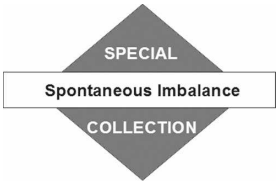

Decoupling of Balanced and Unbalanced Motions and Inertia–Gravity Wave Emission: Small versus Large Rossby Numbers

VLADIMIR ZEITLIN

Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, Institut Pierre-Simon Laplace, Paris, France

(Manuscript received 28 March 2007, in final form 19 November 2007)

ABSTRACT

This paper provides a brief review of recent results on decoupling of fast [inertia–gravity wave (IGW)] and slow (vortex) motions at small Rossby numbers obtained in the framework of the geostrophic adjustment of localized perturbations. Special attention is paid to the IGW emission and its interpretation in the context of "spontaneous imbalance." Several mechanisms that lead to spontaneous IGW emission and, thus, to violations of fast–slow splitting at large Rossby numbers are reviewed: Lighthill radiation, symmetric/inertial instability, and ageostrophic shear (Rossby–Kelvin) instability. New results on the saturation of symmetric instability and on the existence of Rossby–Kelvin instability in continuously stratified fluid are presented.

1. Introduction

The question of "spontaneous" emission of inertia–gravity waves (IGWs), which is being actively discussed in the literature, starting from the pioneering paper (O'Sullivan and Dunkerton 1995), in the context of realistic (e.g., Plougonven and Snyder 2007) or idealized (e.g., Dritschel and Viúdez 2007) numerical simulations, is a question of the limits of decoupling of fast (waves) and slow (vortices) motions in geophysical fluid dynamics and, therefore, a question of the limits of balanced models based on such decoupling. In what follows we will review some recent results on fast–slow decoupling and the absence of spontaneous IGW emission at small Rossby numbers Ro. We will, however, point out some phenomena that can be confused with spontaneous emission at small Ro. We then display dynamical mechanisms leading to violation of fast–slow decoupling and emission of IGWs at increasing Rossby numbers and attempt to classify the emission mechanisms on the basis of the fundamental parameters: Rossby and Burger numbers.

The natural context for studying fast–slow decoupling is geostrophic adjustment of localized disturbances. This fundamental process, ubiquitous in the atmosphere and oceans, consists of the relaxation of any perturbation toward the state of geostrophic equilibrium (balance). Such process is an obvious source of IGW because any initially unbalanced disturbance is getting rid of its unbalanced part by emitting waves. This IGW emission is primary. As will be shown below, a much weaker secondary emission also exists and may be quantified. At least at small Rossby numbers, the balanced (vortex) and unbalanced (IGW) parts of the flow are well identified and are well separated dynamically. However, the presence of the secondary IGW emission means that this separation is not complete. At large Rossby numbers, the dynamical separation of balanced and unbalanced motions may be strongly violated because of qualitatively new phenomena, such as ageostrophic instabilities of balanced flows or Lighthill radiation, both leading to IGW emission.

The paper is organized as follows: We first review the fast–slow decoupling in the framework of the geostrophic adjustment at small Rossby numbers in a hierarchy of GFD models of increasing complexity (rotating shallow water, multilayer rotating shallow water, and continuously stratified primitive equations), showing the limits of fast–slow separation in each regime, with special emphasis on the unbalanced part which, under some circumstances, may be misinterpreted as "spontaneously emitted" IGWs. We then give an overview of new dynamical mechanisms arising at large

Corresponding author address: Vladimir Zeitlin, Ecole Normale Supérieure, 24 rue Lhomond, 75231 Paris Cedex 05, France.
E-mail: zeitlin@lmd.ens.fr

DOI: 10.1175/2008JAS2481.1

Rossby numbers and their role in fast–slow coupling and IGW emission.

The presentation is mostly based on the works of the author with collaborators and is not an exhaustive review. The relevant references may be found in the papers cited below.

2. Nonlinear geostrophic adjustment and fast–slow splitting at small Ro

In this section we give a brief review of the analytic results on nonlinear geostrophic adjustment of localized initial disturbances at small Rossby numbers.

a. Analytic approach: The method

The method (Reznik et al. 2001; Zeitlin et al. 2003; see Reznik and Zeitlin 2007 for a review) consists in considering an initial-value problem for an arbitrary localized initial disturbance treated by perturbation theory in Rossby number Ro. The technique uses asymptotic expansions in Ro and time averaging. Smallness of the Rossby number is always assumed, which does not necessarily mean that nonlinearity is small. In fact, the nonlinearity may be arbitrary, provided that the geostrophic balance is verified at the lowest order. The typical spatial scales are either of the order of or much greater than the Rossby deformation radius R_d. Multiple time scales are used: $t_0 \sim (1/f)$, $t_1 \sim (1/\mathrm{Ro}f), \ldots$, where f is the mean value of the Coriolis parameter [the midlatitude tangent plane approximation is being used; for adjustment in the equatorial tangent plane, see Le Sommer et al. (2004)]. In this approach, the slow dynamics appears from the removal of resonances in the fast one by fast-time averaging.

The nonlinear geostrophic adjustment of localized (finite energy) initial perturbations was considered by this method within the following models of increasing complexity:

- rotating shallow water (RSW),
- 2-layer RSW (2RSW) with a flat bottom and rigid lid boundary conditions, and
- hydrostatic primitive equations (HSPEs) with a flat bottom and rigid lid boundary conditions.

The equations of the models are given in the appendix for completeness. The following key parameters govern the dynamics in all models:

- the Rossby number $\mathrm{Ro} = (U/fL)$, where U and L are typical velocity and horizontal scales, respectively;
- the nonlinearity λ [i.e., the normalized typical deviation of the isopycnal (ocean) or isentropic (atmosphere) surfaces from the equilibrium positions]; and
- the Burger number $\mathrm{Bu} = (R_d^2/L^2)$, where R_d is the deformation radius.

These parameters obey the following relation, which guarantees the geostrophic balance in the lowest order in Ro:

$$\lambda \mathrm{Bu} = O(\mathrm{Ro}). \tag{1}$$

The following dynamical regimes are thus possible:

- the quasigeostrophic (QG) regime, in which $\mathrm{Ro} = \epsilon(\text{small})$, $\mathrm{Bu} = O(1)$ with small deviations of the isopycnal (isentropic) surfaces, and
- the frontal geostrophic (FG) regime, in which $\mathrm{Ro} = \epsilon(\text{small})$, $\mathrm{Bu} = O(\mathrm{Ro})$ with large deviations of the isopycnal (isentropic) surfaces.

Note that a typical value of Ro for synoptic motions in the atmosphere is ~0.1 (Holton 1992).

b. Fast–slow splitting in RSW

In this subsection we present the results (Reznik et al. 2001) of the above-described approach applied to the RSW model. They corroborate the results of earlier studies on fast–slow splitting (Warn et al. 1995; Embid and Majda 1996; Babin et al. 1997a), with notable differences of (i) considering localized initial conditions, (ii) going farther in asymptotic expansion in Ro, and (iii) considering the FG regime as well as the standard QG. The advantage of using the open boundary conditions and working with localized initial disturbances in all models below is that, in addition to allowing us to assess the radiation efficiency of the adjustment process, we get rid of most of the wave–wave resonances [studied in detail in Embid and Majda (1996) and Babin et al. (1997a)]. The only resonances which do give contributions below are those corresponding to so-called "catalytic" interactions (Lelong and Riley 1991).

1) THE QG REGIME

The balanced motions are described by the "improved" quasigeostrophic potential vorticity (PV) equations for the slow part of the elevation of the free surface in the first order in Ro $\overline{h} = \overline{h}_0 + \epsilon\overline{h}_1$:

$$\frac{D}{D\tau}[\nabla^2\overline{h} - \overline{h} - \epsilon\overline{h}(\nabla^2\overline{h} - \overline{h}) - \epsilon\nabla\overline{h}\cdot\nabla(\nabla^2\overline{h} - \overline{h}) - 2\epsilon J(\partial_x\overline{h}, \partial_y\overline{h})] = 0. \tag{2}$$

Here,

$$\frac{D}{D\tau}(\ldots) := \partial_\tau(\ldots) + J\left[\overline{h} - \epsilon\frac{(\nabla\overline{h})^2}{2}, \ldots\right], \tag{3}$$

where the slow-time derivative is $\partial_\tau = \partial_{t_1} + \epsilon\partial_{t_2}$, $J[\ldots,\ldots]$ denotes Jacobian, and the standard QG equation corresponds to $\epsilon = 0$.

The fast component is the inertia–gravity wave packet formed by the initial perturbation and described by the homogeneous wave equation for the fast component of the free surface elevation $\tilde{h}_0$, namely,

$$-\frac{\partial^2\tilde{h}_0}{\partial t^2} - \tilde{h}_0 + \nabla^2\tilde{h}_0 = 0 \tag{4}$$

in the lowest order, and an inhomogeneous linear initial-value problem in the next order, namely,

$$-\frac{\partial^2\tilde{h}_1}{\partial t^2} - \tilde{h}_1 + \nabla^2\tilde{h}_1 = \mathcal{F}^{(s)}(\mathbf{x})\mathcal{F}_0^{(f)}(\mathbf{x};t) + \mathcal{F}_1^{(f)}(\mathbf{x};t)\mathcal{F}_2^{(f)}(\mathbf{x};t), \tag{5}$$

where the script $\mathcal{F}$ with corresponding superscripts denote some slow and fast spatially localized functions, respectively. They are uniquely defined from the initial conditions and the lowest-order results (see Reznick et al. 2001 for details). The fast–slow (balanced–unbalanced) separation is unique and well defined. The initialization problem (i.e., the problem of projection of the initial conditions onto the slow manifold) is completely solved order by order in Ro.

Thus, in the leading order in Ro the standard QG dynamics for the balanced part of the initial perturbation is recovered, and IGWs run out of the location of the perturbation. Balanced and unbalanced motions are noninteracting. Once the first-order correction is taken into account, which is necessary for applying slow dynamics at times longer than 1/Rof, the standard QG for balanced motion is "improved." This may be interpreted as "waves" giving corrections to the slow manifold, although there is no fast-motion drag in the slow-motion dynamics. However, the first-order correction [the first term on the rhs of (5)] to the fast motion gives a secondary IGW emission due to fast–slow interaction.

