guckenheimer & Holmes (1983)).

- ∘ In a (supercritical) Hopf bifurcation, the real parts of the eigenvalues of a fixed point increase through zero from negative to positive, causing the stationary point to lose stability and a stable periodic orbit to be created. We shall see that this is the initial mechanism by which periodic orbits appear in the double-gyre circulation as δ_i is increased.
- ∘ In a saddle-node bifurcation, a pair of stationary fixed points, or a pair of periodic orbits, one stable and one unstable, are created simultaneously from the same locus

of phase space as R is increased or decreased (this may be viewed as the collision and subsequent annihilation of a saddle and a stable orbit (or node) in reverse). We shall see that this is the mechanism by which multiple orbits first appear following the initial Hopf bifurcation in our model of the double-gyre circulation.

∘ In a period-doubling bifurcation, which can involve only asymmetric orbits, an asymmetric "parent" orbit loses stability and in the process a new stable (and asymmetric) orbit is created having twice the period of the original orbit. A period-doubling bifurcation is often followed by an infinite cascade of period-doubling bifurcations which culminates in the creation of a chaotic attractor. Only the first period-doubling of such cascades will be indicated in the bifurcation diagrams that follow (Figs. 5, 6, and 7). It is by this route that chaos first appears in the double-gyre circulation model.

∘ Finally, a symmetry-breaking (or pitchfork) bifurcation occurs when a symmetric orbit loses stability, creating a pair of stable asymmetric solutions. This is the symmetric analog of a period-doubling bifurcation. A symmetry-breaking bifurcation is also often followed by two infinite period-doubling cascades begining with the two asymmetric solutions created in the symmetry-breaking bifurcation. This is also the mechanism by which asymmetric steady state solutions first arise in the symmetrically forced double-gyre circulation.

We now discuss the Shilnikov Phenomenon. Figs. 5, 6, and 7 illustrate in an idealized way three basic bifurcation scenarios that have been discussed in the literature by Glendinning & Sparrow (1984), Gaspard, Kapral, & Nicolis (1984), Tresser (1984) and Glendinning (1984). These describe the so-called "principal" periodic orbits near a homoclinic to a saddle-focus in the presence of inversion symmetry and correspond to the following regimes in the parameter δ: $1/2 < \delta \ll 1$ (case I; Fig. 5), $1/2 \ll \delta < 1$ (case II; Fig. 6), and $\delta > 1$ (case III; Fig. 7). In particular, Figs. 5-7 illustrate the bifurcation histories of the principal periodic orbit families of both symmetric and asymmetric kind. In these diagrams, we indicate the *half* period of the symmetric orbits, and the *full* period of the asymmetric solutions in order to make the illustrations appear symmetric and easier to understand. We stress that not all the periodic solutions that exist in the phase space are represented—in fact, for cases I and II the vast (in fact infinite) majority are not. The principal orbits that are described, however, play a fundamental role in the creation of many if not most of the other periodic solutions that exist in the vicinity of the homoclinic.

In Fig. 5, case I, we see that when R is much less than zero that there is only a single, stable, asymmetric orbit (which may, for example, have been created in a Hopf bifurcation). As R is increased, a series of (forward) saddle-node bifurcations occurs, creating both asymmetric solutions and symmetric solutions, until the pair of homoclinics at $R = 0$ is reached. Then, as R is increased further, a sequence of (reverse) saddle-node bifurcations occur, until all that remains is a single, stable, symmetric orbit. Note that in any open interval in R that contains the origin there are an infinite number of saddle-node, period-doubling, and symmetry-breaking bifurcations. This entire process, that is, traversing from a single asymmetric orbit to a symmetric orbit, in either cases I, II, or III, is what we collectively refer to in this paper as "Shilnikov Phenomenon".

Most importantly, note that although this process involves an infinite number of forward and backward saddle-node bifurcations, one may think of these bifurcations as involving only two families of periodic orbits by continuously tracing the curves of Figs. 5, 6, and 7. These two families, the *principal* families, may be thought of as having deformed into the pair of homoclinics at $R = 0$. Of course, this interpretation requires one to accept the idea that principal orbits can have multiple representations at a given value of R.

It is thought that for values of R sufficiently close to zero, each of the period-doubling and symmetry-breaking bifurcations illustrated in Fig. 5 lies at the beginning of a complete period-doubling cascade, which terminates in the creation of a chaotic attractor. Moreover, as we have mentioned previously, it is known that there is a countably infinite set of multiple-pulse homoclinics in any R-interval containing the origin, and each of these is thought to possess all of the same associated structure we have so far described for the "primary" homoclinic. Thus, there exist an enormous number of solutions and bifurcations occurring that are not shown in Fig. 5. Obviously, there is the possibility of great variability in the temporal dynamics, including chaos, for case I (and case II).

The question of what *attractors* one is likely to see in simulations, e.g. chaotic attractors and/or stable periodic orbits, depends a great deal on the specific example studied, and to the best of our knowledge there are few general statements which can be made. Likewise, the exact bifurcation histories of individual solutions created in the various period-doubling cascades is also not known generally, although some specific case histories for some orbits have been described in the literature (see Glendinning & Sparrow (1984)).

Figure 7 illustrates the bifurcation diagram for case III SP that is, $\delta > 1$. (Case II is addressed below.) In this case, the bifurcation can be characterized as a "gluing bifurcation", that is, a pair of stable asymmetric solutions simply deform smoothly into a pair of homoclinic solutions at $R = 0$, and then a single, stable symmetric orbit emerges as R is further increased (thus the two asymmetric orbits appear to have been "glued" together). Note that this entire sequence involves only three periodic orbits and one pair of homoclinic orbits, and is thus vastly simpler than case I or II SP.

A possible bifurcation diagram for case II SP ($1/2 \ll \delta <\approx 1$) is illustrated in Fig. 6. This diagram is based on Fig. 3.8 of Glendinning & Sparrow (1984) (except that we have added the symmetric orbit curve resulting from the inversion symmetry), and it is this sort of scenario that we believe is occurring in our reduced-gravity QG model of double-gyre circulation. We emphasize, however, that less is known about case II than for either case I or III SP, and significant variations Note that the diagram resembles the diagram for case I very close to the homoclinic, that is, there is an infinite series of saddle-node bifurcations alternating on either side of the homoclinic as the period of the principal orbits approach infinity. Further away from the homoclinic are pairs of outlying saddle-node bifurcations (sometimes called "S-shaped bifurcations" in the literature), which can form if some of the lower period alternating saddle-node bifurcations are shifted to one side and away from the homoclinic as δ increases.

We remark that as δ is increased through one, i.e., into case III SP, a finite number of these S-shaped bifurcations may persist, but the infinite set of alternating saddle-nodes approaching the homoclinic orbits must shrink and disappear. In general, as Glendinning & Sparrow (1984) point out, the bifurcation diagram for any system with variable δ must continuously deform as δ is increased from the regime of case I to case III, so that the bifurcation diagram for case II will be an interpolation between the other two cases. on Fig. 6 are possible. This concludes our brief sketch of the Shilnikov Phenomenon.

4. The use of spectra and phase projections for study of bifurcation

In the next section we present evidence for SP in the double-gyre circulation using two types of dynamical data—collections of power spectra plotted together, which provide us with a global picture of the dynamics as a function of δ_i, and projections of individual dynamical trajectories in phase space generated by "time-delay embedding" of time series. In this section we explain why these data types are useful and complementary, and how they are obtained from the model

The first data type—collections of power spectra— consist of a series of power spectra calculated from time series obtained from the model over an interval of inertial scales δ_i. (Recall that the inertial scale is the natural bifurcation parameter for this problem.) The particular quantity measured to produce these series is discussed at the end of this section. After the spectrum of each series is computed, the spectra were plotted together as a two-dimensional color-mapped plot for the entire collection with the inertial scale along the horizontal axis and spectral component on the vertical axis (e.g. see Fig. 8a and b). Note that, contrary to conventional practice, the spectra in Figs. 8a and 8b were plotted with temporal *period* on the vertical axis, instead of frequency. This was done so that the resulting plots can be directly compared with the SP bifurcation diagrams of Figs. 5, 6, and 7, which are also plotted with temporal period on the vertical axis.

The advantage of this approach is that it essentially furnishes us, at least in this particular case, with a bifurcation diagram for the dynamics of the system without having to perform the difficult task of isolating and tracking individual periodic solutions. The latter task is often undertaken for systems of coupled ordinary differential equations, but is extremely difficult to carry out for partial differential equations and entails the use of specialized and dedicated numerics.

Generally, power spectra of chaotic time series associated with chaotic evolution exhibit a *broadband* structure, i.e. energy at infinitely many frequencies, making it hard to discern the frequency characteristics of individual orbits. In the case studied in this paper, however, the chaotic attractor of the double-gyre circulation appears to be tightly compressed in phase space around what we suspect are the principal orbits involved in a SP bifurcation, causing the spectra to have a strong component at the principal orbit's frequencies that stand out against the broadband background. The reason for this compression appears to be because the parameter δ (see (3.3)) is close to one. This issue is further explored when the attractor projections are examined in section 5.

The second type of data we consider in this paper are three-dimensional projections of individual phase trajectories (see Fig. 9). These provide us with direct information about attractor structure, and are obtained by the widely employed technique of time-delay embedding of time series. That is, given a time series $q(t)$, where $q(t)$ is some time series derived from the system, we choose a delay time τ and plot the triplet $(q(t), q(t-\tau), q(t-2\tau))$ in three-dimensional phase space. If the dynamics that produced the time series in the first place are also three-dimensional, then this approach generally produces topologically accurate and very useful projections of the dynamics. For detailed discussions of time-delay embeddings, see for example, Sauer, Yorke, and Casdagli (1991).

We now take up the question of how time series were obtained. At very least, the time series should be choosen such that they contain information about the symmetry of the flow because the SP bifurcation diagrams with which we are to compare the model dynamics display a change in orbital symmetry as a principal characteristic. After some experimentation, we found that summing the stream function $\psi_i(t)$ over two areas located symmetrically about $y = 0$ (the zero wind stress curl line) seems to provide useful time series for our purposes. Thus we define

$$\alpha(t) = \sum_i \psi(x_i, y_i)(t), \quad 0 < x_i, |y_i| \le \frac{L}{4}. \tag{4.1}$$

This quantity clearly contains information about the symmetry of the flow—if a flow is symmetric about the y axis and consists of two counter-rotating mirror image gyres, then this quantity is identically zero. If a flow is asymmetric at any given time t, then $\alpha(t)$ will (typically) be nonzero, and its sign will invert under the transformation $y \to -y$.

