

Gravity Wave Emission by Shear Instability

CARSTEN EDEN AND MANITA CHOUKSEY

Institut für Meereskunde, Universität Hamburg, Hamburg, Germany

DIRK OLBERS

*Alfred Wegener Institut für Polar- und Meeresforschung, Bremerhaven, and MARUM,
Universität Bremen, Bremen, Germany*

(Manuscript received 2 February 2019, in final form 27 May 2019)

ABSTRACT

Gravity wave emission by geostrophically balanced flow is diagnosed in numerical simulations of lateral and vertical shear instabilities. The diagnostic method in use allows for a separation of balanced flow and residual wave signal up to fourth order in the Rossby number (Ro). While evidence is found for a small but finite gravity wave emission from balanced flow in a single-layer model with large lateral shear and large Ro , a vertically resolved model with moderate velocity amplitudes appropriate to the interior ocean hardly shows any wave emission. Only when static instabilities generated by the shear instability of the balanced flow are allowed can a gravity wave signal similar to the ones reported in earlier studies be detected in the vertically resolved case. This result suggests a relatively small role of spontaneous wave emission in the classical sense of Lighthill radiation, and emphasizes the role of convective or symmetric instabilities during frontogenesis for the generation of internal gravity waves in the ocean and atmosphere.

1. Introduction

Gravity waves are a ubiquitous feature of the ocean and are generated by a variety of processes. One of them, the generation by spontaneous wave emission (e.g., Vanneste 2013), has recently attracted considerable attention. The process refers to the generation of gravity waves by balanced flow spontaneously in the absence of any external forcing, which appears to happen during baroclinic or barotropic instability (e.g., Plougonven and Snyder 2007; Hien et al. 2018; Chouksey et al. 2018), and is thought to be similar to sound generation in aerodynamics (Lighthill 1978; Ford et al. 2000), a process referred to as Lighthill radiation. Although it is intuitively clear to a physical oceanographer what balanced flow and gravity waves are, the exact definition of slow balanced flow and fast gravity waves is surprisingly difficult, since it is complicated by the presence of the nonlinear terms coupling both motions with each other. It is known that such nonlinearities can generate wavelike structures, which are nevertheless part of the balanced flow, the so-called ageostrophic balanced modes or slaved modes

(e.g., Warn et al. 1995; McIntyre and Norton 2000; Kafabad and Bartello 2018). The wavelike structures often seen in numerical simulations of lateral and vertical shear instabilities of the balanced flow could therefore be misinterpreted as spontaneous wave emission by the balanced flow, where in fact they are part of the balanced flow without the occurrence of spontaneous emission.

The most familiar model for the balanced mode is the result of the quasigeostrophic approximation, a first-order asymptotic expansion in the limit of small Rossby (Ro) and Froude number. Since the zero-order geostrophic velocity is free of horizontal divergence in that expansion, any significant horizontal divergence or vertical velocity seen in model simulations or observations is often interpreted as a gravity wave signal. It is, however, clear that the first-order ageostrophic vertical velocity—which can be calculated, for example, from the so-called omega equation (Hoskins et al. 1978)—although of first order in Ro and thus usually small, could also contribute to the observed vertical velocity. The ageostrophic vertical and horizontal first-order velocities in the quasigeostrophic approximation correspond to the ageostrophic balanced mode or slaved wave mode (to first order in Ro), and it is shown in Warn et al. (1995) how this concept of balanced

Corresponding author: Carsten Eden, carsten.eden@uni-hamburg.de

DOI: 10.1175/JPO-D-19-0029.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

models and wave modes can be extended to higher orders in Ro . This concept allows in principle to uniquely define the slow balanced flow and to differentiate the “true” fast gravity wave signal from that to arbitrary precision in terms of orders of Ro . The limit for infinitive order is called the slow manifold (and the difference to the full state the fast manifold), but it remains unclear if this limit exists (Lorenz 1980; Leith 1980; Lorenz 1986; Ford et al. 2000). In fact, for simple analogs to the equation of motions of the ocean and atmosphere it can be demonstrated that such an asymptotic limit does not exist (Vanneste 2013). In that case, the expansion in Ro is expected to diverge at large orders of Ro , although the lower orders of the expansion would still be useful for practical analysis.