A preliminary conclusion from this analysis is that although fast–slow (balanced–unbalanced) splitting is not complete, the effect of secondary emission is small at small Ro. To have such secondary emission, the fast component should be present in initial conditions. Its appearance, thus, is not spontaneous, although it may be confused with such if the initial conditions are not properly posed. It is to be noted that as in all other models/regimes considered below, the consistency of the approach relies on nongeneration of scales essentially smaller than L (leading to the violation of the small Rossby number hypothesis) in times less than $1/\mathrm{Ro}^2 f$ because of the intrinsic dynamics of the improved QG equation.

2) THE FG REGIME

The slow component of the pressure field evolves in the second slow time t_2 and obeys the equation

$$\partial_{t_2}\bar{h}_0 - J\left[\bar{h}_0, \bar{h}_0\nabla^2\bar{h}_0 + \frac{(\boldsymbol{\nabla}\bar{h}_0)^2}{2}\right] = 0. \tag{6}$$

In this case of small Burger numbers, the fast motion consists of inertial oscillations with $\omega \simeq f$. The envelope of the inertial oscillations $\mathcal{A}_0$ evolves according to the Schrödinger-like modulation equation

$$\partial_{t_1}\mathcal{A}_0 + J[\bar{h}_0, \mathcal{A}_0] - \frac{i}{2}[\nabla^2(\bar{h}_0\mathcal{A}_0) - \mathcal{A}_0\nabla^2\bar{h}_0] = 0, \tag{7}$$

where the coefficients depend on the balanced component. Hence, there is no drag of the fast motion upon the slow one, but the slow motion guides the envelope of inertial oscillations. This means that the fast–slow separation is, again, not complete. The waves "feel" the slow motion as they are guided and possibly focused at some locations determined by the slow field. Again, the manifestations of the fast component (e.g., by amplification in some locations prescribed by the slow field) are not spontaneous because they should be present in the initial conditions at some level. However, the focusing phenomenon may intensify the inertial oscillations field. If, for example, it was initially beyond the detection threshold, focusing may make it detectable in some locations, which may be confused with spontaneous emission.

c. *Fast–slow splitting in 2RSW*

In this section we present the results (Zeitlin et al. 2003) of the application of the general method in the case of two-layer RSW model, which is the simplest one taking into account the baroclinic effects.

1) THE QG REGIME

The slow and the fast component of motion are (almost) decoupled. The slow dynamics is described by the improved quasigeostrophic PV equations for the slow parts of the pressure in respective layers in the first order in Ro: $\overline{\pi}_i = \overline{\pi}_i^{(0)} + \epsilon\overline{\pi}_i^{(1)}$, $\overline{\eta} = \eta^{(0)} + \epsilon\eta^{(1)}$, with $\overline{\pi}_2 - \overline{\pi}_1 = \overline{\eta}$:

$$\frac{D_i}{D\tau}\{\nabla^2\overline{\pi}_i + (-1)^{i+1}\overline{h}_{i+1}\overline{\eta} + \epsilon(-1)^{i+1}\overline{h}_{i+1}\overline{\eta} \times [\nabla^2\overline{\pi}_i + (-1)^{i+1}\overline{h}_{i+1}\overline{\eta}] - \epsilon\boldsymbol{\nabla}\overline{\pi}_i \cdot \boldsymbol{\nabla}[\nabla^2\overline{\pi}_i + (-1)^{i+1}\overline{h}_{i+1}\overline{\eta}] - 2\epsilon J[\partial_x\overline{\pi}_i, \partial_y\overline{\pi}_i]\} = 0, \tag{8}$$

where $\overline{h}_i = (H_i/H_0)$ ($H_0 = H_1 + H_2$, H_i; $i = 1, 2$) are unperturbed depths and

$$\frac{D_i}{D\tau}(\ldots) := \partial_\tau(\ldots) + J\left[\overline{\pi}_i - \epsilon\frac{(\nabla\overline{\pi}_i)^2}{2}, \ldots\right], \quad i = 1, 2. \tag{9}$$

The fast component is the IGW packet formed by the initial conditions and obeying the homogeneous wave equation

$$-\frac{\partial^2\tilde{\eta}^{(0)}}{\partial t^2} - \tilde{\eta}^{(0)} + \nabla^2\tilde{\eta}^{(0)} = 0 \tag{10}$$

at the lowest order. At the next order, the correction to $\tilde{\eta}^{(0)}$ obeys an inhomogeneous equation of the same type as (5) with the rhs defined entirely by the fast and slow motions of the previous approximation and the initial conditions. The same qualitative conclusions as in the QG regime in RSW, thus, hold in this case. Note that the phenomenon of classical baroclinic instability is intrinsically present in Eqs. (8). Thus, if the single typical spatial scale hypothesis is not violated during the development of the baroclinic instability, and hence (8) is still applicable, no IGW emission is expected at time scales up to $1/\mathrm{Ro}^2 f$.

2) THE FG REGIME

Fast motion consists of inertial oscillations with $\omega \simeq f$. They split out from the slow motion and exert no drag upon this latter. The slow motion concerns both vortex and wave components and consists of the evolution of the barotropic (P) and baroclinic (η) pressure fields and the slow modulation of the envelope of the inertial oscillations $\mathcal{A}$. Two subregimes are to be distinguished because an additional parameter, the ratio of the layer's depths $d = H_1/H_2$, appears in the problem.

Layers with depths of the same order $d = O(1)$:

$$\partial_{t_1}\eta = \frac{1}{1 + d^{-1}}J[\eta, P], \tag{11}$$

$$\partial_{t_1}\nabla^2 P + \frac{1}{1 + d^{-1}}(J[P, \nabla^2 P] + \nabla\cdot\{(1 - \eta) \times (d^{-1} + \eta)J[\eta, \nabla\eta]\}) = 0, \tag{12}$$

$$(1 + d^{-1})\partial_{t_1}\mathcal{A} + J[P - \eta^2 - (d^{-1} - 1)\eta, \mathcal{A}] + \frac{i}{2}\{\nabla^2[P - \eta^2 - (d^{-1} - 1)\eta] + (\boldsymbol{\nabla}\eta)^2\}\mathcal{A} - \frac{i}{2}\nabla^2[(1 - \eta)(d^{-1} + \eta)\mathcal{A}] = 0. \tag{13}$$

Thin upper layer $d = O(\epsilon^2)$:

$$\partial_{t_2}\eta + J[P, \eta] + J\left[\eta, (1 - \eta)\nabla^2\eta - \frac{(\nabla\eta)^2}{2}\right] = 0, \tag{14}$$

$$\partial_{t_2}\nabla^2 P + J[P, \nabla^2 P] - J\left[\eta, (1 - \eta)\nabla^2\eta + \frac{(\nabla\eta)^2}{2}\right] = 0, \tag{15}$$

$$\partial_{t_1}\mathcal{A} - J[\eta, \mathcal{A}] - \frac{i}{2}\nabla^2\mathcal{A} + \frac{i}{2}[\nabla^2(\eta\mathcal{A}) - \mathcal{A}\nabla^2\eta] = 0. \tag{16}$$

The modulation of the wave field is thus guided by the pressure field in both subregimes, whereas the pressure evolves without feeling any influence of the waves. The same conclusion as in the FG regime in RSW is valid for the second subregime (layers of essentially different depth), namely, that guiding of the envelope of the fast motion by the slow one may give amplification of the former in specific locations. The dynamical separation of slow and fast components is, thus, incomplete, in spite of the absence of any fast-motion drag on the slow one. The manifestations of the fast motion are, nevertheless, not spontaneous in the sense that the presence of the fast component in the initial conditions is required. An even more drastic "non-decoupling" of slow and fast components arises in the first subregime (layers of comparable depth). Although there is still no drag of the fast motion on the slow one, the envelope of the inertial oscillations evolves at the same slow time scale as the underlying (guiding) slow field. Again, the manifestations of the fast field on the background of the slow one are not spontaneous, although the focusing of the inertial oscillations field due to guiding may lead to confusion with such, as in the RSW case.

d. Fast–slow splitting in HSPE

In this section the results (Zeitlin et al. 2003) of the application of the general method to hydrostatic primitive equations are presented. Again, they corroborate the earlier results (Embid and Majda 1998; Babin et al. 1997b), with the distinction of treating localized initial conditions, and thus wave emission, including higher-order corrections in Ro, and treating the FG regime.

1) THE QG REGIME

An arbitrary initial perturbation can be decomposed in a unique way into a slow and fast component. The latter is an IGW field emitted by initial disturbance and

obeying, at the leading order, the three-dimensional wave equation

$$\partial_{tt}^2\partial_z\left(\frac{1}{N^2}\,\partial_z\tilde{p}^{(0)}\right)+\partial_z\left(\frac{1}{N^2}\,\partial_z\tilde{p}^{(0)}\right)+\nabla_h^2\tilde{p}^{(0)}=0 \quad (17)$$

with well-defined initial conditions. Here, $N(z)$ is the vertical profile of the Brunt–Väisälä frequency. The first-order correction to the fast wave field is governed by the inhomogeneous Eq. (17), with the rhs defined by self-interaction of the leading-order wave field and its interaction with the leading-order balanced field.