We remark that for the full range of parameters that we consider in this article, the maximal zonal penetration of the separated midlatitude jet is about a quarter of the horizontal extent of the domain and thus the measure $\alpha(t)$ is biased to characterize behavior in the vicinity of this midlatitude jet. The necessity for this bias is related to some subtle aspects of mass conservation and the effective dynamical dimension of the double-gyre circulation. We defer discussion of this issue to section 6. Finally we remark that all of the embeddings and the power spectra presented in this paper were generated from reasonably long samples (typically 8 gyre-turnaround times) of $\alpha(t)$ measured after the flow was allowed to relax (16 gyre-turnaround times) to an attractor. All statements in the following are thus to be taken as pertaining to the *time-asymptotic* dynamics of the double-gyre circulation.

5. Evidence for Shilnikov Bifurcations in the Double-Gyre Circulation

We begin by examining plots generated from 61 power spectra (generated from time series $\alpha(t)$ as described in the foregoing section) for inertial scales extending from $\delta_i = 1.25$ to 1.55 in increments of .005. These spectra are plotted together in Fig. 8a and b, and, as we now describe, provide a global picture of the bifurcations occurring as the inertial scale is varied. Note that the spectra in Fig. 8a are plotted with the base 10 logarithm of the period on the vertical axis, so as to compress the range in period and to make the structure at smaller periods more visible. In addition, the spectral energy at zero frequency (infinite period) is also included in the topmost data point at the top of Fig. 8a, because, as we will describe later, this component contains useful information about the symmetry of the flow. Figure 8b, on the other hand, is plotted with a linear period scale, and is included for ease in identifying period-doubling relationships.

We first draw attention to the yellow curves beginning in the lower left corners of Fig. 8a and b, beginning at an inertial scale of about $\delta_i = 1.28$. (The power spectra are color-coded such that blue-black corresponds to the lowest energy levels and yellow corresponds to the highest energy levels). These are labeled "fundamental (asymmetric branch)" and "harmonics" in the figure. At about $\delta_i = 1.374$ the time series become chaotic and the spectra become broadband. Interestingly, the yellow curves appear to persist despite the broadband structure. This apparently occurs, because, as we will illustrate and discuss later, the chaotic attractor in this regime is tightly compressed in phase space, i.e. it has the form of a slightly thickened periodic orbit, and so its power spectrum retains a clear characteristic frequency.

Notice that these curves appear to diverge in period as they approach an inertial scale value just slightly less than $\delta_i = 1.395$ (the vertical locus corresponding to this value is optimistically labeled as the "homoclinic point" in the figure). We conjecture that this divergence is associated with the formation of a pair of homoclinic orbits slightly prior to $\delta_i = 1.395$. Assuming that this is the case, we will denote the exact (but unknown) value of δ_i as δ_i^{hom}. Notice also that a similar set of curves exists in $\delta_i > \delta_i^{hom}$, which are visibly more diffuse in appearance and harder to discern (the uppermost of these is labeled "fundamental (symmetric branch)". These curves appear to diverge in period as $\delta_i \to \delta_i^{hom}$ from above.

We now conjecture that the uppermost curves correspond to the fundamental period of families of exact periodic orbits of the double-gyre circulation. (The curves below the fundamental, e.g., those labeled "harmonics" in $Rh < \delta_i^{hom}$ are the super-harmonics of the fundamental, and they lie below the fundamental because we have plotted period instead of frequency on the vertical axis.) This will be directly confirmed for $\delta_i < 1.374$ (and also at $\delta_i = 1.405$), where the flow is not chaotic, by examining phase projections. We will also show that all the attractors in the region $\delta_i < \delta_i^{hom}$) are asymmetric (implying that there are really pairs of asymmetric attractors), whereas in the region $Rh > \delta_i^{hom}$ the attractors are symmetric or very nearly so.

All of the phase projections we shall discuss in this paper, and generated in the way discussed in the foregoing section, are collected together in Fig. 9. All of these projections were generated from the same time series used to generate Fig. 8 so that we may establish direct correlations between the figures: note that the letters along the bottom of Fig 8 and the letters indexing each projection in Fig. 9 correspond. The delay-time τ used to generate these projections was equal to 1/128 in Sverdrup gyre-turnaround time units and was chosen on the basis of mutual redundancy analysis and false nearest neighbor tests—it was also found simply to give the best visual images of the dynamics.

Projections a-f in Fig. 9 clearly display closed trajectories and therefore directly confirm the conjecture that the uppermost curve in Fig. 8 corresponds to the fundamental period of families of periodic orbits for $\delta_i < 1.374$. Next, projections a-h display *asymmetric* attractors, with both left-hand and right-hand versions visible. For example, projections c and d are almost mirror images—not exact mirror images since they occur at slightly different values of δ_i. On the other hand, projections i-l appear to be symmetric in structure.

That *all* the attractors we have sampled in $\delta_i < 1.395$ (including those not included in Fig. 9), chaotic or otherwise, are asymmetric can be verified by examining the zero-frequency component of the power spectra. Recall that this component is included along the top edge of Fig. 8a. If this component is nonzero, then the orbit must be asymmetric, i.e. the time series that generated it is not symmetric about the time axis. It can be seen in Fig. 8a that there appears to be significant and nearly constant zero-frequency energy everywhere in $\delta_i < \delta_i^{hom}$. At $\delta_i = \delta_i^{hom}$ (around 1.395), the zero-frequency component becomes suddenly darker. Fig. 10 plots this component by itself, showing that it drops suddenly more than three orders of magnitude near $\delta_i = \delta_i^{hom}$, indicating a sudden dramatic change in orbital symmetry. Beyond this point, the component begins to increase slowly and irregularly as the inertial scale is increased.

To make the physical interpretations of these dynamical structures clear, figures 11, 12, and 13, demonstrate the connection between orbital symmetry, as seen in the phase-space embeddings, and the double-gyre circulation. In these figures, the streamfunction for the full domain is color coded so that blue-black corresponds to the minimum and yellow to the maximum values. In the approximation of the reduced-gravity QG model, the difference between this upper layer streamfunction and the streamfunction of the bottom layer (which presently is identically zero) is proportional to temperature at the base of the upper layer and therefore blue-black in these figures corresponds to lower temperatures (cold) and red-yellow corresponds to higher temperatures (warm).

When the phase-space attractors are asymmetric, the circulation patterns in physical space are such that the separated and eastward flowing midlatitude jet oscillates asymmetrically about the zero wind stress curl line ($y = 0$). Figures 11 and 12, which correspond to the asymmetric projections g and h in Fig 9, respectively, demonstrate this. In Fig. 11, cold-core rings are seen to dominate, while in Fig. 12 warm-core rings dominate. Further visual inspection of the flow patterns seems to suggest that the shedding of eddies only occurs on the dominant side. On the other hand, when the attractors are symmetric (or nearly so), the oscillation of the midlatitude jet is symmetric about $y = 0$ (or nearly so), so that when eddy shedding occurs, cold-core and warm-core rings are equally preferred (or nearly so). This is demonstrated by Fig. 13, which corresponds to the symmetric projection l of Fig. 9.

Finally, we note that the fast oscillations, e.g., seen very clearly in projections e–h of Fig. 9 correspond to mesoscale variability. For the representative set of physical parameters mentioned in section 2, this mesoscale variability has a time scale of about 110 days whereas the overall oscillation occurs on a time scale that varies between roughly 1 and 8 years. Furthermore, it is clear from Fig. 9 (and Fig. 8) that the amplitude of this mesoscale variability is quite small compared to the amplitude of the overall (interannual) oscillation. Now, we get back to presenting the evidence for SP.

The Hopf bifurcation which created the asymmetric branch in $\delta_i < \delta_i^{hom}$ is indicated in Fig. 8 (see the key included in the figure). This is the point at which an asymmetric steady-state became unstable, leading to the onset of periodic oscillations in the gyres. We confirmed that a Hopf bifurcation, and not a bifurcation such as a saddle-node, was responsible for the birth of the asymmetric periodic oscillation by examining many phase projections near the bifurcation, and this is also evidenced in part by Fig. 8. In this figure, note that the super-harmonics of the newly created orbit become visible only somewhat beyond the initial appearance of the fundamental period indicating that the orbit was initially almost purely fundamental (sinusoidal). This is what one expects for a Hopf bifurcation, and not, for example, in a saddle-node bifurcation. In Fig. 9 it can also be seen that the orbits with the lowest inertial scale are the simplest in structure, reflecting their initial elliptical structure.

Having conjectured, and in some cases demonstrated that the yellow curves of Fig. 8 correspond to families of periodic orbits, we now point out that the overall features of Fig. 8 in fact conforms very closely to the features of case II SP. First of all, the observed trend that both sets of curves corresponding to periodic orbits appear to diverge as they approach $\delta_i = \delta_i^{hom}$, and that this event accompanies a dramatic change in attractor symmetry, strongly hints at the presence of a pair of homoclinic orbits near $\delta_i = 1.395$. The existence of these is of course a crucial prerequisite for SP. The existence of the homoclinic is also suggested by the sequence of projections in Fig. 9: Projections a-h show that as δ_i increases these trajectories pass nearer to and spend increasingly longer periods of time in the vicinity of the origin in these plots (The homoclinic fixed point in this problem must be located at the origin because this point is associated with a symmetric steady-state on which the measure $\alpha(t)$ is identically zero.) At projection i, in particular, note that the (now symmetrical) chaotic attractor is very tightly compressed in phase space, and actually resembles a projection of a pair of homoclinic orbits.

These projections also exhibit dynamical behavior very consistent with the existence of a saddle-focus type fixed point at the steady-state—another crucial feature of SP. To see this, first note that as δ_i is increased to 1.395 the trajectories show an increasing tendency to exit the vicinity of the origin along a specific axis which is inclined at about 70 degrees to the horizontal in the plots (this is particularly evident in projections g-l). This behavior suggests that the fixed point at the origin possesses a real, positive eigenvalue (the "saddle" part). Secondly, while exiting the origin, the trajectories also show a tendency, which increases with δ_i, to oscillate around this inclined axis. When $\delta_i \approx 1.395$, these oscillations are found to decay while the trajectory is exiting. This is clearly exhibited in the time series at $\delta_i = 1.395$ which is plotted in Fig. 17, corresponding to projection i. This behavior suggests that the steady-state also possesses a complex pair of eigenvalues with negative real part (the "focus" part).

A third cardinal characteristic of SP in the presence of an inversion symmetry exhibited in Fig. 8 is that the symmetric principal orbit branch on one side of the homoclinic has twice the period of the (asymmetric) principal orbit branch on the other side. This can be seen to be the case in Fig. 8b, if points on the uppermost yellow curves on opposite sides of but near to and equidistant to the suspected homoclinic point are compared.