In practice, however, the differentiation of waves and balanced flow in numerical models turns out to be rather difficult even at low order, such that the issue of the actual existence of the slow manifold appears to be of less concern. McIntyre and Norton (2000) applied a method to construct balanced models valid to higher order in Ro to a numerical model of the shallow water equations on a (half) sphere. They find that the balanced model simulation agrees surprisingly well with a model simulation using the full equations including the waves, such that they conclude that spontaneous wave emission remains unimportant in that simulation. To our knowledge, similar higher-order balance models have not yet been applied to numerical simulations of vertically resolved models. Therefore, we perform here a similar analysis as in McIntyre and Norton (2000) but for simulations of lateral shear instability in an idealized configuration applicable to the ocean for a range of values of Ro using a novel decomposition method at the discrete level of the model valid up to fourth order in Ro , and also extend this to the more realistic case of a vertically resolved model having also vertical shear instability.

In the next section, the theoretical background of the concept to differentiate waves from the balanced flow is outlined, while in section 3 the method is applied to numerical models of lateral and vertical shear instability. The last section provides a summary and discussion of the results.

2. Theoretical background

The theoretical background for the method to diagnose the model simulations following Warn et al. (1995) and Kafiabad and Bartello (2018) is outlined in this section. The focus is here on the analytical formulation, but since it is found that the numerical evaluation of the balanced modes needs exact consistency also on the discrete level of the numerical models, all relations on the discrete level which are actually used for the diagnosis are given in the appendix.

a. The model

We consider the equations of motions in hydrostatic and Boussinesq approximation for a constant Coriolis parameter f and constant background stability frequency N . The resulting primitive equations are scaled using $1/f$, L , H , and U as time scale, horizontal and vertical length scale, and horizontal velocity scale, respectively. The vertical velocity scale follows from the continuity equation, the pressure scale from geostrophic balance, and the buoyancy perturbation scale from hydrostatic balance. This yields¹

$$\begin{aligned}\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p &= -Ro(\mathbf{u} \cdot \nabla \mathbf{u} + w \partial_z \mathbf{u}), \\ \partial_t b + w &= -Ro(\mathbf{u} \cdot \nabla b + w \partial_z b)\end{aligned}\quad (1)$$

together with the diagnostic relations $\partial_z p = b$ and $\nabla \cdot \mathbf{u} + \partial_z w = 0$. The only parameter in Eq. (1) controlling the nonlinear terms is the Rossby number $Ro = U/(Lf)$ by setting $Ro = Fr$, where Fr is the Froude number $Fr = U/(NH)$. By this setting, the (unscaled) first baroclinic Rossby radius $L_r = NH/f$ becomes $L_r = L$. The case $Ro \ll 1$ is appropriate to the stratified interior ocean, while a larger Ro approaching unity is more appropriate to the so-called submesoscale flow which occurs for instance close to or in the mixed layer of the ocean. The scaled linearized kinematic surface and bottom boundary conditions become

$$\partial_t p|_{z=0} = -g \int_{-h}^0 \nabla \cdot \mathbf{u} dz, \quad w|_{z=-h} = 0 \quad (2)$$

with the scaled water depth h , the parameter $g = L_0^2/L^2 \gg 1$ and the (unscaled) barotropic Rossby radius $L_0 = \sqrt{H(9.81 \text{ m s}^{-2})/f}$. Vertically integrating the buoyancy equation and the hydrostatic relation yields

$$\partial_t p + M(\nabla \cdot \mathbf{u}) = Ro \int_z^0 dz' (\mathbf{u} \cdot \nabla b + w \partial_z b) \quad (3)$$

with the integral operator $M = g \int_{-h}^0 dz + \int_z^0 dz' \int_{-h}^{z'} dz''$. Equation (3) can be used to replace the hydrostatic relation and the buoyancy equation in Eq. (1).

b. The vertical modes

The operator M yields the vertical or baroclinic modes by its scaled eigenvalues c_n^2 and eigenfunctions ϕ_n . Taking ∂_{zz} of the eigenvalue equation $M(\phi_n) = c_n^2 \phi_n$ yields

¹ All vectors are horizontal. The vector \mathbf{u} denotes anticlockwise rotation of the vector \mathbf{u} by $\pi/2$, that is, $\mathbf{u}^\perp = (-v, u)$ for $\mathbf{u} = (u, v)$.

$$c_n^2 \partial_{zz} \phi_n + \phi_n = 0, \tag{4}$$

which is the equivalent eigenvalue equation formulated as differential equation. Considering $\partial_z M(\phi)$ at $z = -h$ and $M(\phi)$ and $\partial_z M(\phi)$ at $z = 0$, the boundary conditions to the eigenvalue problem are derived as

$$\phi_n|_{z=0} + g \partial_z \phi_n|_{z=0} = 0, \quad \partial_z \phi_n|_{z=-h} = 0. \tag{5}$$

