Linear Baroclinic Instability in Extended Regime Geostrophic Models

K. SHAFER SMITH

Department of Physics, University of California, Santa Cruz, Santa Cruz, California

GEOFFREY K. VALLIS*

Department of Ocean Science, University of California, Santa Cruz, Santa Cruz, California

(Manuscript received 16 September 1997, in final form 29 May 1998)

ABSTRACT

The linear wave and baroclinic instability properties of various geostrophic models valid when the Rossby number is small are investigated. The models are the ''L_1'' dynamics, the ''geostrophic potential vorticity'' equations, and the more familiar quasigeostrophic and planetary geostrophic equations. Multilayer shallow water equations are used as a control. The goal is to determine whether these models accurately portray linear baroclinic instability properties in various geophysically relevant parameter regimes, in a highly idealized and limited set of cases. The L_1 and geostrophic potential vorticity models are properly balanced (devoid of inertio-gravity waves, except possibly at solid boundaries), valid on the β plane, and contain both quasigeostrophy and planetary geostrophy as limits in different parameter regimes; hence, they are appropriate models for phenomena that span the deformation and planetary scales of motion. The L_1 model also includes the ''frontal geostrophic'' equations as a third limit. In fact, the choice to investigate such relatively unfamiliar models is motivated precisely by their applicability to multiple scales of motion.

The models are cast in multilayer form, and the dispersion properties and eigenfunctions of wave modes and baroclinic instabilities produced are found numerically. It is found that both the L_1 and geostrophic potential vorticity models have sensible linear stability properties with no artifactual instabilities or divergences. Their growth rates are very close to those of the shallow water equations in both quasigeostrophic *and* planetary geostrophic parameter regimes. The growth rate of baroclinic instability in the planetary geostrophic equations is shown to be generally less than the growth rate of the other models near the deformation radius. The growth rate of the planetary geostrophic equations diverges at high wavenumbers, but it is shown how this is ameliorated by the presence of the relative vorticity term in the geostrophic potential vorticity equations.

1. Introduction

The large-scale circulation of the mid- and high-latitude atmosphere and ocean is characterized by a small Rossby number and velocities close to geostrophic balance. Although it is true that the primitive equations, which do not explicitly employ such a balance, are more commonly used for forecasting and climate studies, much of our conceptual understanding of the circulation has been attained by exploiting the simplifications that can then be made in the equations of motion.

The two classic simplified sets of equations that have been most commonly used for theoretical and conceptual studies are the quasigeostrophic (QG) and the planetary geostrophic (PG) equations. Both are valid for low Rossby number flow. The former requires scales near the deformation radius and simultaneously much smaller than the planetary scale, while the latter is valid for scales that are large compared to the deformation radius and on the order of the planetary scale. Typically, for the atmosphere, the QG equations are valid for scales of order one to a few thousand kilometers, and the PG equations are valid for nearly global scales (excluding the equatorial region where the Rossby number may not be small). In the ocean the QG equations are valid for scales of order tens to hundreds of kilometers, and the PG equations again are valid for much larger scales. Furthermore, the large separation in spatial scale between the deformation radius and the radius of the planet in the ocean yields an additional parameter regime, namely, the so-called frontal geostrophic (FG) regime. In this regime large variations in the height field (or stratification) are allowed, but the Coriolis parameter is not allowed to vary significantly. A balanced set of equations valid in this regime was asymptotically derived by Cushman-Roisin (1986). Whether this regime

* Current affiliation: Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey.

Corresponding author address: K. Shafer Smith, Department of Physics, University of California, Santa Cruz, Santa Cruz, CA 95064. E-mail: shafer@physics.ucsc.edu

exists or is important in the atmosphere is less likely, due to the lack of a significant scale separation between the deformation radius and the planetary radius.

Many of the important circulation patterns in the atmosphere or ocean span these parameter regimes. For example, although baroclinic instability may preferentially occur near the deformation scale, there may be a significant instability at larger scales [e.g., the Green modes, found by Green (1960)], which might more properly be described with a model that is valid in the PG regime. In any case, the nonlinear interactions of eddies at the deformation scale (leading to a cascade of energy to larger scales), and eddy–mean flow interaction certainly span the parameter regime from deformation scale to planetary scale, although the flow is in near geostrophic balance at all scales. An ideal model for conceptual studies of the circulation would contain both the QG and PG (and possibly FG) regimes, while exploiting the smallness of the Rossby number. While formal accuracy with respect to the primitive equations (or shallow water equations, in an idealized setting) should be roughly maintained over the parameter range of interest (as a function of the small parameter exploited in the approximation), it is (we believe) more important that it be valid over a broad parameter regime than that the model have high-order accuracy with respect to that small parameter.

Two ''geostrophic'' models (by geostrophic model we mean merely that it is based on the smallness of the Rossby number) have been proposed that (we explicitly show) do in fact span both QG and PG regimes. These are the L_1 model (Salmon 1983) and the simpler geostrophic potential vorticity (GPV) equations (Vallis 1996; see also Bleck 1973). That is, both models include both the QG and the PG equations in the appropriate limit in parameter space. Each model is thus valid for $O(1)$ variations in the layer thickness (provided the variations occur on a sufficiently large horizontal scale) and the Coriolis parameter, and neither model neglects the contribution of relative vorticity in the advection of potential vorticity. Both models, in fact, conserve the same form of potential vorticity and, additionally, the L_1 model conserves global energy and local mass (though at the price of a more complex solution algorithm). We will also show (in appendix C) that the L_1 model contains the FG equations as a third limit in parameter space.

In this paper, as one step in exploring whether these models might indeed be useful tools, we explore their linear wave and stability properties. As a control we compute the linear properties of the shallow water (SW) equations, from which the balanced models are derived, and those of the familiar QG and PG equations. All of the models are cast in multilayer form on a differentially rotating zonal channel with variable Coriolis parameter.

In section 2 we describe the models studied and develop their linearized representations in each configuration. In section 3 we investigate the dispersion, eigenfunctions, and instability properties of the models at hand. A discussion of the results and conclusions are presented in section 4. In addition, the two-layer L_1 equations are derived from Hamilton's principle in appendix A and the growth rate for three-layer linear PG is derived analytically in appendix B. Finally, it is shown explicitly in appendix C that L_1 contains FG as a limit.

2. Model formulations

We begin with a presentation of the SW equations, from which the other models are derived. This set is nondimensionalized, cast in a two-layer setting, and linearized about a locally geostrophic, vertically sheared zonal velocity, following, for example, Kuo (1978). The approximate models are presented in two-layer (and, for the GPV equations, multilayer), nondimensionalized form and are linearized about a basic-state equivalent to that used for the SW system. Single-layer forms are described, where possible, as limits of the multiple-layer forms. Furthermore, QG and PG are shown to be formal limits of GPV and L_1.

For a single layer, the SW equations are given by

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u} + f\hat{\mathbf{z}}\times\mathbf{u} + g\nabla h = 0, \qquad (2.1)$$

$$\frac{\partial h}{\partial t} + \nabla\cdot(\mathbf{u}h) = 0, \qquad (2.2)$$

where $f = f_0 + \beta_0 y$ is the Coriolis parameter for which the β-plane approximation has been employed, ∇ is the horizontal gradient operator, $\mathbf{u} = (u, v)$ is the two-dimensional horizontal fluid velocity, and h is the fluid layer thickness.

The equations are written in nondimensional form as

$$\epsilon\frac{\partial \mathbf{u}}{\partial t} + \epsilon\mathbf{u}\cdot\nabla\mathbf{u} + f\hat{\mathbf{z}}\times\mathbf{u} + \nabla\eta = 0, \qquad (2.3)$$

$$\epsilon F\frac{\partial \eta}{\partial t} + \nabla\cdot(\mathbf{u}h) = 0, \qquad (2.4)$$

where η is the surface height variation, written in terms of the nondimensional height as

$$h = 1 + \epsilon F\eta. \qquad (2.5)$$

The Coriolis parameter is now

$$f = 1 + \epsilon\beta y, \qquad (2.6)$$

and the nondimensional numbers are the Rossby number (ϵ), the squared Froude number (F), and the ''β'' parameter, defined, respectively, as

$$\epsilon \equiv \frac{U}{f_0 L}, \qquad F \equiv \frac{f_0^2 L^2}{g' H_0} = \left(\frac{L}{L_d}\right)^2, \qquad \beta \equiv \frac{\beta_0 L^2}{U}. \qquad (2.7)$$

Here U and L are velocity and length scales for the problem at hand, H_0 is the mean layer thickness (here the same for each layer in the two-layer case), and g' is the reduced

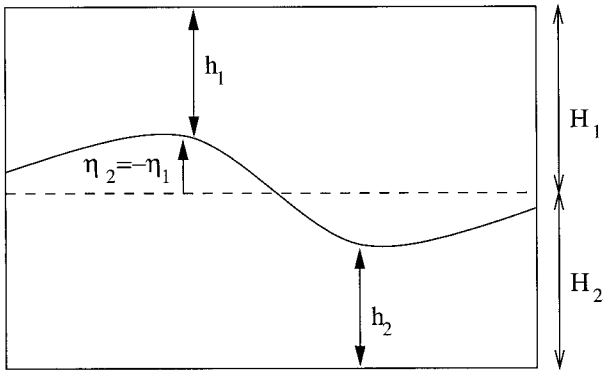

FIG. 1. Model domain for a two-layer system under a rigid lid.

gravity, equal to g in the single-layer case and to $g(\Delta\rho/\rho_0)$ in the multiple-layer case (which will follow). For example, for two layers, with ρ_1 the density of the upper layer, ρ_2 the density of the lower layer, and ρ_0 as their mean, $\Delta\rho = \rho_2 - \rho_1$. Finally, L_d is the radius of deformation, equal to $\sqrt{g'H_0}/f_0$.

a. Two-layer shallow water equations

In multiple density layer formulations, the thickness of the nth layer is written

$$h_n = H_n + \epsilon F\eta_n, \tag{2.8}$$

where η_n is the displacement of the nth interface and H_n is the nondimensional mean thickness (scaled by H_0) of the nth layer (see Fig. 1). Layer-wise quantities will be denoted with the subscript n, where n increases with depth of the layer. A rigid lid and flat bottom are imposed, and model-specific equations of motion are written separately for each layer in terms of two-dimensional velocity and pressure. Coupling of the motion in each layer occurs via the relation of the interface displacement to the pressure difference across the interface, which follows from the requirement of pressure continuity across the interface.