As noted previously, in projection i the chaotic attractor appears to be very compressed such that it closely resembles a pair of homoclinic orbits, and this is apparently the reason why there exist clearly discernable yellow curves in Fig. 8 even where the power spectrum is broadband. We suspect that this compression occurs because, as we conjecture later on other grounds, the parameter δ is quite close to one. To understand why we might expect compression of the attractor in this case, recall that, as discussed in section 3, as δ approaches one the infinite alternating sequence of saddle-node bifurcations approaching the homoclinic must disappear. This implies that when δ is close to one, the onset of chaos associated with these bifurcations must be closely confined in parameter space to the vicinity of the homoclinic.

We remark that although this compression phenomenon has not be definitively studied for SP to the best of our knowledge, a similar phenomenon is known to exist in the real eigenvalue analogue of SP, the "homoclinic explosion" route to chaos. For example, similarly highly compressed chaotic attractors associated with homoclinic orbits were found by Luce (1995) (see Fig. 6 of that paper for example), and shown to be compressed specifically because the analogue of δ was close to one. If the system of equations (21) from Luce (1995) are integrated, with $\mu = -4$, $\nu = 5.5$, $k = 4\pi^2$, and with R ranging from 43.96 to 44.02, and the power spectra are plotted together, one obtains the plot shown in Fig. 14. This plot is quite similar in overall form to Fig. 8.

We now discuss some evidence for certain other local bifurcations in Fig. 8. First, note once again the discontinuities that the yellow curves corresponding to the periodic orbit families in $\delta_i < \delta_i^{hom}$ exhibit—that is, points where there appear to be discontinuous jumps in period (the discontinuities close to the suspected homoclinic are actually most clearly seen in the harmonics of the asymmetric families). These jumps, as we demonstrate later with a known example, are exactly what we can expect to see in the spectral plots if pairs of saddle-node bifurcations ("S-shaped bifurcations") are present. We have thus labeled some of these discontinuities as saddle-node bifurcation pairs in Fig. 8—six are clearly visible in Fig. 8a by looking at the first harmonic of the asymmetric branch, while four more appear to be barely distinguishable in the second harmonic (just into the chaotic regime). While there are more of these bifurcation events as δ_i approaches the homoclinic point, they are not visible on the scale of the plots in Fig. 8.

Next, as described in section 3, in case I and II SP we generally expect period-doubling bifurcations to occur in between the pairs of saddle-node bifurcations and symmetry-breaking bifurcations, especially when close to the homoclinic point. We can verify that some period-doubling bifurcations do in fact occur in the double-gyre sequence. For example, in the power spectra of Fig. 8, isolated "islands" of spectral power at higher periods can be seen, visible as blotches of lighter color at periods larger than the fundamental (these are labeled as "fundamentals from period-doubling" in the figure). Fig. 8b has a linear period scale, and can be used to quickly verify that these islands occur at multiples of two, four, and eight times the fundamental. Additional islands can also be seen in the harmonics, as should be the case. It is verified directly from the phase projections that these are due to period-doubling of the principal orbit. Projection f in Fig. 9 is an example where the period-doubling can be seen explicitly, and corresponds to the most visible period-doubling island in Fig. 8. As another example, the trajectory in Projection e is also seen to be slightly period-doubled (although this is hard to detect visually), and is associated to a correspondingly weaker looking island. More detailed considerations suggest that the period-doubling is the mechanism for the onset of chaos around $\delta_i = 1.374$.

To validate our interpretation of the spectral features in terms of saddle-node bifurcations and period-doublings, we briefly consider an example system in which the existence of SP has been established. To this end, we consider system (18) of Glendinning & Sparrow (1984):

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= z \\ \dot{z} &= -z - by + cx - x^2, \end{aligned} \tag{5.1}$$

with $b = 0.8$. This system displays both case III SP, and also S-shaped and period-doubling bifurcations. While the reader is referred to Glendinning & Sparrow (1984) for details of the dynamics, Fig. 15 compares one of that paper's figures to our method of spectral analysis of this system for $1.69 \leq c \leq 1.71$. Fig. 15a is an approximate reproduction of Fig. 4.7 (ii) of Glendinning & Sparrow (1984), whereas Fig. 15b plots our type of spectral graphic. The color-coded spectra are shown with the bifurcation parameter c on the x-axis and the period (linear scale) on the y-axis. In this comparison, the correspondence between the jump in the dominant period at c between 1.694 and 1.695 in Fig. 15b and the saddle-node at about the same value of c in Fig. 15a is clear. (On zooming in, a second jump is also clear in our spectral analysis at a value of c close to 1.706.) The island indicating a period-doubled orbit in Fig. 15b beginning at a c value of about 1.699 is not seen in Fig. 15a, because Glendinning & Sparrow (1984) obtained their figure using continuation techniques which ignored this orbital branch.

Finally, guided by our knowledge that a series of saddle-node and symmetry-breaking bifurcations occur on the symmetric branches for case I and II SP, we have (very optimistically) labeled as symmetry-breaking bifurcation pairs what appears to be a systematic pattern of irregularities in the fundamental of the symmetric branch in Fig. 8. Specifically, we suspect that these irregularities result from other bifurcations that directly change the chaotic attractor structure, such as period-doubling cascades, each of which follows a symmetry-breaking bifurcation involving the principal periodic orbit. Our conjecture here is supported somewhat by the irregular behavior of the zero-frequency component beyond the homoclinic point in Fig. 10, because such irregular behavior would be consistent with the occurence of a series of local symmetry-breaking events which contiguously and only slightly break the symmetry of orbits (unlike a global homoclinic event which produces a discontinuous change in symmetry). Moreover, the overall increase in this irregularity as we leave the vicinity of the homoclinic is also what we expect, because in general we expect bifurcations on the principal orbit to produce large changes in period further from the homoclinic.

Summarizing the evidence presented so far for the double-gyre circulation, we see 1) that there is evidence suggestive of a homoclinic bifurcation involving a saddle-focus, 2) that a change from asymmetric to symmetric attractors with a corresponding doubling of period occurs at the suspected homoclinic, and 3) that there appears to exist saddle-node, period-doubling, and possibly symmetry-breaking bifurcations arranged in a way which is consistent with SP. Based on this we propose that Shilnikov Phenomenon is occurring in the double-gyre circulation.

We now take up the more specific question of what kind of SP—case I, II, or III, is presently occuring in the double-gyre circulation? We notice from Fig. 8 that the two branches of period orbits do not noticeably appear to cross back and forth across the suspected homoclinic point at $\delta_i = \delta_i^{hom}$. This seems to rule out case I SP (see Fig. 5). Likewise, the ubiquitous chaos for inertial scales above $\delta_i = 1.374$ appears to rule out case III. We thus suspect case II SP.

To strengthen the evidence for case II SP, we attempted to measure the value of parameter δ for the double-gyre circulation near the suspected homoclinic point. Recall that equation (3.3) defined δ to be the ratio of the attracting rate to the repelling rate of the saddle-focus, and that for case II SP δ lies in the range $1/2 \ll \delta <\approx 1$. We therefore attempted to compute δ by measuring the exponential rate of decay of oscillations and the overall exponential growth during periods of time that a (near-homoclinic) phase trajectory was in the vicinity of the origin and where the evolution is dominated by the linear eigenvalues of the fixed point.

To carry out this measurement, we examined the time series, shown in Fig. 16, corresponding to projection i in Fig. 9 (δ_i=1.395) and two other time series very close to the suspected homoclinic (δ_i=1.393, and 1.394). We first note that each of these chaotic series is seen to repeatedly exhibit episodes of decaying oscillations accompanied by overall exponential growth when the trajectory "dwells" close to the steady-state solution. The ratio of the decay rate of these oscillations to the overall exponential growth rate is therefore the quantity of interest to us for measuring δ.

Fig. 17 shows the result of fitting a nonlinear function containing exponentially decaying oscillations and exponential growth to the longest of these dwelling episodes, where the trajectory is thought to be making its closest approach to the fixed point. One segment is drawn from the series in Fig. 16b and the other from Fig. 16c. The function fit was

$$\alpha_{\text{fit}}(t) = \beta e^{\rho t} + \gamma e^{-\lambda t} cos(\omega t + \phi) + \epsilon cos(\omega t + \phi'), \tag{5.2}$$

which describes a linear combination of exponential growth (the saddle part), decaying Rossby waves (the focus), and also a non-decaying background Rossby wave field. The last of these was found to be necessary for obtaining a robust fit if the integral of the stream function is conserved, an issue which is taken up in the next section. The values for δ obtained for the segments shown in Fig. 17a and b were $\delta = \frac{\lambda}{\rho} \approx .85$, and .91 respectively. These values clearly favor case II SP. If the integral of the streamfunction is not conserved (see the next section for further discussion of issues related to conserving or not conserving this integral), the value of δ obtained close to the homoclinic was also approximately .9.

Finally, for the set of values chosen for the physical parameters in section 2, the time scale corresponding to ω obtained in the above nonlinear fit is about 112 days. In this context we note the following (1) Using the Rossby wave dispersion relation, the above frequency, ω, corresponds to a wavelength of about 250 km, and the existence of oscillations on this scale was independently confirmed by examining movies of the flow. This aspect of the nonlinear fit, therefore, validates the assertion that the oscillatory behavior in the vicinity of the saddle-focus is associated with Rossby waves. (2) That the oscillatory time scale obtained in the above nonlinear fit also corresponds to the characteristic time scale of mesoscale variability (e.g., seen very clearly in projections f, g, and h of Fig. 9) points to the fact that the origin of this Rossby wave field is directly related to mesoscale variability. (3) While the amplitude of the mesoscale variability is always small compared to the lower-frequency (interannual) oscillation, the former is further suppressed close to δ_i^{hom}, at values of forcing close to the suspected homoclinic. Referring to Fig. 2 this seems consistent with the manner in which the homoclinic orbit(s) exit the vicinity of the fixed point.

6. Dimensionality of the dynamics and conservation of ψ.

When mass is exactly conserved in the reduced-gravity QG model, the effects of Kelvin waves are parameterized in a system that does not explicitly allow for such waves (e.g., see Milliff & McWilliams (1994)). If for convenience, we rewrite (2.4) as

$$\left(\nabla^2 - \frac{L^2}{L_R^2}\right)\frac{\partial \psi}{\partial t} = A,$$

then, the streamfunction $\psi(t)$ can be decomposed as

$$\psi(x,y,t) = \psi_I(x,y,t) + c(t)\psi_B(t),$$

with

$$\left(\nabla^2 - \frac{L^2}{L_R^2}\right)\frac{\partial \psi_I}{\partial t} = A, \quad \psi_{IB} = 0,$$

$$\left(\nabla^2 - \frac{L^2}{L_R^2}\right)\frac{\partial \psi_B}{\partial t} = 0, \quad \psi_{BB} = 1,$$

and,

$$c(t) = \frac{\int \psi_I dxdy}{\int \psi_B dxdy}.$$

Such an implementation results in a propagation of the effects of Kelvin waves around the basin with infinite speed.