The slow field experiences no drag from the fast field and obeys the improved QG equation

$$\frac{D}{D\tau}\Bigg[\partial_z\left(\frac{1}{N^2}\,\partial_z\overline{p}\right)+\nabla_h^2\overline{p}-\epsilon 2J[\partial_x\overline{p},\partial_y\overline{p}]+\frac{\epsilon}{N^2}\,(\partial_{zz}^2\overline{p}\nabla_h^2\overline{p}-(\partial_{zx}^2\overline{p})^2-(\partial_{zy}^2\overline{p})^2-\sigma(\partial_z\overline{p})^2$$
$$-\nabla_N\overline{p}\cdot\nabla\left\{N^2\left[\nabla_h^2\overline{p}+\partial_z\left(\frac{1}{N^2}\,\partial_z\overline{p}\right)\right]\right\}\Bigg)\Bigg], \quad (18)$$

where $D/D\tau$ is the advective derivative corresponding to the "full" velocity field $\overline{\mathbf{v}}^{(0)}+\epsilon\overline{\mathbf{v}}^{(1)}$:

$$\frac{D}{D\tau}\ldots=\partial_\tau\ldots+J\left[\overline{p}-\frac{\epsilon}{2}\,\boldsymbol{\nabla}\overline{p}\cdot\nabla_N\overline{p},\ldots\right]. \quad (19)$$

Here, $\boldsymbol{\nabla}_N=(\boldsymbol{\nabla}_h,(1/N^2)\partial_z)$, $\sigma=(1/2N^2)(\log N^2)''$, and the full pressure field $\overline{p}=\overline{p}^{(0)}+\epsilon\overline{p}^{(1)}$ and full slow time $\partial_\tau=\partial_{t_1}+\epsilon\partial_{t_2}$ were introduced. These equations extend the domain of validity of the standard QG equation to the times of the order $(1/\mathrm{Ro}^2 f)$.

The same conclusions as in the RSW and 2RSW cases concerning fast–slow splitting and IGW emission hold in this case.

2) The FG regime

The envelope of the inertial oscillations representing the fast motion in this regime obeys the three-dimensional modulation equation

$$\partial_{t_1}\partial_{zz}^2\mathcal{M}+J[\overline{p},\partial_{zz}^2\mathcal{M}]+i\,\frac{\nabla_h^2\overline{p}}{2}\,\partial_{zz}^2\mathcal{M}-\frac{i}{2}\,(\partial_z\overline{p})\nabla_h^2\mathcal{M}$$
$$-i[\partial_z(\partial_x+i\partial_y)\overline{p}][\partial_z(\partial_x-i\partial_y)\mathcal{M}]=0, \quad (20)$$

in which we introduce the primitive with respect to z of the modulated amplitude $\mathcal{A}$ of the inertial oscillations

$$\partial_z\mathcal{M}=\mathcal{A}, \quad (21)$$

and the slow component of the pressure field is denoted by $\overline{p}$. The inertial oscillations exert no drag upon the slow component, which obeys the following equations:

$$\int_{-1}^{0}dz(\partial_{t_1}\nabla^2\overline{p}+J[\overline{p},\nabla^2\overline{p}])=0 \quad \text{and} \quad (22)$$

$$\partial_{t_1 z}\overline{p}+J[\overline{p},\partial_z\overline{p}]=0. \quad (23)$$

The time scales of the slow evolution of the inertial oscillation envelope and of the slow vortex field are, thus, the same, and there is no time-scale separation between the modulation of inertial oscillations and the guiding pressure field, unlike in the RSW FG regime. The FG regime in HSPEs, thus, resembles the FG regime with layers of comparable depth in 2RSW, with similar conclusions.

e. Discussion

We have seen in this section that although traditional balanced QG and FG models are proved to be self-consistent because of the absence of fast-motion drag on the slow motion (however, corrections leading to the improved versions are necessary to safely apply the models at longer time scales), the fast motion itself does feel the slow component. Both secondary emissions and guiding phenomena can lead to manifestations of wave activity in some locations at some times. To distinguish them from truly spontaneous emissions (see below), an accurate analysis of initial conditions and thorough identification of dynamical regimes are needed. The persistence of the quasi-inertial oscillations highlighted in the FG regimes in all models has a simple physical explanation: the group velocity of quasi-inertial waves is close to zero, and they are difficult to evacuate. This was repeatedly observed in both direct numerical simulations (e.g., Bouchut et al. 2004) and in experimental studies of the geostrophic adjustment (Stegner 2007).

3. Obstacles for the fast–slow splitting at large Ro and IGW emission

a. Lighthill radiation

The following observation is almost trivial: any motion that stirs the fluid at frequencies larger than the IGW threshold frequency f cannot avoid generating waves. This is also true for vortex motions, but the

Rossby number should necessarily exceed unity in this case. Wave emission by vortices in the acoustic context was considered in the pioneering paper by Lighthill (1952). We remind below of the Lighthill radiation mechanism in nonrotating fluids in the technically simplest case of point vortices (following Gryanik 1983) and then make extension to the rotating case (following Zeitlin 2007). A more elaborated study with similar results was done in Ford et al (2000). The singular character of point vortices is not important in this context; basically the same results can be obtained for distributed vorticity configurations, like the elliptic Kirchhoff vortex (Zeitlin 1991). The extension from shallow-water to full primitive equations is also rather straightforward (Plougonven and Zeitlin 2002).

1) The nonrotating case

If the free surface of shallow water is "frozen" (i.e., $\partial_t h = 0$), the motion obeys the two-dimensional Euler equations, which are equivalent to the equation for vorticity $\zeta = \partial_x v - \partial_y u$:

$$\partial_t \zeta + J[\eta, \zeta] = 0, \quad \zeta = \nabla^2 \eta, \tag{24}$$

where η is the streamfunction: $\partial_x \eta = v$, $\partial_y \eta = -u$. The elementary solutions of this equation in the unbounded domain are point vortices. A pair of vortices with intensities $\kappa_{1,2}$ at a distance a is the simplest nontrivial exact solution of the vorticity equation. The vortices rotate around the center of vorticity with a constant angular velocity $\Omega = (\kappa_1 + \kappa_2)/2\pi a^2$, where a is the distance between the vortices. The following multipolar expansion of the complex velocity potential of the vortex pair at $r \gg a$ may be easily obtained:

$$\phi_2|_{r\to\infty} = i\frac{\kappa_1 + \kappa_2}{2\pi}\log(re^{i\theta}) - \frac{i}{4\pi}\frac{\kappa_1\kappa_2}{\kappa_1 + \kappa_2} \times \left(\frac{a}{r}\right)^2 e^{2i(\Omega t - \theta)} + \ldots. \tag{25}$$

The vortex-pair solution is relevant for the full shallow-water equations, provided the distance a between the vortices is much smaller than the typical wavelength λ of the gravity waves. In the context of the gravity wave emission by the pair, the frequency of the wave is necessarily Ω. Therefore, the wavelength is $\lambda = (\Omega/c_0)$, where $c_0 = \sqrt{gH_0}$ is the phase velocity of the gravity waves, and the characteristic Froude (Mach) number of the pair is $M = (a/\lambda) = (a\Omega/c_0)$.

The gravity waves are solutions of the equation

$$\partial_{tt} h - c_0^2 \nabla^2 h = 0. \tag{26}$$

In the absence of rotation the waves are strictly dispersionless. The physical picture of the presumable wave emission by the vortex pair is as follows: Close to the origin, where the pair is situated (the inner region), the vortex-pair solution is a good approximation for small M. Further off (the intermediate region) the vortex solution should be matched with the solution of the shallow-water (SW) equation with a nonfrozen free surface. Because the time-dependent part of the vortex solution rapidly decays with the distance to the origin, it is reasonable to take this full solution at a small amplitude (i.e., in the form of gravity waves). In the outer region far from the vortex, the solution is the outgoing gravity wave field carrying the energy toward infinity. If matching satisfying these conditions is possible, then wave emission by the vortex is taking place.

Consider the wave solution in the outer region. In polar coordinates, the solution of the wave equation (26) can be sought in the form

$$h(r, \theta, t) = \sum_n h_n(r) e^{in(\Omega t - \theta)}, \tag{27}$$

and in terms of the new variable $\rho = (n\Omega/c_0)r$, Eq. (26) takes the form of the Bessel equation:

$$h_n'' + \frac{1}{\rho}h_n' + \left(1 - \frac{n^2}{\rho^2}\right)h_n = 0. \tag{28}$$

The pressure is proportional to the time derivative of the complex potential for the wave field. Hence, the matching condition is

$$gh^{(\text{waves})}|_{r\to 0} = -\partial_t \phi_2{}^{(\text{vortex})}|_{r\to\infty}. \tag{29}$$

Only the $n = 2$ harmonics should be present in the wave field, and the following choice of its amplitude ensures the matching (29):

$$gh^{(\text{waves})} = -\frac{i}{8\pi^3 c_0}\frac{\kappa_1\kappa_2(\kappa_1 + \kappa_2)^2}{a^4} H_2^{(2)}\left(\frac{2\Omega}{c_0} r\right) e^{2i(\Omega t - \phi)}. \tag{30}$$

Here, $H_n^{(2)}$ is the Hankel function of the second kind, a solution of (28) corresponding to the outward energy flux.

2) The effects of rotation

In the rotating case at large Burger numbers, the vortex dynamics are still described by the standard vorticity equation. Hence, if the same reasoning as above is to be applied to the rotating case, the vortex part is the same if its characteristic scale is much less than the deformation radius. The ratio of the vortex pair fre-

quency to $f/2$ gives the Rossby number of the system: $\mathrm{Ro}_2 = 2\Omega/f$. For the wave part, the linearized RSW equations are equivalent to the single equation for h:

$$-\frac{\partial^2 h}{\partial t^2} - \frac{1}{R_d^2} h + \nabla^2 h = 0. \tag{31}$$

If solutions of this system are sought in the form $h(r, \theta, t) = \Sigma_n h_n(r) e^{in(\Omega t - \theta)}$, the Bessel equation again results, with a modified independent variable ρ:

$$\rho = \frac{n\Omega}{c_0} r\sqrt{\mathrm{Ro}_n^2 - 1}; \quad \mathrm{Ro}_n = \frac{n\Omega}{f}, \tag{32}$$

where the Rossby number Ro_n corresponding to a given wave mode is introduced. An immediate consequence of this formula is that propagating waves exist only for $\mathrm{Ro}_n > 1$. A second conclusion is that the same procedure as in the nonrotating case above (up to a constant) will allow us to match the near wave field with the far field of the vortex at $a \ll r \ll R_d$, and thus obtain the wave radiation from the vortex. Therefore, the necessary condition of the wave radiation is $\mathrm{Ro} > 1$; $\mathrm{Bu} \gg 1$ is a sufficient condition. The condition $M \ll 1$ of the nonrotating case becomes $\mathrm{Bu} \gg \mathrm{Ro}_2^2$.