For a system with two layers, the interface displacement is

$$\eta_n = \Delta_n(p_n), \qquad n = 1, 2, \tag{2.9}$$

where Δ_n is the layer difference operator, defined in a two-layer setting for any layer-wise argument, ϕ_n, as

$$\Delta_n(\phi_n) \equiv (-1)^n(\phi_2 - \phi_1), \qquad n = 1, 2,$$

where the subscript on the *argument* is retained merely as a reminder that the operator works only on layer-wise quantities. Thus, in this two-layer case, $\eta_1 = p_1 - p_2 = -\eta_2$.

The two-layer nondimensional SW equations are then

$$\epsilon\frac{\partial \mathbf{u}_n}{\partial t} + \epsilon\mathbf{u}_n\cdot\nabla\mathbf{u}_n + f\hat{\mathbf{z}}\times\mathbf{u}_n + \nabla p_n = 0, \tag{2.10}$$

$$\epsilon F\frac{\partial \eta_n}{\partial t} + \nabla\cdot(h_n\mathbf{u}) = 0. \tag{2.11}$$

We linearize the motion about a symmetric vertical shear in the zonal velocity, which we assume to be in local geostrophic balance with a basic-state pressure. In terms of SW variables, the basic state is

$$u_n(x, y, t) = U_n + u_n'(x, y, t), \tag{2.12}$$

$$v_n(x, y, t) = v_n'(x, y, t), \tag{2.13}$$

$$p_n(x, y, t) = P_n(y) + p_n'(x, y, t), \tag{2.14}$$

where for two layers, $U_n = (-1)^{n+1}U_0$. Geostrophic balance then requires

$$U_n = -\frac{1}{1 + \epsilon\beta y}\frac{\partial P_n(y)}{\partial y}, \tag{2.15}$$

which upon integration yields

$$P_n(y) = -U_nG, \tag{2.16}$$

$$G = y + \frac{\epsilon\beta}{2}y^2. \tag{2.17}$$

Finally, we seek wave solutions of the resulting linearized equations of the form

$$\phi(x, y, t) = \tilde{\phi}(y)e^{ik(x-ct)}, \tag{2.18}$$

where ϕ is an arbitrary dependent variable. Under this substitution, the equations transform to coupled, linear, variable coefficient ODEs in which the dependent variables are functions of y only (denoted with a $\sim$). These are then finite differenced and solved as algebraic generalized matrix eigenvector problems. The growth rate is given by kc_i, where c_i is the imaginary part of the wave speed.

For the SW system, the linearized wave equations are

$$i\epsilon(U_n - c)\tilde{u}_n - \frac{f}{k}\tilde{v}_n + i\tilde{p}_n = 0, \tag{2.19}$$

$$i\epsilon(U_n - c)\tilde{v}_n + \frac{f}{k}\tilde{u}_n + \frac{1}{k}\frac{\partial\tilde{p}_n}{\partial y} = 0, \tag{2.20}$$

$$i\epsilon F(U_n - c)\Delta_n(\tilde{p}_n) + i\epsilon\tilde{u}_n + \frac{1}{k}\frac{\partial}{\partial y}(\overline{h}_n\tilde{v}_n) = 0, \tag{2.21}$$

where $\overline{h}_n$ is the mean layer thickness of the basic-state solution,

$$\overline{h}_n = 1 + (-1)^n\epsilon FU_sG, \tag{2.22}$$

and where U_s is the shear velocity, given for the two-layer case by

$$U_s = U_1 - U_2. \tag{2.23}$$

The boundary conditions for the SW example are

$$\tilde{v}_n(y = -½, ½) = 0. \tag{2.24}$$

b. Geostrophic potential vorticity equations

The GPV equations (Vallis 1996; see also Bleck 1973) are derived by first assuming that geostrophic

potential vorticity on fluid parcels is conserved layerwise, and that all velocities are determined by local geostrophic balance. Hence,

$$\left(\frac{\partial}{\partial t} + \mathbf{u}_n \cdot \boldsymbol{\nabla}\right) q_n = 0, \tag{2.25}$$

where

$$\mathbf{u}_n = \frac{1}{f}\hat{\mathbf{z}} \times \boldsymbol{\nabla} p_n, \tag{2.26}$$

$$q_n = \frac{f + \epsilon\zeta_n}{h_n}, \tag{2.27}$$

$$\zeta_n = \hat{\mathbf{z}} \cdot \boldsymbol{\nabla} \times \mathbf{u}_n = \frac{1}{f}\left(\nabla^2 p_n - \frac{\epsilon\beta}{f}\frac{\partial p_n}{\partial y}\right), \tag{2.28}$$

and where p_n is related to h_n via (2.8)–(2.9). The potential vorticity conservation statement [(2.25)] is identical to the equivalent SW statement, except for the approximation to the velocity field. Because the velocity field is now determined from the height field alone, one obtains two coupled prognostic relations for p_n (one for each layer), the solutions of which are used to diagnose $\mathbf{u}_n$. For the single-layer representation, remove the subscripts and replace p with η.

1) Parameter space limits

Two consistent limits may be taken from these equations. First, we assume $\beta \sim F \sim O(1)$ and take the limit of (2.25)–(2.28) as $\epsilon \to 0$; this is equivalent to the assumption of small variations in the height field and Coriolis parameter. In this case, we obtain QG directly, given by

$$\mathbf{u}_n = \hat{\mathbf{z}} \times \boldsymbol{\nabla} p_n, \tag{2.29}$$

$$q_n = \nabla^2 p_n - F\Delta_n(p_n) + \beta y. \tag{2.30}$$

If instead we assume $\beta \sim F \sim O(1/\epsilon)$ and take the limit of (2.25)–(2.28) as $\epsilon \to 0$—equivalent to the assumption of $O(1)$ variations in the height field and Coriolis parameter—we obtain PG, written as

$$\mathbf{u}_n = \frac{1}{f}\hat{\mathbf{z}} \times \nabla p_n, \tag{2.31}$$

$$q_n = \frac{f}{h_n}. \tag{2.32}$$

Thus, GPV contains as limits the two approximations most commonly employed in geophysical fluid dynamics: the quasigeostrophic and planetary geostrophic equations.

2) Linear wave equation

We linearize the GPV equations about the basic state [(2.12)] to get

$$\frac{D_n^0}{Dt}[\zeta_n \overline{h}_n - fF\Delta_n(p_n)] + \left(\frac{\beta\overline{h}_n}{f} + U_s F f\right)\frac{\partial p_n}{\partial x} = 0, \tag{2.33}$$

where the linear advective derivative operator is defined as

$$\frac{D_n^0}{Dt} = \frac{\partial}{\partial t} + U_n\frac{\partial}{\partial x},$$

and where $\overline{h}_n$ and U_s are defined by (2.22) and (2.23), respectively. We then seek a wave solution of the form [(2.18)], and the resulting eigenvalue problem is

$$(U_n - c)[\tilde{\zeta}_n \overline{h}_n - fF\Delta_n(\tilde{p}_n)] + \left(\frac{\beta\overline{h}_n}{f} + U_s F f\right)\tilde{p}_n = 0, \tag{2.34}$$

where

$$\tilde{\zeta}_n = \frac{1}{f}\left(\frac{\partial^2 \tilde{p}_n}{\partial y} - \frac{\epsilon\beta}{f}\frac{\partial \tilde{p}_n}{\partial y} - k^2\tilde{p}_n\right). \tag{2.35}$$

Furthermore, the boundary conditions are

$$\tilde{p}_n(y = -\tfrac{1}{2}, \tfrac{1}{2}) = 0, \tag{2.36}$$

which ensures no normal flow at the side walls.