In a separate control experiment, wherein the Kelvin wave parameterization (KWP) is neglected, we find that the signature of the westward-propagating Rossby wave field is related almost entirely to the relaxation phase of the oscillation of the midlatitude jet after it has meandered away fartherest from its antisymmetric state. On the other hand, in the simulations that we have presented in this article, a KWP was used, and in these simulations, we find that in addition to the Rossby wave field associated with the relaxation phase of the midlatitude jet, the parameterized-Kelvin wave gives rise to an additional background field of Rossby waves. These Rossby waves originate at the eastern boundary of the domain and propagate westward in a north-south symmetric fashion as, e.g., in the study of Milliff & McWilliams (1994).

Thus, we suspect that while in the case without KWP, energy wholly remains in the gyre modes, with KWP, a part of this energy is diverted into the Kelvin-Rossby wave production mechanism. Indeed, we find that when KWP is neglected it is not necessary to include the term $\epsilon cos(\omega t + \phi)$, which accounts for the background Rossby wave field, in the nonlinear fit (5.2) to obtain a robust fit, and that the values obtained for δ are found to be considerably less dependent on the way $\alpha(t)$ is obtained, i.e. the areas over which $\alpha(t)$ is obtained can be varied greatly with little effect on the values of δ. With KWP, consistent measurements for δ necessitate that the patches be significantly localized over the aforementioned region. In contrast to this, the global bifurcation picture obtained via power spectra, i.e. Fig. 8, is found to change little in form with or without KWP and regardless of patch size.

These facts lead us to suspect that (1) with KWP, the *dynamical* dimension of the entire system is increased relative to the non-KWP case due to the Kelvin-Rossby wave mechanism, and (2) the dynamics of the gyres themselves are still well characterized as exhibiting the SP bifurcation both with and without KWP. The latter fact accounts for the invariance of the global bifurcation picture, while the former explains the need to localize the measurement for $\alpha(t)$ with KWP. We are presently examining this issue in greater depth.

7. Conclusion

It is now well recognized that simple models of wind-driven circulation in midlatitude oceans can exhibit a rich variety of variability including those on the mesoscale and interannual time scales. Most previous studies, with the exception of Meacham (2000), have considered steady states and periodic solutions and their local bifurcations and have suggested the existence of a multiplicity of such simple solutions or a sequence of Hopf bifurcations (leading to a strange attractor) as the primary cause of such variability. Complementary to these studies, we show that other solution types like, e.g., homoclinic orbits (formed by a collision of periodic solutions and steady states in phase space), and their associated global bifurcations, can play a central role in organizing the variability of the wind-driven circulation.

In particular, by considering the double-gyre circulation as represented by a reduced-gravity QG approximation, we have shown that its dynamics exhibits a structure resembling a bifurcation scenario called Shilnikov Phenomenon. This interpretation of the dynamics provides an underlying coherence to the double-gyre circulation despite the fact that its dynamics are seen to change drastically at some parameters with a small change in the forcing. Thus, while one may have concluded that the model is structurally unstable in the sense of variability as a function of parameters, it can be better understood from a global perspective. We remark that it would be very difficult to extract these details of the dynamical structure by direct observation of the flow patterns. This also demonstrates the utility of the dynamical systems tools and concepts used here for the analysis of ocean dynamics.

The primary implication of this study to the double-gyre circulation is that if global bifurcations and dynamical chaos occur in an approximation like the reduced-gravity QG model, they are likely to be of importance in more realistic models, and that, therefore, homoclinic bifurcation mechanisms like SP may explain in part the variability of wind-driven circulation. In the future we hope to give a more comprehensive overview of global bifurcations of the double-gyre circulation and resolve some of the questions raised in this paper about dynamical dimension.

BTN was supported in part by the US Department of Energy's Climate Change Prediction Program (CCPP). Computer resources for this work were made available by IGPP and CNLS at the Los Alamos National Lab.

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow, Part 1. *J. Comput. Phys.*, **1**, 119–143.

Berloff P.S. and McWilliams J.C., 1999: Large-scale, low-frequency variability in wind-driven ocean gyres. *J. Phys. Oceanogr.*, **29**, 1925–1949.

Berloff P.S. and Meacham S.P., 1997: The dynamics of an equivalent-barotropic model of the wind-driven circulation. *J. Mar. Res.*, **55**, 407–451.

Berloff P.S. and Meacham S.P., 1998: The dynamics of a simple baroclinic model of the wind-driven circulation. *J. Phys. Oceanogr.*, **28**, 361–388.

Bryan K., 1963: A numerical investigation of a nonlinear model of a wind-driven ocean. *J. Atmos. Sci.*, **20**, 594–606.

Cessi, P., and G.R. Ierley, 1995: Symmetry-breaking multiple equilibria in quasi-geostrophic wind-driven flows. *J. Phys. Oceanogr.*, **25**, 1196–1205.

Constantin, P., Foias C, Temam, R., 1985: Attractors representing turbulent flows. *Mem. Am. Math. Soc.*, **53**, 1–67.

Dijkstra HA, Katsman CA, 1997: Temporal variability of the wind-driven quasi-geostrophic double gyre ocean circulation : basic bifurcation diagrams. *Geophys. Astrophys. Fluid Dyns.*, **85**, 195-232.

Gaspard P., Kapral R., and Nicolis G., 1984: Bifurcation phenomena near homoclinic systems: A two-parameter analysis. *J. Stat. Phys.*, **35**, 697–727.

Ghil M. and Childress S., 1987: Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics. Springer-Verlag, 485pp..

Ghil M., Feliks Y., and Sushama L.U., 2000: Baroclinic and barotropic aspects of the wind-driven ocean circulation. preprint.

Gill A.E., 1982: Atmosphere-Ocean Dynamics. Academic Press, 666pp..

Glendinning P. and Sparrow C., 1984: Local and global behavior near homoclinic orbits. *J. Stat. Phys.*, **35**, 645–696.

Glendinning P., 1984: Bifurcations near homoclinic orbits with symmetry. *Phys. Lett.*, **103A**, 163–166.

Greatbatch R.J. and Nadiga B.T., *Four gyre circulation in a barotropic model with double gyre wind forcing*, J. Phys. Ocean., (to appear)

Guckenheimer, J., and Holmes, P., 1983: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer-Verlag, 453pp..

Holland W.R. & Haidvogel D.B., 1981: On the vacillation of an unstable baroclinic wave field in an eddy resolving model of the oceanic general circulation. *J. Phys. Oceanogr.*, **11**, 557–568.

Ierley G.R. and Sheremet V.A., 1995: Multiple solutions and advection-dominated flows in the wind-driven circ.. *J. Mar. Res.*, **53**, 703–737.

Jiang, S., F.-F. Jin and M. Ghil, 1995: Multiple-equilibria, periodic and aperiodic solutions in a wind-driven, double-gyre shallow-water model. *J. Phys. Oceanogr.*, **25**, 764–786.

Le Provost C. and Verron J., 1987: Wind-driven ocean circulation transition to barotropic instability. *Dyns. of Atmos. and Oceans*, **11**, 175–201.

Luce B.P., 1995: Homoclinic explosions in the complex Ginzburg-Landau equation. *Physica D*, **84**, 553–581.

Marshall J.C., 1984: Eddy-mean-flow interaction in a barotropic ocean model. *Quarterly J. Meteo. Soc*, **110**, 573–590.

McCalpin, J.D., and D. B. Haidvogel, 1996: Phenomenology of the low-frequency variability in a reduced-gravity, quasi-geostrophic double-gyre model. *J. Phys. Oceanogr.*, **26**, 739–752.

Meacham S.P. and Berloff P.S., 1997: Barotropic, wind-driven circulation in a small basin. *J. Mar. Res.*, **55**, 523–563.

Meacham S.P., 2000: Low-frequency variability in the wind-driven circulation. *J. Phys. Oceanogr.*, **30**, 269–293.

Milliff R.F., and McWilliams J.C., 1994: The evolution of boundary pressure in ocean basins. *J. Phys. Oceanogr.*, **24**, 1317–1338.

Moro B., 1988: On the nonlinear Munk model. 1. Steady Flow. *Dyns. of Atmos. and Oceans*, **12**, 259–287.

Nadiga, B.T., and D.D. Holm, 1997: Low-frequency variability in the double-gyre circulation. *Fourth SIAM conference on applications of dynamical systems, May 18-22*, **Abstracts**, 20.

Pedlosky J., 1996: Ocean Circulation Theory. Springer, 453pp..

W.H. Press, B.P. Flannery, S.A.Teukolsky, and W.T. Vetterling, 1992: Integration of ordinary differential equations. *Numerical Recipes in Fortran 77*, Cambridge University Press, 708–716.

Primeau FW, 1998: Multiple equilibria of a double-gyre ocean model with super-slip boundary-conditions. *J. Phys. Oceanogr.*, **28**, 2130–2147.

Salmon, R., and L.D. Talley, 1989: Generalization of Arakawa's Jacobian. *J. Comput. Phys*, **83**, 247–259.

Sauer T. ,Yorke J.A. , and Casdagli M., 1991: Embedology. *Journal of Statistical Physics*, **65**, 579–616.

Scott, R.B., and D.N. Straub, 1998: Small viscosity behaviour of a homogeneous, quasi-geostrophic, ocean circulation model. *J. Mar. Res.*, **56**, 1225–1284.

Shilnikov L.P., 1965: A case of the existence of a denumerable set of periodic motions. *Sov. Math. Dokl*, **6**, 163–166.

Spall M.A., 1996: Dynamics of the Gulf Stream/deep western boundary current crossover. Part II: Low-frequency internal oscillations. *J. Phys. Oceanogr.*, **26**, 2169–2182.

Speich S., Dijkstra H., & Ghil G., 1995: Successive bifurcations in a shallow water model applied to the wind-driven ocean circulation. *Nonlin. Proc. Geophys.*, **2**, 241–268.

Tresser C., 1984: About some theorems by L.P. Shilnikov. *Ann. Inst. Henri Poincaré*, **40**, 441–461.

Veronis G., 1966: Wind-driven ocean circulation—Part II: Numerical solution of the nonlinear problem. *Deep Sea Res.*, **13**, 31–55.

Wiggins S., 1990: Introduction to applied nonlinear dynamical systems and chaos. Springer-Verlag, 573–591.

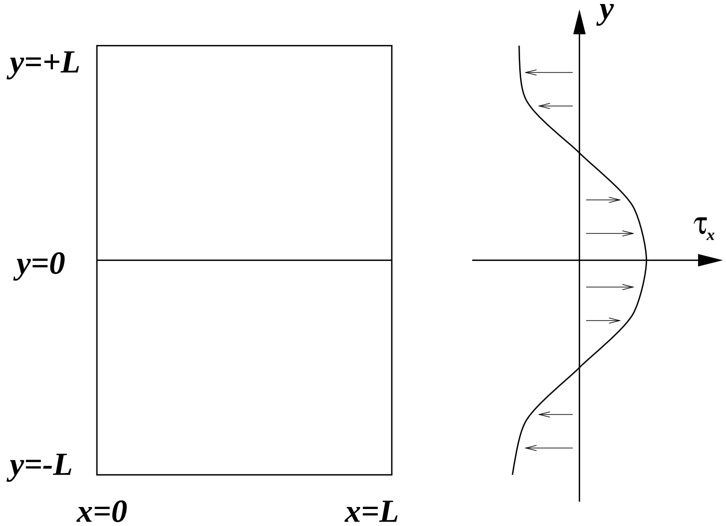

FIGURE 1. Layout of the model.