Note that by the surface boundary condition, the eigenvectors ϕ_n are not the ones for a rigid lid, but modified by the free surface and given by $\phi_n = A_n \cos(z + h)/c_n$ with $\tan h/c_n = c_n/g$. Approximate solutions of c_n for $g \gg 1$ correspond to the different zero crossings of the tangent function and are $c_0^2 \approx gh$ and $c_n^2 \approx h^2/(n\pi)^2$ for $n > 0$. The constant amplitude A_n is chosen to satisfy the orthogonality condition $\int_{-h}^0 \phi_m \phi_n dz = \delta_{n,m}$.

The eigenfunctions of the operator M allow to diagonalize the linear part of the system. The variables p and \mathbf{u} are therefore expressed in terms of the eigenfunctions or vertical modes as

$$p = \sum_n p_n(x, y, t) \phi_n(z), \quad \mathbf{u} = \sum_n \mathbf{u}_n(x, y, t) \phi_n(z) \tag{6}$$

and $b = \sum_n p_n \partial_z \phi_n$ and $w = \sum_n c_n^2 \nabla \cdot \mathbf{u}_n \partial_z \phi_n$, which follows from the hydrostatic relation and the continuity equation. Using the expansion Eq. (6) in Eq. (7) and in the momentum equation from Eq. (1), multiplying with ϕ_m and integrating over depth yields

$$\partial_t \mathbf{u}_n + \mathbf{u}_n + \nabla p_n = \text{Ro} \mathbf{n}_n^n, \quad \partial_t p_n + c_n^2 \nabla \cdot \mathbf{u}_n = \text{Ro} n_p^n \tag{7}$$

with $\mathbf{n}_u^n = -\int_{-h}^0 dz \phi_n (\mathbf{u} \cdot \nabla \mathbf{u} + w \partial_z \mathbf{u})$ and $n_p^n = \int_{-h}^0 dz \phi_n \int_z^0 dz' (\mathbf{u} \cdot \nabla b + w \partial_z b)$. Equation (7) has in principle the same algebraically structure as a single-layer model, since the different vertical modes n are only coupled by the nonlinear terms \mathbf{n}_u and n_p . The equivalent nonlinear terms for a single-layer model with layer thickness p_n and layer velocity \mathbf{u}_n are $\mathbf{n}_u^n = -\mathbf{u}_n \cdot \nabla \mathbf{u}_n$ and $n_p^n = -\nabla \cdot p_n \mathbf{u}_n$, which would replace the right hand sides in Eq. (7).

c. The spectral representation

Assuming a double-periodic domain and applying the Fourier ansatz

$$\mathbf{u}_n(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\mathbf{k} \hat{\mathbf{u}}_n(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{x}}, \quad p_n(\mathbf{x}, t) = \int_{-\infty}^{\infty} d\mathbf{k} \hat{p}_n(\mathbf{k}, t) e^{i\mathbf{k} \cdot \mathbf{x}} \tag{8}$$

with wavenumber vector $\mathbf{k} = (k_x, k_y)$, yields in Eq. (7) after multiplication with $e^{-i\mathbf{k} \cdot \mathbf{x}}$ and integration over \mathbf{x} ,

$$\partial_t \hat{\mathbf{z}} - i \mathbf{A} \cdot \hat{\mathbf{z}} = \text{Ro} \hat{\mathbf{n}}, \quad \mathbf{A} = \begin{pmatrix} 0 & -i & -k_x \\ i & 0 & -k_x \\ -c_n^2 k_x & -c_n^2 k_y & 0 \end{pmatrix}, \tag{9}$$

with the state vector $\hat{\mathbf{z}} = (\hat{\mathbf{u}}, \hat{p})^T$, the vector of the nonlinearities $\mathbf{n} = (\mathbf{n}_u, n_p)^T$, and its Fourier transform $\hat{\mathbf{n}} = (2\pi)^{-2} \int d\mathbf{x} \mathbf{n} e^{-i\mathbf{k} \cdot \mathbf{x}}$. The index n denoting the baroclinic mode number is omitted in Eq. (9) for simplicity. The matrix \mathbf{A} has eigenvalues ω^s and right and left eigenvectors $\mathbf{q}^s = (q_u^s, q_v^s, q_p^s)^T$ and $\mathbf{p}^s = (p_u^s, p_v^s, p_p^s)$, respectively. The index s denotes here either the slow linear geostrophic mode (for $s = 0$) or the fast linear gravity wave mode (for $s = \pm$). For the latter there are two linear waves modes with $s = +$ and $s = -$. The eigenvalues ω^s are given by the geostrophic mode $\omega^0 = 0$ and the fast mode $\omega^\pm = \pm \sqrt{1 + c_n^2 |\mathbf{k}|^2}$. The latter is also the dispersion relation of long gravity waves in the system.