3) N-layer linear wave equation

In an N-layer configuration, the eigenvalue equation is of the same form as (2.34), with the following changes to the parameter functions H_n, Δ_n, and U_n:

$$\Delta_n(\phi_n) \to \Delta_n^s(\phi_n) = \begin{cases} \phi_1 - \phi_2, & n = 1, \\ 2\phi_n - \phi_{n+1} - \phi_{n-1}, & n = 2, \ldots, N-1, \\ \phi_N - \phi_{N-1}, & n = N, \end{cases}$$

$$U_s \to U_n^s = \Delta_n^s(U_n),$$

$$\overline{h}_n \to \overline{h}_n^s = H_n - \epsilon FGU_n^s, \tag{2.37}$$

where H_n is the rest thickness of the nth layer and H_0 is now a reference thickness (see Fig. 2). For simplicity, the density difference between each of the neighboring layers is considered constant, yet the mean thicknesses of each layer are left as parameters.

c. *Salmon's L_1 equations*

The barotropic L_1 dynamics were originally derived by Salmon (1983, 1985) from the SW Lagrangian. The name L_1 is Salmon's notation, where L is the SW Lagrangian, and the subscript 1 means the $O(1)$ expansion of L in terms of the Rossby number; the $O(0)$ expansion is termed L_0 and, in fact, yields PG. The method of approximation was motivated by the desire to preserve the symmetries of the Lagrangian, hence, to maintain

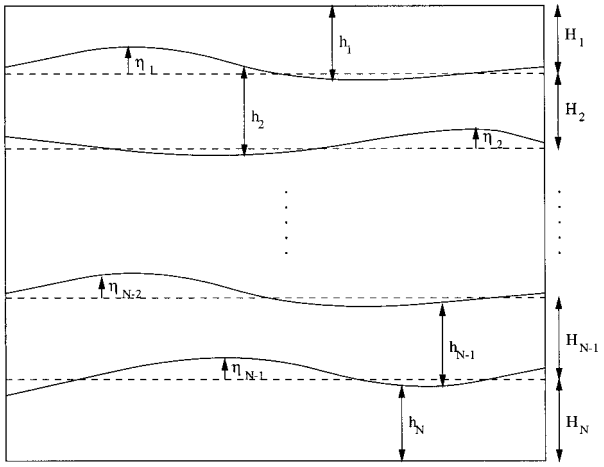

FIG. 2. Model domain for an N-layer system under a rigid lid.

all of the conservation properties of the SW equations. Specifically, the approximation is made by replacing the fluid velocity, where it appears in the conjugate momentum term, with its geostrophic value.

The two-layer, rigid lid L_1 equations are derived from the two-layer SW Lagrangian in appendix A, for which the result is

$$\epsilon\frac{\partial}{\partial t}\mathbf{u}_{g,n} + (f + \epsilon\zeta_{g,n})\hat{\mathbf{z}} \times \mathbf{u}_n + \boldsymbol{\nabla}B_n = 0, \quad (2.38)$$

$$\frac{\partial h}{\partial t} + \boldsymbol{\nabla}\cdot(h\mathbf{u}_n) = 0, \quad (2.39)$$

where

$$B_n = p_n + \frac{\epsilon}{2}|\mathbf{u}_{g,n}|^2 + \frac{\epsilon}{F}\hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\times\left(\frac{h_n\mathbf{u}_{a,n}}{f}\right) \quad (2.40)$$

is the Bernoulli functional, $\epsilon\mathbf{u}_{a,n} = \mathbf{u}_n - \mathbf{u}_{g,n}$ is the ageostrophic velocity, $\zeta_{g,n} = \hat{\mathbf{z}}\cdot\boldsymbol{\nabla}\times\mathbf{u}_{g,n}$ is the geostrophic relative vorticity, and $\mathbf{u}_{g,n}$ is identical to the velocity in the GPV equations, given by (2.26). The equations admit no time derivative of the ageostrophic velocity, which eliminates gravity waves from the solution [save for a particular brand of distorted Kelvin waves at solid boundaries; Allen et al. (1990)]. In taking the variation of L_1, two surface integrals arise that vanish only if the tangential components of $\mathbf{u}_{a,n}$ vanish at the horizontal boundaries of the domain (see appendix A), in complete analogy with the barotropic case (Salmon 1983; Allen and Holm 1996).

By taking variations of the Lagrangian, it is shown in Salmon (1983) that a form of potential vorticity is conserved in the barotropic L_1 model. Alternatively, beginning from (2.38)–(2.40) and referring to the derivation of SW potential vorticity conservation (Pedlosky 1987), one can see immediately that the form conserved is identical to that of the GPV model [(2.27)]. In fact, the L_1 equations can be closed in nearly the same way as the GPV equations, except for the addition of the ageostrophic velocities in the former equations.

1) PARAMETER SPACE LIMITS

The L_1 dynamics include QG, PG, *and* FG as parameter space limits. The QG limit [$\beta \sim F \sim O(1)$ and $\epsilon \to 0$] follows easily since the potential vorticity equation is the same as that of GPV, with the addition of $O(\epsilon)$ ageostrophic velocities. Thus, the QG limit of L_1 is obtained in the same way as for the GPV equations. An alternate derivation begins with (2.38)–(2.40) and proceeds by assuming $\epsilon \ll 1$ and expanding the fields asymptotically. The details are somewhat tedious but straightforward, and in any case they parallel exactly the derivation of QG given by Pedlosky (1987, section 3.12). The result is (2.29)–(2.30).

The PG limit is obtained by taking the straightforward limit $\epsilon \to 0$ of (2.38)–(2.40) while assuming concurrently that ϵF and $\epsilon\beta$ remain $O(1)$. The result is

$$\mathbf{u}_n = \frac{1}{f}\hat{\mathbf{z}} \times \boldsymbol{\nabla}p_n, \quad (2.41)$$

$$\frac{\partial h_n}{\partial t} + \boldsymbol{\nabla}\cdot(h_n\mathbf{u}_n) = 0, \quad (2.42)$$

where h_n is given in terms of p_n by (2.8)–(2.9). This is equivalent to the representation (2.32)–(2.32).

Because no direct numerical comparisons of the *linear* models can be performed in the FG limit (we explain why in section 4a), the demonstration that L_1 contains FG is relegated to appendix C.

2) WAVE EQUATION

The basic state [(2.12)–(2.16)] is a solution of the two-layer L_1 dynamics. Linearization about this solution and subsequent substitution of a wave solution of the form (2.18) then yields

$$(U_n - c)i\frac{\partial\tilde{p}_n}{\partial y} + \frac{f^2}{k}\tilde{v}_{a,n} - \frac{f}{F}\frac{\partial A}{\partial x} = 0, \quad (2.43)$$

$$(U_n - c)k\tilde{p}_n - \frac{f^2}{k}\tilde{u}_{a,n} - \frac{f}{F}\frac{\partial A}{\partial y} = 0, \quad (2.44)$$

$$F[-U_n(\tilde{p}_1 - \tilde{p}_2) - c\Delta_n(\tilde{p}_n)] - \frac{i}{k}\frac{\partial}{\partial y}(\tilde{v}_{a,n}\overline{h}_n) + \overline{h}_n\left(\tilde{u}_{a,n} - \frac{\beta}{f^2}\tilde{p}_n\right) = 0, \quad (2.45)$$

where

$$\frac{\partial A}{\partial x} = -\hat{h}_n\left(k\tilde{v}_{a,n} - i\frac{\partial \tilde{u}_{a,n}}{\partial y}\right) - i\tilde{u}_{a,n}\frac{\partial \hat{h}_n}{\partial y}, \quad (2.46)$$

$$\frac{\partial A}{\partial y} = \hat{h}_n\left(i\frac{\partial \tilde{v}_{a,n}}{\partial y} + k\frac{\partial^2 \tilde{u}_{a,n}}{\partial y^2}\right) + \frac{\partial \hat{h}_n}{\partial y}\left(i\tilde{v}_{a,n} - \frac{2}{k}\frac{\partial \tilde{u}_{a,n}}{\partial y}\right) - \frac{1}{k}\tilde{u}_{a,n}\frac{\partial^2 \hat{h}_n}{\partial y^2}, \quad (2.47)$$

and $\hat{h}_n = \overline{h}_n/f$. The boundary conditions are then

$$\tilde{u}_{a,n}(y = -½, ½) = 0, \quad (2.48)$$

$$-\frac{1}{f}\frac{\partial \tilde{p}_n}{\partial y}(y = -½, ½) + \epsilon\tilde{v}_{a,n}(y = -½, ½) = 0, \quad (2.49)$$

where the first condition ensures no tangential ageostrophic component at the side walls, while the second condition yields no normal flow at the boundaries.

d. Parameter regimes: A cautionary note

It should be pointed out that a QG model is *always* in a QG regime. That is to say, it is an asymptotically derived model, derived under the assumptions that ϵ is asymptotically small and that ϵF and $\epsilon\beta$ are both $O(\epsilon)$. It is simply impossible to choose parameters in the QG equations such that these are not satisfied, essentially because the Rossby number does not appear as a parameter. One might attempt to choose parameters, such as U, L, f_0, and L_d, such that one is outside of the QG regime. However, a QG model that is initialized with such parameter values will nevertheless be in a QG regime, because one does not truly have the freedom to independently choose U, L, and f_0. For example, the deformation radius does appear as a parameter, and the length scale of motions appearing in a QG solution may be much shorter or much longer than it, yet even for scales much longer than the deformation radius the parameter $\epsilon F = \epsilon(L/L_d)^2$ is implicitly small, since ϵ itself is assumed vanishingly small. Similarly, a PG model is always in a PG regime. Neither a QG model nor a PG model can evolve away from its regime of validity.

These considerations should be borne in mind in the following comparisons when ''QG'' and ''PG'' regimes are discussed. In particular, in the figure captions the given value of the Rossby number, ϵ, is not applicable to either a PG or a QG model, since for these models $\epsilon = 0$.

The GPV and L_1 equations, on the other hand, are not asymptotically derived in the same way. Indeed, the Rossby number appears as a parameter in these equations, even though their regime of validity is restricted to small Rossby number. Thus, it is possible to initialize the equations in a regime in which the models are not formally valid, or for the equations to evolve away from the regime of validity. Often, when this happens the method of solution will fail; for example, an operator will cease to become elliptic and it then becomes impossible to diagnose a subsidiary velocity field. Whether one regards this property of the equations as a ''bug'' or a ''feature'' is debatable.

