*Journal of Marine Research, 53,* 703–737, 1995

Journal of
MARINE RESEARCH

Volume 53, Number 5

# Multiple solutions and advection-dominated flows in the wind-driven circulation. Part I: Slip

by Glenn R. Ierley[1] and Vitalii A. Sheremet[1,2]

ABSTRACT

We consider steady solutions of the barotropic quasigeostrophic vorticity equation for a single subtropical gyre with dissipation in the form of lateral friction. Solutions are governed by two parameters: inertial boundary-layer width; and viscous boundary-layer width. Numerical computations for slip conditions indicate a wedge-shaped region in this two-dimensional parameter space, where three solutions coexist. One of these is a viscous solution with weak recirculation; one a solution of intermediate recirculation; and one a strongly nonlinear recirculation gyre. Parametric scalings based on elementary solutions are numerically corroborated as the first and third of these solutions are continued away from the vicinity of the wedge. The multiplicity of solutions is anticipated by a severely truncated Fourier modal representation paralleling Veronis (1963). The Veronis work was originally applied to predict the possibility of multiple solutions in Stommel's (1948) bottom friction model of the circulation. Paradoxically, it appears the solutions are, in that case, unique.

## 1. Introduction

We present a family of steady solutions of the forced dissipative barotropic beta-plane potential vorticity equation, one of the most basic models of the wind-driven ocean circulation. Save for the brief consideration of bottom friction in Section 5, all of the computations and figures presented refer to a single subtropical gyre of square aspect ratio with stress-free boundary conditions throughout. Given

1. Scripps Institution of Oceanography, La Jolla, California, 92093-0225, U.S.A.
2. On leave from: Laboratory of Geophysical Fluid Dynamics, P. P. Shirshov Institute of Oceanology, Russian Academy of Sciences, Krasikova St., 23, Moscow 117218, Russia.

the lengthy history of similar studies, it is surprising that anything could be said to warrant further attention to a problem that seems so painfully elementary, particularly in comparison to present primitive equation general circulation models. However, careful computation reveals the presence of multiple solutions to the quasigeostrophic equations caused by the north-south symmetry breaking of the beta-effect. In the language of either algebraic geometry or dynamical systems, this is a "cusp."

Veronis (1963) first adumbrated this possibility in a model with bottom friction by using a severe Fourier truncation of the governing equations to obtain a cubic equation for the steady state amplitudes. Making suitable adjustments to the present case of lateral diffusion, the slip condition equivalent of Veronis' predicted location of the "cusp" retains a surprising qualitative validity in spite of the extreme truncation, the essential reasons for which we discuss later in the paper. Ours is hardly the first realization of multiple equilibria in barotropic models of circulation. Sakai (1986), for example, exhibits such solutions for a model of the Kuroshio. His is a local study, rather than full basin, with specified inflow and outflow. As we do here, he compares the qualitatively correct prediction of a low order model with the results of a highly resolved computation.

How has it been that a generic feature of the quasigeostrophic equations[3] would remain unremarked in standard expositions of large-scale circulation theory?

Stommel's seminal description appeared in 1948 and that, along with Hidaka's complementary higher order friction model in 1949 (with subsequent explicit solution in Munk, 1950), admitted attack by linear and weakly nonlinear analysis. These analytic methods produced solutions which were, if not wholly realistic even given the generously large error bars in the then-prevailing observational picture, an apparently satisfactory zeroth order picture of great physical transparency. By the time the Veronis paper appeared in 1963, however, numerical models, as in the landmark paper of Bryan (1963), began to offer a competing source of model development which, though more "realistic" (e.g., in resolving time-dependent instability), was simultaneously less accessible to the qualitative and intuitive insights characteristic of analysis.

With the increasing attention to numerical models, the notion of studying exact steady solutions, to say nothing of crude approximations of steady solutions, acquired a faint aura of quaintness even by the time of this exchange recorded at a 1972 symposium organized by the National Academy of Sciences following a lecture by Steven Orszag on spectral methods for numerical modeling of fluid flows:[4]

3. Multiple solutions can also arise as a *spontaneous* symmetry breaking. For example, equal forcing of a subtropical/subpolar gyre pair can give rise to an unsymmetrical response but this is a more tenuous feature since there is no reason to single out exact symmetry of forcing as a preferred condition. In contrast, the influence of the explicit symmetry breaking by the beta-effect is hard to escape.

4. Source: "Numerical Models of Ocean Circulation" (page 299), proceedings of a symposium organized by the Ocean Science Committee of the Ocean Affairs Board of the National Academy of Sciences. Held at Durham, New Hampshire, October 17–20, 1972.

> *Charney:* George (Veronis), you tried to model the wind-driven ocean circulation with how many Fourier components including the Gulf Stream? Come on now.[5]
> *Veronis:* Well, Okay. Look, I admit, but after all, it wasn't a very good calculation. The only thing I can do is take refuge in saying I was pioneering.

which bears witness to the decisive tilt in favor of computer modeling as the ultimate arbiter of "truth." Quite apparently, if computer simulation failed to reveal any influence of Veronis' multiple equilibria, they could be dismissed as an artifact of excessive truncation. Another datum: Veronis (1963) was cited in the literature six times in the interval 1965–1969, twice more in 1975 and one last time in 1979.

While the flood of computer simulations may have seemed an advance to which the analytic queries of theoreticians would add only a trickle of occasional tabulations, paradoxically, a precise determination of both unstable steady solutions and their accompanying linear eigenmodes is substantially *more* demanding than stepping the time-dependent solution forward, so one could argue that an interregnum of numerical neglect of foundational theoretical issues was inevitable. Computers in 1963 *were* a powerful new tool. They simply weren't powerful enough to address analytical questions that, at the time, would anyway have seemed nugatory.

With the hindsight of thirty years, it was not only the advent of dynamical systems, with its attention to attractors, fractal sets, homoclinic orbits, and various other oddments, that would turn one's attention back to a basic problem like multiple steady state equilibria but also some powerful results on "approximate inertial manifolds" or their closely related cousins of GFD provenance, slow manifolds. These theorems justify the application of the machinery of dynamical systems to precisely the evolution equations of atmospheric and oceanic circulation models. (For an accessible discussion in the context of the related two-dimensional Navier-Stokes equations see Temam, 1991.)

Beyond that, the cusp "catastrophe" is robust. By robust we mean it is a generic aspect of the solution, i.e., that its qualitative consequences persist in variations of the original problem which may include the addition of dynamical complications such as topography and stratification. (For a useful introduction to the subject see Gilmore, 1981 or the terser, more abstract, Arnol'd, 1984.) To make a central point consider Figure 1. The vertical axis is chosen as a characteristic feature of the solution. In this paper we shall adopt the log of the global maximum of the streamfunction. The horizontal log axes are the controlling parameters for the system, i.e., the inertial and viscous boundary-layer thicknesses, $\delta_I$ and $\delta_M$ respectively. We have omitted axis calibration to focus upon the qualitative character of the bifurcation. (More quantitative aspects are described later but the surface is essen-

5. By an entertaining bit of irony, Charney himself would appeal to a model every bit as crude as Veronis (1963) when modeling multiple equilibria as an explanation of atmospheric blocking in Charney and DeVore (1978).

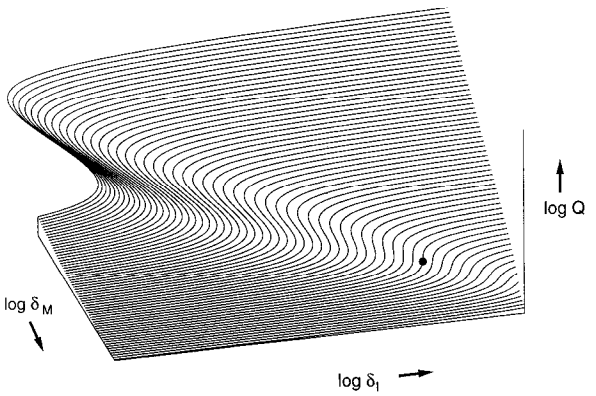


Figure 1. A schematic bifurcation plot to illustrate the occurrence of multiple solutions of (2.1) over a finite region in the control parameter space ($\delta_M$, $\delta_I$). Vertical axis indicates the log of the maximum of the streamfunction (or, equivalently, maximum transport). The arrows indicate increasing values of the respective coördinates.

tially quantitatively correct barring some mild liberties with smoothing to create an easily interpolated and rendered surface.)

As the reader can see, for a range of the parameters, there are three possible steady solutions. (The terminus of that region, the "cusp," is marked with a solid black dot.) This much is generally familiar about bifurcation curves, though not so in the context of the wind-driven circulation problem. Less apparent, though equally important, is the *magnitude* of the circulation associated with the "upper branch." This, as we shall see, can be considerable, say 50 $ms^{-1}$ *everywhere* in the basin. In addition, for fixed inertial boundary-layer width, in the limit as the viscous boundary-layer width tends to zero, the magnitude of this *steady* transport *does not* saturate, it increases without bound.

How does our interest in the latter asymptotic limit mesh with present day GCM computations?

A near-universal defining characteristic of numerical and theoretical models on gyre scales and larger is the use of enhanced dissipation to account for the action of unresolved scales of motion with eddy resolving general circulation models (ERGCMs), for example, commonly operating in the range of $10^2 - 10^3 m^2 s^{-1}$. The conventional wisdom is that instabilities in the aggregate act on the large scales much as disorder at the microscopic scale acts to produce molecular viscosity. However appealing this analogy, one cannot regard its application as satisfactory if two calculations at explicit resolutions differing, say, by an order of magnitude and each with a suitably tuned explicit dissipation give results differing at first order in predicted features of the large-scale circulation. While for the purpose of interpolation or short term prediction there is nothing invidious about setting an Austauch coefficient to that value producing best agreement with past and present observation,

as a tool for substantial extrapolation outside the range of observation (climate change, for example), models not shown to converge in the large scale for decreasing eddy viscosity coupled to higher explicit resolution are questionable at best as the results remain hostage to wholly unknown, and possibly gross, inadequacies in the subgrid-scale modeling.

Apart from the deficiencies of subgrid-scale modeling, and while hopelessly far removed from the realm of direct simulation, one can at least imagine a gedanken experiment with the viscous coefficient of the conventional lateral friction term now to be regarded as tending to its molecular value. In *this* limit, it is of interest to know whether or not the time-dependent solution does "run away." At issue is the existence of what we might call a sensible mechanical balance. By this, we mean that the magnitude of the large-scale (mean) circulation remains sensibly constant with changes in the boundary-layer Reynolds number. That is, the cascade of instabilities cuts off at some viscously controlled wavenumber, $k_{\text{visc}}$, and the large-scale solution becomes asymptotically independent of $k_{\text{visc}}$. Lack of a mechanical balance means that the amplitude of the mean circulation remains controlled by $k_{\text{visc}}$, diverging in the limit of $k_{\text{visc}} \to \infty$. In this regard, it is instructive to recall that the mean downstream velocity for simple turbulent pipe flow *never* ceases to depend upon the molecular viscosity, $\nu$. Turbulence does, however, slow the rate of growth: while the laminar cross-stream average grows as $\nu^{-1}$, in the fully turbulent regime both experiment and theory indicate attainment of a weaker asymptotic scaling of $\log(1/\nu)$. Wishful thinking aside, that oceanographic flows should behave any differently is neither obvious nor, in our view, likely.

To recapitulate, a "proper" eddy viscosity would be one which, as reduced, both exposes progressively smaller scales explicitly to be resolved and leaves the mean large-scale flow sensibly unchanged. It is dubious that any present GCM has this property or that any form so simple as a scalar coefficient in front of some convenient number of powers of the Laplacian could achieve that end. A second, and deeper, issue is whether the large-scale flow is ultimately controlled by the actual value of the molecular viscosity. It is naïve to suppose that one can provide a wholly satisfactory resolution of the first issue without also addressing the second.

While these concerns inform our interests and establish a context, their resolution lies far over the horizon. In this paper we content ourselves with examining the narrow issue of characterizing the steady state solutions, fixed points in the language of dynamical systems, for single gyre barotropic flow. (Even this is not so elementary as it appears, as we have discovered from our ongoing work with no-slip boundary conditions.)