We therefore see that nonstationary vortex motions do spontaneously emit waves via the Lighthill radiation mechanism. This happens, however, in the region of parameters $\mathrm{Ro} > 1$, $\mathrm{Bu} \gg 1$, far from that considered in the previous section. In addition, in this parameter regime it is the cyclogeostrophic (gradient wind) balance, rather than geostrophic one, that is relevant for localized vortex systems. Note also that the intensity of the radiation is weak, $\sim M^4$. An important consequence of Lighthill radiation is its backreaction (Gryanik 1983; Zeitlin 1991; Plougonven and Zeitlin 2002), leading to a slow change of vortex parameters due to the radiative energy loss.

b. *High Ro and baroclinicity: Symmetric/inertial instabilities of geostrophic motions*

We show in this and the following subsections that in the presence of baroclinic effects, even in their simplest form, the balanced motions at high Ro may be subject to instabilities with respect to unbalanced perturbations. The subsequent development of such instabilities gives rise to secondary IGW emission. The first example is symmetric/inertial instability. We consider it in the "dry" case. Although it is known that dry symmetric instability is hard to achieve in midlatitudes (Bennets and Hoskins 1979), it gives an interesting example of ageostrophic instability of the balanced motion arising at large Ro. A peculiar feature of the inertial instability is its relation with the fast modes trapped by the slow one, which are destabilized beyond the threshold of the instability.

1) Trapped waves and symmetric instability in 2RSW

We consider first the (dry) symmetric instability in the framework of the 2RSW model. By definition, this is an instability which is quasi one-dimensional. Hence, following Le Sommer et al. (2003), we consider the 2RSW equations with no y dependence:

$$\partial_t u_i + u_i \partial_x u_i - f v_i + \rho_i^{-1} \partial_x \pi + \delta_{i2} g' \partial_x \eta = 0, \tag{33a}$$

$$\partial_t v_i + u_i (f + \partial_x v_i) = 0, \quad \text{and} \tag{33b}$$

$$\partial_t [H_i - (-1)^{i+1}\eta] + \partial_x \{u_i [H_i - (-1)^{i+1}\eta]\} = 0, \quad i = 1, 2. \tag{33c}$$

Here, δ_{ij} is Kronecker delta and g' is the reduced gravity: $g' = (\rho_2 - \rho_1)/\rho_2$.

Stationary solutions corresponding to the geostrophic equilibria are the exact solutions of the full nonlinear equations and are given by

$$v_i = \frac{1}{f\rho_i} \partial_x \pi + \delta_{i2} \frac{g'}{f} \partial_x \eta, \quad i = 1, 2. \tag{34}$$

In the case of barotropic and/or baroclinic pressure distributions with compact support variations, the balanced jet solutions arise.

A new phenomenon appears at large enough Ro because of the baroclinic effects: a part of initial perturbation can be trapped inside the jet in the form of fast oscillations. The first consequence of this fact is that geostrophic adjustment becomes incomplete if the initial perturbation has a projection on such trapped modes. But even more drastic effects may arise because for anticyclonic shears with high enough Rossby numbers the eigenfrequencies of the trapped modes can become imaginary, thus leading to instability. This instability, which is called symmetric because of the absence of y dependence, is an instability of a balanced flow with respect to unbalanced perturbations and thus destroys the fast–slow decoupling.

To demonstrate this mechanism, let us consider a balanced jet. The jet is characterized by its velocity V_{ig} in each layer, the barotropic pressure Π, and the variable interface height

$$V_{ig} = (\delta_{i1} + \delta_{i2} r) \partial_x \Pi + \delta_{i2} \mathrm{Bu} \partial_x h_{2g}, \, i = 1, 2, \tag{35}$$

where we use the nondimensional form of equations and introduce the Burger number $\mathrm{Bu} = [g' \min(H_1, H_2)/f^2 L^2]$

and the ratio of densities $r = (\rho_1/\rho_2)$. Note that $h_{1g} + h_{2g} = 1$.

Linearizing about the balanced jet state and introducing a new variable $U = h_{2g}u_2 = -h_{1g}u_1$ allows us to reduce the system to a single equation:

$$\mathrm{Bu}\partial_{xx}^2 U - \left[\left(\frac{rh_{2g} + h_{1g}}{h_{1g}h_{2g}}\right)(\partial_{tt}^2 + 1) + r\partial_{xx}^2 \Pi\left(\frac{1}{h_{1g}h_{2g}}\right) + \mathrm{Bu}\,\frac{\partial_{xx}^2 h_{2g}}{h_{2g}}\right]U = 0. \tag{36}$$

It is easy to check that in the absence of the jet, the equation is reduced to that of free inertia–gravity waves at the interface.

The trapped modes necessarily have frequencies below the inertial frequency f. By introducing the following auxiliary functions

$$F(x) = \frac{rh_{2g} + h_{1g}}{h_{1g}h_{2g}}, G(x) = r\partial_{xx}^2 \Pi\left(\frac{1}{h_{1g}h_{2g}}\right) + \mathrm{Bu}\,\frac{\partial_{xx}h_{2g}}{h_{2g}}, \tag{37}$$

and the Fourier transform of $U(x, t) = \int d\omega \tilde{U}(\omega, x)e^{-i\omega t}$ + c.c. (where c.c. stands for complex conjugate), we get, for each of the Fourier components $\tilde{U}(\omega, x)$,

$$Bu\partial_{xx}^2 \tilde{U} - (F(1 - \omega^2) + G)\tilde{U} = 0. \tag{38}$$

After multiplication by $\tilde{U}^*$ and integration over the whole domain, the following estimate for the frequencies ω of the localized modes results:

$$\omega^2 = 1 + \frac{\mathrm{Bu}\int |\partial_x \tilde{U}|^2\,dx + \int G|\tilde{U}|^2\,dx}{\int F|\tilde{U}|^2\,dx}. \tag{39}$$

Here F is, by definition, positive whereas G may be negative, particularly in the anticyclonic regions where $\partial_{xx}^2 \Pi < 0$. Hence, subinertial frequencies and trapped waves are possible if the anticyclonic shear is strong enough.

To illustrate the existence of the trapped modes and their transformation into unstable modes with increasing anticyclonic shear, we choose the simplest case of the barotropic jet with no displacement of the interface. Thus, Eq. (38) for $\tilde{U}$ becomes

$$\partial_{xx}^2 \tilde{U} + \frac{1}{\mathrm{Bu}}\{\omega^2 H_e^{-1} - [H_e^{-1} + (H_1H_2)^{-1}r\partial_{xx}^2 \Pi]\}\tilde{U} = 0. \tag{40}$$

This is the quantum mechanical Shrödinger equation

$$\partial_{xx}^2 \psi + [E - V(x)]\psi = 0 \tag{41}$$

for a particle having the energy $E = \omega^2(H_e\mathrm{Bu})^{-1}$ and moving in the potential $V(x) = \mathrm{Bu}^{-1}[H_e^{-1} + (H_1H_2)^{-1}r\partial_{xx}^2 \Pi]$. It is known that in the case of a potential well there are both propagating solutions corresponding to the continuous spectrum $\omega^2 \geq 1$ and localized (trapped in the well) solutions corresponding to the discrete spectrum $\min [V(x)]H_e < \omega^2 < 1$. The potential well is the region of anticyclonic shear. Therefore the trapped modes are localized there, oscillating at subinertial frequencies. It is worth noting that small Burger numbers makes the trapping easier.

If the potential is deep enough, that is, for strong enough anticyclonic shears (i.e., large enough Ro), nonoscillatory unstable modes with $\omega^2 < 0$, appear, producing the instability.

2) A Lagrangian approach to two-layer RSW equations

It is possible (Le Sommer et al. 2003) to eliminate the pressures $\pi_1 = \pi$, $\pi_2 = \pi + g'\eta$ in both layers in (33a)–(33c):

$$\frac{\partial \pi_1}{\partial x} = \left(\frac{h_1}{\rho_1} + \frac{h_2}{\rho_2}\right)^{-1}\left[f(h_1\upsilon_1 + h_2\upsilon_2) - \frac{\partial}{\partial x}(h_1u_1^2 + h_2u_2^2) - \frac{gh_2}{\rho_2}\frac{\partial}{\partial x}(\rho_1h_1 + \rho_2h_2)\right], \tag{42}$$

$$\frac{\partial \pi_2}{\partial x} = \left(\frac{h_1}{\rho_1} + \frac{h_2}{\rho_2}\right)^{-1}\left[f(h_1\upsilon_1 + h_2\upsilon_2) - \frac{\partial}{\partial x}(h_1u_1^2 + h_2u_2^2) + \frac{gh_1}{\rho_1}\frac{\partial}{\partial x}(\rho_1h_1 + \rho_2h_2)\right]. \tag{43}$$