3. Results

The linearized wave equations derived in the previous section are finite differenced and solved numerically as generalized algebraic eigenvalue problems. The goal of the numerical study is the qualitative assessment of each model's ability to capture the salient features of the SW results, for a parameter range that covers the spectrum from deformation to planetary scales. We present the the growth rates and frequencies, as well as the eigenfunctions for the pressure and horizontal velocities in each layer at the wavelength of maximum instability. In addition, we investigate the short-wave divergence of a three-layer PG model relative to three-layer QG and GPV results.

a. Baroclinic instability in the two-layer models

Here we do essentially the problem formulated by Phillips (1954) for the models under consideration. In particular, we calculate the growth rates (kc_i) in units of the basic-state zonal velocity (U_0) as a function of zonal wavenumber (k) for the gravest meridional mode of the potential vorticity mode solutions. Furthermore, U_0 is chosen in each case such that the shear velocity $2U_0$ is twice the critical shear necessary for baroclinic instability to occur in the QG model with positive β, which corresponds to $U_0 = \beta/F$ (Pedlosky 1987). The eigenfunctions are normalized such that the maxima of the upper-layer pressures for each model are 1, and the zonal wavenumbers k are scaled by $F^{1/2}$, so that a value of 1 on the abscissa of the plots corresponds to a wavelength equal to the radius of deformation.

In the first case, we consider a parameter set that is essentially in the atmospheric QG domain, with values of $F = 10$, $\beta = 3$, and $\epsilon = 0.05$. In this case, we find that all of the models are in near complete agreement (Fig. 3). Note that $\epsilon F = 0.5$ is rather large, but for two-layer QG, one must have $F > (k^2 + \pi^2)/2$ in order for baroclinic instability to occur (Pedlosky 1987). So for $k = 0$, $F_{\min} = 4.93$, and in order to see a complete cycle, F must be large enough for baroclinic instability to occur at a range of k values. In any case, the data itself verifies that this set of parameter values does represent a QG regime.

In general, variation of the eigenfunctions in the lower layer will always be greater, because the fields are normalized by the maximum *upper*-layer pressure, and because the coupling terms for each model are different and all are relatively sensitive to the nondimensional parameters.

In a regime that roughly represents spatial scales in the atmosphere just beyond the QG regime, but not in

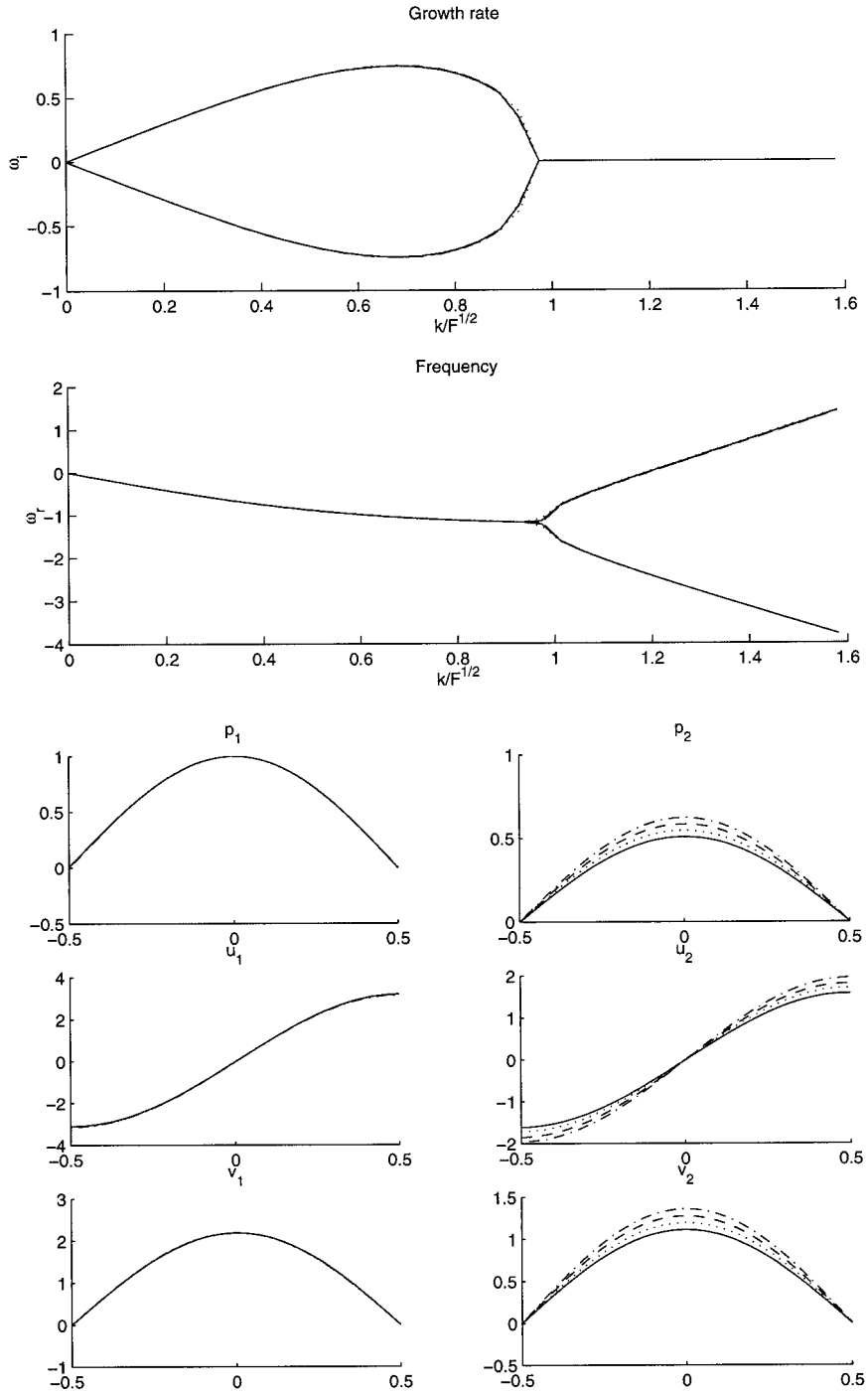

FIG. 3. Eigenvalues and eigenfunctions for SW (solid line), L_1 (dashed line), GPV (dash–dot line), and QG (dotted line) in a quasigeostrophic parameter regime ($\beta = 3$, $F = 10$, and $\epsilon = 0.05$).

the PG regime, we find a slight variation between all of the models (Fig. 4). Notice that in the upper layer, L_1 eigenfunctions correspond very closely to those of SW, but that in the lower layer, again, all of the models disagree. The QG is the farthest off, with its inability to capture any of the asymmetry present due to the relatively strong meridional variation in dynamic topography and Coriolis parameter.

We will not present comparisons in the FG regime, because all of the *linearized* models presented here

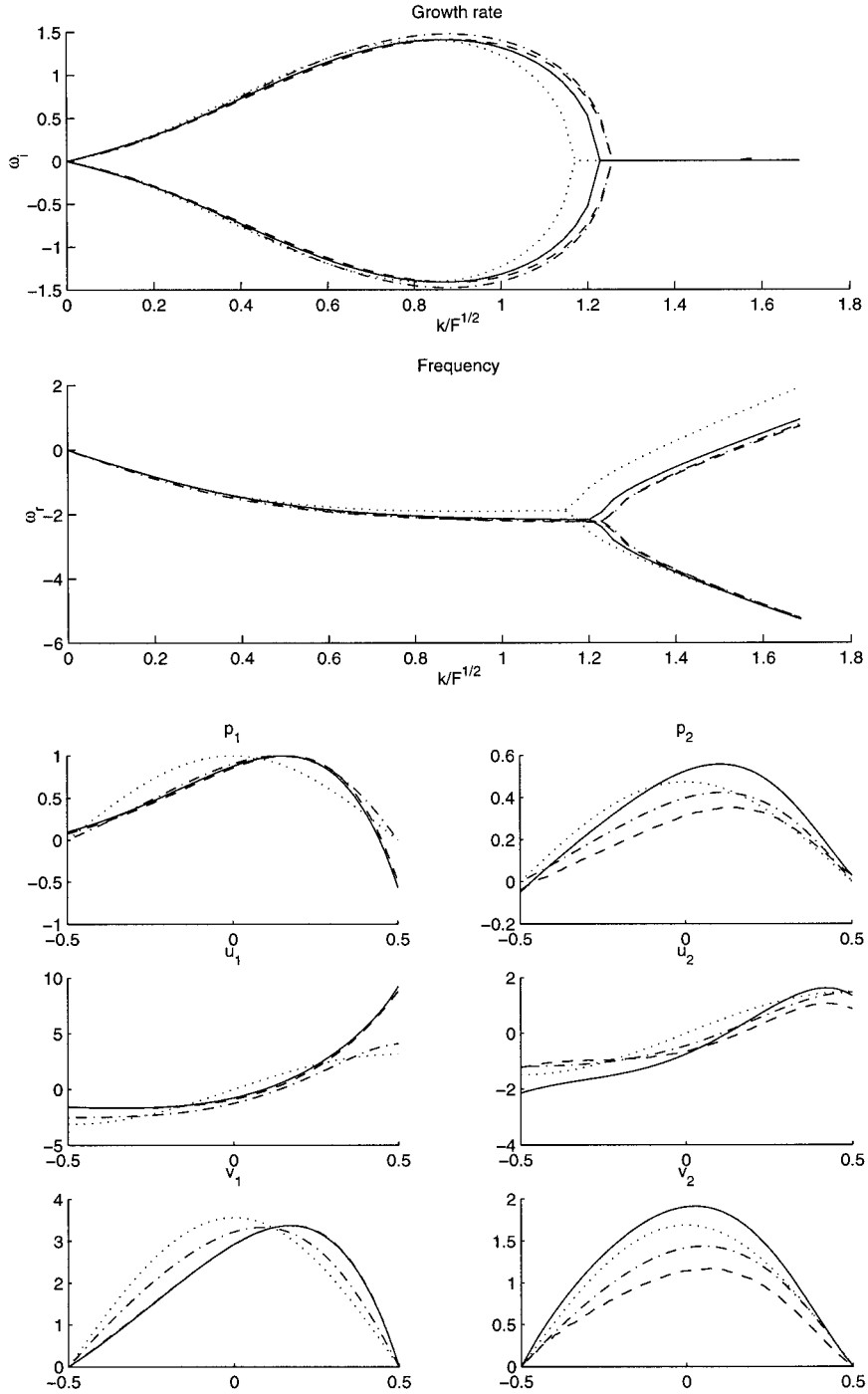

FIG. 4. Eigenvalues and eigenfunctions for SW (solid line), L_1 (dashed line), GPV (dash–dot line), and QG (dotted line) in a parameter regime just beyond that of quasigeostrophy ($\beta = 17$, $F = 26$, and $\epsilon = 0.02$).