In the next section we define the nondimensional problem to be solved. The discussion of parameters and scaling is for lateral friction with stress-free (slip) boundary conditions. In Section 3 we turn to a description of the steady state solution space. We do not approach the problem with mathematical rigor, indeed all of the

detailed results proceed found from computation. Strictly speaking we cannot prove that our numerical results are correct. They are internally consistent with appropriately small *ex post facto* error estimates based on spectral decay. Further corroboration emerges in the interpretation of those results by appeal to two limiting analytic solutions. The numerical results approaching the relevant limits are found to be consistent with analytic scalings. Section 4 outlines the constraints imposed by the integrated vorticity budget upon the class of allowed $q(\psi)$ relations for flows dominated by advection. In Section 5, we review Veronis' original low order truncation of a bottom friction model and extend the model to the case of lateral friction. We give a brief synopsis of bottom friction computations for the full partial differential equation (PDE). The low order model with lateral friction is in qualitative agreement with the full PDE. Such agreement is not found in the case of bottom friction. We suggest an explanation for the disparity. We conclude in Section 6 with a brief summary of the relation of this work to others both in oceanography and in general fluid mechanics.

## 2. Mathematical formulation of the problem

We consider an idealized model of the wind-driven circulation in a barotropic ocean of constant depth. With the shallow water, quasigeostrophic, beta-plane and rigid lid approximations the problem reduces to solving a two-dimensional vorticity equation, usually referred to in the geophysical literature as the Barotropic Vorticity Equation (BVE)[6]

$$\nabla^2\psi_t + \delta_I^2 J(\psi, \nabla^2\psi) + \psi_x = \delta_M^3 \nabla^4\psi + \nabla \times \tau. \tag{2.1}$$

where $\psi$ is a streamfunction; $J(\psi, \zeta) = \psi_x\zeta_y - \psi_y\zeta_x$ is the advection operator. (This and succeeding equations are all nondimensionalized.) The $x$-axis is directed to the east; the $y$-axis is directed to the north. The zonal and meridional velocity components are obtained from the streamfunction with

$$u = -\psi_y, \qquad v = \psi_x. \tag{2.2}$$

The relative vorticity $\zeta$ and the streamfunction are related by Laplace's equation

$$\nabla^2\psi = \zeta. \tag{2.3}$$

The model basin is square ($0 \le x \le 1$, $0 \le y \le 1$). For simplicity we assume that there is only a zonal component of the wind stress $\tau^{(x)} = -(1/\pi)\cos \pi y$, so that the curl of the wind stress takes the form

$$\nabla \times \tau = -\sin \pi y, \tag{2.4}$$

which mimics the annually averaged wind stress distribution over the North Atlantic

6. Notation as in (Pedlosky, 1987).

and ensures zero Sverdrup flow at the northern and southern boundaries of the basin. Subject to the approximations noted above, the surface wind stress is represented by a body force in the two-dimensional equation (2.1).

The pair of equations (2.1) and (2.3) are solved subject to zero mass flux kinematic boundary conditions

$$\psi = 0 \quad \text{at} \quad x = 0, \quad x = 1, \quad y = 0, \quad y = 1 \tag{2.5}$$

and slip (no-stress) dynamical conditions

$$\zeta = 0 \quad \text{at} \quad x = 0, \quad x = 1, \quad y = 0, \quad y = 1. \tag{2.6}$$

We choose this simple form since it gives a nonsingular solution in the limit of vanishing viscosity and hence averts certain formal difficulties associated with other possible choices.

We have chosen as a length scale the basin dimension $L$; and as a time scale $T = 1/\beta L$, the inverse of the Coriolis parameter variation over the meridional extent of the basin. The velocity scale in the interior, $U$, and a characteristic value of wind stress, $\tau_0$, are connected by Sverdrup relation (Sverdrup, 1947): $U = \tau_0/\rho_0\beta LH$. ($\rho_0$ is the constant density of water and $H$ is the constant depth of the ocean.) The chosen nondimensionalization is convenient since the maximum nondimensional value of the Sverdrup transport at the western boundary is then unity.

In nondimensional form problem (2.1)–(2.6) has only two parameters:

$$\delta_I = \sqrt{\frac{U}{\beta}} \Big/ L, \tag{2.7}$$

a measure of nonlinearity determined as the width of the inertial boundary layer (or, equivalently, the intensity of the wind forcing) and

$$\delta_M = \sqrt[3]{\frac{A_L}{\beta}} \Big/ L, \tag{2.8}$$

a diffusive parameter based upon the width of purely viscous boundary layer. $A_L$ is the coefficient of horizontal eddy viscosity. As $\delta_I$ is increased to equality with $\delta_M$ an inertial-viscous western boundary layer forms (see, for example, Kamenkovich, 1977, Ch. 6 or Pedlosky, 1987, Ch. 5).

In application to the real ocean the values of the parameters $\delta_I$, $\delta_M$, if not rigorously determined, are small, but a key point is that the structure of the solution of the BVE is determined by their ratio or, for convenience, the Reynolds number for the boundary current

$$R = \left(\frac{\delta_I}{\delta_M}\right)^3. \tag{2.9}$$

We will mostly use nondimensional variables throughout the paper; therefore, it is useful to bear in mind the following values typical for the subtropical gyre in the North Atlantic. The meridional extent is $L = 2000$ km. With $\beta = 2 \times 10^{-13}$ s$^{-1}$ cm$^{-1}$, the time scale is then $T = 2.5 \times 10^4$ s. A typical value of the Sverdrup transport for the Gulf Stream is $G = 30$ Sv $= 3 \times 10^{13}$ cm$^3$ s$^{-1}$ suggesting a scale of the streamfunction of $\Psi = G/H = UL = 6 \times 10^8$ cm$^2$ s$^{-1}$, where we have assumed the effective depth of the Gulf Stream to be $H = 500$ m.[7] The canonical basin interior velocity is $U = 3$ cm s$^{-1}$. To provide the stated value of the Sverdrup transport $G$ we need a wind stress of $\tau_0/\rho_0 = 6$ cm$^2$s$^{-2}$ which is larger than the typical observed value 1 cm$^2$ s$^{-2}$. However, if we take into account the aspect ratio of the North Atlantic, 2000 km/6500 km and the factor of $1/\pi$ we easily produce a more reasonable value 0.6 cm$^2$ s$^{-2}$.

The dimensionless inertial boundary-layer width is $\delta_I = 1.94 \times 10^{-2}$ (equivalent to a dimensional value of $\sqrt{\pi}\delta_I L \simeq 70$ km) with the maximum velocity of the boundary current $U/\delta_I \simeq 150$ cm s$^{-1}$. One cannot give an unambiguous estimate of the horizontal eddy viscosity, $A_L$, but with reference to the purely frictional Munk theory with no-slip boundary conditions the maximum value of the streamfunction is found at a distance $2\pi/\sqrt{3}(A_L/\beta)^{1/3}$ from the coast. If we take the observed value of about 200 km and subtract the width of the shelf, we get a value of about 130 km, which corresponds to $A_L = 1.0 \times 10^7$ cm$^2$ s$^{-1}$. (This is less than the overestimated value $3.3 \times 10^7$ cm$^2$ s$^{-1}$ given by Munk himself.) With this value of lateral viscosity and 2000 km for the meridional extent of the basin we get $\delta_M = 1.84 \times 10^{-2}$. This gives an estimate for the Reynolds number of $R = 1.17$ which, in the context of the Munk solution, would be correct only if the nonlinearity were small, $\delta_I \ll \delta_M$. Our estimate of $\delta_I$, however, is determined independently, from the amplitude of wind forcing. The fact that even with this estimate $R$ is larger than one means that the nonlinear terms are significant in the Gulf Stream. The main point is that, regarding the BVE as a model of the mean circulation, we are interested in solutions where the inertial layer is as large as, or larger than, the viscous layer.

## 3. Multiple steady solutions

In this work we concentrate on steady solutions of the problem (2.1)–(2.6), their existence and uniqueness. (For an abstract treatment on the existence of solutions in a model with bottom friction only, see Barcilon *et al.,* 1988.) In general, the steady solution $\psi_0$ of (2.1) may be unstable and thus inaccessible to traditional time stepping

7. Strictly speaking, for a barotropic model confined to the upper 500 m of the water column, we ought to consider a one-and-a-half layer model. This adds a term $-F\psi_t$ to the left-hand side of (2.3) where $F = f_0^2 L^2/(g'H)$. Steady solutions are unaffected by this addition insofar as advection is concerned. One could argue whether or not to modify the dissipative term, i.e., is $\nabla^2 q$ or $\nabla^4\psi$ the more fundamental, but the qualitative outline presented here is unaffected by such considerations. A more compelling argument can be made that linear stability is appreciably affected. Because our main concern in this paper is with the steady solution we do not further argue the point here.

computations. We apply Newton's method to a Chebyshev series expansion for the solution $\psi_0$. This approach is little used in studies of the general circulation so, for the benefit of the interested reader, we have provided a few guidelines in an Appendix.

While oceanic interest attaches to a certain parameter range for the BVE, in this paper we consider the global range of solutions over the entire $(\delta_I, \delta_M)$-plane since particular features in one region, largely inexplicable if seen in isolation, are greatly clarified if viewed as one member of a continuous family.

We begin by noting the existence of analytical approximations in two limiting cases:

1. The solution with slip boundary conditions in the linear case ($\delta_I = 0$, $\delta_M \ll 1$), first given in Welander (1964), may be written as

$$\psi_M(x,y) = \left((1-x) - e^{-\xi/2}\left(\cos\frac{\sqrt{3}}{2}\xi - \frac{1}{\sqrt{3}}\sin\frac{\sqrt{3}}{2}\xi\right)\right)\sin \pi y + O(\delta_M), \quad (3.1)$$

where we have introduced the stretched variable $\xi = x/\delta_M$. This solution has a boundary-layer character and consists of a weak ($v \sim O(1)$) southward Sverdrup flow in the interior of the basin (Sverdrup, 1947) and an intense northward current ($v \sim O(1/\delta_M)$) along the western boundary (an analog of the Gulf Stream) which closes the pattern of circulation.

2. The gyre solution in the highly nonlinear case ($\delta_I \gg \delta_M, \delta_M \ll 1$).

$$\psi_{IG} = \frac{1}{\pi^5\sigma_M^3}\sin \pi x \sin \pi y + \frac{1}{150\pi^5\delta_M^3}(\sin 3\pi x \sin \pi y + \sin \pi x \sin 3\pi y) + \cdots \quad (3.2)$$

This solution, which derives ultimately from Lamb (1932), represents a single rapidly rotating gyre occupying the whole basin. Odd harmonics in $x$ beyond the first are systematically taken into account by averaging the wind curl over the streamlines of the fundamental, $\sin \pi x \sin \pi y$ (V. Barcilon, private communication). These produce a slight curvature to the $q(\psi)$ relation for large $R$. Though a little crude to neglect terms whose formal order is the same as that of the fundamental, that the first term of (3.2) alone immediately satisfies

$$J(\psi, \delta_I^2\nabla^2\psi) = 0 \quad (3.3)$$

is qualitatively correct in its implication that advection dominates the dynamics for this solution.

Before Briggs (1980) nobody paid serious attention to this solution since it gives unrealistically high velocities ($O(50\ \mathrm{ms}^{-1})$ for the North Atlantic) throughout the gyre. Restored to dimensional terms, the streamfunction is proportional to

$(\delta_I^2/\delta_M^3)\beta L^3$. This upper branch of solutions, with large amplitude even in the vicinity of unit boundary layer Reynolds number and sharply increasing amplitude away from the fold, cannot be ignored merely because it is inconvenient. One can question, however, the extent to which it represents *a characteristic* feature of the solutions of the BVE. Though problems most amenable to analytic inquiry may rely upon simplifications such as forcing in the form of the gravest eigenmode of the Helmholtz operator, we suggest that the existence of strongly nonlinear solutions of the upper branch is typical, at least for the slip conditions considered. As noted in Cessi and Thompson (1990), Cessi (1991) and Marshall (1993), the combined influences of forcing and geometry are critical in determining whether recirculation is present. The double gyre problem is thus one obvious extension to test the merit of our conjecture, for which see Cessi and Ierley (1995), where an even richer spectrum of Briggs-type solutions is exhibited. For the moment, however, it is enough to note that recirculation *is* a dominant feature of the mean circulation in the North Atlantic so our restriction here to circumstances which favor its appearance is not without purpose.