By using (42) and (43), the system is reduced to four equations for four independent variables u_2, h_2, υ_2, and υ_1 [i.e., lower (heavier)-layer variables plus upper-layer jet velocity]:

$$\frac{\partial u_2}{\partial t} + u_2\frac{\partial u_2}{\partial x} - f\upsilon_2 + \frac{\rho_1}{\rho_2h_1 + \rho_1h_2}\left[f(h_1\upsilon_1 + h_2\upsilon_2) - \frac{\partial}{\partial x}(h_1u_1^2 + h_2u_2^2) + \frac{g\Delta\rho}{\rho_1}h_1\frac{\partial h_2}{\partial x}\right] = 0, \tag{44}$$

$$\frac{\partial h_2}{\partial t} + u_2\frac{\partial h_2}{\partial x} + h_2\frac{\partial u_2}{\partial x} = 0, \tag{45}$$

$$\frac{\partial \upsilon_2}{\partial t} + u_2\frac{\partial \upsilon_2}{\partial x} + fu_2 = 0, \quad \text{and} \tag{46}$$

$$\frac{\partial \upsilon_1}{\partial t} + u_2\frac{\partial \upsilon_1}{\partial x} + (u_1 - u_2)\frac{\partial \upsilon_1}{\partial x} + fu_1 = 0, \tag{47}$$

where

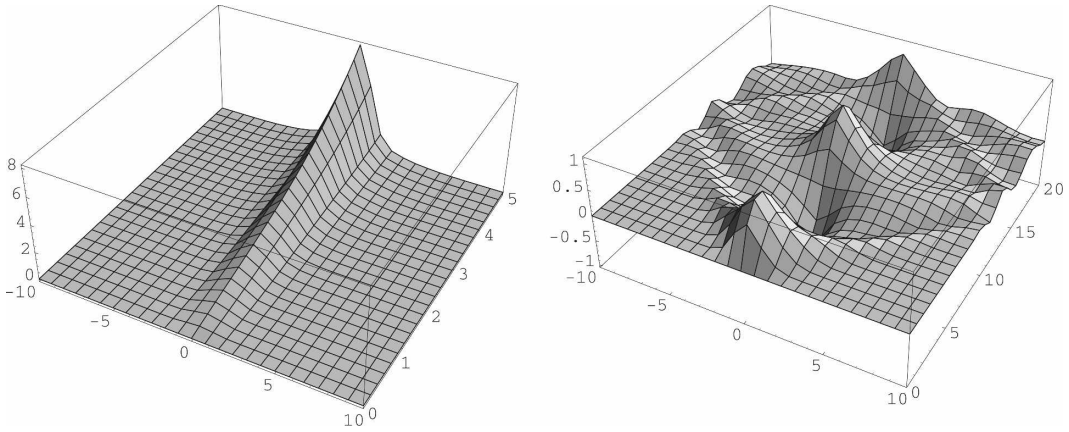

FIG. 1. (left) Linear and (right) nonlinear simulations of the evolution of symmetric instability in 2RSW with initial condition $\sim\exp(-x^2)$ for the Lagrangian displacements $\phi(x, t)$. Spatiotemporal evolution of $\phi(x, t)$ in the domain $-10 < x < 10$, $0 < t < 5$ with periodic boundary conditions in x is shown.

$$u_1 = \frac{h_2 u_2}{h_2 - H}, \quad h_1 = H - h_2. \tag{48}$$

We start from the system (44)–(48), taken for simplicity in the limit $r \to 1$, and introduce the Lagrangian coordinate $X(x, t)$ corresponding to the positions of the fluid particles in the lower layer. In terms of displacements ϕ with respect to initial positions $X(x, t) = x + \phi(x, t)$, the corresponding Lagrangian derivative is $(d/dt) = (\partial/\partial t) + u_2(\partial/\partial x)$. The dependence of the height variable h_2 on the Lagrangian labels and the transformation of its derivatives are obtained via mass conservation in the lower layer: $h_{2I}\, dx = h_2[X(x, t), t]dX$. Equation (23) expresses the conservation of the geostrophic momentum in the lower layer and allows us to eliminate v_2 in terms of ϕ and its initial value; thus,

$$v_2(x, t) + f\phi(x, t) = v_{2I}(x). \tag{49}$$

For simplicity, we will consider again the particular case of the barotropic initial flow with $h_{2_I} = H_2 = \text{const}$, $v_{2_I} = v_{1_I} = v_I(x)$. By introducing the notation $\alpha_1 = (H_1/H_0)$, $\alpha_2 = (H_2/H_0)$, $\alpha_1 + \alpha_2 = 1$, we have, after an obvious nondimensionalization:

$$\ddot{\phi} + \phi\frac{\alpha_1 + \phi'}{1 + \phi'} + \frac{\alpha_1 + \phi'}{1 + \phi'}(v_1 - v_I) - \frac{\alpha_2}{1 + \phi'}\left[\frac{\dot{\phi}^2}{(\alpha_1 + \phi')}\right]' - \gamma\frac{1}{(1 + \phi')^4}\phi'' = 0, \tag{50}$$

$$\dot{v}_1 - \frac{1}{\alpha_1 + \phi'}\dot{\phi}v_1' - \frac{\alpha_2}{\alpha_1 + \phi'}\dot{\phi} = 0, \tag{51}$$

where $\gamma = \text{Bu}$, and $\epsilon = \text{Ro}$ was absorbed in v_1. The equations are to be solved with initial conditions $\phi(x, 0) = 0$, $\dot{\phi}(x, 0) = u_{2_I}$, $v_1(x, 0) = v_I$.

3) NONLINEAR SATURATION OF SYMMETRIC INSTABILITY AND SECONDARY IGW EMISSION

Equations (50) and (51) may be directly solved numerically. We present here simulations with a special profile $v_I = -2\cosh^{-2}(2x)$, allowing us to get explicit eigenfunctions of the linearized problem and to compare linear and fully nonlinear simulations in Fig. 1. The nonlinear saturation of the symmetric instability with secondary emission of IGWs is visible in the right panel of Fig. 1.

Thus, at large Ro and small Bu the part of the IGW spectrum is trapped in the anticyclonic shear of the balanced flow. At large enough Ro, the gravest modes destabilize and lead to the development of symmetric–inertial instability, which in turn leads to a secondary emission of IGWs that may be considered spontaneous.

4) TRAPPED MODES AND SYMMETRIC INSTABILITY IN HSPE

The Lagrangian approach is also efficient for studying symmetric–inertial instability in the continuously stratified case. For this, we consider the HSPE without y dependence, taken between the flat top ($z = H$) and bottom ($z = 0$). Following Plougonven and Zeitlin (2005), we rewrite the HSPE in terms of Lagrangian positions $X(x, z, t)$, $Z(x, z, t)$ of the fluid parcels initially situated at x, z in the vertical plane:

$$\frac{\partial(X, \ddot{X})}{\partial(x, z)} - \frac{\partial(X, fM_I)}{\partial(x, z)} + \frac{g}{\theta_r}\frac{\partial(\theta_I Z)}{\partial(x, z)} = 0, \tag{52a}$$

$$\frac{\partial(X, Z)}{\partial(x, z)} = 1. \tag{52b}$$

Here, M_I is the initial value of the geostrophic momentum $M = v + fX$, which is a Lagrangian invariant in the y-independent ("symmetric") configurations; v is the y component of velocity; θ_I is the initial value of potential temperature (or density, up to a sign, in the oceanic context), which is also a Lagrangian invariant in HSPE; and θ_r is a normalization constant. Note that the thermal wind balance can be rewritten in the form

$$f\frac{\partial M}{\partial Z} = \frac{g}{\theta_r}\frac{\partial \theta}{\partial X}, \tag{53}$$

and therefore, a potential Φ such that

$$M = f^{-1}\frac{\partial \Phi}{\partial X}, \quad \theta = \frac{\theta_r}{g}\frac{\partial \Phi}{\partial Z} \tag{54}$$

may be introduced for balanced (adjusted) states. In fact, Φ is geopotential "extended" by $f^2X^2/2$: $\Phi = \phi + f^2(X^2/2)$.

We introduce the deviations of the particle positions from the balanced state

$$X = \overline{X} + \chi, \quad Z = \overline{Z} + \zeta, \tag{55}$$

so that Eq. (52a) becomes

$$\frac{\partial(\overline{X} + \chi, \ddot{\chi})}{\partial(x, z)} - \frac{\partial(\overline{X} + \chi, fM_I)}{\partial(x, z)} + \frac{g}{\theta_r}\frac{\partial(\theta_I, \overline{Z} + \zeta)}{\partial(x, z)} = 0. \tag{56}$$

It is more convenient to use as independent variables the positions of the particles in the adjusted state $(\overline{X}, \overline{Z})$ rather than the initial positions (x, z). When this change of variables is made in (33), two terms which express the thermal wind relation in the adjusted state cancel out. Furthermore, it is advantageous to express the gradients of $\overline{v}$ and θ in the adjusted state through the geopotential $\overline{\phi}$, making use of geostrophic and hydrostatic balances, respectively. Eq. (56) then becomes

$$\left(\frac{\partial^2}{\partial t^2} + f^2 + \frac{\partial^2\overline{\phi}}{\partial \overline{X}^2}\right)\frac{\partial \chi}{\partial \overline{Z}} + \frac{\partial^2\overline{\phi}}{\partial \overline{X}\partial \overline{Z}}\left(-\frac{\partial \chi}{\partial \overline{X}} + \frac{\partial \zeta}{\partial \overline{Z}}\right) - \frac{\partial^2\overline{\phi}}{\partial \overline{Z}^2}\frac{\partial \zeta}{\partial \overline{X}} = 0. \tag{57}$$

The incompressibility condition (52b) gives

$$\frac{\partial \chi}{\partial \overline{X}} + \frac{\partial \zeta}{\partial \overline{Z}} + \frac{\partial(\chi, \zeta)}{\partial(\overline{X}, \overline{Z})} = 0. \tag{58}$$

After appropriate rescalings, we obtain

$$\left(\frac{\partial^2}{\partial t^2} + 1 + \mathrm{Ro}\frac{\partial^2\overline{\phi}}{\partial \overline{X}^2}\right)\frac{\partial \chi}{\partial \overline{Z}} + \mathrm{Ro}\frac{\partial^2\overline{\phi}}{\partial \overline{X}\partial \overline{Z}}\left(-\frac{\partial \chi}{\partial \overline{X}} + \frac{\partial \zeta}{\partial \overline{Z}}\right) - \mathrm{Bu}\frac{\partial^2\overline{\phi}}{\partial \overline{Z}^2}\frac{\partial \zeta}{\partial \overline{X}} = 0, \tag{59a}$$

$$\frac{\partial \chi}{\partial \overline{X}} + \frac{\partial \zeta}{\partial \overline{Z}} + \delta\mathrm{Ro}\frac{\partial(\chi, \zeta)}{\partial(\overline{X}, \overline{Z})} = 0, \tag{59b}$$

where $\delta = O(1)$ is the ratio of the typical across-jet velocity to the along-jet one. We consider intense jets and presume $\mathrm{Ro} \sim 1$, $\mathrm{Bu} \sim 1$, expand the variables in δ, and get at leading order

$$\frac{\partial \chi^{(0)}}{\partial \overline{X}} + \frac{\partial \zeta^{(0)}}{\partial \overline{Z}} = 0 \Rightarrow \chi^{(0)} = -\frac{\partial \psi^{(0)}}{\partial \overline{Z}}, \quad \zeta^{(0)} = \frac{\partial \psi^{(0)}}{\partial \overline{X}}. \tag{60}$$