are then in complete agreement. To understand why, note that linearized QG contains a time derivative of the relative vorticity ($\nabla^2\eta$), which is absent in linearized FG, whereas linearized FG contains a correction to the velocity field (with coefficient $\epsilon\beta$), which is absent in linearized QG. Now both terms are contained in both GPV and L_1, and the term that is absent in QG (the velocity correction proportional to $\epsilon\beta$) is

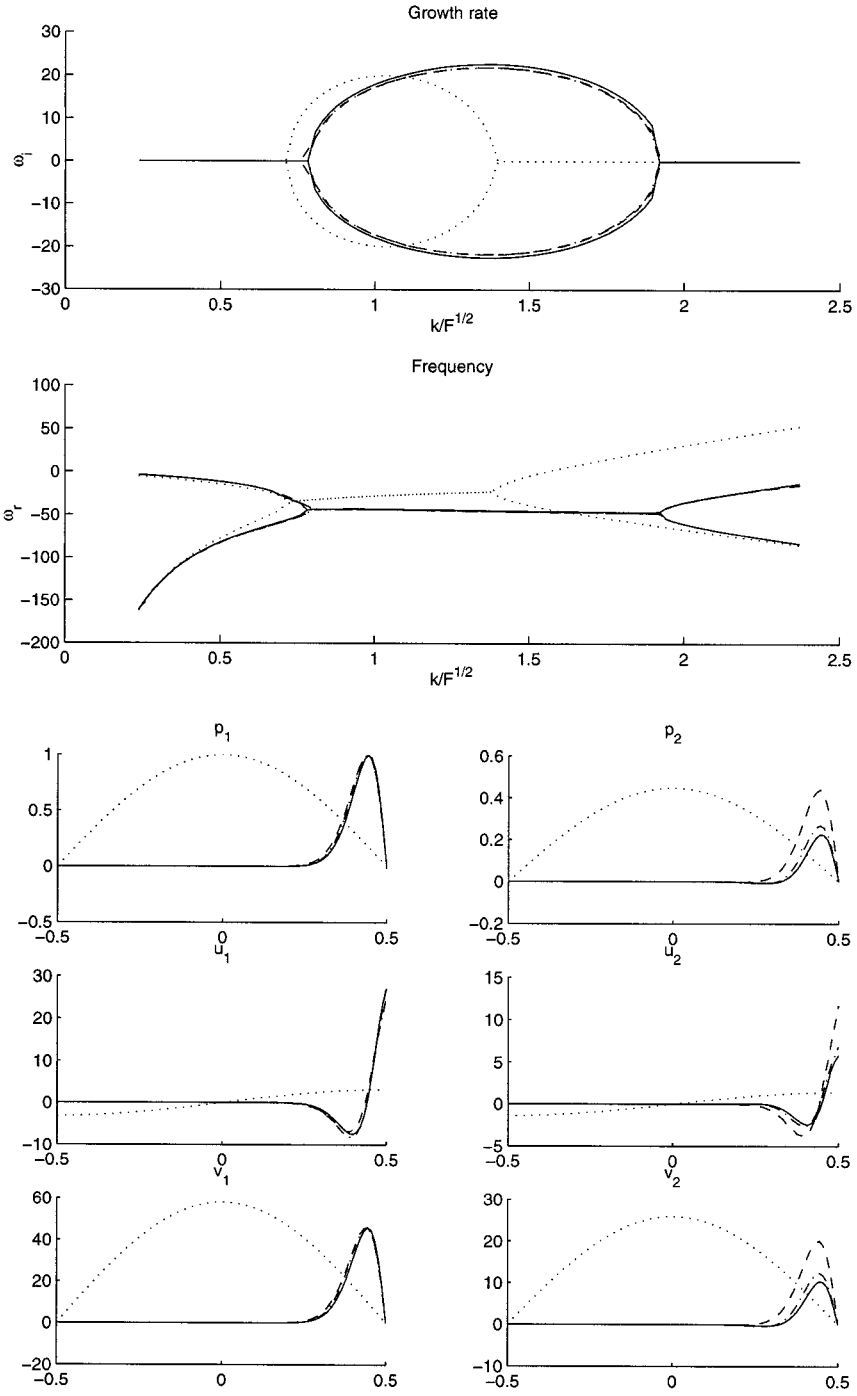

FIG. 5. Eigenvalues and eigenfunctions for SW (solid line), L_1 (dashed line), GPV (dash–dot line), and QG (dotted line) in a planetary geostrophic regime ($\beta = 3913$, $F = 1779$, and $\epsilon = 1.7 \times 10^{-4}$).

necessarily small in the FG limit. Hence, all three linear models should coincide in the FG limit, because when linearized, FG and QG yield the same equations. In fact, this further explains the agreement of the models in the first parameter set (Fig. 3), which tends toward the FG regime.

Finally, we compare the models in the PG limit (Fig. 5). In this case we find close agreement between

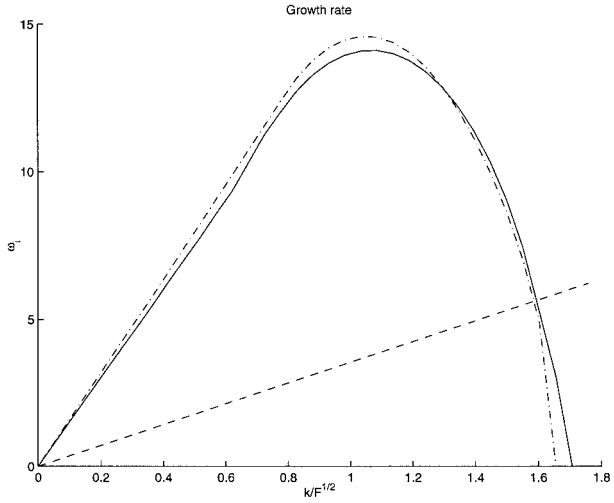

FIG. 6. Growth rates for three-layer GPV (solid line), QG (dotted line), and PG (dashed line) with quasigeostrophic scales of $\beta = 494$, $F = 808$, and $\epsilon = 0.000\,275$.

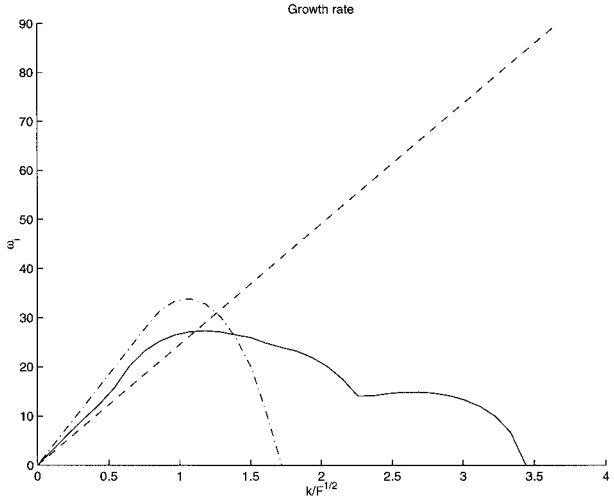

FIG. 7. Growth rates for three-layer GPV (solid line), QG (dotted line), and PG (dashed line) with planetary geostrophic scales of $\beta = 7897$, $F = 4310$, and $\epsilon = 6.88 \times 10^{-5}$.

SW, GPV, and L_1 for all of the eigenvalues and eigenfunctions, and a large disagreement with QG (as one should expect).

b. *Divergent growth rate in the planetary geostrophic model*

Baroclinic instability occurs in a PG model only if there are three or more active fluid layers in the system (Colin de Verdière 1986). The growth rate increases linearly with the wavenumber k, yielding an ultraviolet divergence, an artifact of the neglect of the inertial terms in the momentum equations. In a PG circulation model this divergent instability can be quelled by frictional and viscous terms (e.g., Samelson and Vallis 1997), provided that the instability is not too strong. That is, in order that the coefficients providing the small-scale damping be realistically small, the growth rate of the PG instability at the wavenumber of maximum instability in the real system should preferably be equal to or smaller than the growth rate itself of the real system.