It is useful first to demonstrate the monotonic transition from the linear to the highly nonlinear case by increasing the Reynolds number $R$ for fixed $\delta_M$ (or, in equivalent dimensional terms, increasing the forcing at fixed eddy viscosity) since the dimensionless solution in the two limiting cases depends only on the diffusion parameter (see (3.1) and (3.2)). Figure 2 shows a sequence of six streamfunctions for $\delta_M = 0.06$ as a function of increasing Reynolds number. The linear Munk solution ($R = 0$) is symmetrical relative to $y = 1/2$. Nonlinearity breaks the symmetry; the maximum of the streamfunction migrates from $y = 1/2$ to the northwest corner of the basin, where the recirculation gyre forms ($R = 0.2$). As nonlinearity increases, the recirculation gyre grows in size and intensity ($R = 1$) and eventually extends to the eastern coast ($R = 1.45$). Then it fills the basin ($R = 2$) and ultimately approaches a basin-filling gyre solution as $R \to \infty$. In fact, the solution is quite close to this condition even for $R = 5$–$10$. Insofar as symmetries realized by the sequence in Figure 2 are concerned, there is a visual degree of similarity to the equivalent sequence displayed in Veronis (1966b) but one should be cautious not to overextend the analogy given the potentially singular character of the problem with lateral friction. Indeed a similarly appealing, but misleading, analogy is the identification of the numerical sequence of Veronis (1966b) as one tending to the Fofonoff (1954) unforced, inviscid solution (discussed in Section 4). As Zimmerman (1993) notes, the Veronis sequence is one of increasingly nonlinear solutions, but these do not recover the Fofonoff solution in the limit.

Maximum transport, $Q$, (maximum value of the streamfunction) as a function of Reynolds number, $R$, offers a useful characterization of the solutions of the BVE. The transport is shown in Figure 3a as a function of $R$ for several values of $\delta_M$. The cases presented in Figure 2 are marked by small circles on the curve corresponding to

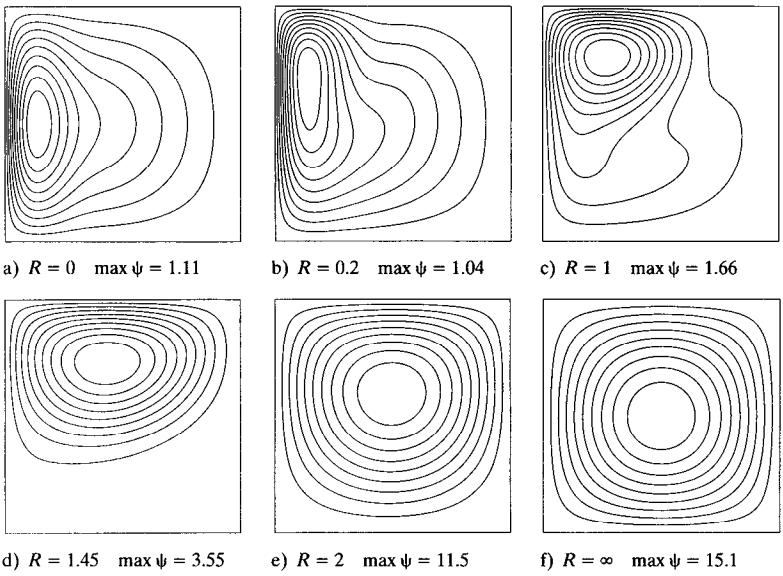


Figure 2. A sequence of the streamfunctions illustrating the monotonic transition from the linear Munk solution ($R = 0$) to the highly nonlinear gyre solution ($R = \infty$) for fixed viscosity, $\delta_M = 0.06$, and increasing Reynolds number, $R$. The contour interval is $CI = 0.1 \times \max\psi$.

$\delta_M = 0.06$. In the linear case ($R = 0$) the streamfunction has a maximum, $Q = 1 + \exp(-2\pi/3\sqrt{3}) - O(\delta_M) = 1.30 - O(\delta_M)$, in the midpoint of the western boundary current at the latitude $y = 1/2$ where the Sverdrup transport attains its maximum value. In the highly nonlinear case as the solution approaches the basin-filling gyre solution $Q \to 1/(\pi^5\delta_M^3)$ (with additional corrections as indicated by successive terms in 3.2).

At any $\delta_M$ the transition from linear to nonlinear circulation occurs for $R \sim 1$, i.e., when the thickness of the inertial boundary layer, $\delta_I$, becomes comparable to the thickness of the viscous boundary layer, $\delta_M$. A physical argument can be made as follows: In steady state, global vorticity balance requires the input of negative vorticity by the wind to equal the diffusion of positive vorticity through the boundary of the basin. In the linear case, due to the beta-effect, an intense boundary layer of width $\delta_M$ forms near the western boundary. The high vorticity gradients in this region provide sufficient diffusion of positive vorticity from the boundary into the area enclosed by the streamlines passing through the interior to balance the wind. As the thickness of the inertial boundary layer approaches that of the viscous boundary layer, there forms an inertial-viscous boundary layer, the characteristic thickness of which cannot be less than $\delta_I$ (Kamenkovich, 1966). This hinders effective diffusion of vorticity into the interior as compared with the previous case in which all streamlines thread the viscous boundary layer. In order to balance the global vorticity budget,

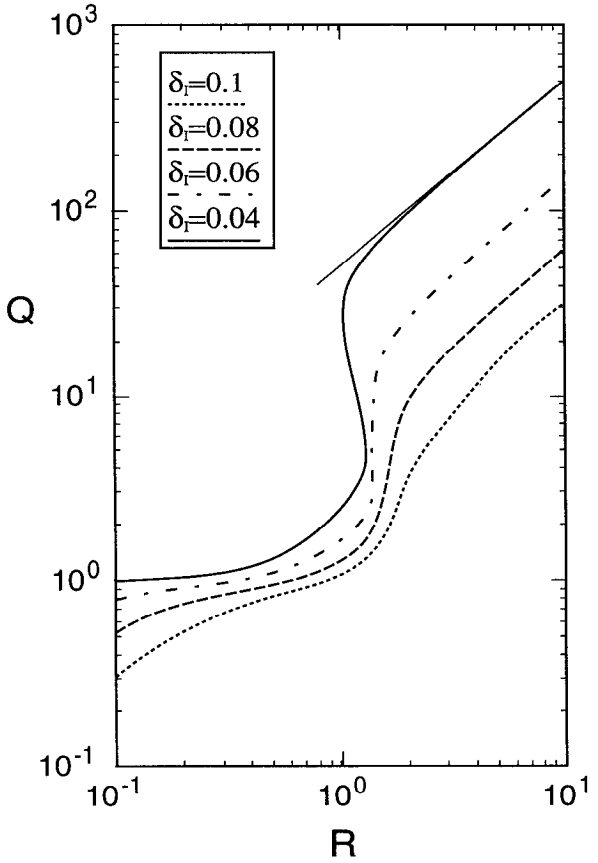


Figure 3. (a) The dependence of the maximum of the streamfunction, $Q$, on the Reynolds number $R$ for several values of viscosity, $\delta_M$. *For* $\delta_M < \delta_{M_C} = 0.0555$ the function $Q(R)$ is multivalued corresponding to the 'S'-shaped transition from the linear to the highly nonlinear case. The small circles indicate cases shown in Figures 2 and 4. The triangles mark the three different solutions shown in Figure 4a,b,c. The secondary (Sverdrup) maximum of the streamfunction is labeled with $Q_{Sv}$. The asymptotic limit (3.2) is indicated with a light straight line for the most inviscid transect ($\delta_M = 0.02$). Solutions were tracked along the middle branch using a method outlined in the Appendix. Plotted curves here and 3(b) were prepared using a cubic spline fit in log coördinates with arc length as the independent variable.

large vorticity gradients are established by increasing velocities of the particles in the whole basin, which implies a breakdown of the boundary-layer solution and a transition to the basin-filling gyre type solution.

In the viscous regime, say $\delta_M = 0.1$ or larger, the broad features of the solution change relatively little in the transition from linear to nonlinear. As viscosity decreases, however, the difference between linear and highly nonlinear cases both in

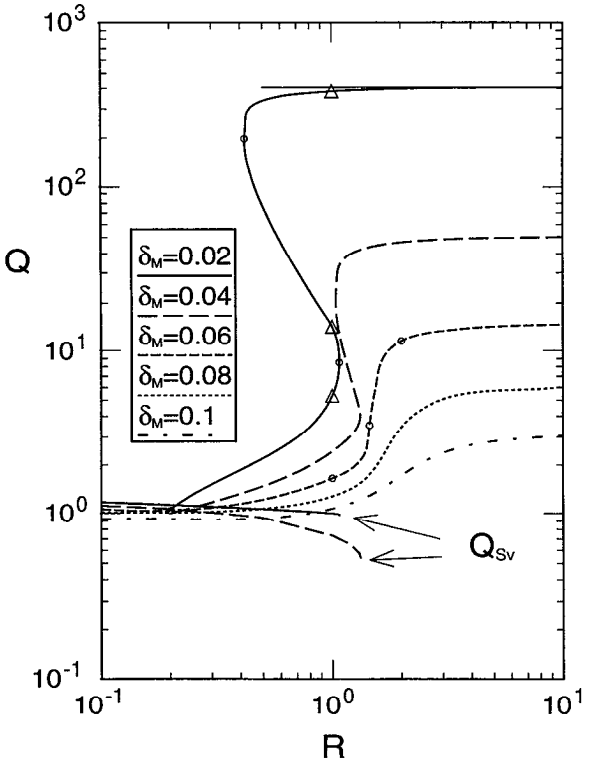


Figure 3. (b) The dependence of $Q$ on $R$ but for several fixed values of the nonlinearity, $\delta_I$. The 'S'-shaped (multivalued) transition takes place for $\delta_I < \delta_{I_C} = 0.06207$. Again, the limiting solution is indicated by the light straight line at the top of the figure.

structure *and* amplitude of respective streamfunctions becomes marked. The case we have just illustrated is transitional between viscous and inviscid and, not coincidentally, this transect in $\delta_I$ at fixed $\delta_M$ passes marginally close to the "cusp" we have thus far indicated only schematically. On transects lying above the cusp we observe a monotonic evolution from Sverdrup maximum to recirculation maximum. For transects in the cusp region, however, with increasing Reynolds number *a second* maximum appears coincident with the onset of recirculation. For the range of Reynolds numbers of interest, this second maximum becomes the more useful bifurcation parameter as it exceeds the Sverdrup value. (Even in the cusp region, for sufficiently large $R$, the highly nonlinear case returns to a single maximum with value, as above, of $1/(\pi^5\delta_M^3)$.)

The most arresting feature of this problem is that, for values of $\delta_I$ and $\delta_M$ lying in an open wedge, the solution becomes *multivalued.* In other words, for the same values of the parameters there may exist three different steady solutions. This possibility is realized in the 'S'-shaped curves of Figure 3a for values of $\delta_M$ of 0.04 and 0.02. Our

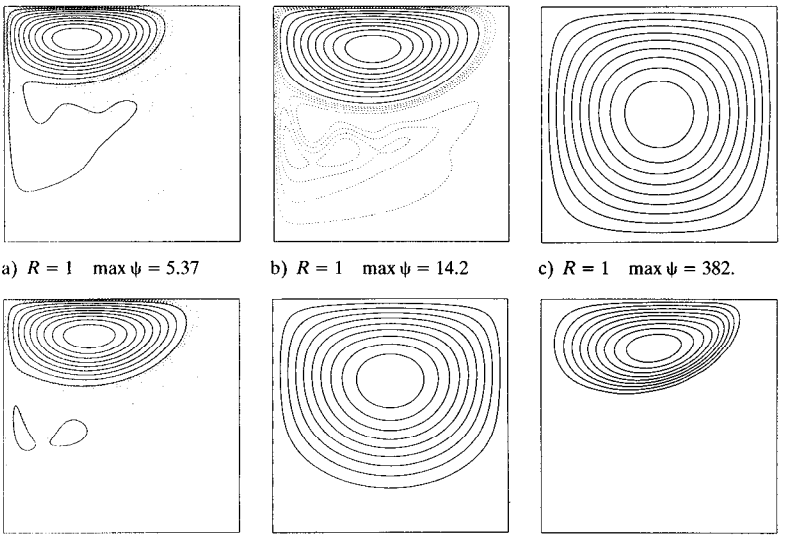


Figure 4. A sequence of the streamfunctions illustrating the 'S'-shaped (multivalued) transition from the linear to the highly nonlinear case for fixed viscosity, $\delta_M = 0.02$. The upper row displays three different steady solutions for the same Reynolds number, $R = 1$, corresponding to the (a) low, (b) middle and (c) high branches of solutions. In the lower row are: (d) the steady solution at the low nose point; (e) the steady solution at the high nose point; and (f) the neutral eigenfunction $\psi_{\text{Rec}}$ at the low nose point. The contour interval of the solid isolines is $0.1 \times \max\psi$. The dotted isolines in (a), (b) and (d) are drawn with a contour interval of 0.2 in order to reveal the Sverdrup interior.

most precise estimate for the tip of the wedge is $\delta_{M_C} = 0.0555$ and $\delta_{I_C} = 0.06207$. We term the branches of the solution low, middle and high according to the respective global maximum of the streamfunction. Solutions corresponding to all three branches for $\delta_M = 0.02$ and $R = 1$ are shown in Figure 4a,b,c. The first two are similar, differing only in the intensity of the recirculation.