Equation (59a) can then be rewritten in terms of $\psi^{(0)}$:

$$\left(\frac{\partial^2}{\partial t^2} + 1 + \frac{\partial^2\overline{\phi}}{\partial \overline{X}^2}\right)\frac{\partial^2\psi^{(0)}}{\partial \overline{Z}^2} - \frac{\partial^2\overline{\phi}}{\partial \overline{X}\partial \overline{Z}}\frac{\partial^2\psi^{(0)}}{\partial \overline{X}\partial \overline{Z}} + \frac{\partial^2\overline{\phi}}{\partial \overline{Z}^2}\frac{\partial^2\psi^{(0)}}{\partial \overline{X}^2} = 0. \tag{61}$$

The boundary conditions consist of the absence of vertical displacements for parcels at the top and bottom boundaries. This implies that $\psi^{(0)}$ is constant on the boundaries. Because there is no overall displacement of the fluid layer in the $\overline{X}$ direction, these constants are equal. They can be both set to zero; $\psi^{(0)}(\overline{X}, 0) = \psi^{(0)}(\overline{X}, 1) = 0$. $\psi^{(0)}$ should also remain bounded as $\overline{X} \to \pm\infty$. Equation (61) closely resembles the homogeneous part of the Sawyer–Eliassen equation (e.g., Holton 1992) except for the term with the double-time derivative, which makes (61) prognostic. As in the two-layer case above, we take the simplest example of a barotropic jet. The nondimensional geopotential describing a balanced barotropic jet is

$$\overline{\phi} = \Phi(\overline{X}) + \frac{\overline{Z}^2}{2}. \tag{62}$$

If an unbalanced fast component is added to (62), the equation for its evolution is [cf. (61)]

$$(1 + \Phi'' + \partial_{tt})\partial_{\overline{Z}\,\overline{Z}}\psi^{(0)} + \partial_{\overline{X}\,\overline{X}}\psi^{(0)} = 0. \tag{63}$$

It allows a separation of variables; given the boundary conditions, the vertical eigenfunctions are $\sin(n\pi\overline{Z})$ and

$$\psi^{(0)}(\overline{X}, \overline{Z}, t) = \sum_n \sin(n\pi\overline{Z})\psi_n^{(0)}(\overline{X}, t). \tag{64}$$

We look for $\psi_n^{(0)}(\overline{X}, t)$ with a time dependence of the form $e^{-i\omega t}$ and get a Sturm–Liouville problem on $\{-\infty, +\infty\}$. We denote by $\hat{\psi}_{n\omega}(x)$ the horizontal eigenfunction with vertical wavenumber n and frequency ω. The equation for $\hat{\psi}_{n\omega}(\overline{X})$ has the form of Shrödinger equation of a particle in a well:

$$\partial_{\overline{X}\,\overline{X}}\hat{\psi}_{n\omega} - n^2\pi^2(1 + \Phi'' - \omega^2)\hat{\psi}_{n\omega} = 0. \qquad (65)$$

The potential $(1 + \Phi'') \to 1$ as $\overline{X} \to \infty$; hence, for a given n, we have first, a continuous spectrum of solutions with $\omega > 1$ (this part of the spectrum is doubly degenerate, with two independent solutions for each eigenvalue ω, and corresponds to leftward and rightward propagating IGWs), and second, a discrete spectrum of solutions with subinertial frequencies: $\sqrt{\min(1 + \Phi'')} < \omega < 1$. This part of the spectrum is nondegenerate, and consists of solutions exponentially decaying outside the region where $(1 + \Phi'' - \omega^2) < 0$, and oscillating inside that region; they are trapped in the anticyclonic part of the jet.

For a jet with relative vorticity lower than -1 ($-f$ in dimensional variables), modes with $\omega^2 < 0$ arise in the trapped mode spectrum, thus giving, as in the two-layer case, the symmetric instability.

The nonlinear and/or viscous saturation of the dry symmetric instability is not sufficiently understood in the continuously stratified case (cf., e.g., Thorpe and Rotunno 1989). It is, however, plausible that it will lead to IGW emission in the stratified background.

c. *High Ro and baroclinicity: Nonsymmetric unbalanced shear instabilities of balanced flows*

It is known that in the simplest 2RSW model the balanced adjusted states (35) may be unstable with respect to unbalanced perturbations, in addition to the classical baroclinic instability which is the core feature of these models (cf., e.g., Sakai 1989). Unlike the previously treated symmetric instability, these instabilities are "non-symmetric", but arise at significant Rossby numbers as well. The example treated in the pioneering work of Sakai (1989) was a linear interface profile in the channel, giving the balanced constant-velocity flow in each layer (the Phillips model). The difference between this case and all the cases treated previously in the present paper is the presence of the lateral boundary. As is well known, the boundary changes the spectrum of linear perturbations and introduces the nondispersive boundary Kelvin waves. Their presence, although important for the overall momentum and mass balance (cf. Reznik and Zeitlin 2007), does not radically change either the adjustment scenario for small Rossby numbers or the equations for balanced motion, as was shown in Reznik and Grimshaw (2002). However, conjugated with baroclinicity, the Kelvin waves produce a new ageostrophic instability (Sakai 1989). Remarkably, the growth rates of this instability, called Rossby–Kelvin (RK) because of the nature of the waves at its origin, are higher than that of the standard baroclinic one. Below, we show the results of the analysis, similar to Sakai (1989), but in the general case of layers of nonequal depth. We show in Fig. 2 the unstable regions in the phase space of the model. In addition to the classical long-wave baroclinic instability arising at small Rossby number, three other unstable regions are distinctively visible. The RK instability arises at intermediate Ro for sufficiently long waves. The Kelvin–Helmholtz (KH) instability corresponds to the upper unstable zone of high Ro. Note that Bu may be taken to be of order 1 in the calculations. The corresponding growth rates are presented in Fig. 3.

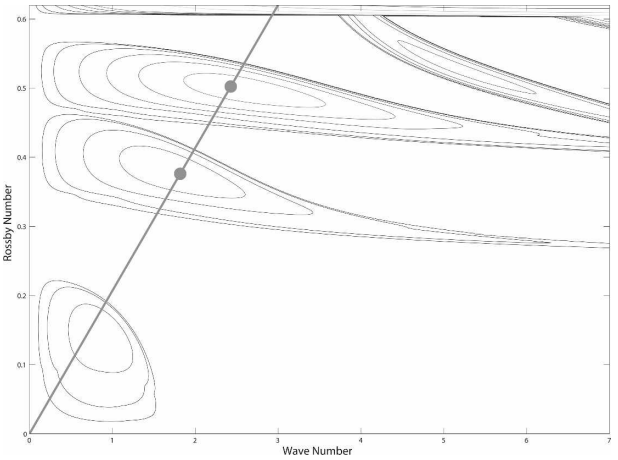

FIG. 2. Unstable regions in the k–Ro plane for the two-layer RSW flow in the channel with constant opposite velocities in the upper and lower layer with aspect ratio $d = 0.7$. Isolines of the growth rate are shown at intervals of 0.02 with the lowest value of 0.01. Calculations are made by the high-resolution collocation method. The section of the growth rates across the indicated line is shown below in Fig. 3.

There are different kinds of linear waves in the system: Rossby waves propagating because of the equivalent beta-effect produced by the inclined interface between the layers, Kelvin waves appearing because of the presence of lateral boundaries, and IGWs. The dynamical origins of the displayed instabilities are direct resonances between these waves propagating in each layer (Sakai 1989): respectively, Rossby–Rossby for the standard baroclinic instability, Rossby–Kelvin for the RK instability, and Kelvin–Kelvin, IGW–Kelvin, or IGW–IGW for KH instability. The wave patterns corresponding to the maximum growth rates of RK instabilities in Fig. 3 (indicated by dots) are presented in Fig.

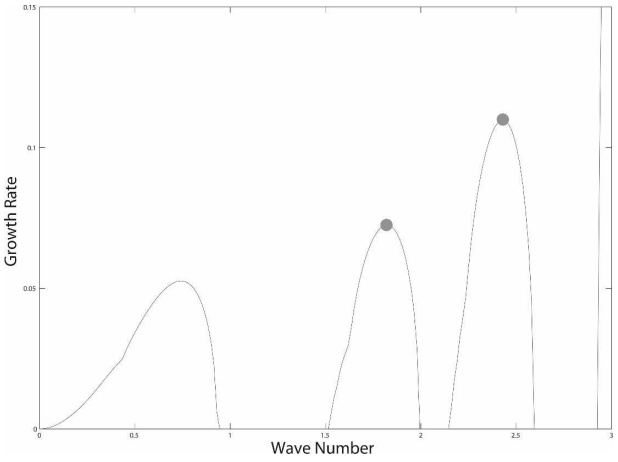

FIG. 3. The growth rates along the section shown in Fig. 2. Note that the growth rate of the RK instability exceeds that of the baroclinic one.

4. One clearly sees the Rossby wave and the Kelvin wave patterns, respectively.

Although "pure" RK instability is produced by the presence of boundaries, other ageostrophic instabilities persist in the absence of boundaries. Apart from the classical KH one, characterized by the highest Rossby numbers, the Rossby–gravity instability (which can be seen as descending from the KH region branch in Fig. 2 and corresponding to the resonance between the Rossby and the IG wave) is, in principle, not sensitive to the presence of lateral boundaries. (Its typical growth rates are of the same order as for RK; not shown).

To test whether the RK instability persists in a more realistic framework of continuously stratified fluid, the corresponding calculations were performed in the channel configuration using the mesoscale atmospheric model WRF (Skamarock et al. 2005), allowing for both operational and idealized simulations (Gula et al. 2007). To be close to the previously described two-layer case, the WRF runs were done periodically in the x-direction channel on the f plane with the flat bottom without a boundary layer. The differences from the two-layer model are background stratification and a free surface (and not a rigid lid) upper boundary condition. The unperturbed front is given by the following profile of potential temperature:

$$\theta = \theta_0 + \Theta z + \delta\theta\left(1 + \tanh\left(\frac{z - z_0 - Sy}{\delta z}\right)\right), \qquad (66)$$

where z_0 is the average height of the front, Θ defines the background stratification, S is the slope of the front, $\delta\theta$ is the potential temperature jump across the front, and δz is the thickness of the frontal zone (cf. Fig. 5). The related jet is found via the thermal wind relation.