The simplicity of the GPV model allows facile extension to the multilayer case, and, given its excellent comparison with the shallow water model across a broad region of parameter space, we use such a model to represent the growth rate of the real system in a comparison with that of the PG model. The three-layer PG growth rate is derived analytically in appendix B, and the growth rates for the GPV model are obtained numerically. (Numerical evaluation of the PG growth rates agreed with the analytic expression.) In these comparisons, we select the *largest* growth rate at each wavenumber, as opposed to the growth rate corresponding to the gravest meridional modes (as in the previous sections). We choose a linear, symmetric velocity profile,

$$U_1 = U_0, \qquad U_2 = 0, \qquad U_3 = -U_0,$$

but here we scale U_0 as $2\beta/F_0$. In this case, the growth rate is (see appendix B)

$$\omega_i = 0.204\left(\frac{\beta}{F}\right)k. \tag{3.1}$$

The growth rates for the three-layer GPV, PG, and QG models are calculated for scales corresponding to a quasigeostrophic regime (Fig. 6) and a planetary geostrophic regime (Fig. 7). In the PG regime, the growth rates of the GPV and PG models have the same slope at small wavenumbers, as should be expected. In the QG regime, it is found that at the wavelength of maximum growth in the GPV model (which here represents the real system), the GPV growth rate is larger than that of the PG model (Fig. 6). Because the regime is set by the nondimensional parameters chosen for the calculation, there is no a priori reason to expect that the growth rates of PG and QG (and hence GPV, which mimics QG in this regime by design) should have the same slope at any wavenumber. For phenomena indigenous to the PG regime, relative vorticity, and hence eddies, are important factors in determining the dynamics, while in a PG regime they are, by definition, negligible. Hence, the fact that the PG growth rate is relatively small at the most unstable wavenumbers in a QG regime means that frictional terms used in a PG-based model do not need to be excessively large and can safely be set only to act on scales much smaller than those at which the instabilities are driving the dynamics.

We also see, of course, the large distortion produced by the QG model when compared to the more accurate model in a regime well beyond the QG range of applicability (Fig. 7).

4. Summary and conclusions

The primary goal of this paper has been to investigate whether two relatively unfamiliar geostrophic models sensibly and accurately portray linear baroclinic instability. To this end the models were cast in a multilayer form on a differentially rotating channel, linearized, and the resulting eigenfunctions and eigenvalues calculated numerically. In addition, we have discussed the behavior of these models in various asymptotic limits. There are four important points which emerge.

1) Both the GPV model and the L_1 model contain the QG and the PG equations as appropriate asymptotic limits in parameter space; essentially, QG results for scales comparable to the deformation scale (more precisely when $(\epsilon(L/L_d)^2 \ll 1)$ and the PG equations result for planetary scales. In addition, the L_1 equations contain frontal geostrophic dynamics, valid in a (geostrophic) regime that is large compared to the deformation radius but small compared to the planetary scale.
2) Both the GPV and L_1 models can be integrated outside of their regime of validity. Indeed, the L_1 and GPV might in a nonlinear integration evolve away from their region of validity. This is impossible for the asymptotically derived QG and PG equations.
3) Both the L_1 and the GPV equations have very accurate linear stability properties. They are noticeably more accurate than either the QG or the PG equations, especially when the parameters are not in a strict QG or PG regime. The stability properties of GPV and L_1 are, in fact, almost identical (although completely different solution algorithms were used for each).
4) The growth rates of a QG model compare very well to those of the more accurate models when in a QG regime. The PG equations, however, have (as is well known) an ultraviolet divergence. That is, the growth rate increases linearly with wavenumber. However, at wavenumbers small or comparable to the deformation scale the growth rate is relatively small, compared to the actual growth rate or the growth rate of a QG model. The inclusion of the relative vorticity in the expression for potential vorticity (as in GPV), or a scale-dependent friction, will eliminate this divergent growth rate.

The advantage of L_1 over GPV is that the unforced inviscid dynamics also conserve energy, although the solution algorithm is more complicated. Either of these models would be appropriate for investigations that involve small Rossby number motion at either or both deformation and the planetary scales. Mundt et al. (1997) have previously shown that the GPV dynamics perform very well in nonlinear integrations of the wind-driven circulation, but if strict energy conservation of the advective dynamics is required then L_1 would be preferable. However, the solution algorithm is sufficiently complex, and the time step may be limited by the presence of boundary Kelvin waves, that the numerical efficiency may not be superior to that of the primitive equations.

Acknowledgments. The authors are grateful to Darryl Holm for very helpful discussions regarding the formulation of the L_1 models, to Rupert Ford for comments about the relationship of the L_1 and FG models, and to Michael Mundt for help with the numerics. We also wish to thank the anonymous reviewers for their helpful comments. This work was funded by the NSF.

APPENDIX A

Two-Layer L_1 Dynamics

In this appendix we derive the two-layer L_1 equations with a rigid lid from the shallow water Lagrangian.

The Lagrangian is written

$$L = T - \overline{V}, \tag{A.1}$$

where T is the kinetic energy of the system and $\overline{V}$ is the generalized potential energy, which includes constraints. For a two-layer shallow water system the kinetic energy is given by

$$T = \sum_{n=1}^{2} \int dx_n\, dy_n h_n \rho_n \left(\mathbf{u}_n \cdot \mathbf{R}_n + \frac{1}{2}|\mathbf{u}_n|^2 \right), \tag{A.2}$$

where $\mathbf{R}$ is defined so that $\boldsymbol{\nabla} \times \mathbf{R} = f\hat{\mathbf{z}}$, and the potential energy per unit horizontal area is

$$\begin{aligned} \upsilon &= \int \rho g z\, dz \\ &= \int_0^{h_2} \rho_2 g z\, dz + \int_{h_2}^{h_1+h_2} \rho_1 g z\, dz \\ &= \frac{g}{2}(\rho_1 h_1^2 + \rho_2 h_2^2 + 2\rho_1 h_1 h_2). \end{aligned} \tag{A.3}$$

The imposed constraint is the rigid lid, which implies that $h_1 + h_2 - H = 0$. The constraint is enforced by the Lagrange multiplier λ, and the resulting generalized potential energy is

$$\overline{V} = \int d\mathbf{x}_1\, d\mathbf{x}_2\, \delta(\mathbf{x}_1 - \mathbf{x}_2)[\upsilon + (h_1 + h_2 - H)\lambda], \tag{A.4}$$

where the delta function is necessary to lock together the two coordinate systems, and $d\mathbf{x}_n$ is defined as $dx_n dy_n$. The two-layer shallow water Lagrangian is then

$$L_{sw} = \sum_{n=1}^{2} \rho_n \int d\mathbf{x}_n h_n \left[(\mathbf{R}_n + \mathbf{u}_n) \cdot \frac{\partial \mathbf{x}_n}{\partial \tau_n} - \frac{1}{2}|\mathbf{u}_n|^2 - \frac{g}{2} h_n^2 \right] - \int d\mathbf{x}_1 \, d\mathbf{x}_2 \cdot \delta(\mathbf{x}_1 - \mathbf{x}_2)[\rho_1 h_1 h_2 + (h_1 + h_2 - H)\lambda], \quad \text{(A.5)}$$

where

$$\partial/\partial\tau_n = \partial/\partial t + u_n \partial/\partial x_n + v_n \partial/\partial y_n \quad \text{(A.6)}$$

is the advective derivative. The L_1 Lagrangian is obtained by replacing $\mathbf{u}_n$ with its geostrophic value, $\mathbf{u}_{g,n}$ [defined in (2.26)], yielding

$$L_1 = \sum_{n=1}^{2} \rho_n \int d\mathbf{x}_n h_n \left[(\mathbf{R}_n + \mathbf{u}_{g,n}) \cdot \frac{d_n \mathbf{x}_n}{d\tau} - \frac{1}{2}|\mathbf{u}_{g,n}|^2 - \frac{g}{2} h_n^2 \right] - \int d\mathbf{x}_1 \, d\mathbf{x}_2 \, \delta(\mathbf{x}_1 - \mathbf{x}_2)[\rho_1 g h_1 h_2 + (h_1 + h_2 - H)\lambda]. \quad \text{(A.7)}$$

We vary the Lagrangian in physical space to obtain (note that the Lagrange multiplier is *not* varied)

$$\delta L_1 = \rho_1 \int d\mathbf{x}_1 \left[(\mathbf{R}_1 + \mathbf{u}_{g,1}) \cdot \mathbf{u}_1 - \frac{1}{2}|\mathbf{u}_{g,1}|^2 - g(h_1 + h_2) - \frac{\lambda}{\rho_1} \right] \delta h_1 + h_1[(\mathbf{R}_1 + \mathbf{u}_{g,1}) \cdot \delta\mathbf{u}_1 + \mathbf{u}_{a,1} \cdot \delta\mathbf{u}_{g,1}] + \rho_2 \int d\mathbf{x}_2 \left[(\mathbf{R}_2 + \mathbf{u}_{g,2}) \cdot \mathbf{u}_2 - \frac{1}{2}|\mathbf{u}_{g,2}|^2 - g\left(h^2 + \frac{\rho_1}{\rho_2} h_1 \right) - \frac{\lambda}{\rho_2} \right] \delta h_2 + h_2[(\mathbf{R}_2 + \mathbf{u}_{g,2}) \cdot \delta\mathbf{u}_2 + \mathbf{u}_{a,2} \cdot \delta\mathbf{u}_{g,2}], \quad \text{(A.8)}$$

where $\mathbf{u}_{a,n} = \mathbf{u}_n - \mathbf{u}_{g,n}$ is the ageostrophic velocity in each layer. Now we must relate $\delta\mathbf{u}_{g,n}$ to δh_n; hence, we must first relate p_n to h_n. The Lagrange multiplier λ is the pressure excess imposed by the rigid lid and is related to the individual pressures in each layer by

$$p_1 = \rho_1 g(h_1 + h_2 - z) + \lambda, \quad \text{(A.9)}$$

$$p_2 = \rho_1 g h_1 + \rho_2 g(h_2 - z) + \lambda, \quad \text{(A.10)}$$

where z is the vertical coordinate ($z = 0$ at the bottom) and h_n in dimensional form is

$$h_n = H_n + (-1)^n \eta. \quad \text{(A.11)}$$

By substituting (A.11) into (A.9)–(A.10), then taking the horizontal gradient of each expression and subsequently eliminating p, we find

$$\nabla p_n = \nabla[p_{3-n} + (-1)^n g(\rho_2 - \rho_1)\eta]. \quad \text{(A.12)}$$