It is useful to project Eq. (9) on the fast or slow modes using the eigenvectors of \mathbf{A} . The fast or slow mode amplitude $g^s = \mathbf{p}^s \cdot \hat{\mathbf{z}}$ is then governed by

$$\partial_t g^s - i \omega^s g^s = \text{Ro} \mathbf{p}^s \cdot \hat{\mathbf{n}} = -i \text{Ro} I^s(g^0, g^\pm), \quad s = 0, \pm, \tag{10}$$

where the case $s = 0$ corresponds to the slow geostrophic mode with amplitude g^0 , and $s = \pm$ to the (two) fast gravity wave mode with g^+ and g^- . For $\text{Ro} = 0$, that is, in the linear case, Eq. (10) describes the fast oscillation in time of the gravity waves $g^\pm = e^{i\omega^\pm t}$, and the stationary solution for the slow mode $g^0 = \text{const}$. The nonlinear terms in $\hat{\mathbf{n}}$ couple the modes g^s at different \mathbf{k} , but also the slow and fast modes which is expressed here by the nonlinear interaction integral $I^s(g^0, g^\pm)$.

d. The interaction integral

The exact form of the interaction integral I^s is in principle not relevant here since it will not be evaluated directly—which is also possible, see, for example, [Eden et al. \(2019\)](#)—but calculated from numerical models. For the expansion into orders of the small parameter Ro which is performed below, however, we need to know about the general structure of the interaction integral. For the single-layer model with $\mathbf{n}_u = -\mathbf{u} \cdot \nabla \mathbf{u}$ and $n_p = -\nabla \cdot p \mathbf{u}$, the nonlinear terms become

$$\hat{\mathbf{n}}(\mathbf{k}) = -i \int d\mathbf{k}_1 \int d\mathbf{k}_2 \mathbf{N} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}), \quad \mathbf{N} = \begin{Bmatrix} [\hat{\mathbf{u}}(\mathbf{k}_1) \cdot \mathbf{k}_2] \hat{u}(\mathbf{k}_2) \\ [\hat{\mathbf{u}}(\mathbf{k}_1) \cdot \mathbf{k}_2] \hat{v}(\mathbf{k}_2) \\ \hat{p}(\mathbf{k}_2) \hat{\mathbf{u}}(\mathbf{k}_1) \cdot (\mathbf{k}_1 + \mathbf{k}_2) \end{Bmatrix}, \tag{11}$$

and the integral I^s

$$I^s(g^0, g^\pm) = i\mathbf{p}^s \cdot \hat{\mathbf{n}} = \int d\mathbf{k}_1 \int d\mathbf{k}_2 \sum_{s_1, s_2=0, \pm} g^{s_1}(\mathbf{k}_1) g^{s_2}(\mathbf{k}_2) C_{\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2}^{s, s_1, s_2} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}) \quad (12)$$

for the modes $s = 0, \pm$. The so-called interaction coefficients C result from the projection of \mathbf{N} on the eigenvector \mathbf{p}^s and are given in [Eden et al. \(2019\)](#). For the primitive equation model the structure of the integral I^s is similar to Eq. (12), but a double sum over the vertical modes will add and the interaction coefficient are different [and also given for the rigid lid case in [Eden et al. \(2019\)](#)]. In any case, the slow and fast mode amplitudes g^s will show up quadratically in a double sum over the

fast and slow modes s , which is true for all kinds of models with quadratic nonlinearities.