The simulations were initiated with a slightly perturbed front with perturbation at various wavenumbers k and with various Ro. A breeding procedure was applied to isolate the most unstable mode, if any, for each set of parameters. As a result, the unstable modes corresponding to those previously identified in the two-layer model were found, with very close growth rates. In particular, the RK instability was clearly seen. Figure 6 shows pressure and velocity fields vertically averaged below and above the front, respectively. A typical structure of the Rossby wave (geostrophic wind along the isobars) can be clearly seen below the front, whereas a typical Kelvin wave pattern (wind parallel to the lateral boundaries with pressure extrema on the boundaries) can be seen above the front. The growth rate of the instability was estimated from the kinetic energy plots. The growth rates along the line Ro = $k/5$ in the parameter space, corresponding to the section in Fig. 3 for the two-layer model, is presented in Fig. 7. The two are remarkably close. The subsequent evolution of the instability leads to its saturation, accompanied by secondary IGW emission (Gula et al. 2007).

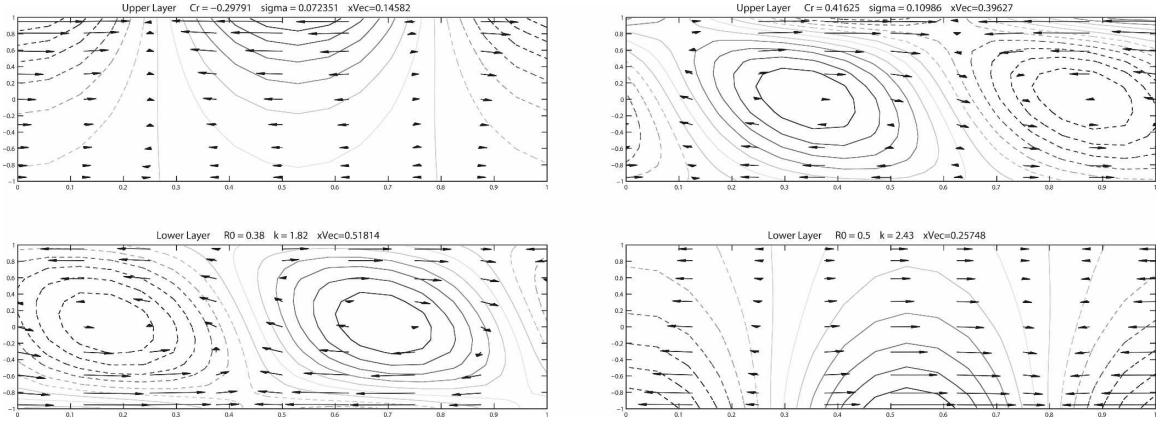

FIG. 4. Growing wave patterns for RK instability: (left) $k = 0.4$, Ro = 0.4; (right) $k = 2.4$, Ro = 0.5.

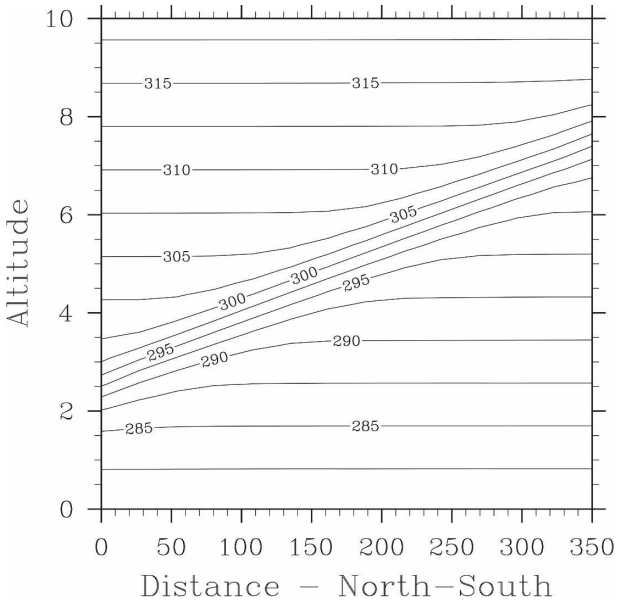

FIG. 5. The distribution of the potential temperature of the front in continuously stratified fluid.

d. Discussion

Thus, at high Ro and Bu the fast–slow splitting is violated because of the Lighthill radiation, although the intensity of the emitted wave field is weak. Note that our analysis of this effect was essentially two-dimensional, and the question of three-dimensional instabilities proper to such regimes arises (cf., e.g., Deloncle et al. 2007).

At high Ro and small or moderate Bu, (i) trapped subinertial waves appear; (ii) growth of trapped modes (symmetric/inertial instability) takes place in sufficiently strong anticyclonic shears (its nonlinear stage, saturated or not, gives rise to the IGW emission); and (iii) shear RK, KH, and related instabilities arise at sufficiently strong shears, thus giving rise, at the nonlinear stage, to the IGW emission.

Symmetric and shear instabilities are fast, as compared to slow baroclinic instability; therefore they violate the fast–slow decoupling. They also produce spontaneous secondary IGW emissions. The ageostrophic shear instabilities are not specific to the layered models; they are present as well in continuously stratified realistic configurations.

4. Summary and discussion

The following tentative classification thus arises in the parameter space of the atmospheric and oceanic flows with regard to decoupling balanced (slow) and unbalanced (fast IGW) motions and spontaneous IGW emission.

- At small Ro and small or $O(1)$ Bu,
  1) corrections to standard balanced models at longer times do exist and may be interpreted as wave modifications of the slow manifold, but there is no

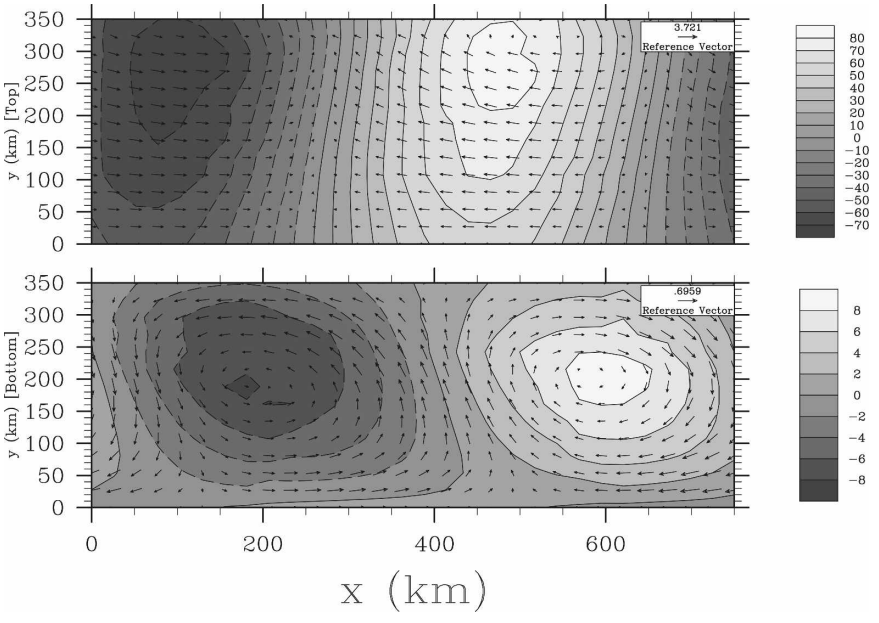

FIG. 6. Vertically averaged pressure and velocity distributions for the RK instability in continuously stratified fluid for Ro = 0.5 and k = 2.4 (top) above and (bottom) below the front.

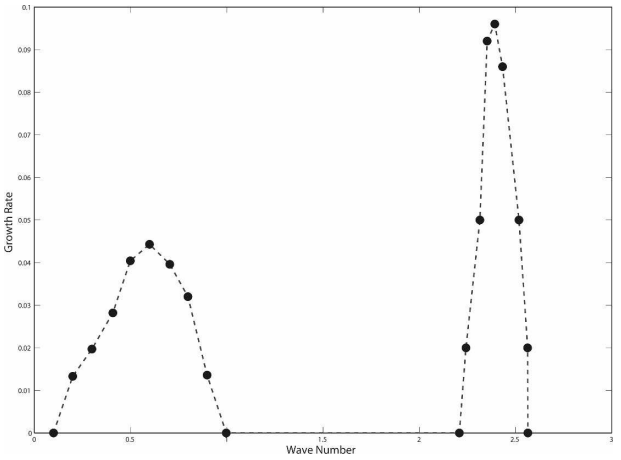

FIG. 7. The growth rates of the (left) baroclinic and (right) RK instabilities of the front in continuously stratified fluid.

fast-motion drag on the slow motion; primary IGW emission corresponding to evacuation of unbalanced part of a given initial disturbance is not sensitive to the balanced part;

2) secondary IGW emission at Bu ~1 exists because of fast–slow interactions, but the presence of initial fast field is necessary; fast–slow coupling leads to guiding and possible focusing of quasi-inertial waves by the balanced flow at Bu $\ll 1$.

- At large Ro and small or $O(1)$ Bu,
  1) trapping of internal waves in anticyclonic shears of the balanced flow takes place;
  2) symmetric and shear instabilities (of different types) arise as a source of secondary IGW emission.
- At large Ro and large Bu,
  1) the IGW emission via the Lighthill radiation mechanism takes place;
  2) the backreaction of this radiation slowly changes the balanced motion.