We may now substitute (A.11) back into (A.12) (the constant terms will not contribute) to get

$$\nabla p_n = \nabla[p_{3-n} + g(\rho_2 - \rho_1)h_n]. \quad \text{(A.13)}$$

Therefore, the geostrophic velocity may be written as

$$\mathbf{u}_{g,n} = \frac{1}{\rho_n f} \hat{\mathbf{z}} \times \nabla(p_{3-n} + g\Delta\rho h_n), \quad \text{(A.14)}$$

so that the variation is

$$\delta\mathbf{u}_{g,n} = \frac{g'_n}{f} \hat{\mathbf{z}} \times \nabla(\delta h_n), \quad \text{(A.15)}$$

where $g'_n = g\Delta\rho/\rho_n$. Then, using the identity for the divergence of a vector product, we may write the terms $h_1\mathbf{u}_{a,1} \cdot \delta\mathbf{u}_{g,1}$ and $h_2\mathbf{u}_{a,2} \cdot \delta\mathbf{u}_{g,2}$ as

$$h_n \mathbf{u}_{a,n} \cdot \delta\mathbf{u}_{g,n} = \nabla \cdot \left[\left(\frac{g'_n}{f} h_n \mathbf{u}_{a,n} \right) \times \hat{\mathbf{z}} \delta h_n \right] - \left[\hat{\mathbf{z}} \cdot \nabla \times \left(\frac{g'_n}{f} h_n \mathbf{u}_{a,n} \right) \right] \delta h_n. \quad \text{(A.16)}$$

Now the variation of the Lagrangian becomes

$$\delta L_1 = \rho_1 \int d\mathbf{x}_1 \left[(\mathbf{R}_1 + \mathbf{u}_{g,1}) \cdot \mathbf{u}_1 - \frac{1}{2}|\mathbf{u}_{g,1}|^2 - gH - \frac{\lambda}{\rho_1} - g'_1 \hat{\mathbf{z}} \cdot \nabla \times \left(\frac{h_1 \mathbf{u}_{a,1}}{f} \right) \right] \delta h_1 + h_1[(\mathbf{R}_1 + \mathbf{u}_{g,1}) \cdot \delta\mathbf{u}_1] + \rho_2 \int d\mathbf{x}_2 \left[(\mathbf{R}_2 + \mathbf{u}_{g,2}) \cdot \mathbf{u}_2 - \frac{1}{2}|\mathbf{u}_{g,2}|^2 - gH + g'_2 h_1 - \frac{\lambda}{\rho_2} - g'_2 \hat{\mathbf{z}} \cdot \nabla \times \left(\frac{h_2 \mathbf{u}_{a,2}}{f} \right) \right] \delta h_2 + h_2[(\mathbf{R}_2 + \mathbf{u}_{g,2}) \cdot \delta\mathbf{u}_2] + \sum_{n=1}^{2} \rho_n \int d\mathbf{x}_n \nabla \cdot \left[\left(\frac{g'_n}{f} h_n \mathbf{u}_{a,n} \right) \times \hat{\mathbf{z}} \delta h_n \right], \quad \text{(A.17)}$$

where H has been substituted for $h_1 + h_2$. The integral over the divergence may be transformed to a surface integral by Gauss's theorem as

$$\sum_{n=1}^{2} \rho_n \int d\mathbf{x}_n \boldsymbol{\nabla} \cdot \left[\left(\frac{g_n'}{f} h_n \mathbf{u}_{a,n} \right) \times \hat{\mathbf{z}} \delta h_n \right] = \sum_{n=1}^{2} \rho_n \oint ds_n \left[\frac{g_n'}{f} h_n \mathbf{u}_{a,n} \times \hat{\mathbf{z}} \delta h_n \right] \cdot \hat{\mathbf{s}}, \quad \text{(A.18)}$$

where $\hat{\mathbf{s}}$ is the unit normal vector on the horizontal boundary of the domain. This vanishes identically if $(\mathbf{u}_{a,n} \times \hat{\mathbf{z}}) \cdot \hat{\mathbf{s}} = 0$, that is, if the tangential component of the ageostrophic velocities are zero at the horizontal boundaries.

Finally, we apply Hamilton's principle,

$$\delta \int L \, d\tau = 0,$$

in the manner prescribed by Allen and Holm (1996), who derive the statement

$$\frac{\partial}{\partial \tau_n} \left(\frac{1}{h_n} \frac{\partial L}{\partial \mathbf{u}_n} \right) + \frac{1}{h_n} \sum_j \frac{\partial L}{\partial (\mathbf{u}_n \cdot \hat{\mathbf{e}}_j)} \boldsymbol{\nabla} (\mathbf{u}_n \cdot \hat{\mathbf{e}}_j) - \boldsymbol{\nabla} \left(\frac{\partial L}{\partial h_n} \right) = 0, \quad \text{(A.19)}$$

where $\hat{\mathbf{e}}_1 = \hat{\mathbf{x}}$ and $\hat{\mathbf{e}}_2 = \hat{\mathbf{y}}$. Application of (A.19) to (A.17) yields

$$\frac{\partial \mathbf{u}_{g,n}}{\partial t} + \mathbf{u}_n \cdot \boldsymbol{\nabla} (\mathbf{R}_n + \mathbf{u}_{g,n}) + \sum_j (\mathbf{R}_n + \mathbf{u}_{g,n}) \cdot \hat{\mathbf{e}}_j \boldsymbol{\nabla} (\mathbf{u}_n \cdot \hat{\mathbf{e}}_j) + \boldsymbol{\nabla} \left[-(\mathbf{R}_n + \mathbf{u}_{g,n}) \cdot \mathbf{u}_n + \frac{1}{2} |\mathbf{u}_{g,n}|^2 - \frac{\lambda}{\rho_n} + g_n' \hat{\mathbf{z}} \cdot \boldsymbol{\nabla} \times \left(\frac{h_n \mathbf{u}_{a,n}}{f} \right) - \{n - 1\} g_2' h_1 \right] = 0. \quad \text{(A.20)}$$

Because the pressure and height terms are now the arguments of the gradient operator, we may use (A.13) to write for $n = 1$

$$\boldsymbol{\nabla} \frac{\lambda}{\rho_1} = \boldsymbol{\nabla} \frac{p_1}{\rho_1},$$

and for $n = 2$

$$\boldsymbol{\nabla} \left(\frac{\lambda}{\rho_2} - g_2' h_1 \right) = \frac{1}{\rho_2} \boldsymbol{\nabla} [p_1 - g \Delta \rho (H_1 - \eta)] = \boldsymbol{\nabla} \frac{p_2}{\rho_2}. \quad \text{(A.21)}$$

Finally, we use the vector identity

$$\boldsymbol{\nabla} (\mathbf{A} \cdot \mathbf{B}) = \mathbf{A} \cdot \boldsymbol{\nabla} \mathbf{B} + \sum_j (\mathbf{B} \cdot \hat{\mathbf{e}}_j) \boldsymbol{\nabla} (\mathbf{A} \cdot \hat{\mathbf{e}}_j) + \mathbf{A} \times \boldsymbol{\nabla} \times \mathbf{B}$$

to rewrite (A.20) in vorticity form as (2.38).

The continuity equations are derived by first noting that the height h_n in each layer is the coordinate transformation Jacobian from label space, (a_n, b_n), to physical space, (x_n, y_n),

$$h_n = \frac{\partial (a_n, b_n)}{\partial (x_n, y_n)}. \quad \text{(A.22)}$$

Application of the advective derivative (A.6) to $1/h_n$ then yields

$$\frac{\partial}{\partial \tau_n} \left(\frac{1}{h_n} \right) = \frac{\partial}{\partial \tau_n} \left[\frac{\partial (x_n, y_n)}{\partial (a_n, b_n)} \right].$$

The left-hand side is

$$\frac{\partial}{\partial \tau_n} \left(\frac{1}{h_n} \right) = -\frac{1}{h_n^2} \frac{\partial h_n}{\partial \tau_n}$$

while the right-hand side becomes (recalling that $\partial x_n / \partial \tau_n = u_n$ and $\partial y_n / \partial \tau_n = v_n$)

$$\frac{\partial}{\partial \tau_n} \left[\frac{\partial (x_n, y_n)}{\partial (a_n, b_n)} \right] = \frac{\partial (u_n, y_n)}{\partial (a_n, b_n)} + \frac{\partial (x_n, v_n)}{\partial (a_n, b_n)} = \frac{\partial u_n}{\partial x_n} \frac{\partial (x_n, y_n)}{\partial (a_n, b_n)} + \frac{\partial v_n}{\partial y} \frac{\partial (x_n, y_n)}{\partial (a_n, b_n)} = \frac{1}{h_n} \left(\frac{\partial u_n}{\partial x_n} + \frac{\partial v_n}{\partial y_n} \right). \quad \text{(A.23)}$$

Equating the two sides then yields

$$\frac{\partial h_n}{\partial \tau} + h_n \left(\frac{\partial u_n}{\partial x_n} + \frac{\partial v_n}{\partial y_n} \right) = 0, \quad \text{(A.24)}$$

which is equivalent to (2.39) upon substitution of the definition (A.6) for the advective derivative into (A.24).