In the problem thus far, we have regarded the basin aspect ratio as a secondary parameter but it is illuminating to turn for the moment to aspect ratio as a controlling parameter for the location of the cusp. As noted earlier, transects in $\delta_I$ for moderate to large values of $\delta_M$ are monotonic: there is a unique steady solution characterized by a single maximum of $Q$ which migrates smoothly from $y = 1/2$ to the northwest corner of the basin (Fig. 2) as the recirculation becomes the dominant feature. Transects at smaller $\delta_M$ show an 'S'-shaped transition from linear to nonlinear behavior: there are multiple steady solutions over a finite range in $R$ coincident with a change *in topology,* that is, *a separate* local secondary maximum, $Q_{Rec}$, associated with a compact recirculation gyre arises in the northwest corner in addition to the mid-basin Sverdrup maximum, $Q_{Sv}$. (See Fig. 4a.) This topological distinction

between a single maximum and the appearance of double maxima is strongly reminiscent of the breakdown of a boundary-layer approximation examined in Ierley (1987)—a reduced version, that is, of (2.1). In that paper, it is suggested that the nonexistence of an outflow boundary-layer solution for $R > R_c = 0.029$, if it does not predict, at least presages the appearance of a second maximum, that associated with recirculation.

This putative relationship seems to contradict experience. In solving the full equation (2.1) with conventional values of aspect ratio, a boundary-layer type solution is seen to persist to substantially greater $R$, e.g., 0.2 in Figure 2, before recirculation is observed. In Ierley (1987), a sequence of regional steady solutions is computed at progressively increasing aspect ratio. The import of these is that the delayed onset of recirculation, $R_c = 0.029$ from the boundary-layer analysis vs. $R_c \approx 0.20$ in conventional gyre models, arises from boundary-layer corrections in the aspect ratio, $\alpha$, of north-south gyre scale to viscous boundary layer width. That is, in the limit of $\alpha \to \infty$, recirculation does commence at $R_c = 0.029$. Analysis by Mallier (1994) has worked in the direction of extending this asymptotic result to finite values of $\alpha$ but progress hinges upon successful treatment of certain Painlevé logarithmic "resonances" which, even for the simpler case of bottom drag, present a formidable challenge.[8] Nonetheless, while a continuous transformation from the boundary-layer limit to the fully two-dimensional solutions discussed here cannot be analytically demonstrated, we believe that there is such a transformation connecting the two. The existence of a cusp is thus anticipated not only in Veronis' direct low order modal truncation, which is one way to project the full equation (2.1) down onto a tractable subspace of modest dimension, but also in the (singular) projection of the problem down onto a parametric ODE which describes the western boundary layer. The latter, in a word, "inherits" the cusp, which manifests itself by the disappearance of the solution at a critical value of the boundary-layer Reynolds number.

In Figure 3a the Sverdrup maximum as a function of the Reynolds number $R$ is shown by the additional curve labeled $Q_{Sv}$ for $\delta_M = 0.04$, $\delta_M = 0.02$. For small Reynolds numbers the recirculation gyre maximum is less than the Sverdrup maximum but is nonetheless readily identifiable as an isolated knot located in the northwest corner of the basin. As the Reynolds number increases the recirculation gyre grows to dominate the Sverdrup maximum $Q_{Sv}$ but $Q_{Sv}$ can be smoothly tracked up to the edge of a "fold" in the solution space. The particular parameters at which the solution surface first folds back on itself trace out a curve in the parameter space. We speak of a given point along this curve as the "low nose." Similarly, the solution surface folding back on itself a second time produces the "high nose." The folds are

8. Note that the transformation from boundary-layer form to the two-dimensional PDE, (2.1), is an unfolding: in the boundary-layer expansion at leading order, there is only a single parameter, the *ratio* of inertial to viscous scale (or, equivalently, the boundary-layer Reynolds number); as a PDE the problem occupies the two-dimensional parameter space we have outlined in this paper. The PDE solutions as represented in Figure 5 thus collapse onto a single line for $\alpha \to \infty$.

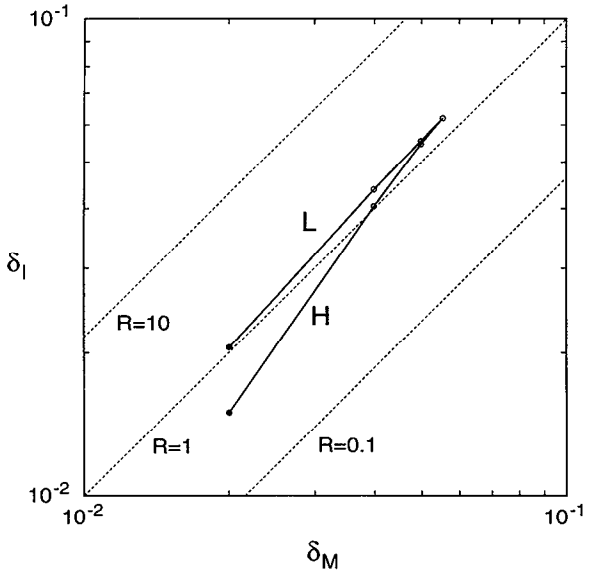


Figure 5. The location of the region in the ($\delta_M$, $\delta_I$)-plane where the three multiple solutions exist. 'L' indicates the projection of the low nose points and 'H' the high nose points (or in other words projections of the folds of the surface shown in Fig. 1). The small circles indicate the solutions at the low and high nose points in some experiments. The dark small circles at the end of the 'L' and 'H' lines correspond to the solutions shown in Figures 4d and 4e. The cusp point is at $\delta_{M_C} = 0.0555$, $\delta_{I_C} = 0.06207$, or $R_C = 1.3987$ with $Q_C = 3.46$ marked by a circle at the tip.

easily seen in the schematic surface shown in Figure 1 as the locus of points where the local surface normal is orthogonal to the vertical axis. The projection of these points onto the horizontal ($\delta_I$, $\delta_M$)-plane forms the wedge-shaped region shown in Figure 5 inside of which multiple steady solutions exist. (The highly nonlinear case returns to a single maximum.) The bounding curves are marked 'L' and 'H', corresponding to the location of the low and high nose points respectively. The curve 'L' indicating the breakdown of the boundary layer type solution closely approaches the line $R = 1$ while the curve 'H' deviates more and more from the curve 'L' with decreasing viscosity. From a transect at $\delta_M = 0.02$ we show solutions from the edge of the folds (at the points indicated with open circles in Fig. 5). The solution at 'L' appears in Figure 4d and that at 'H' in Figure 4e.

This situation is a typical example of the cusp catastrophe (Arnol'd, 1984, Ch. 2). It arises when a three-dimensional surface given by the equation $y = x^3 - 3px$ is projected onto the plane $(p, y)$, where one sees a semi-cubic parabola $y = \pm 2p^{3/2}$ with a cusp at the origin. In application to our case we can make the following qualitative correspondences: $x \leftrightarrow \log(Q)$, $y \leftrightarrow \log(R)$, $p \leftrightarrow (\log(\delta_{M_C}) - \log(\delta_M))$. The two branches of the semi-cubic parabola correspond to the 'H' and 'L' curves.

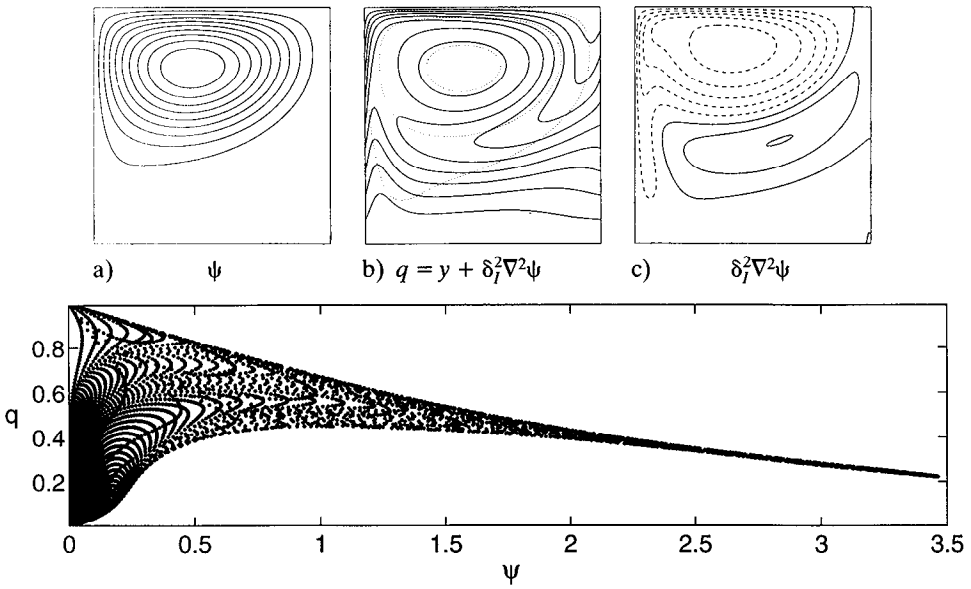


Figure 6. The solution at the cusp point $\delta_{M_C} = 0.0555$, $\delta_{I_C} = 0.06207$, or $R_C = 1.3987$. (a) The streamfunction $\psi$. $Q_C = \max\psi = 3.46$. The contour interval is $CI = 0.1 \times \max\psi$. (b) The potential vorticity $q = y + \delta_I^2\nabla^2\psi$. $CI = 0.1$. The dotted lines show streamlines of $\psi = 0.2$, 1 and 3. (c) The relative vorticity $\delta_I^2\nabla^2\psi$. $CI = 0.1$. Dashed lines indicate negative values. (d) The scatterplot $q$ vs. $\psi$ produced by evaluating $q$ and $\psi$ at evenly spaced points in physical space.

Inside the small region separated by 'H' and 'L' curves there are three steady solutions. Outside this region there is only one solution and on the curves, two. Approaching either the 'H' or the 'L' curve from the interior of the wedge, two solutions coalesce and here, according to Whitney's singularity theory (1955), a *fold* singularity takes place. Approaching the tip of the wedge, all *three* solutions coalesce to form a so-called *cusp* singularity.

Whitney proved an important theorem that the cusp is topologically stable, which implies that a change of variables affects neither the location of the cusp point nor the scaling law of '3/2' for the separation of the two curves as a function of the distance from the cusp point. (This can be demonstrated by Taylor expansions in the vicinity of the cusp point.) Numerical experiments in the immediate vicinity of the cusp (with $\delta_M = 0.054$, 0.055, 0.0555 and 0.056) confirm that, indeed, the difference in Reynolds numbers between the low and high nose points scales as $\Delta R = R_L - R_H \sim p^{3/2}$, while the difference in the maximum values of the streamfunction between the high nose point and low nose point scales as $\Delta Q = Q_H - Q_L \sim p^{1/2}$.

The solution corresponding to the terminal point of the cusp is shown in Figure 6. The spatial structure of the streamfunction (Fig. 6a) is something intermediate in

transition from the Munk solution to the basin-filling gyre solution, with a maximum value of the streamfunction of $Q_C = 3.46$. Figure 6b gives the potential vorticity $q = y + \delta_I^2 \nabla^2 \psi$ with several superimposed typical streamlines shown by dotted lines. When the friction and forcing are not important in the balance of terms, fluid particles tend to conserve their potential vorticity, as in the recirculation gyre, where the isolines of $\psi$ and $q$ almost coincide. This conservation is implicit in a portion of the scatterplot of $q$ vs. $\psi$ (Fig. 6d). The recirculation gyre corresponds to larger values of $\psi$ where a tight functional relation is seen though not, strictly speaking, a linear one. The further from the center of the gyre, the greater the scatter. We return to these points in Section 5. The relative vorticity $\zeta$ is presented in Figure 6c.