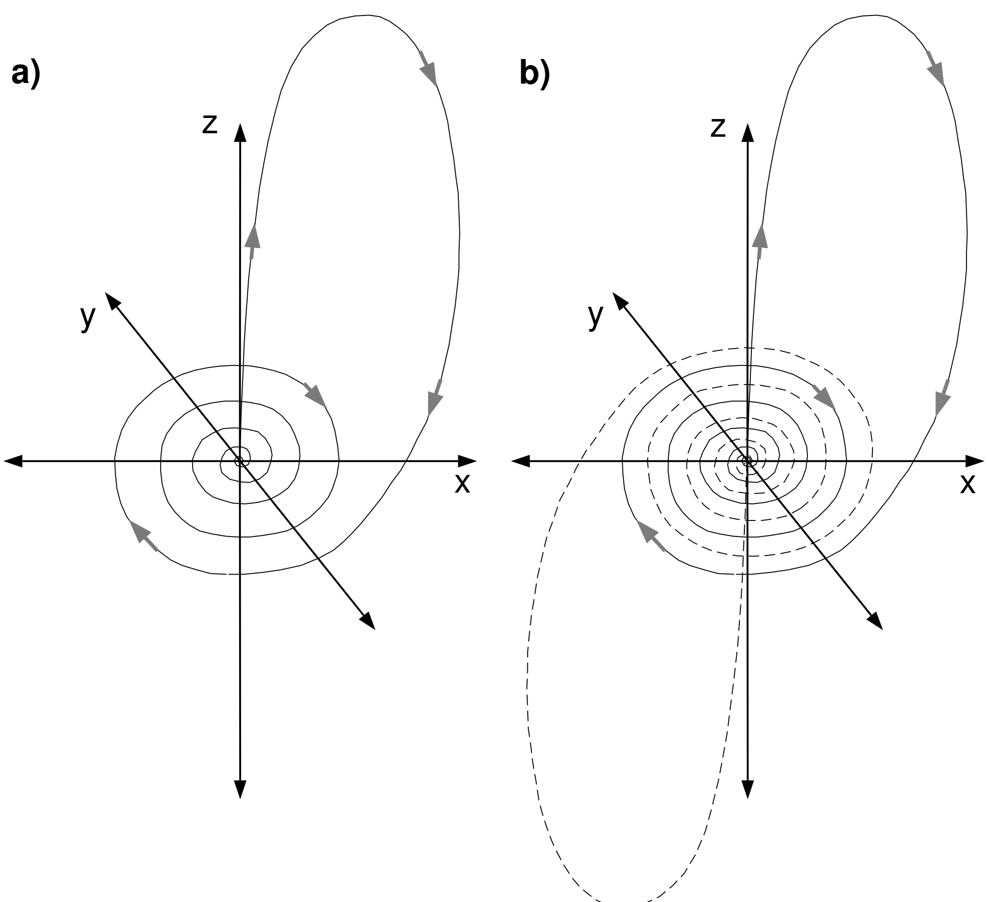

FIGURE 2. Homoclinic orbits to a saddle-focus; a) A single homoclinic, b) A pair of homoclinics in the presence of inversion symmetry.

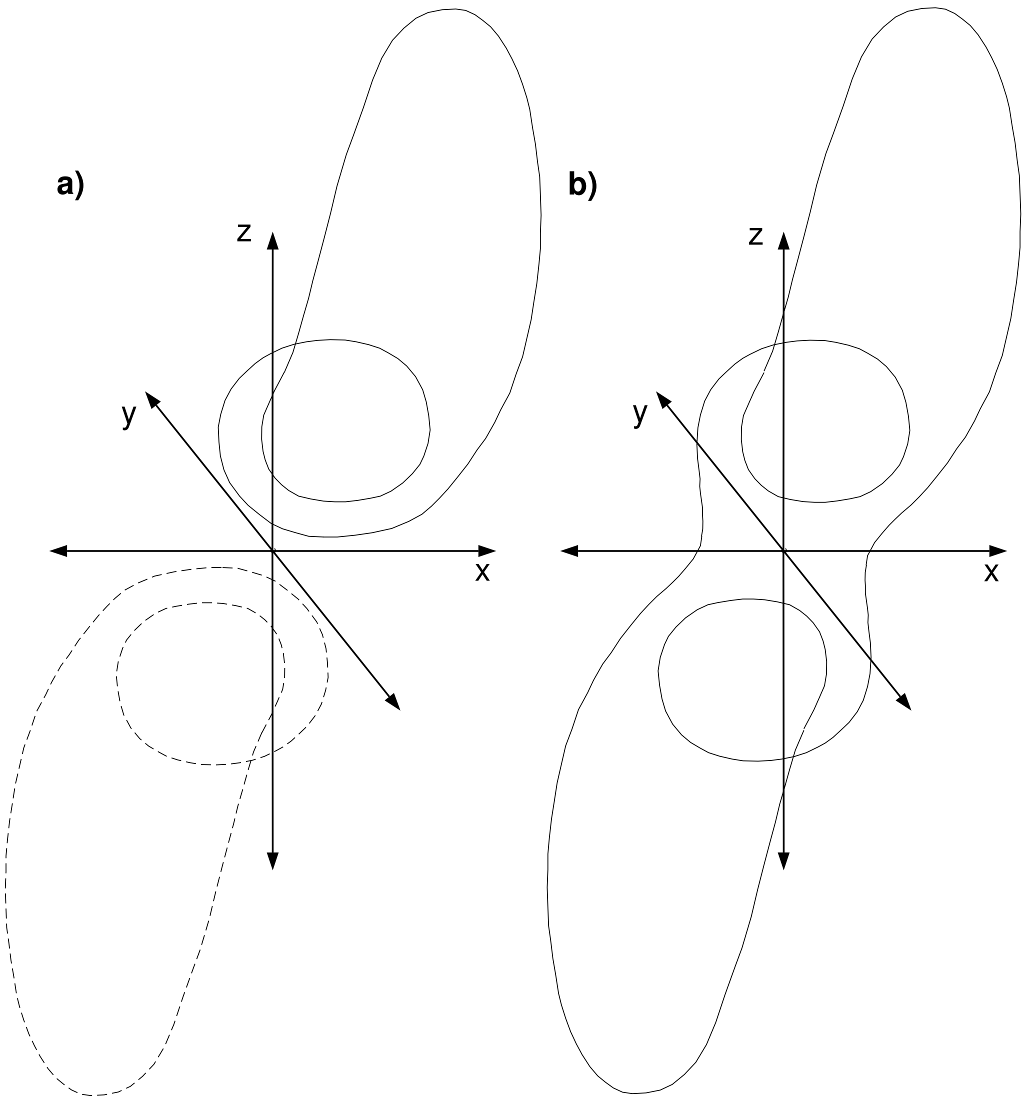

FIGURE 3. Periodic orbits in the presence of inversion symmetry. a) A pair of asymmetric orbits, b) A single symmetric orbit.

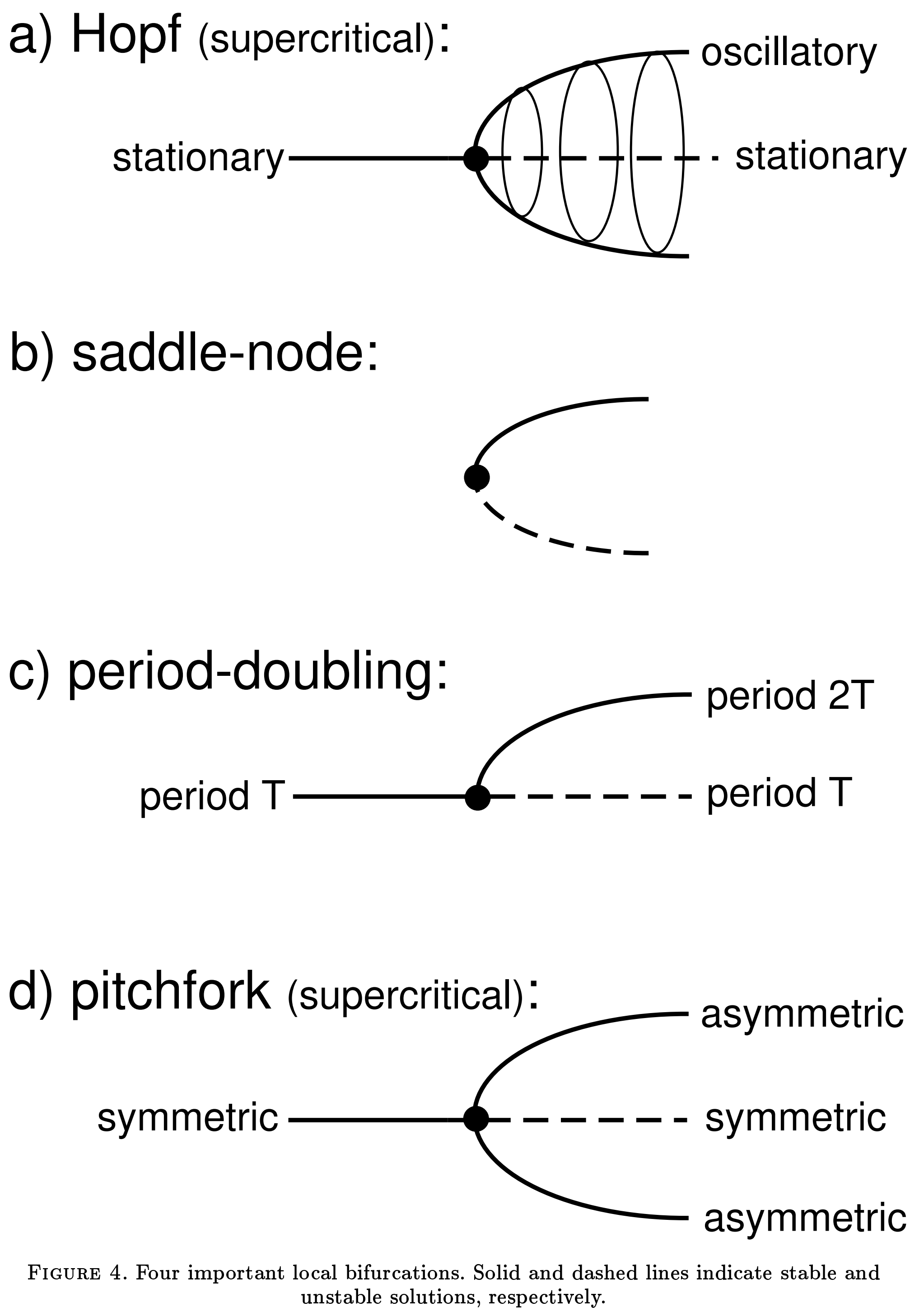

FIGURE 4. Four important local bifurcations. Solid and dashed lines indicate stable and unstable solutions, respectively.