Although it is important to identify the concrete mechanisms leading to the IGW emission at high Ro, in general there is no surprise that fast–slow decoupling, well established at small Ro, ceases at large Ro because the Rossby number, by definition, is a ratio of the slow (vortex turnover) to fast (inertial) time scales. Hence, the question of spontaneous imbalance is, in fact, a question of how the balanced, small-Ro flow could organize the high-Ro regions during its evolution. The processes that are able to create small-scale and/or high-velocity structures locally in the small-Ro flow, apart from external forcing and (steep) topography, are still ill understood. One way to address this problem is to thoroughly study the improved QG–FG equations and possible (cascade) processes which would be able to produce, starting from initial configurations with a single typical scale, regions of localized high vorticity (=high Ro) within the time scale of validity of these equations. It is worth noting that the conclusion based on the improved QG models that no IGW emission is expected up to times $\sim (1/\mathrm{Ro}^2 f)$ is consistent with the results of Plougonven and Snyder (2007), in which the typical Rossby numbers of initial configuration were rather high Ro $\sim$ 0.4, and the zones of high local Rossby numbers IGW emission were observed after ~6 days of evolution of the baroclinic instability.

Acknowledgments. The presented results are based on the work done in collaboration with J. Gula, J. Le Sommer, S. B. Medvedev, R. Plougonven and G. M. Reznik, which is gratefully acknowledged. Special thanks to P. Lelong and T. Dunkerton for organizing the "Spontaneous Imbalance" Workshop with its stimulating atmosphere.

APPENDIX

A Hierarchy of the GFD Models

a. The RSW model

Equations of motion:

$$\partial_t \mathbf{v} + \mathbf{v}\cdot\nabla\mathbf{v} + f\hat{\mathbf{z}}\cdot\mathbf{u} + g\nabla h = 0, \tag{A1}$$

$$\partial_t h + \nabla\cdot(\mathbf{v}h) = 0. \tag{A2}$$

Here, $\mathbf{v}(\mathbf{x}, t)$ is the velocity of the fluid column and $h(\mathbf{x}, t)$ is the free-surface elevation.

b. The 2RSW model

The unperturbed depths of the upper and lower layers are H_1 and H_2, respectively; $H_1 + H_2 = H$. Equations of motion are

$$\partial_t \mathbf{v}_i + \mathbf{v}_i\cdot\nabla\mathbf{v}_i + f\hat{\mathbf{z}}\cdot\mathbf{v}_i + \frac{1}{\rho_i}\nabla\pi_i = 0, \quad i = 1, 2; \tag{A3}$$

$$\partial_t[H_i - (-1)^{i+1}\eta] + \nabla\cdot\{\mathbf{v}_i[H_i - (-1)^{i+1}\eta]\} = 0, \quad i = 1, 2, \tag{A4}$$

with no summation over i; $i + 1$ is understood modulo 2; $\mathbf{v}_i = [u_i(x, y, t), v_i(x, y, t)]$ are velocity fields in each layer; ρ_i are the densities of the layers; η is the vertical displacement of the interface; and π_i are defined with the help of the full pressure fields P_i in each layer; that is,

$$P_i = -\rho_i g z + (i - 1)(\rho_1 - \rho_2)gH_1 + \pi_i. \tag{A5}$$

From the dynamical boundary condition on the interface in the limit $\rho_1 \to \rho_2$, it follows that

$$(\rho_2 - \rho_1)g\eta = \pi_2 - \pi_1. \tag{A6}$$

c. Hydrostatic stratified primitive equations

$$\partial_t \mathbf{v}_h + \mathbf{v} \cdot \nabla \mathbf{v}_h + f\hat{\mathbf{z}} \cdot \mathbf{v}_h + \nabla_h \phi = 0, \tag{A7}$$

$$\phi_z = g\frac{\theta}{\theta_r}, \tag{A8}$$

$$\nabla_h \mathbf{v}_h + w_z = 0, \quad \text{and} \tag{A9}$$

$$(\partial_t + \mathbf{v} \cdot \nabla)\theta = 0. \tag{A10}$$

Here $\mathbf{v} = (\mathbf{v}_h, w)$ is the three-dimensional fluid velocity, ϕ is the geopotential, g is the gravitational acceleration constant, and θ_r is a reference value for the potential temperature. The index h denotes the horizontal part. The meaning of z and θ is different for the oceanic and atmospheric applications. In the oceanic context, z is physical height and θ is density up to a sign. For the atmospheric interpretation of equations, θ is potential temperature and z is pseudoheight: $z = z_r[1 - (P/P_r)^{(\gamma-1)/\gamma}]$.

REFERENCES

Babin, A., A. Mahalov, and B. Nicolaenko, 1997a: Global splitting and regularity of rotating shallow-water equations. *Eur. J. Mech.,* **15B,** 725–754.

——, ——, ——, and Y. Zhou, 1997b: On the asymptotic regimes and the strongly stratified limit of rotating Boussinesq equations. *Theor. Comput. Fluid Dyn.,* **9,** 223–251.

Bennets, D. A., and B. J. Hoskins, 1979: Conditional symmetric instability—A possible explanation for frontal rainbands. *Quart. J. Roy. Meteor. Soc.,* **105,** 945–962.

Bouchut, F., J. Le Sommer, and V. Zeitlin, 2004: Frontal geostrophic adjustment, slow manifold and nonlinear wave phenomena in one-dimensional rotating shallow water. Part 2. Numerical simulations. *J. Fluid Mech.,* **514,** 35–63.

Deloncle, A., J. M. Chomaz, and P. Billant, 2007: Three-dimensional stability of a horizontally sheared flow in a stably stratified fluid. *J. Fluid Mech.,* **570,** 297–305.

Dritschel, D. G., and A. Viúdez, 2007: The persistence of balance in geophysical flows. *J. Fluid Mech.,* **570,** 365–383.

Embid, P. F., and A. J. Majda, 1996: Averaging over fast gravity waves for geophysical flows with arbitrary potential vorticity. *Comm. Part. Diff. Equations,* **21,** 619–658.

——, and ——, 1998: Low Froude number limiting dynamics for stably stratified flow with small or finite Rossby numbers. *Geophys. Astrophys. Fluid Dyn.,* **87,** 1–50.

Ford, R., M. E. McIntyre, and W. A. Norton, 2000: Balance and the slow quasimanifold: Some explicit results. *J. Atmos. Sci.,* **57,** 1236–1254.

Gryanik, V., 1983: Emission of sound by linear vortex filaments. *Izv. Atmos. Ocean Phys.,* **19,** 150–153.

Gula, J., R. Plougonven, and V. Zeitlin, 2007: Ageostrophic instabilities of a front in stratified rotating fluid. *Proc. 18th Congrès Français de Mécanique,* Grenoble, France, Association Français de Mécanique. [Available online at http://irevues.inist.fr//cfm2007, s.9 "Vortex dynamics."]

Holton, J., 1992: *An Introduction to Dynamic Meteorology.* 3rd ed. Academic Press, 511 pp.

Lelong, M. P., and J. J. Riley, 1991: Internal wave–vortical mode interactions in strongly stratified flows. *J. Fluid Mech.,* **232,** 1–19.

Le Sommer, J., S. B. Medvedev, R. Plougonven, and V. Zeitlin, 2003: Singularity formation during the relaxation of jets and fronts towards the state of geostrophic equilibrium. *Comm. Nonlinear Sci. Numer. Simul.,* **8,** 415–442.

——, G. M. Reznik, and V. Zeitlin, 2004: Nonlinear geostrophic adjustment of long-wave disturbances in the shallow-water model on the equatorial beta-plane. *J. Fluid Mech.,* **515,** 135–170.

Lighthill, M. J., 1952: On sound generated aerodynamically. I. General theory. *Proc. Roy. Soc. London,* **211A,** 564–571.

O'Sullivan, D., and T. J. Dunkerton, 1995: Generation of inertia–gravity waves in a simulated life cycle of baroclinic instability. *J. Atmos. Sci.,* **52,** 3695–3716.

Plougonven, R., and V. Zeitlin, 2002: Internal gravity wave emission from a pancake vortex: An example of wave–vortex interaction in strongly stratified flows. *Phys. Fluids,* **14,** 1259–1268.

——, and ——, 2005: Lagrangian approach to geostrophic adjustment of frontal anomalies in a stratified fluid. *Geophys. Astrophys. Fluid Dyn.,* **99,** 101–135.

——, and C. Snyder, 2007: Inertia–gravity waves spontaneously generated by jets and fronts. Part I: Different baroclinic life cycles. *J. Atmos. Sci.,* **64,** 2502–2520.

Reznik, G. M., and R. Grimshaw, 2002: Nonlinear geostrophic adjustment in the presence of a boundary. *J. Fluid Mech.,* **471,** 257–283.

——, and V. Zeitlin, 2007: Asymptotic methods with application to the fast–slow splitting and the geostrophic adjustment. *Nonlinear Dynamics of Rotating Shallow Water: Methods and Advances,* V. Zeitlin, Ed., Elsevier, 47–120.

——, ——, and M. Ben Jelloul, 2001: Nonlinear theory of geostrophic adjustment. Part 1. Rotating shallow-water model. *J. Fluid Mech.,* **445,** 93–120.

Sakai, S., 1989: Rossby–Kelvin instability: A new type of ageostrophic instability caused by a resonance between Rossby waves and gravity waves. *J. Fluid Mech.,* **202,** 149–176.

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the advanced research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp.

Stegner, A., 2007: Experimental reality of geostrophic adjustment. *Nonlinear Dynamics of Rotating Shallow Water: Methods and Advances,* V. Zeitlin, Ed., Elsevier, 323–378.

Thorpe, A. J., and R. Rotunno, 1989: Nonlinear aspects of symmetric instability. *J. Atmos. Sci.,* **46,** 1285–1299.

Warn, T., O. Bokhove, T. G. Shepherd, and J. Vallis, 1995: Rossby number expansions, slaving principles, and balance dynamics. *Quart. J. Roy. Meteor. Soc.,* **121,** 723–739.

Zeitlin, V., 1991: On the back-reaction of acoustic radiation for distributed two-dimensional vortex structures. *Phys. Fluids,* **3A,** 1677–1680.

——, 2007: Fundamentals of rotating shallow water in the geophysical fluid dynamics perspective. *Nonlinear Dynamics of Rotating Shallow Water: Methods and Advances,* V. Zeitlin, Ed., Elsevier, 1–46.

——, G. M. Reznik, and M. Ben Jelloul, 2003: Nonlinear theory of the geostrophic adjustment. Part 2. Two-layer and continuously stratified models. *J. Fluid Mech.,* **491,** 207–228.