As implied by Figure 5, transects in the $(\delta_M, \delta_I)$-phase plane for fixed nonlinearity $\delta_I$ track the transition from linear to highly nonlinear solutions. When $\delta_I < \delta_{I_C}$, we encounter multiple solutions. Figure 3b, where $Q$ as a function of $R$ for several values of $\delta_I$ is shown, illustrates this. In contrast with the transects of Figure 3a which asymptote to constant $Q$ as $R \to \infty$, here the limiting nondimensional solution diverges as $R \to \infty$. As one can see by reference to (2.1), an increase in $\delta_I$ can be compensated structurally, that is, by making $\nabla^2 \psi$ bear a tighter functional relation to $\psi$, thus decreasing the size of the Jacobian in proportion while leaving the amplitude of $\psi$ unchanged. No such purely structural adjustment compensates for reduction of $\delta_M$—rather $\psi$ increases globally to provide for a balance of forcing against dissipation. (Note that the *dimensional* streamfunction which, as noted earlier, scales as $(\delta_I^2/\delta_M^3)\beta L^3$, diverges along transects of both fixed $\delta_I$ *and* of fixed $\delta_M$).

Although we concern ourselves principally with a characterization of steady solutions of the BVE, a brief word is warranted about the time-dependent behavior of runs in the domain of the upper branch. Naturally, these large $R$ upper branch steady solutions may be linearly unstable and thus what Figure 1, for example, says about the expected *time mean* circulation is not clear. It *could* be the instabilities are such that the average of time-dependent solutions does not grow with increasing $R$ even though the steady solutions do but this appealing, perhaps intuitive, speculation appears to be wrong, at least for the forcing considered here. Experiments with a variety of initial conditions all converge rapidly to the upper branch solution suggesting that, with stress-free boundaries, the upper branch is *stable* to linear and even arbitrary nonlinear disturbances as $R \to \infty$. We conjecture that the upper branch is a "global attractor": no matter what the initial condition, all solutions end up there after sufficiently long time. A rigorous demonstration of a similar global attractor appears in Foias *et al.* (1983) which treats the two-dimensional Navier-Stokes equation in a doubly periodic domain with forcing chosen as the lowest eigenmode of the Helmholtz problem on the same domain, viz., $-\sin \pi x \sin \pi y$. Though the two problems are not identical, they are not far apart either. It is straightforward to adapt the arguments in Foias *et al.* to stress-free boundary conditions. On the upper branch, the influence of $\beta$ becomes arbitrarily small with

increasing $\psi$. Finally, our forcing is similar to theirs in having no internal nodes and being wholly convex. One expects a global attractor such as that demonstrated in Foias *et al.* to extend over a finite neighborhood of related problems. It is plausible that our problem falls within that neighborhood. While, as we have noted, the forcing chosen for this problem is far from generic owing to its proximity to the lowest eigenmode of the Helmholtz problem, we would risk misleading the reader if we failed to make the further observation that, intriguing though a global attractor may be, such a time-dependent response is exceptional even among the class of problems characterized by special forcing. Double gyre problems admit an important sinuous instability of the midlatitude eastward jet. Additionally, flows in bounded domains are typically rendered unstable by the wall-trapped shear instabilities characteristic of the Orr-Sommerfeld equation with no-slip boundary conditions. Periodic boundary conditions and slip boundary conditions have a number of attractive mathematical and numerical virtues but these should not be conflated with *physical* arguments about their merits or defects.

## 4. The recirculation gyre and the Fofonoff free inertial mode

The Fofonoff free inertial mode is a special solution of the unforced inviscid problem ($\nabla \times \tau = 0, \delta_M = 0$)

$$J(\psi, y + \delta_I^2 \nabla^2 \psi) = 0 \tag{4.1}$$

Eq. (4.1) will be satisfied provided the potential vorticity and streamfunction are related by the functional equation

$$y + \delta_I^2 \nabla^2 \psi = F[\psi] \tag{4.2}$$

where $F$ is an arbitrary function. Fofonoff (1954) chose a simple linear relation,

$$y + \delta_I^2 \nabla^2 \psi = k\psi + y_0. \tag{4.3}$$

While he considered both positive and negative slope, $k$, it is the positive slope which has come to be considered the standard solution. (As shorthand we shall, for the remainder of this paper, mean $k > 0$ when we refer to the "Fofonoff solution" without further qualification.) Appropriate to our single wind gyre case is the choice $y_0 = 0$. This solution consists of a uniform westward flow in the basin interior with velocities, $u = -1$, $v = 0$, and inertial boundary layers of width $\delta_I$ with velocities $O(1/\delta_I)$ (an accelerating western boundary layer, a current going east along the northern boundary, and a decelerating eastern boundary layer) closing the pattern of the circulation.

Comparing the Fofonoff solution to the solution of the full (forced and damped) problem one can see that the recirculation gyre formed in the northwest corner of the basin is also characterized by a strong correlation between $\psi$ and $q$, a regional

dynamical balance characterized to leading order by $J(\psi, q) = 0$. Figure 6 shows a typical dependence. (A more complete discussion, incorporating the higher order corrections required to lift the degeneracy at leading order, appears in Cessi, *et al.*, 1987.) To the degree that the tight $q$-$\psi$ relation associated with recirculation approximates a straight line, one might incline to identify the response as a Fofonoff solution, but the slope is of the wrong sign. (Also, the residual blurring hints at the persistent influence of viscosity in the dynamics of the recirculation.) In a historical sense, these distinctions are hairsplitting. Fofonoff certainly recognized the import of nonlinear dynamics so that it would not be amiss to term *any* choice of $F[\psi]$ a "Fofonoff solution" but, whatever the nomenclature, it *is* important to distinguish between solutions of positive and negative slope; the dynamics of the two are completely different. For the Fofonoff flow ($dq/d\psi > 0$) the relative vorticity is significant only within the inertial boundary layers, while in the recirculation gyre ($dq/d\psi < 0$) it is important everywhere; our recirculation gyre is thus characterized by high velocities of the fluid particles not only in the northern (eastward flowing) part but also in the southern (westward flowing) part.

The sign of $dq/d\psi$ is determined by the additional balance between the wind forcing and viscous damping. To explain this let us start with the highly nonlinear limit when we can neglect everything in (2.1) except for the balance between the input (averaged along each streamline) of negative vorticity by the wind and the lateral diffusion of vorticity due to viscosity. If the wind curl were exactly the lowest Helmholtz eigenfunction, $\phi_{11} = \sin \pi x \sin \pi y$, of the basin, then (inverting the Laplacian twice or, in the case of circular basin geometry, simply integrating in radius) the relative vorticity $\zeta$ and streamfunction $\psi$ would also be proportional to $\phi_{11}$; therefore, the $q[\psi]$ relation would be exactly a straight line with negative slope. In Figure 7, $q[\psi]$ scatterplots for several values of $R$ at $\delta_M = 0.06$ are superimposed. The limiting theoretical prediction is indicated by a dotted line. One sees that as $R \rightarrow \infty$ the solution approaches (3.2) with $q[\psi]$ an increasingly linear relation of negative slope. (Actually, a slight residual curvature of $q[\psi]$ persists because the wind curl is not *exactly* $\phi_{11}$ but makes an excess contribution of negative vorticity near the edge of the gyre.)

As $R$ decreases the recirculation gyre shrinks in size while its center retreats to the northwest corner of the basin. In absence of the beta-effect an anticyclonic eddy would, due to interaction with its image, travel clockwise along the basin boundary. Only in the northwest corner can this self-propagation be balanced by the westward drift induced by the beta-effect.

As the recirculation gyre gets smaller, the direct wind input of negative vorticity becomes less important in maintaining it. Indeed, the recirculation gyre forms even if the wind curl vanishes in the northwest corner. Instead, the driving agency is then indirect; it is transport of anomalous relative vorticity by the western boundary current as in Cessi *et al.* (1987) where a recirculation gyre is induced by prescribing

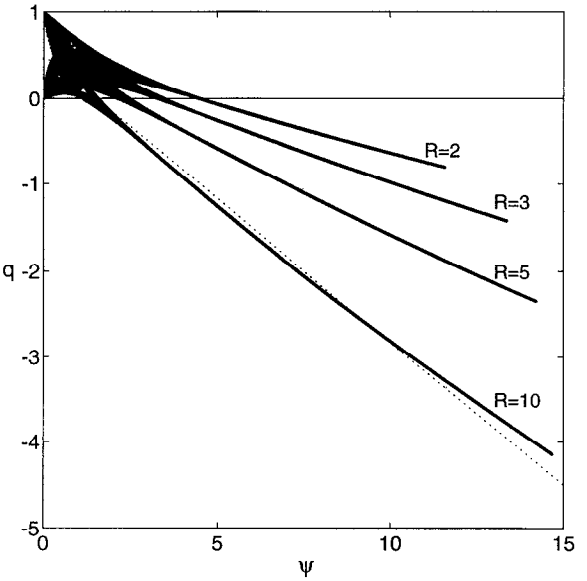


Figure 7. Superimposed scatterplots $q$ vs. $\psi$ for several values of $R$ at $\delta_M = 0.06$. The dotted straight line shows the limiting solution (3.2) for $R = 10$.

anomalously negative values of potential vorticity along the western and northern boundaries.

With indirect driving, to leading order, the gyre induced has a uniform distribution of potential vorticity $q = \bar{q}$ everywhere except in narrow boundary layers close to the edge of the gyre and hence the slope of the $q[\psi]$ relation is zero. This is a well-known result (Batchelor, 1956) about homogenization of vorticity inside a closed streamline in a rapidly rotating gyre. It is familiar in large scale circulation theory from the work of Rhines and Young (1982). The value of $\bar{q}$ is given by a so-called 'velocity weighted average' defined as

$$\bar{q} = \frac{\oint q_B \mathbf{u}_B \cdot d\mathbf{l}}{\oint \mathbf{u}_B \cdot d\mathbf{l}} \tag{4.4}$$

where the integration is performed along the boundary (subscript $B$) of the gyre. In our case the order of magnitude estimate of terms in (2.1) for $\delta_M = 0.06$ shows that the wind input is equally important as indirect (boundary) forcing in driving the recirculation gyre, which is also confirmed by the perceptibly negative, i.e., nonzero, slope of the $q[\psi]$ relation in Figure 6.

Impressed perhaps as much by the simplicity of the Fofonoff solution, as by any other consideration, authors sometimes rather loosely suggest that the linear solution and the Fofonoff inertial flow represent the low and high amplitude limits of the

wind-driven ocean circulation problem but this cannot hold for lateral friction, as one can easily demonstrate. If we start with the Fofonoff solution and perturb it by adding weak forcing and lateral friction then, by applying the integral vorticity constraint to a typical streamline of the Fofonoff solution

$$\delta_M^3 \int\int_\psi \nabla^2 q \, dxdy + \int\int_\psi \nabla \times \tau \, dxdy = 0, \tag{4.5}$$

we discover that the terms in (4.5) have the same sign and thus the integral vorticity constraint cannot be satisfied. The conclusion is, as Pedlosky is careful to emphasize (Ch. 5, Section 11), that adding lateral friction and wind forcing does more than slightly modify the Fofonoff flow, it completely changes the solution. Resonant excitation of the Fofonoff free mode is simply not possible in the case of lateral friction.

With bottom friction, $\delta_S \nabla^2 \psi$, it is indeed possible to treat small bottom friction and wind forcing as a perturbation of the Fofonoff flow as in Niiler (1966). By analogy with the lateral friction, one can introduce a parameter $R_b = \delta_I / \delta_S$ to characterize inertial and dissipative effects in a boundary layer. As $R_b$ varies from small to intermediate, $R_b \sim O(1)$, the recirculation grows in the northwest corner and quickly extends to the eastern boundary. During this stage the $q$ vs. $\psi$ scatterplot evolves from that typical of Sverdrup flow to the well defined linear $q[\psi]$ relation of positive slope, $k$, characteristic of the canonical Fofonoff solution. However, as $R_b$ increases further, basin velocities grow until eventually the inertial boundary-layer thickness $\delta_I \sqrt{\psi_{\max}}$ becomes comparable to the basin scale. The slope then changes sign, as indicated in Figure 8. In the limit as $R_b \to \infty$, the solution ultimately approaches $2/(\pi^3 \delta_S) \sin \pi x \sin \pi y$. Thus, with bottom friction, the classical Fofonoff solution is only an intermediate state in transition to a large amplitude solution in the form of a basin-filling gyre. (Again, vide Zimmerman (1993), the limit taken here is not the particular one that yields the Fofonoff solution, hence its appearance as a transitional feature.)