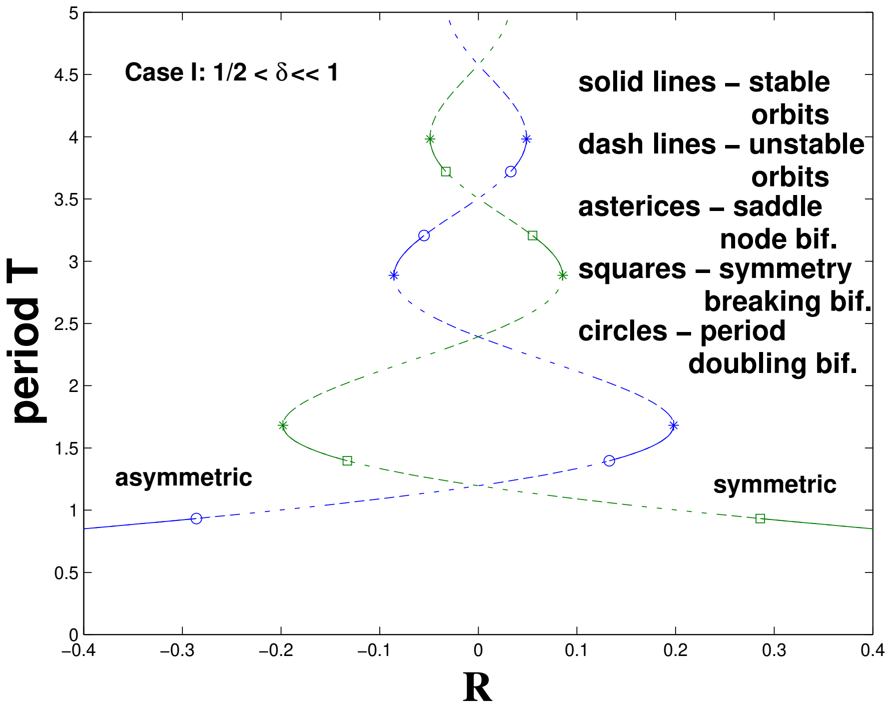

FIGURE 5. Bifurcation diagram for a case I Shilnikov Phenomenon ($1/2 < \delta \ll 1$), indicating the *half* period of the symmetric principal orbit and the full period of the asymmetric principal orbit. The principal orbits are asymptotic to a pair of homoclinic orbits at $R = 0$.

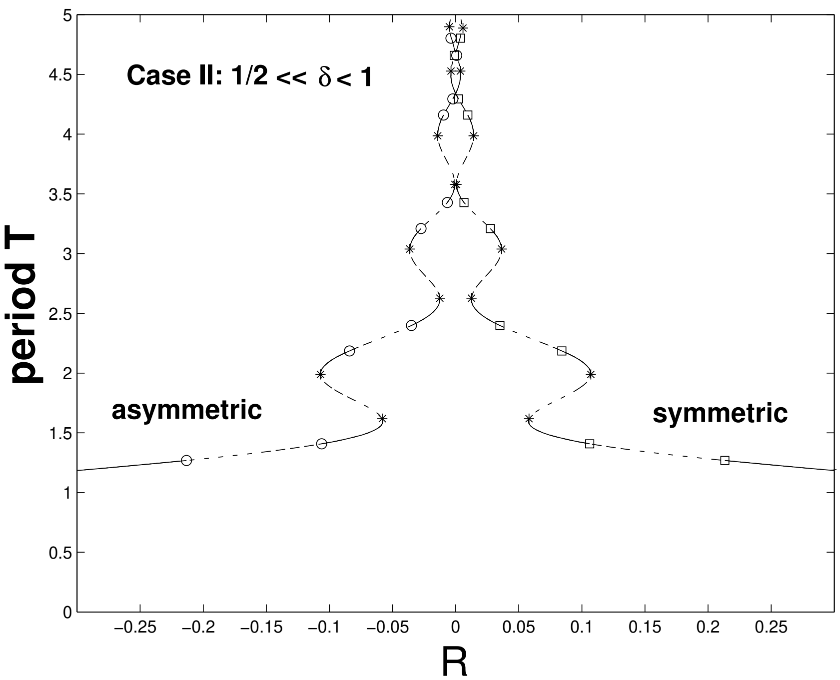

FIGURE 6. Bifurcation diagram for a case II Shilnikov Phenomenon ($1/2 \ll \delta < 1$).

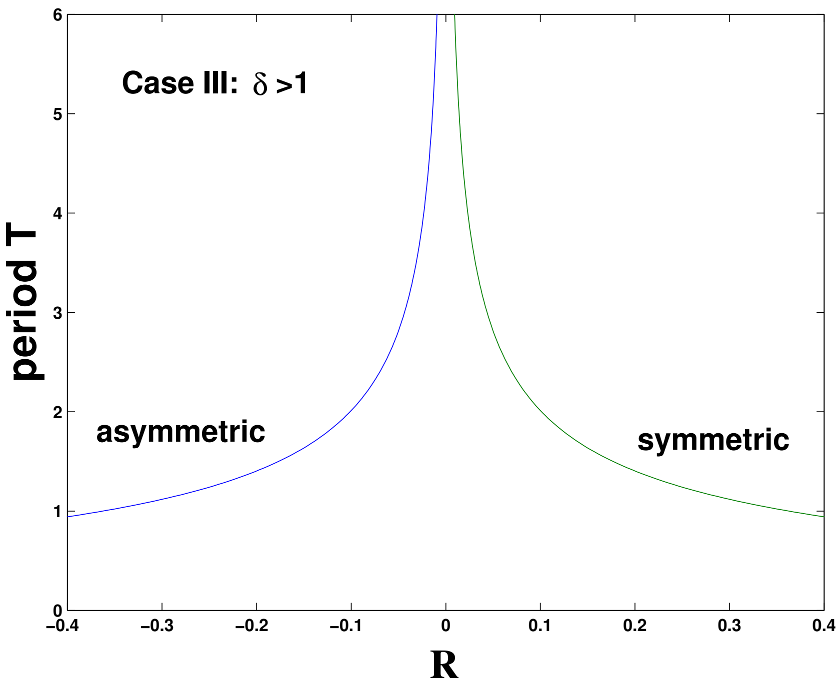

FIGURE 7. Bifurcation diagram for a case III Shilnikov Phenomenon ($\delta > 1$).

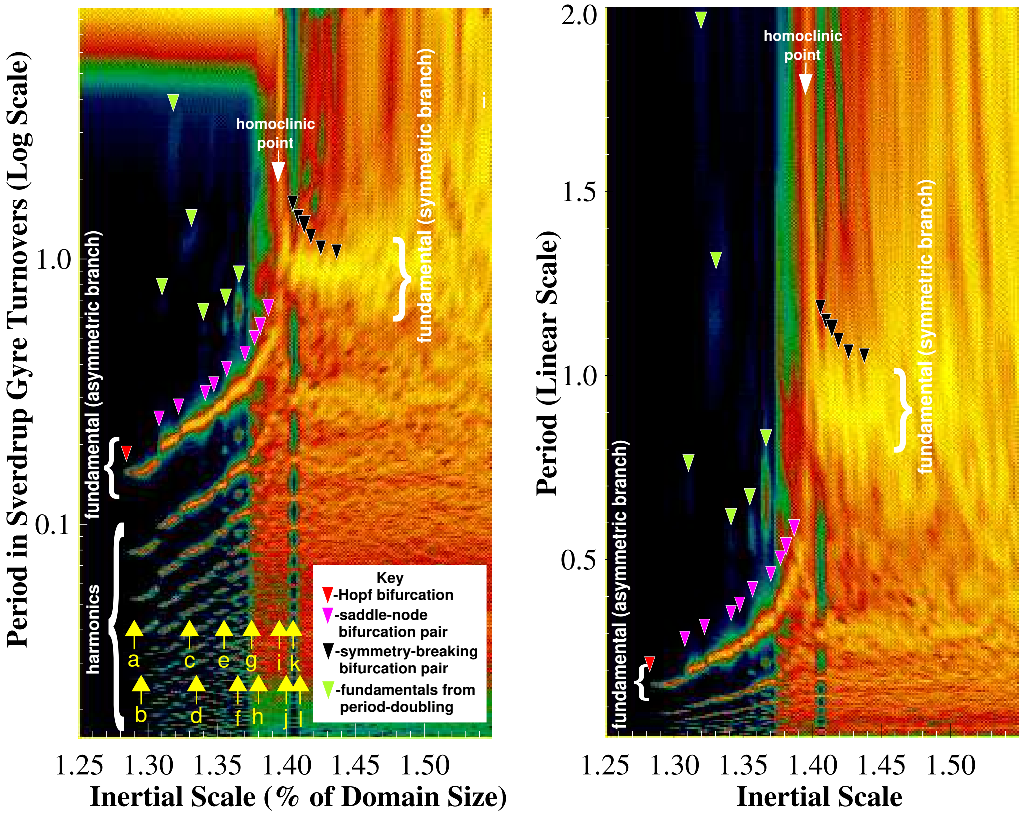

FIGURE 8. Power spectra of the time-asymptotic circulations, calculated from the measured time series $\alpha(t)$, plotted together as a function of the inertial scale. The spectra are color-coded such that blue-black corresponds to the lowest energy levels and yellow corresponds to the highest energy levels. The figure on the left uses a logarithmic scale for the period, includes a larger range than the right-hand figure, and includes the zero-frequency component along the top (for determining the symmetry of the circulation). The figure on the right uses a linear scale for the period (to assist in identifying period-doubling).