By resolving and rearranging the double sum of the slow and fast modes s_1, s_2 , the integral I^s in Eq. (12) can therefore also be written as

$$I^s(g^0, g^\pm) = I^s(g^0, 0) + I^s(0, g^\pm) + K^s(g^0, g^\pm) \quad (13)$$

with

$$\begin{aligned} I^s(g^0, 0) &= \int d\mathbf{k}_1 \int d\mathbf{k}_2 g^0(\mathbf{k}_1) g^0(\mathbf{k}_2) C_{\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2}^{s, 0, 0} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}), \\ I^s(0, g^\pm) &= \int d\mathbf{k}_1 \int d\mathbf{k}_2 \sum_{s_1, s_2=\pm} g^{s_1}(\mathbf{k}_1) g^{s_2}(\mathbf{k}_2) C_{\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2}^{s, s_1, s_2} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}), \\ K^s(g^0, g^\pm) &= 2 \int d\mathbf{k}_1 \int d\mathbf{k}_2 \sum_{s_2=\pm} g^0(\mathbf{k}_1) g^{s_2}(\mathbf{k}_2) C_{\mathbf{k}, \mathbf{k}_1, \mathbf{k}_2}^{s, 0, s_2} \delta(\mathbf{k}_1 + \mathbf{k}_2 - \mathbf{k}), \end{aligned} \quad (14)$$

and similar for the primitive equation model but involving sums over all vertical modes. Equation (13) and Eq. (14) are used below to assess the orders of $I^s(g^0, g^\pm)$ in terms of Ro expanding the amplitudes g^s into a power series of Ro.

e. The balanced models

The spectral representation of the primitive equations given by Eq. (10) describes the evolution of the gravity wave amplitudes g^\pm and the slow geostrophic mode g^0 , and how they are coupled by the interaction integral I^s . Due to this coupling, the definition of a slow balanced flow must involve not only the linear geostrophic mode g^0 , but also the linear gravity wave amplitudes g^\pm . A method to uniquely describe the balanced part of the flow in this framework to any order in Ro is described by [Warn et al. \(1995\)](#). The method originates from the problem of balanced model initialization in numerical weather prediction, pioneered by [Machenhauer \(1977\)](#) and [Baer and Tribbia \(1977\)](#). The task in that problem is to provide an initial condition for a model for which a subsequent integration shows no or only a weak fast adjustment by gravity wave activity. Such an initial condition is then considered to be a balanced state.

[Warn et al. \(1995\)](#) assume accordingly a state in which the gravity waves are initially vanishing and then weakly growing in amplitude, such that $g^\pm = O(\text{Ro})$ and expand the gravity wave amplitudes accordingly as

$$g^\pm = \text{Ro} f_1^\pm + \text{Ro}^2 f_2^\pm + \text{Ro}^3 f_3^\pm + \dots \quad (15)$$

The time derivative in Eq. (10) includes the fast gravity wave oscillations with frequency ω^\pm and the slow growth and decay of the amplitudes $g^{0, \pm}$ of both slow and fast mode by the nonlinear interactions. Therefore, a fast and slow time scale is introduced with $T = \text{Ro} t^*$ and $\partial_t = \text{Ro} \partial_T + \partial_{t^*}$. The slow mode g^0 is a function of the slow time T only since $\omega^0 = 0$, and thus Eq. (10) becomes for the geostrophic mode $s = 0$

$$\text{Ro} \partial_T g^0 = -i \text{Ro} I^0(g^0, g^\pm). \quad (16)$$

Since the wave mode g^\pm is assumed to be of first order in Ro, from Eq. (13) and the form of the integrals in Eq. (14), it follows that the first order of Eq. (16) simply becomes

$$\partial_T g^0 = -i I^0(g^0, 0). \quad (17)$$

This first-order balanced model is identical to the familiar (first order) quasigeostrophic approximation as shown by [Leith \(1980\)](#). Only the slow mode amplitude g^0 is involved in this first-order balanced model, and in this way Eq. (17)—which is a spectral representation of the quasigeostrophic potential vorticity equation—is closed. The second order of Eq. (10) for the geostrophic mode $s = 0$ becomes

$$0 = K^0(g^0, f_1^\pm) = I^0(g^0, f_1^\pm) - I^0(g^0, 0) - I^0(0, f_1^\pm), \quad (18)$$

using again Eq. (13) and thus with Eq. (17)

$$\partial_T g^0 = -iI^0(g^0, f_1^\pm) + iI^0(0, f_1^\pm). \quad (19)$$

Together with the diagnostic expression for the auxiliary wave amplitude f_1^\pm , which is derived below, Eq. (19) is a second-order balanced model. Since the expression for f_1^\pm involves the geostrophic mode amplitudes g^0 only, the second-order balanced model is also closed. Third-order yields accordingly a third-order balanced model given by

$$\partial_T g^0 = -iI^0(g^0, f_2^\pm) + iI^0(0, f_2^\pm) - iI^0(0, f_1^\pm) \quad (20)$$

and so on for higher order. Again, the expression for f_2^\pm given below involves the geostrophic mode amplitudes g^0 only. If the asymptotic expansion in Ro would converge, balanced models to arbitrary order of Ro could in principle be constructed, which are closed with respect to the geostrophic mode g^0 . The remaining task to find the auxiliary wave modes f_n^\pm is demonstrated in the next section. We will define the deviation from an actual state from the state given by g^0 together with the auxiliary wave modes f_n^\pm , as the nonbalanced or fast gravity wave mode.