## 5. Veronis model

For comparison with the results presented earlier, it is interesting to parallel Veronis (1963), repeating here with appropriate modification his computation based upon a severely truncated Fourier series representation. (He used bottom friction $-\delta_S \nabla^2 \psi$ and a slightly different distribution of the wind stress curl, $\phi_{11}$.)

Veronis employed a streamfunction consisting of only two Fourier harmonics in each direction

$$\begin{aligned} \psi(x, y, t) &= a(t)\phi_{11} + b(t)\phi_{21} + c(t)\phi_{12} + d(t)\phi_{22} \\ \phi_{nm} &= \sin n\pi x \sin m\pi y \end{aligned} \tag{5.1}$$

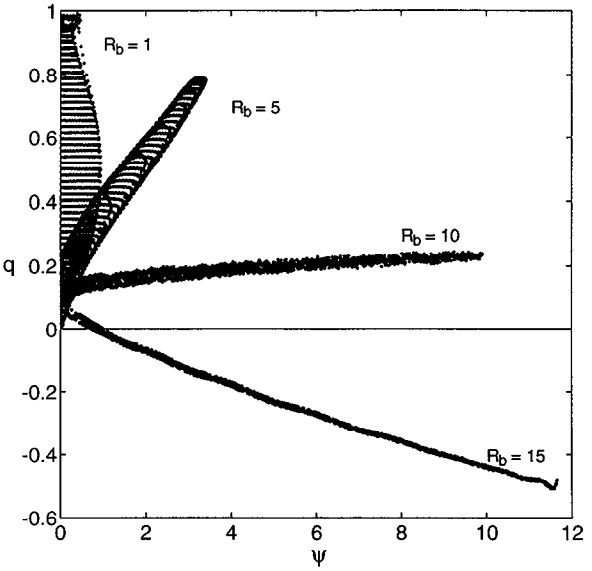


Figure 8. Superimposed scatterplots $q$ vs. $\psi$ for several values of $R_b$ at $\delta_S = 0.005$. In the bottom friction case the solution passes through the intermediate state, the Fofonoff solution, with positive slope, $k = dq/d\psi$.

and yet, when employed with lateral friction, this simple expansion gives qualitatively useful results, reproducing certain gross features of the full solution.

Substituting this expression in (2.1) and keeping only those harmonics gives a set of four algebraic equations for the Fourier amplitudes $(a, b, c, d)$

$$\begin{aligned}
2\pi^2\dot{a} &= -\frac{8}{3}b - 4\pi^4\delta_M^3 a + \frac{4}{\pi} \\
5\pi^2\dot{b} &= \frac{8}{3}a + \frac{9}{4}\pi^4\delta_I^2 ac - 25\pi^4\delta_M^3 b \\
5\pi^2\dot{c} &= -\frac{8}{3}d - \frac{9}{4}\pi^4\delta_I^2 ab - 25\pi^4\delta_M^3 c \\
8\pi^2\dot{d} &= \frac{8}{3}c - 64\pi^4\delta_M^3 d.
\end{aligned} \tag{5.2}$$

The first terms on the right-hand side come from the beta-effect; the terms with a $\delta_I^2$ factor come from the nonlinearity; lateral friction gives the terms with a $\delta_M^3$ factor; and $4/\pi$ is the projection of the wind forcing onto $\phi_{11}$.

Setting $\delta_I$, $\delta_M$, and the forcing to zero we obtain analogs of the two lowest basin modes of Rossby waves, with a combination of $\phi_{11}$, $\phi_{21}$ representing $\psi_{11}$ and a combination of $\phi_{21}$, $\phi_{22}$ representing $\psi_{21}$ (see (3.25.16), Pedlosky, 1987). Compare

also the frequencies

$$\sigma_{11} = \frac{8}{3}\frac{1}{\sqrt{2\cdot 5}}\frac{1}{\pi^2} = 0.085 \qquad \sigma_{21} = \frac{8}{3}\frac{1}{\sqrt{5\cdot 8}}\frac{1}{\pi^2} = 0.043 \tag{5.3}$$

with the exact values $\sigma_{11} = 1/(2\sqrt{2}\pi) \sim 0.1125$ and $\sigma_{21} = 1/(2\sqrt{5}\pi) \sim 0.0711$ for idealized basin modes.

Let us turn to steady solutions. By setting time derivatives to zero and eliminating $b$, $c$ and $d$ we obtain a cubic equation for $a$. In the linear case ($\delta_I = 0$, $\delta_M \ll 1$) the combination ($a \simeq 0$, $b = 3/2\pi$, $c = 0$, $d = 0$) represents the Munk solution (more precisely, only the Sverdrup interior solution since this representation does not resolve the western boundary layer). In the strongly nonlinear case ($\delta_I \gg \delta_M$, $\delta_M \ll 1$) the combination [$a = 1/(\pi^5\delta_M^3)$, $b \simeq 0, c \simeq 0, d \simeq 0$] represents the basin-filling gyre solution (3.2). (It is easy to show that the same value of $a$ holds for the general case with a large number of Fourier harmonics.) Note that it is *incorrect* to use the global vorticity constraint (4.5) applied to the closed streamline $\phi_{11} = 0$ in order to obtain the amplitude $a$ of the approximate solution $\psi \simeq a\phi_{11}$ (see also (3.2)). From this recipe, one obtains $a = 1/(8\pi^3\delta_M^3)$ which is close to the coefficient of the fundamental in (3.2) but nevertheless incorrect. A full accounting of successive terms from all the harmonics in (3.2) satisfies the balance given in (4.5) as

$$\frac{16}{\pi^3}\left[\left(1 + \frac{1}{9} + \frac{1}{25} + \cdots \frac{1}{(2k+1)^2} + \cdots\right) - \frac{\pi^2}{8}\right] = 0. \tag{5.4}$$

One shortcoming of the Veronis model *vis-à-vis* bottom friction is that, at the indicated level of truncation, the model cannot capture the Fofonoff solution. This means that without forcing and friction the only solution to (5.2) is the trivial one of no flow ($a, b, c, d = 0$). However, we know that there is an infinite family of steady solutions given by (4.2). It is the last equation in (5.2) (which is, by the way, the least accurate) that selects the trivial solution. If we were to disregard this equation, the combination ($a$, $b = 0$, $c = -32/(27\pi^4\delta_I^2)$, $d = 0$), with $a$ arbitrary, would then represent the Fofonoff solution with $c$ corresponding to $-k$ and $a$ related to $y_0$ (see (4.3)).

The most notable result is, as we have remarked, that a low order model correctly anticipates the existence of multiple solutions for the case of lateral friction. While one could trace out the boundaries of the predicted region in the ($\delta_M$, $\delta_I$)-plane of multiple solutions by tracking a zero discriminant of the cubic equation, it would be too much to expect even crude agreement at that level as the true solutions develop a richer structure than can be captured by so extreme a truncation as that considered here. At the cusp itself, however, the exact solution is not so convoluted. There also, we can simplify algebra a bit. We reduce the cubic equation to the canonical form,

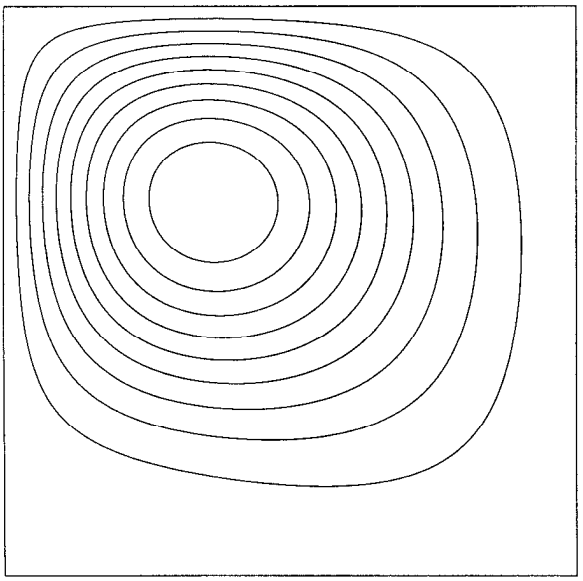

Figure 9. The solution at the cusp point according to the Veronis model. Compare this with the exact solution in Figure 6a. $CI = 0.1 \times \max \psi$.

$y^3 + py + q = 0$, and set $p, q = 0$ to obtain

$$\delta_{M_C}(\mathrm{V}) = \left(\frac{8}{3}\frac{1}{2 \cdot 5 \cdot \sqrt{8}\pi^4}\right)^{1/3} = 0.0989; \quad \delta_{I_C}(\mathrm{V}) = 0.1423 \tag{5.5}$$

and finally the Fourier amplitudes at the cusp itself:

$$\begin{aligned} a &= \frac{4}{\pi}\frac{1}{3 \cdot ((8/3)^2 + 4\pi^4\delta_{M_C}(\mathrm{V}))} = 1.1253 \\ b &= 0.31831; \quad c = -0.45015; \quad d = -0.19894. \end{aligned} \tag{5.6}$$

This solution is shown in Figure 9. It is a crude representation of the exact solution shown in Figure 6a but the likeness is good enough that the prediction that a cusp exists is correct. Moreover, since one knows that if a cusp exists there *will* be a fold catastrophe emanating from the cusp, the moral of Veronis (1963) is that one can often anticipate the effects of a variety of added dynamical effects using models of only modest resolution, which are entirely adequate in the vicinity of the cusp, before troubling to do more tedious computations at far smaller values of $\delta_M$ and $\delta_I$.

Finally, with bottom friction $-\delta_S \nabla^2 \psi$ only, i.e., the Stommel problem, the Veronis model predicts a cusp branching from

$$\delta_{S_C}(\mathrm{V}) = \left(\frac{8}{3}\frac{1}{\sqrt{2 \cdot 5} \cdot \sqrt{8}\pi^4}\right)^{1/3} = 0.0302. \tag{5.7}$$

If lateral friction is taken as indicative, then one anticipates the actual $\delta_{S_C}$ will be smaller. Paradoxically, Veronis (1966a,b) used $\delta_S = 0.05$ in numerical experiments on the full BVE, which would certainly lie above the cusp, and thus was exploring a regime ostensibly having a unique steady (stable) solution.

So far we have completed only an initial numerical exploration of the bottom friction case. We have traced the monotonic evolution of steady solutions from linear to highly nonlinear for $\delta_S$ as small as 0.008, i.e., there are no multiple solutions at least for $\delta_S$ larger than this value. Unfortunately, and paradoxically in view of the general thrust of this paper, in application to the case of bottom friction, we can only echo Veronis' earlier quoted remark—the truncated model is too crude. It fails for bottom friction because the transition from linear to nonlinear behavior proceeds via the Fofonoff solution which, as noted above, is missed in the truncated set.

One has to be cautious about extending the conclusion about uniqueness (and stability) to smaller values of $\delta_S$. The case of bottom friction case is technically demanding since, in an ideal fluid, singularities crop up like weeds. By contrast, with lateral friction singularities are exceptional. (This tendency toward regularity is used to analytic advantage by Barcilon *et al.* (1988) who consider the existence of weak solutions to the bottom friction case by adding small lateral friction and then passing to zero.)

As posed here, in a square domain, direct numerical solution is hampered at smaller values of $\delta_S$ owing to corner singularities (not previously discussed, insofar as we are aware). An intuitive explanation of these proceeds from the observation that, absent bottom friction, the standard Fofonoff solution is characterized by negative values of relative vorticity. Consider, for example, the northwest corner of the basin with its northward flowing western boundary current, which conserves potential vorticity. The apex of the corner is a stagnation point of the flow. It takes infinite time for a fluid particle to negotiate the corner. Adding even slight bottom friction proportional to the relative vorticity, $-\delta_S\zeta$, produces, relative to neighboring negative values, an anomalous zero value for the relative vorticity in the corner. For small values of $\delta_S$ the resulting sharp gradient in relative vorticity may trigger instability since, in contrast with the lateral friction, small-scale wavenumbers are not efficiently damped. This particular mathematical delicacy can be remedied numerically by moving to a smooth basin, free of corners.

It would appear one should be able to parallel the argument given in Section 3, using breakdown of the boundary-layer problem for bottom friction to infer the existence of a cusp in the corresponding PDE. Possibly the breakdown has instead to do with the formation of corner singularities but an alternate explanation is that a cusp in the case of bottom friction might be attained only for more extreme basin aspect ratio, a suggestion perhaps best explored along the lines of Ierley (1987) with a regional model in a sequence of experiments of progressively increasing aspect ratio.