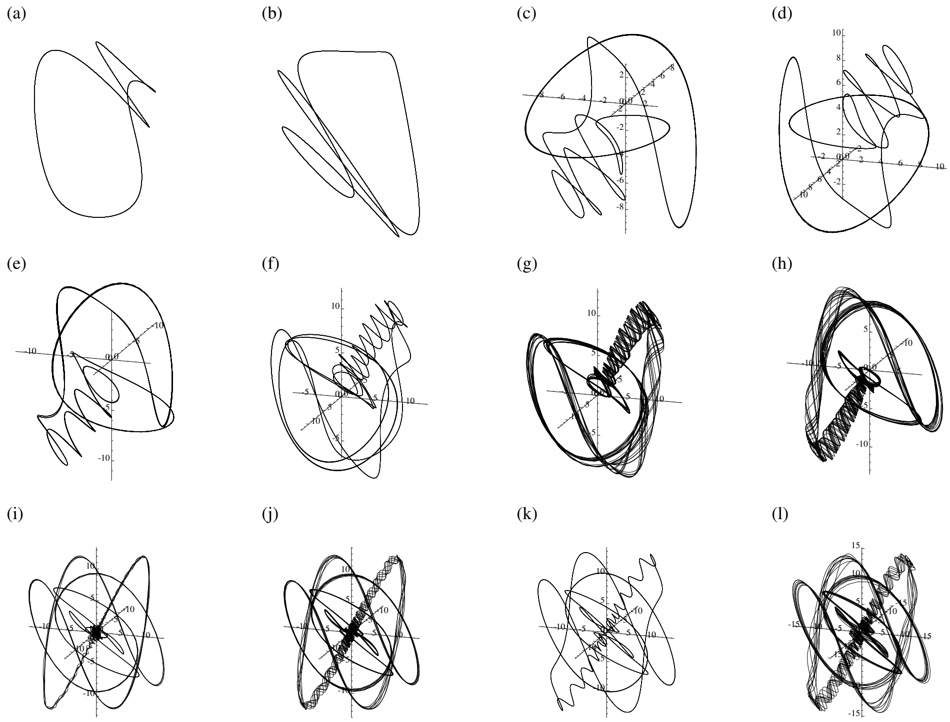

FIGURE 9. Phase trajectories reconstructed from time delay-embeddings of the time series $\alpha(t)$. The letters labeling each projection correspond to the same-labeled points in Fig. 8a. Projections a-h show asymmetric attractors for $\delta_i < \delta_i^{hom}$. In such cases, both left-hand and right-hand versions are present and map onto each other under the inversion symmetry of the system. Projections i-l display apparently symmetric attractors for $\delta_i > \delta_i^{hom}$ (there may be slight asymmetry here in some cases). The coordinate-axes for the first two embeddings are not shown since the origin of the three-dimensional phase space is far removed from the location of those trajectories.

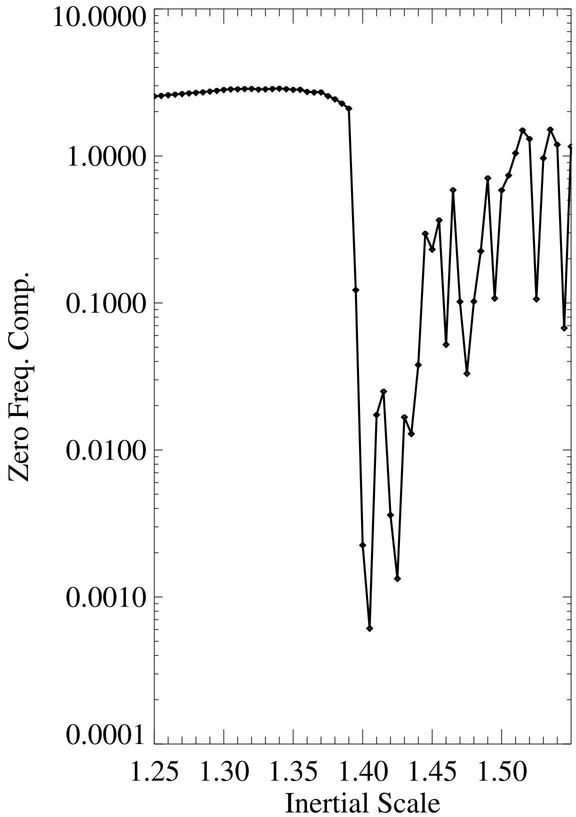

FIGURE 10. The zero-frequency component of the spectrum of $\alpha(t)$ as a function of the inertial scale. Note the sudden drop at around $\delta_i = \delta_i^{hom} \approx 1.395$, indicating a sudden transition from asymmetric to symmetric circulations. Also note the irregular behavior of this component for $\delta_i > \delta_i^{hom}$, which may be due to local symmetry-breaking bifurcations.

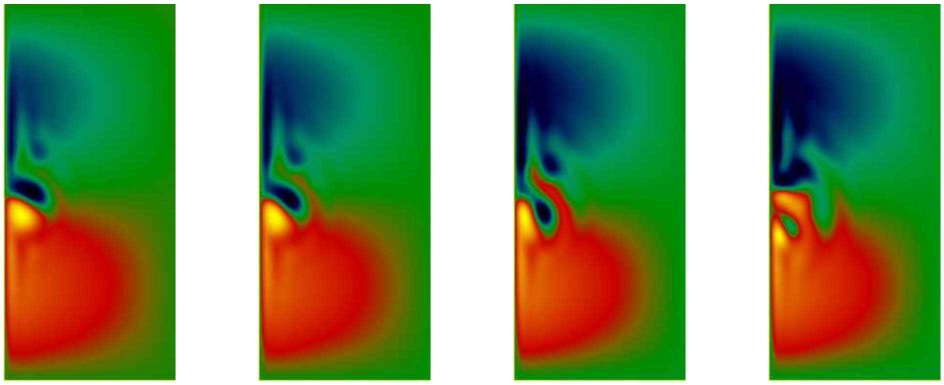

FIGURE 11. Snapshots of the streamfunction at $\delta_i = 1.375$, corresponding to an asymmetric attractor (projection g in Fig. 9). The streamfunction for the full domain is color coded so that blue-black corresponds to the lower values (and hence cold temperatures) and yellow to the higher values (warm temperatures). The flow is seen to be asymmetric as well—in particular, only cold-core rings are seen to be shed.

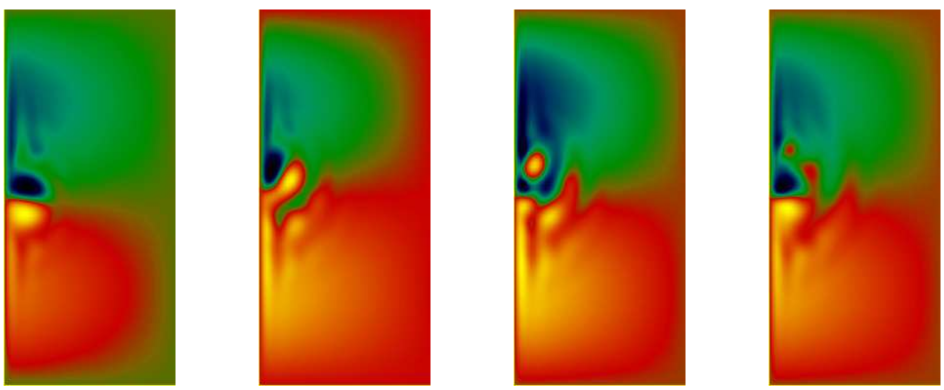

FIGURE 12. Snapshots of the streamfunction at $\delta_i = 1.380$, corresponding to an asymmetric attractor (projection h in Fig. 9). Here, only warm-core rings are seen to be shed.

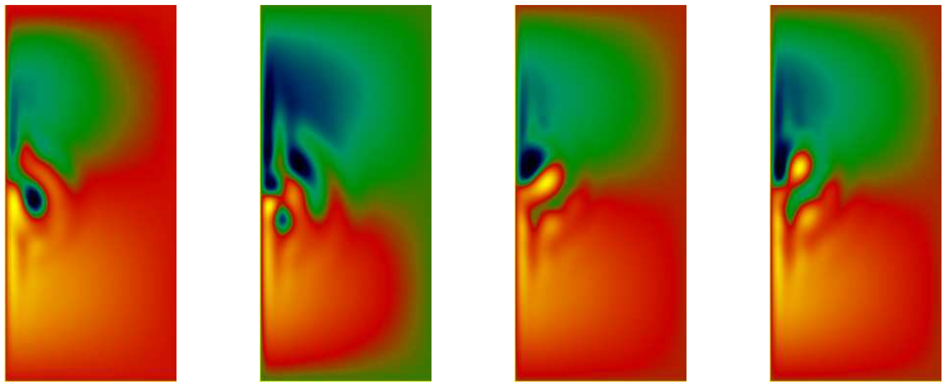

FIGURE 13. Snapshots of the streamfunction at $\delta_i = 1.410$, corresponding to an symmetric or very nearly symmetric attractor (projection l in Fig. 9). Here, the system sheds both cold-core and warm-core rings symmetrically (or very nearly symmetrically).

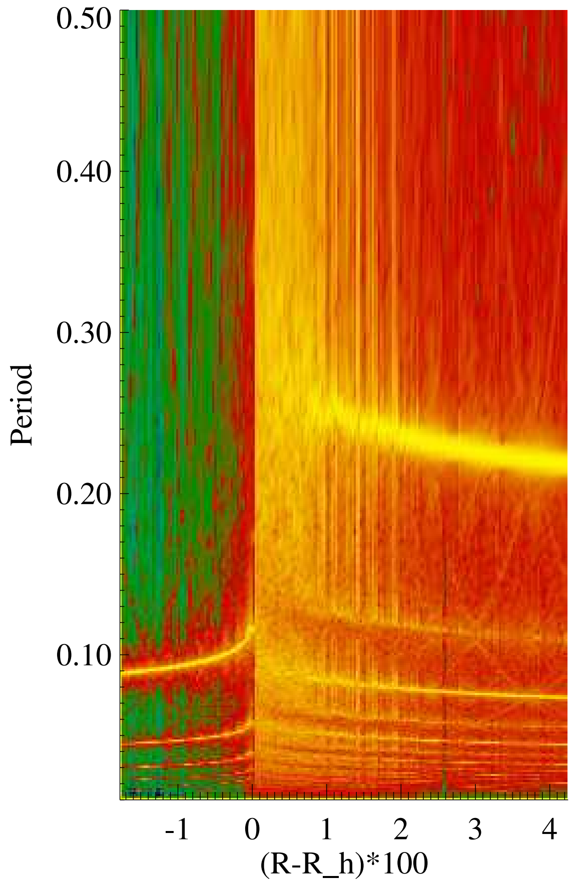

FIGURE 14. Power spectra of the time-asymptotic behavior of time series from system (21) of Luce (1995), with $\nu = 5.5$, $\mu = -4$, and $k = 4\pi^2$. This system possesses a homoclinic at R=R_h =43.9778. Note the overall similarity of this plot to Fig. 8, particularly the divergence at (R-R_h)*100 = 0 (where the homoclinic explosion occurs). Unlike Fig. 8, saddle-node, period-doubling, and symmetry-breaking bifurcations are absent.