APPENDIX B

Three-Layer Planetary Geostrophic Growth Rate

The N-layer linearized PG equations are obtained from the N-layer linearized GPV equations (2.33) by taking the limit described in section 2a(1). The result is

$$\frac{d_n^0}{dt} \Delta_n^s (p_n) - \left(\frac{\beta \overline{h}_n^s}{F_0 f^2} + U_n^s \right) \frac{\partial p_n}{\partial x} = 0, \quad \text{(B.1)}$$

where Δ_n^s, $\overline{h}_n^s$, and U_n^s are described in (2.37). A wave solution of the form (2.18) is substituted into (B.1) to get the algebraic set of equations

$$[(c - U_n) \Delta_n^s + R_n] \tilde{p}_n = 0, \quad \text{(B.2)}$$

where

$$R_n = \frac{\beta \overline{h}_n^s}{F_0 f^2} + U_n^s.$$

For three layers, the operator in (B.2) is a 3×3 matrix,

the determinant of which must be zero for nontrivial solutions to exist. This condition yields a quadratic characteristic equation in the eigenvalue c. Instability thus occurs when the radicand is negative, that is, when $B^2 < 4AC$, where

$$A = -\frac{\beta}{F_0 f^2},$$

$$B = R_1(U_2 + U_3) + R_2(U_1 + U_3) + R_3(U_1 + U_2) - R_1R_2 - 2R_1R_3 - R_2R_3,$$

$$C = U_1R_2R_3 + 2R_1U_2R_3 + R_1R_2U_3 - U_1U_2R_3 - U_1R_2U_3 + R_1U_2U_3 - R_1R_2R_3, \quad \text{(B.3)}$$

and the imaginary part of the eigenvalue, c_i, is in that case

$$c_i = \frac{1}{2A}\sqrt{4AC - B^2}. \quad \text{(B.4)}$$

Note that the terms A, B, and C are functions of the meridional coordinate y, which is now a parameter. Therefore, motion is essentially decoupled in y (hence shocks are possible in the inviscid solution), and if *any* value of y within the domain yields an instability, then the fluid is unstable. Stability is, hence, dependent on the size of the domain relative to the typical horizontal length scale. In order to arrive at the value listed in (3.1), we select the maximum growth rate in the domain.

APPENDIX C

The FG Limit of L_1

That L_1 contains FG is perhaps unsurprising, given that a model closely related to L_1 is shown in its derivation by Salmon (1985) to satisfy an exact analogue of the FG equation in geostrophic coordinates. We nevertheless explicitly demonstrate the limit here in order to unambiguously show the validity of L_1 in this subset of parameter space. For the sake of simplicity, the following pertains to a single-layer formulation.

The FG regime is indigenous to the extratropical ocean only, where the Rossby number is small. The regime is valid for phenomena whose length scale, L, is small compared to the planetary scale, $L_\beta = f_0/\beta_0$, but whose *squared* external length scale, L^2, is large compared to the *squared* external radius of deformation, L_d^2 (where $L_d = \sqrt{g'H_0}/f_0$). Variations in the height field, but not the Coriolis parameter, are thus assumed to be $O(1)$. Additionally, one must assume that the evolution timescale is $O(\epsilon)$ smaller than the advective timescale [which is the natural result of the spatial scaling assumptions; see Cushman-Roisin (1986)]. In terms of the nondimensionalization used in this paper, the FG regime implies that $\beta \sim O(1)$ and $F \sim O(1/\epsilon)$. A proper asymptotic expansion in ϵ of the SW equations under these assumptions yields the FG model, which we write here as

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}_1) = 0, \quad \text{(C.1)}$$

where

$$\mathbf{u}_1 = J(\nabla h, h) - \beta y\hat{\mathbf{z}} \times \nabla h, \quad \text{(C.2)}$$

and the full velocity field is written as

$$\mathbf{u} = \hat{\mathbf{z}} \times \nabla h + \epsilon\mathbf{u}_1. \quad \text{(C.3)}$$

For the sake of simplicity, the FG limit of the L_1 equations is derived for a single layer under a free surface. To facilitate the asymptotics, the equations are nondimensionalized in a slightly different way than in all of the above treatments. The difference is that the velocity is scaled geostrophically, $U \sim gH_0/(f_0L)$, so that $F = 1/\epsilon$ exactly, while the timescale is *chosen* as T, rather than scaled by other parameters (in the previous cases, T is scaled as the advective timescale L/U, and U is chosen). This introduces the parameter $\omega = 1/f_0T$, which, due to the balance found a posteriori in the FG regime (Cushman-Roisin 1986), is essentially $O(\epsilon^2)$. Furthermore, with this scaling in the FG parameter regime, $\epsilon \ll 1$ and $\beta \sim O(1)$, and the L_1 equations for a single layer become

$$\omega\frac{\partial}{\partial t}\mathbf{u}_g + (f + \epsilon\zeta_g)\hat{\mathbf{z}} \times \mathbf{u} + \nabla B = 0 \quad \text{(C.4)}$$

$$\epsilon\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}) = 0, \quad \text{(C.5)}$$

where

$$\mathbf{u}_g = \frac{1}{f}\hat{\mathbf{z}} \times \nabla h, \quad \text{(C.6)}$$

$$B = \eta + \frac{\epsilon}{2}|\mathbf{u}_g|^2 + \epsilon\hat{\mathbf{z}} \cdot \nabla \times \left[\frac{h(\mathbf{u} - \mathbf{u}_g)}{f}\right], \quad \text{(C.7)}$$

$$\zeta_g = \frac{1}{f}\left(\nabla^2 h - \frac{\epsilon\beta}{f}\frac{\partial h}{\partial y}\right). \quad \text{(C.8)}$$

If we now expand $\mathbf{u}$ as a series in ϵ,

$$\mathbf{u} = \mathbf{u}_0 + \epsilon\mathbf{u}_1 + O(\epsilon^2),$$

and substitute into (C.4)–(C.8), we get geostrophic balance at $O(1)$,

$$\mathbf{u}_0 = \hat{\mathbf{z}} \times \nabla h, \quad \text{(C.9)}$$

while at $O(\epsilon)$ we find

$$(\beta y + \zeta_0)\hat{\mathbf{z}} \times \mathbf{u}_0 + \hat{\mathbf{z}} \times \mathbf{u}_1 + \frac{1}{2}\nabla(|\mathbf{u}_0|^2) = 0 \quad \text{(C.10)}$$

$$\frac{\partial h}{\partial t} + \nabla \cdot (h\mathbf{u}_1) = 0, \quad \text{(C.11)}$$

where $\zeta_0 = \nabla^2 h$. We can then rewrite (C.10) as

$$(\mathbf{u}_0 \cdot \nabla)\mathbf{u}_0 + \beta y\hat{\mathbf{z}} \times \mathbf{u}_0 + \hat{\mathbf{z}} \times \mathbf{u}_1 = 0, \quad \text{(C.12)}$$

which can be solved for $\mathbf{u}_1$ as

$$\mathbf{u}_1 = -(\mathbf{u}_0 \cdot \nabla)\mathbf{u}_0 - \beta y\hat{\mathbf{z}} \times \nabla h. \quad \text{(C.13)}$$

Then, using (C.9) the advection term can be rewritten as the Jacobian, $J(\nabla h, h)$, hence

$$\mathbf{u}_1 = J(\nabla h, h) - \beta y\hat{\mathbf{z}} \times \nabla h, \quad \text{(C.14)}$$

and the system is now identical to the FG equations, (C.1)–(C.3).

REFERENCES

Allen, J. S., and D. D. Holm, 1996: Extended-geostrophic Hamiltonian models for rotating shallow water motion. *Physica D,* **98,** 229–248.

——, J. A. Barth, and P. A. Newberger, 1990: On intermediate models for barotropic continental shelf and slope flow fields. Part I: Formulation and comparison of exact solutions. *J. Phys. Oceanogr.,* **20,** 1017–1042.

Bleck, R., 1973: Numerical forecasting based on the conservation of potential vorticity on isentropic surfaces. *J. Appl. Meteor.,* **12,** 737–752.

Colin de Verdière, A., 1986: On mean flow instabilities within the planetary geostrophic equations. *J. Phys. Oceanogr.,* **16,** 1981–1984.

Cushman-Roisin, B., 1986: Frontal geostropic dynamics. *J. Phys. Oceanogr.,* **16,** 132–143.

Green, J. S. A., 1960: A problem in baroclinic stability. *Quart. J. Roy. Meteor. Soc.,* **86,** 237–251.

Kuo, H. L., 1978: A two-layer model study of the combined barotropic and baroclinic instability in the tropics. *J. Atmos. Sci.,* **35,** 1840–1860.

Mundt, M. D., G. K. Vallis, and J. Wang, 1997: Balanced models and dynamics for the large-and mesoscale circulation. *J. Phys. Oceanogr.,* **27,** 1133–1152.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics.* Springer-Verlag, 710 pp.

Phillips, N. A., 1954: Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model. *Tellus,* **6,** 273–286.

Salmon, R., 1983: Practical use of Hamilton's principle. *J. Fluid Mech.,* **132,** 431–444.

——, 1985: New equations for nearly geostrophic flow. *J. Fluid Mech.,* **153,** 461–477.

Samelson, R. M., and G. K. Vallis, 1997: A simple friction and diffusion scheme for planetary geostrophic basin models. *J. Phys. Oceanogr.,* **27,** 186–194.

Vallis, G. K., 1996: Potential vorticity inversion and balanced equations of motion for rotating and stratified flows. *Quart. J. Roy. Meteor. Soc.,* **122,** 291–322.