On the other hand, there is reason beyond our exploratory numerical results for the full two-dimensional problem and the known limitations of the truncated model to believe that bottom friction alone does *not* lead to multiple solutions. Several years ago Victor Barcilon (private communication) outlined an analytic demonstration suggesting that Veronis (1966a,b) is exceptional in its use of a wind stress curl that happens also to be the lowest eigenmode of the Helmholtz equation for the same basin. Based on a large Rossby number expansion, carried out as far as a solvability condition at second order, Barcilon's results suggest the solution for bottom friction with $\nabla \times \tau = \sin \pi x \sin \pi y$ is unique.[9] While our wind differs slightly, it requires only a modest amount of good will to imagine the same proof might be extended to include our forcing.

## 6. Discussion

In recapitulation, we have approached the solution of the BVE with three themes in mind:

The first thread is the abstract perspective of functional analysis and the theory of dynamical systems, as in the earlier cited work of Foias *et al.* (1983) illustrating particular circumstances in which a solution of the two-dimensional Navier-Stokes equation is a global attractor. While our model problem differs in some details, it is not implausible that, as our numerical evidence suggests, the BVE with slip boundary conditions has the same behavior. Though most models will exhibit a richer temporal evolution, the pertinence of fixed points and attractors persists, a mode of thought and a language that has not yet made deep inroads in the oceanographic literature.

A second train of thought, bearing on spectral decay and the influence of the dissipation range upon the equilibrated large-scale circulation, dates back to Kraichnan (1967), who pointed out that two-dimensional (Navier-Stokes) turbulence differs radically from behavior we have come to accept as "natural" in three-dimensional flows: in two dimensions enstrophy cascades to smaller scales while energy flows to larger scales. In modifying the scenario adduced by Kraichnan, Rhines (1975) pointed out that the inverse energy cascade on a beta-plane is arrested at the scale $\sqrt{U/\beta}$. While evaluation of this for typical midlatitude Sverdrup interior flows gives a modest scale of, say, ten kilometers, one needs to be cautious. The point here is not whether the observed circulation conforms to theoretical speculation (or, to be scientific about it, *vice versa*) but, rather, how specific models behave in particular limits. For the latter, it is not fair to prejudge the answer by assuming $U$. The model determines that for us—acceptably or not.

The final theme is the accumulated collection of theoretical ideas on strongly nonlinear gyres. The earliest of these, of course, is Fofonoff (1954). The inviscid, unforced idealization was pressed further in the equilibrium statistical mechanics

9. Note that the case of $\nabla \times \tau = 1$ also has a special standing in regard to the posited expansion.

introduced by Salmon *et al.* (1976), which ascribes particular significance to Fofonoff solutions (of either sign of $dq/d\psi$) as realizing a condition of maximum entropy. A particularly incisive paper by Carnevale and Fredericksen (1987) elaborated upon this by showing that, in the limit of a zero wavelength cutoff, the statistical equilibrium collapses to a single realization, a Fofonoff-like solution (though possibly one with a nonlinear $q[\psi]$ relation) which is nonlinearly stable in the sense of Arnol'd. (For the BVE, however, one cannot appeal to any simple selection principle based on equilibrium statistical mechanics to explain the incidence of multiple steady solutions.) Griffa and Salmon (1989) tested the connection between the ideal inviscid solutions predicted by application of statistical mechanics and the solutions actually realized in an interesting sequence of runs with bottom friction. Their central finding is that the tendency toward a Fofonoff flow is crucially dependent upon the geometry of the wind forcing. Flows driven by a wind tending to *cancel* the bottom-drag torque around every closed streamline of the Fofonoff solution (positive $dq/d\psi$) are "energetic, Fofonoff-like, and nearly steady." Not so flows when the wind is in the *same* sense as the bottom-drag torque. Then the solution is turbulent, with a small mean. (It is impossible, of course, to conclude anything about possible differences in the underlying steady state (unstable) solutions that may exist in each of these cases.) In spite of the long observed occurrence of recirculation gyres in models with lateral friction, not until Marshall and Nurser (1986) and Cessi *et al.* (1987) did a theoretical framework develop comparable to that for bottom friction. (Curiously enough, for hard analysis, such as existence and uniqueness of solutions, as noted earlier, the bottom friction case is technically much more the difficult of the two.)

One unsatisfactory aspect of these works on recirculation and the large-scale circulation is the degree to which they partake of answer analysis, i.e., they are constrained either explicitly or implicitly to a range of parameters suggested by observation with particular parameters standing proxy for absent dynamical processes particularly, as has been emphasized here, eddy viscosity (though one might with equal justice focus upon the considerable gaps in our understanding of transfer of momentum across the air-sea interface). The point is that some of these processes may, through their persistent secular influence, play a controlling rôle in determining the grossest order mean dynamical balance at the large scale, e.g., a linear Sverdrup balance of the interior. This is more explicit in the statistical mechanics of Salmon *et al.* which requires that global energy and enstropy be given a *priori.* It is natural to select values considered "reasonable" but while maximum entropy solutions are hypothetical states toward which it is suggested nonlinear terms are tending to drive the flow (much as we have proposed the fixed points can "steer" the time-dependent behavior) models determine for themselves what dynamical regime to inhabit.

Much as in Griffa and Salmon (1989) where that selection turns importantly upon the pattern of wind forcing in relation to integral constraints over closed streamlines,

so we expect that the location of cusps, the multiplicity of folds and other features we assert are characteristic of the BVE will vary. Understanding how *these* elements in turn are affected by forcing, boundary conditions, stratification, etc., will help to clarify the physical origin and selection of particular dynamical regimes. Further progress along these lines requires that one take a serious approach to the mathematics—sensitivity studies, papers on numerical resolution, and the like, generally have other objectives.

The main point of this paper is not merely the existence of multiple solutions but their disparate physical character. One is not surprised that a differential equation observes one scaling in a viscously dominated regime and a second, asymptotic, scaling in the strongly nonlinear limit. The surprise is that, for the BVE, the one exists cheek by jowl with the other, and at a boundary-layer Reynolds number of order one. In that regard, the bottom friction case is the more intuitive as the transition from linear to nonlinear is gradual, rather than abrupt. This distinction highlights the influence of the form of dissipative parameterization. A second thesis advanced here is that even in the time-dependent case, the large amplitude upper branch solution may exert a significant dynamical influence.

One cannot fail from this to appreciate the central rôle of dissipation in setting the grossest dynamical balance of models of the circulation. Indeed we expect such dependence is characteristic of a wide array of present two-dimensional and quasi-two-dimensional ocean circulation models—roughly the oceanographic equivalent of the earlier noted laminar $\nu^{-1}$ (or turbulent $\log(1/\nu)$) dependence of channel flow arising in this case, however, not as a reflection of the correct dependence upon molecular viscosity, but simply as an undesirable artifact of defective models of eddy viscosity. With respect to the gedanken experiment of the diffusion coefficient tending to molecular values, certainly we expect that there are physical processes that do intervene to slow, if not entirely to arrest, the tendency toward a basin-filling pattern of recirculation. We do not, however, hold to a Micawberish optimism that some rich source of dissipative small-scale disorder will simply turn up. The issue *is which* physical processes are critical. Brown and Owens (1981) suggest from observed wave stresses that coupling of mesoscale flow to the internal gravity wave field can provide an equivalent viscosity of order $10^6\ \mathrm{cm^2\ s^{-1}}$. Given the enormous disparity throughout the ocean interior of mean kinetic energy density in the internal wave field compared to the far smaller kinetic energy of the large-scale circulation, one wouldn't be surprised that the former might constrain the amplitude of the latter. In some respects, our focus upon the ideal single gyre overestimates the need for a large eddy viscosity: we preclude, for example, the possibly significant exchange of vorticity between adjacent gyres moreover, when the bottom is flat, an important torque is missing from the right-hand side of the vorticity equation.

Although this paper has considered only steady solutions, thus making any direct assertion about turbulent models tenuous at best, it is motivated by some earlier

direct simulations that suggested similar conclusions. Several years ago, one of us (GI) engaged in a collaborative program of high resolution, low viscosity, computations done on the NCAR Cray computer. A disturbing aspect of those runs was that reduction of lateral friction to even 10 $m^2\ s^{-1}$ (compared to a commonly used GCM value of 1000 $m^2\ s^{-1}$) gave no evidence of saturation at the large scale, that is, the emergence of a well defined structure and amplitude. If anything, the suggestion of runs for both two- and five-layer single gyre models with no-slip conditions at east and west was a trend toward a basin-filling time mean recirculation. Such a model evidently has no sensible oceanic application apart from its use with some large, fixed, value of the eddy viscosity. We view it as entirely likely that present primitive equation GCMs respond in the same way (assuming, of course, that wholly spurious behavior attributable simply to inadequate numerical resolution is eliminated).

While an extensive paper cataloging statistics of the earlier Cray runs might have been prepared, that did not seem a satisfactory synthesis for what we intended to be an elementary process model. However, we did not apprehend any simple analytic framework about which to organize a more insightful note, hence those numerical results have been presented only informally in several seminars (leavened, it must be admitted, by a flashy color videotape of the sort that today serves so readily as proxy for insight). This paper is, at long last, some fruit borne of that earlier effort.

*Acknowledgments.* We are grateful to Rick Salmon who generously provided much-needed support for VS (NSF grant OCE92-16412). We are equally pleased to thank Bill Young both for his subsequent (and continuing) years of graduate support (OCE93-01462) and for his considerable help in the early stages of this work, when he not only directed our attention to George Veronis' 1963 paper, but also encouraged us to track solutions to larger values of the inertial boundary-layer thickness to clarify the connection with solutions for vanishing beta. As the work was assembled in a manuscript, he offered a number of useful suggestions to improve the presentation as did Victor Barcilon, Joe Pedlosky, Rick Salmon and two anonymous referees. We thank one and all for their constructive contributions.

## APPENDIX

Solution for the steady form of (2.1) cannot be found by time stepping at Reynolds numbers past the onset of the first instability. Instead we tackle the two-dimensional nonlinear PDE directly. We have found that an expansion in Chebyshev polynomials efficiently resolves the characteristically sharp boundary-layered structure of the solution. (By efficiency, we mean that convergence as a function of increasing resolution is exponential in the number of degrees of freedom, rather than algebraic, as in the case of finite difference and finite element methods.)

Substitution of this basis set leads to a (large) coupled set of nonlinear algebraic equations. Newton's method is used to determine the solution(s) of each set. Application of Newton's method requires computation of a matrix, $A$, (termed the

"Hessian") and subsequent solution of the standard linear algebra problem:

$$Ax = b. \tag{A1}$$

($A$ simply expresses the first variation of the residual with respect to changes in each of the $N$ coefficients of the solution vector. Although the original expansion form of the solution finds natural expression as a *matrix* of coefficients, if they are regarded simply as algebraic unknowns, it is preferable to suppose the coefficients are rearranged in the form of a single column vector. The distinction is merely one of indexing.)

Numerical experience with application of a spectral version of Newton's method to this problem (and a variety of related two-dimensional advection-diffusion problems) suggests several general observations which we include here since they are not generally noted elsewhere in the literature:

1) Provided the solution is adequately resolved (a judgement which can be made from *ex post facto* examination of spectral decay), standard application of Newton's method is normally found to converge to machine accuracy in about five steps.

2) For certain boundary conditions (typically slip), even if the solution is inadequately resolved, convergence is equally rapid (to an underresolved, but typically qualitatively correct, result).

3) Convergence often fails when the structure of the solution to be found differs grossly from the initial guess, as may be the case when one finds the very first solution. Normally this can be circumvented by a suitable choice of parameters to render the problem nearly linear. Thereafter, solutions found with initial guesses based on converged solutions for nearby parametric values can quickly be produced. As the solution sought is one of increasing nonlinearity, the "basin of attraction" of Newton's method tends to shrink, with the consequence that parametric changes must be successively reduced to ensure convergence.

4) Computation divides naturally into two tasks: forming the matrix $A$; and solution of $Ax = b$. In those problems where it is feasible to compute the elements of $A$ analytically, as here, monitoring of CPU usage shows that about 95% of the total is spent on the second task, i.e., solving $Ax = b$. From the standpoint of efficiency, this is ideal since the latter problem is one for which most computer centers and many workstations are apt to have locally optimized subroutines available, e.g., those from IMSL or LINPACK. In our computations, we have made extensive use of the relevant IBM/ESSL library routines, which our experience suggests make optimal use of the RS/6000 architecture. $A$ is a full nonsymmetric real matrix and not diagonally dominant, so neither iterative nor non-pivoting structured programs, e.g., sparse or banded algorithms, are applicable. The full structure of $A$ is, at first glance, a severe drawback since memory limitations can easily become a constraint on the attainable range of solutions while finite difference and element methods enjoy the advantage of a sparse Hessian matrix with consequent reductions in memory and

CPU time. However, the exponential convergence of the spectral formulation means that neither timing nor memory usage can be compared with finite difference and finite element methods on the basis of an equal number of degrees of freedom.