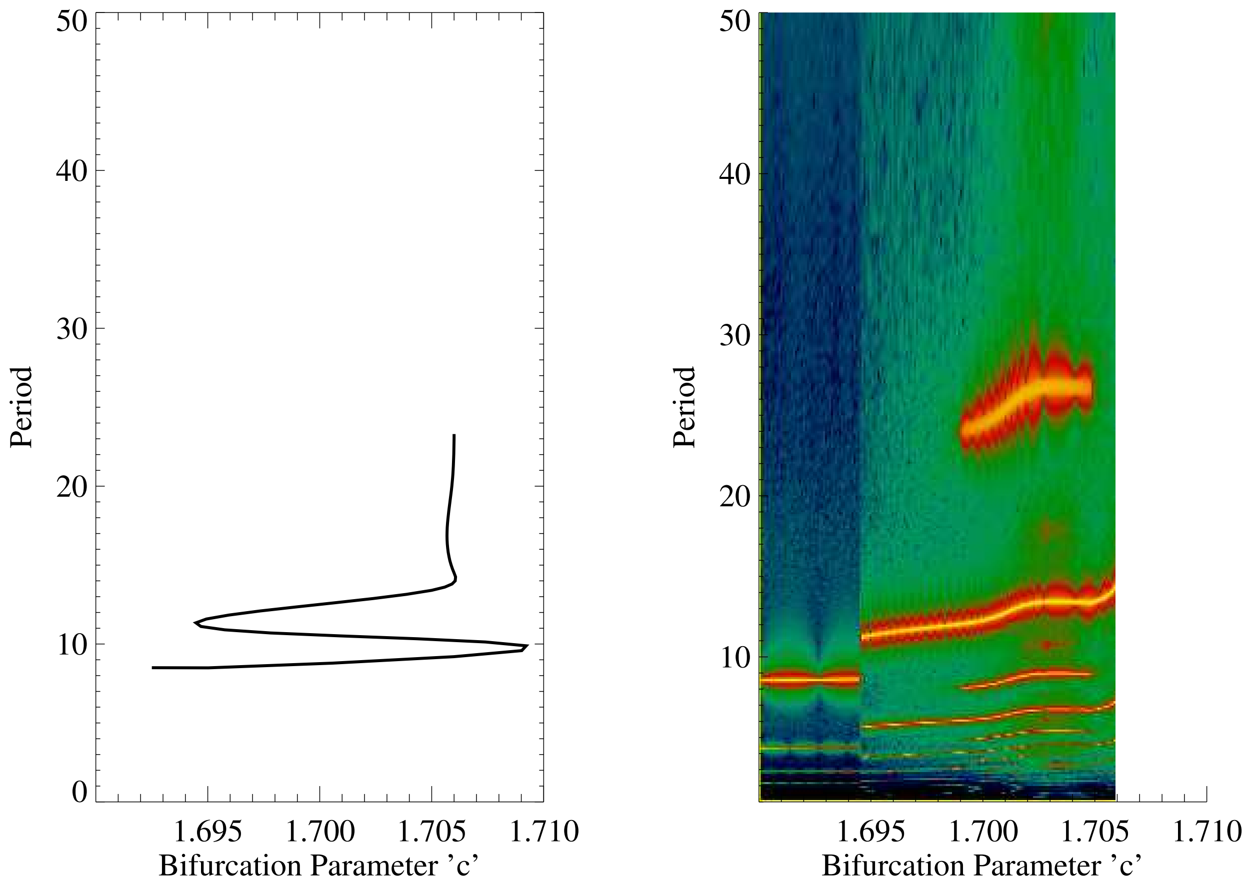

FIGURE 15. Our method of analysis (right) of a simpler ode system in which the Shilnikov phenomenon is better understood further reinforces the analysis of the double-gyre circulation. The figure on the left is an approximate reproduction of Fig. 4.7 (ii) of Glendinning & Sparrow (1984).

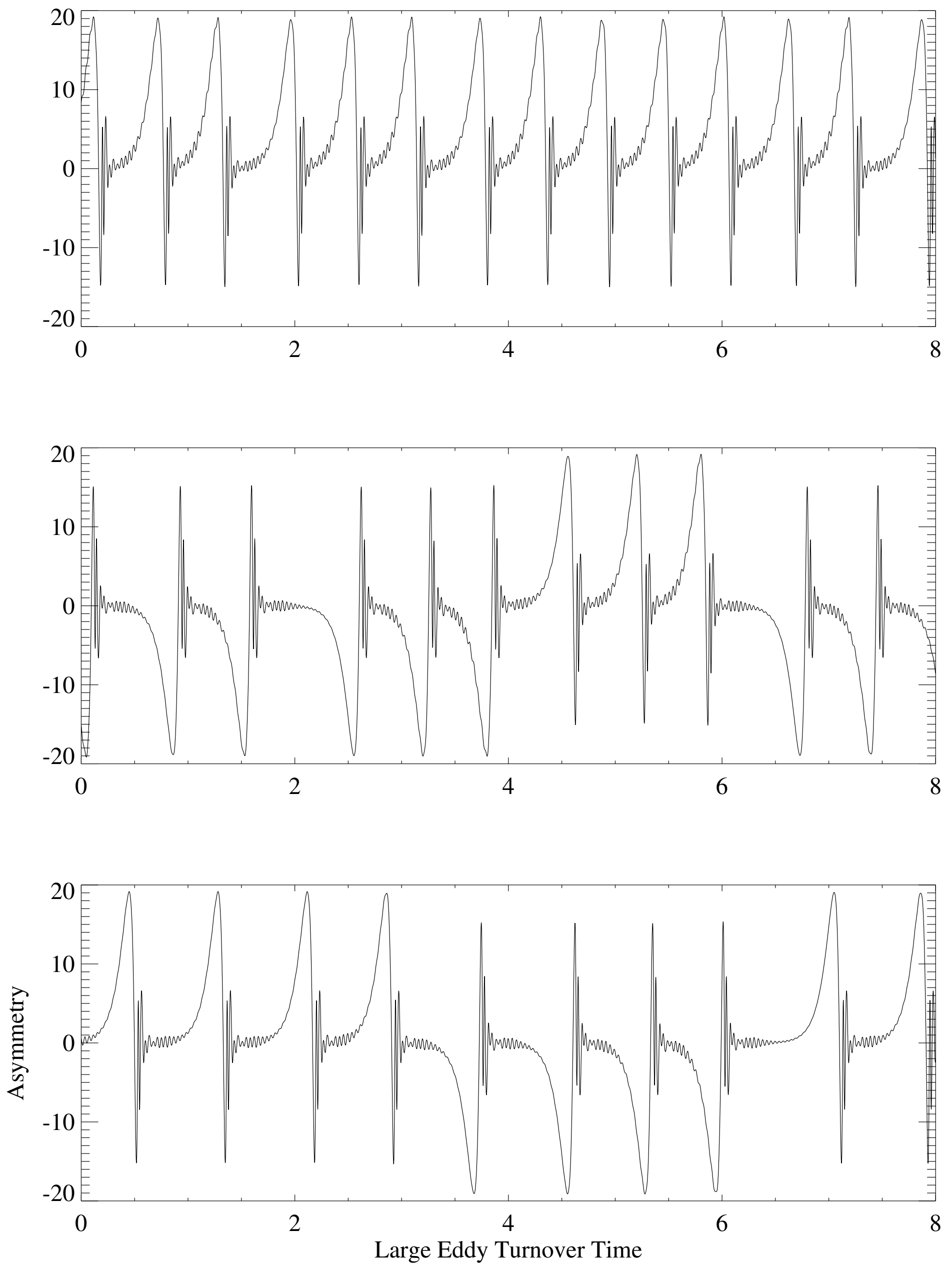

FIGURE 16. Time series of the asymmetry measure at three values of forcing close to the forcing at which the homoclinic is suspected. From top to bottom, the value of the inertial scale δ_i is 1.393 (a), 1.394 (b), and 1.395 (c).

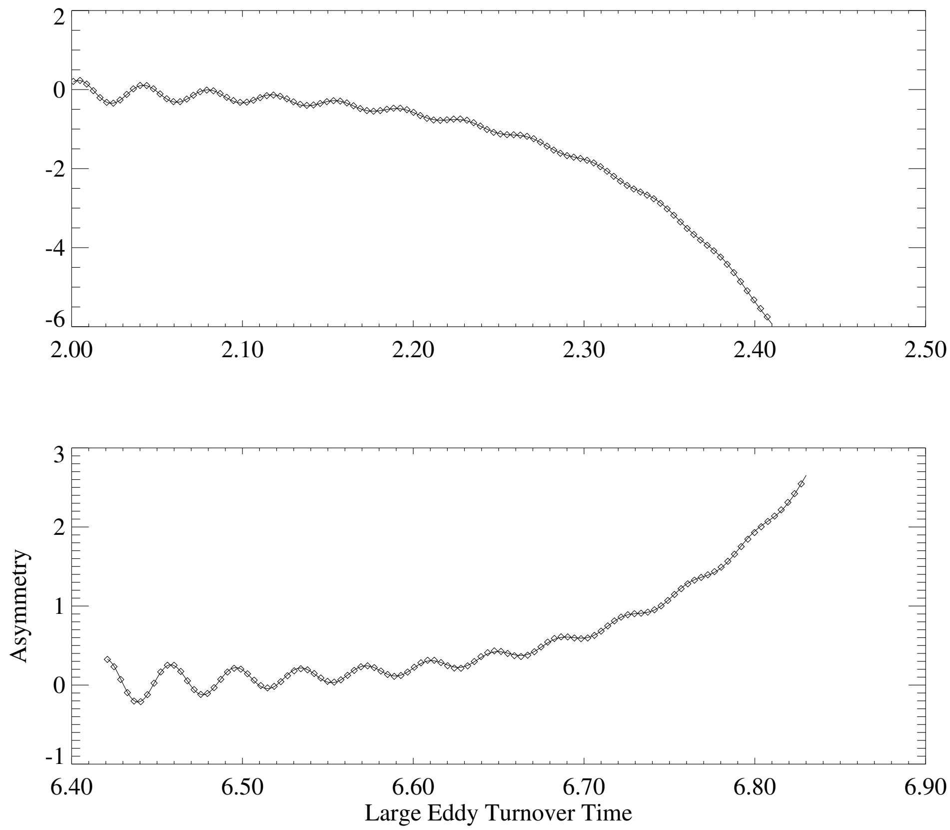

FIGURE 17. Nonlinear fit (continuous lines) of segments of time series $\alpha(t)$ (symbols) in Fig. 16 to equation (5.2) to extract the parameter δ (equation (3.3)). The fits are performed at parameter values corresponding to near-homoclinic circulation, and consist of a duration during which the phase trajectory is close to the saddle-focus fixed point where its evolution is dominated by the linear structure in the vicinity of the saddle-focus. To ensure robustness, two different segments are used. The first segment is chosen from Fig. 16b ($\delta_i = 1.394$) and the second segment is chosen from Fig. 16c ($\delta_i = 1.395$).