5) Typical timing for the solution of the (nonlinearly) coupled set of the nearly 1400 algebraic equations resulting from a 41 × 41 Chebyshev expansion[10] is 65.6 seconds per step on an IBM RS/6000 Model 350. This compares very favorably with the 49.4 seconds per step (CPU time, not elapsed time) we experienced on the UCSD Cray-YMP using putatively optimized IMSL software. Timing scales approximately as the cube of any change in resolution (based on the Gaussian elimination operations count). For this problem, 1400 variables suffice to explore a wide range of parameter space though a few selected runs at a higher resolution of 51 × 51 have proved useful. Memory requirement is essentially the square of the number of variables. Commonly our usage has ranged from 8 Mb to 40 Mb. (Linear stability analysis requires twice this as the streamfunction formulation leads naturally to the generalized eigenvalue problem $Ax = \lambda Bx$.)

6) In contrast with textbook treatments (e.g., Numerical Recipes) which advocate the use of "quasi-Newton" methods, such as that due to Broyden based on a sequence of rank one updates, our experience (based on experiments with the available range of IMSL quasi-Newton algorithms) is that such methods are less efficient for this class of problems and where $A$ has the structural characteristics indicated. Simple application of Newton's method with elementary quadratic minimization in the reciprocal gradient direction is easily superior to any other method tried.

With reference to the occurrence of multiple solutions, a limitation of Newton's method applied to nonlinear equations is that, in order to ensure convergence, it is necessary to provide an initial guess that is sufficiently close to the exact solution. Generally this is done by starting from some known solution and then varying parameters of the problem. For example, one can start from the Munk solution and increase the Reynolds number $R$. Obviously one can be more or less fancy about this, moving a small distance at each step and either discarding previous information or else predicting the new solution based on first (or higher) differences. By such incremental means it is possible to track solutions all the way up to highly nonlinear gyre solutions, but only if $\delta_M > \delta_{M_C}$. When $\delta_M < \delta_{M_C}$, the solutions can be traced along the low branch only up to the (lower) nose point, $R_L$, that is, the turning point of the curve in Figure 3a ($\delta_M = 0.04, 0.02$). Similarly, if one starts from the basin-filling gyre solution, one can descend along the high branch only down as far as the high nose point, $R_H$. For $\delta_M < \delta_{M_C}$ the low and high branches overlap each other in the interval of the Reynolds numbers $R_H < R < R_L$. For example, for $\delta_M = 0.04$, $R_L = 1.3203$ and $R_H = 1.0377$; while for $\delta_M = 0.02$, $R_L = 1.0735$ and $R_H = 0.4206$.

10. The tau method is used to eliminate the last four variables in each direction.

The overlapping of the of the high and low branches suggests the existence of the middle branch but it requires a little care to obtain the solution on the middle branch, i.e., to get around the nose. We adopt an approach based upon the general spectral properties of the solution. Linearizing the BVE (2.1) about the steady solution $\psi_0$ or, in other words taking a functional derivative of (2.1) with respect to $\psi$, we obtain a spectral problem for the perturbations $\psi'(x, y) \cdot \exp(-i\sigma t)$

$$-i\sigma\nabla^2\psi' + \delta_I^2[J(\psi_0, \nabla^2\psi') + J(\psi', \nabla^2\psi_0)] + \psi_x' = \delta_M^3\nabla^4\psi' \qquad \text{(A2)}$$

with appropriate boundary conditions. Solving this spectral problem we obtain a set of eigenfrequencies $\sigma$ and corresponding eigenfunctions $\psi_\sigma$. Normally, the eigenfrequencies $\sigma$ are complex since, in the presence of viscosity, the problem is not self-adjoint. The imaginary part of the frequency gives the rate of growth (or decay) of the perturbation.

The significant eigenmode is that associated with the recirculation. It is a nonoscillatory eigenmode, i.e., its eigenvalue, $\sigma_{Rec}$, lies on the imaginary axis with $Re(\sigma_{Rec}) = 0$. The pattern of the associated eigenfunction, $\psi_{Rec}'$ closely resembles the recirculation gyre. The distinguishing feature of this eigenmode is that as the Reynolds number increases from zero to $R_L$ the eigenvalue drifts along a negative half of the imaginary axis and, at $R = R_L$ (corresponding to the low nose), vanishes identically.

According to the Fredholm theorem for linear operators, the existence of a zero eigenvalue means that the solution of the problem is not unique. The same conclusions can be drawn for a zero eigenvalue of the Fréchet derivative (A2) for our *non*linear problem.[11] Indeed, adding a small amount of the eigenfunction, $\psi_{Rec}'$, corresponding to the zero eigenvalue of the solution $\psi_{0L}$ on the low branch close to the low nose point, we can jump to the solution on the middle branch. The eigenfunction $\psi_{Rec}'$ at the low nose point for $\delta_M = 0.02$ is shown in Figure 4f. Essentially the same structure is found if one computes the difference of the middle and low branch solutions shown in Figure 4b and Figure 4a, respectively. In practice it is that difference, rather than a true neutral mode found by solving the associated eigenvalue problem, that is used to negotiate the fold. Once on the middle branch we can, by *decreasing* the Reynolds number, trace it up to the high nose point, where the middle branch connects with the high branch.

There is, incidentally, a large numerical literature devoted to automated means of following bifurcation branches through their contortions. In many problems there are physically motivated parameterizations with respect to which variation in the bifurcating quality is rendered single-valued. Failing that, any of various intrinsic parameterizations can be used. In this problem, our simple-minded approach is more

11. For a useful introduction to the Fréchet derivative see Finlayson, "The Method of Weighted Residuals and Variational Principles," Academic Press, 1972.

than adequate. Far and away the majority of time is spent plodding along the successive branches, not negotiating the turns between them.

In accord with the usual expectation of alternating stability of branches in bifurcation theory, on the low branch the imaginary part of the eigenvalue associated with the recirculation is negative, $I_m(\sigma_{Rec_L}) < 0$ for $R < R_L$; on the middle branch it is positive, $Im(\sigma_{Rec_I}) > 0$ for $R_H < R < R_L$; and on the high branch it becomes negative again, $Im(\sigma_{Rec_H}) < 0$ for $R_H < R$, i.e., $Im(\sigma_{Rec})$ changes sign in passing a nose point. In other words the middle branch is unstable.

## REFERENCES

Arnol'd, V. I. 1984. Catastrophe theory. Springer-Verlag, New York, 80 pp.

Barcilon, V., P. Constantin and E. S. Titi. 1988. Existence of solutions to the Stommel-Charney model of the Gulf Stream. SIAM J. Math. Anal., *19,* 1355–1364.

Batchelor, G. K. 1956. On steady laminar flow with closed streamlines at large Reynolds number. J. Fluid Mech., *1,* 177–190.

Briggs, W. L. 1980. A new class of steady solutions of the barotropic vorticity equation. Dyn. Atmos. Oceans, *4,* 67–99.

Brown, E. D. and W. B. Owens. 1981. Observations of the horizontal interactions between the internal wave field and the mesoscale flow. J. Phys. Oceanogr., *11,* 1474–1480.

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

Carnevale, G. and J. S. Frederiksen. 1987. Nonlinear stability and statistical mechanics of flow over topography. J. Fluid Mech., *175,* 157–181.

Cessi, P., 1991. Laminar separation of colliding western boundary currents. J. Mar. Res., *49,* 697–717.

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

Cessi, P., G. R. Ierley and W. R. Young. 1987. A model of the inertial recirculation driven by potential vorticity anomalies. J. Phys. Oceanogr., *17,* 1640–1650.

Cessi, P. and L. Thompson. 1990. Geometrical control of the inertial recirculation. J. Phys. Oceanogr., *20,* 1867–1875.

Charney, J. G. and J. G. DeVore. 1978. Multiple flow equilibria in the atmosphere and blocking. J. Atmos. Sci., *36,* 1205–1216.

Fofonoff, N. P. 1954. Steady flow in a frictionless homogeneous ocean. J. Mar. Res., *13,* 254–262.

Foias, C., O. P. Manley, R. Temam and Y. Treve. 1983. Asymptotic analysis of the Navier-Stokes equations. Physica D, *9,* 157–188.

Gilmore, R. 1981. Catastrophe Theory for Scientists and Engineers. J. Wiley, New York, 666 pp.

Griffa, A. and R. Salmon. 1989. Wind-driven ocean circulation and equilibrium statistical mechanics. J. Mar. Res., *47,* 457–492.

Hidaka, K. 1949. Mass transport in ocean currents and lateral mixing. J. Mar. Res., *7,* 132–136.

Ierley, G. R. 1987. On the onset of inertial recirculation in barotropic general circulation models. J. Phys. Oceanogr., *17,* 2366–2374.

Kamenkovich, V. M., 1966. A contribution to the theory of the inertial-viscous boundary layer in a two-dimensional model of ocean currents. Izvestiya, Atmospheric and Oceanic Physics, Vol 2, No 12, 1274–1295.

—— 1977. Fundamentals of Ocean Dynamics. Elsevier Sci. Publ. Co., Amsterdam, 249 pp.
Kraichnan, R. H. 1967. Inertial ranges in two-dimensional turbulence. Phys. of Fluids, Suppl. II, 233–239.
Lamb, H. 1932. Hydrodynamics. 6th ed. Cambridge University Press, Cambridge.
Mallier, R. 1994. On the parametric model of western boundary outflow. Studies Appl. Math., *91,* 17–25.
Marshall, D. 1993. Resonance of Fofonoff's mode in a rotated basin. J. Phys. Oceanogr., *23,* 970–978.
Marshall, J. C. and A. J. G. Nurser. 1986. Steady free circulation in a stratified, quasi-geostrophic ocean. J. Phys. Oceanogr., *16,* 1799–1813.
Munk, W. H. 1950. On the wind-driven ocean circulation. J. Meteorol. *7,* 79–93.
Niiler, P. 1966. On the theory of the wind-driven ocean circulation. Deep-Sea Res., *13,* 597–606.
Pedlosky, J. 1987. Geophysical Fluid Dynamics. (second edition) Springer-Verlag, 710 pp.
Rhines, P. B. 1975. Waves and turbulence on a beta-plane. J. Fluid Mech., *69,* 417–443.
Rhines, P. B. and W. R. Young. 1982. Homogenization of potential vorticity in planetary gyres. J. Fluid Mech., *122,* 289–325.
Sakai, S. 1986. Vortex flow regime in an inflow-outflow model at midlatitude. Deep-Sea Res., *33,* 1107–1125.
Salmon, R., G. Holloway and M. C. Hendershott. 1976. The equilibrium statistical mechanics of simple geostrophic models. J. Fluid Mech., *75,* 691–703.
Stommel, H. 1948. The westward intensification of wind-driven ocean currents. Trans. Am. Geophysical Union, *29,* 202–206.
Sverdrup, H. 1947. Wind-driven currents in a baroclinic ocean: with application to the equatorial currents of the eastern Pacific. Proc. Natl. Acad. Sci. USA, *33,* 318–326.
Temam, R. 1991. Approximation of attractors, large eddy simulations and multiscale methods. Proc. Roy. Soc. A, *434,* 23–39.
Veronis, G. 1963. An analysis of wind-driven ocean circulation with a limited number of Fourier components. J. Atmos. Sci., *20,* 577–593.
—— 1966a. Wind-driven ocean circulation—Part 1. Linear theory and perturbation analysis. Deep-Sea Res., *13,* 17–29.
—— 1966b. Wind-driven ocean circulation—Part 2. Numerical solution of the nonlinear problem. Deep-Sea Res., *13,* 30–55.
Welander, P. 1964. Note on the role of boundary friction in the wind-driven ocean circulation. Tellus *16,* 408–410.
Whitney, H. 1955. Mapping of the plane into the plane. Ann Math., *62,* 374–410.
Zimmerman, J. T. F. 1993. A simple model for the symmetry properties of nonlinear wind driven ocean circulation. Geophys. Astrophys. Fluid Dyn., *71,* 1–15.

Received: *30 June, 1993;* revised: *30 May, 1995*