f. The ageostrophic balanced modes

The auxiliary wave modes f_n^\pm are also called ageostrophic balanced modes or slaved wave modes (Warn et al. 1995; McIntyre and Norton 2000; Kafiabad and Bartello 2018) and are part of the balanced mode since they evolve only slowly as seen below. The lowest order of those modes, f_1^\pm , corresponds to the first-order (ageostrophic) variables in the quasigeostrophic approximation (Leith 1980). The ageostrophic balanced modes are not needed to predict the evolution of the geostrophic variables given by g^0 to the first order, that is, they are not needed and in general unknown in the quasigeostrophic model, but for the higher-order balanced models they are needed and diagnostic relations for f_n^\pm need to be derived. To first order in Ro, Eq. (10) becomes for the gravity wave modes $s = \pm$

$$\partial_{t^*} f_1^\pm - i\omega^\pm f_1^\pm = -iI^\pm(g^0, 0), \quad (21)$$

using again Eq. (13) and Eq. (14) to infer the first order in Ro of the interaction integral I^\pm . The time derivative in Eq. (21) generates the fast wave oscillation with frequency ω^\pm . To suppress this fast oscillation in the balanced state, it is necessary to balance f_1^\pm as

$$f_1^\pm = I^\pm(g^0, 0)/\omega^\pm. \quad (22)$$

This balancing of f_1^\pm yields $\partial_{t^*} f_1^\pm = 0$ in Eq. (21) and thus no waves are generated. Equation (22) is therefore the

diagnostic relation which completes the second-order balance model in Eq. (19). It was shown by Leith (1980), that the ageostrophic balanced mode f_1^\pm corresponds to the first iteration step in the balanced initialization technique by Machenhauer (1977), and that f_1^\pm also corresponds to the ageostrophic variables in the (first order) quasigeostrophic approximation.

Setting $\partial_{t^*} f_n^\pm = 0$ to suppress wave activity in general, Eq. (10) for $s = \pm$ becomes

$$\sum_{n=1}^{\infty} (\text{Ro}^{n+1} \partial_T - i\omega^\pm \text{Ro}^n) f_n^\pm = -i\text{Ro} I^\pm(g^0, 0) - i\text{Ro} I^\pm\left(0, \sum_{n=1}^{\infty} \text{Ro}^n f_n^\pm\right) - i \sum_{n=1}^{\infty} \text{Ro}^{n+1} K^\pm(g^0, f_n^\pm), \quad (23)$$

using again Eq. (13). From the form of the integrals in Eq. (14) it becomes clear that the second, third, and fourth orders are given by

$$\begin{aligned} \partial_T f_1^\pm - i\omega^\pm f_2^\pm &= -iK^\pm(g^0, f_1^\pm) \\ \partial_T f_2^\pm - i\omega^\pm f_3^\pm &= -iI^\pm(0, f_1^\pm) - iK^\pm(g^0, f_2^\pm) \\ \partial_T f_3^\pm - i\omega^\pm f_4^\pm &= -iI^\pm(0, f_1^\pm + f_2^\pm) + iI^\pm(0, f_1^\pm) \\ &\quad + iI^\pm(0, f_2^\pm) - iK^\pm(g^0, f_3^\pm). \end{aligned} \quad (24)$$

From the first line of Eq. (24), f_2^\pm is calculated, from the second line f_3^\pm , etc. Since only slow time derivatives show up, the ageostrophic balanced modes f_n^\pm are only slowly evolving in time as the geostrophic mode. The combination of geostrophic mode amplitude g^0 and f_n^\pm given by $g^0 \mathbf{q}^0 + \sum_i \text{Ro}^i f_i^\pm \mathbf{q}^\pm$ defines the balanced mode in spectral space, and inverse Fourier transform yields the balanced flow in physical space

In practice, we evaluate the interaction integral I^\pm from a numerical model to find f_n^\pm up to fourth order. The term $K^\pm(g^0, f_n^\pm)$ is given by Eq. (13) replacing g^\pm with f_n^\pm and found by evaluating the integrals I^\pm using the numerical model. To find $\partial_T f_1^\pm$, Eq. (22) is analytically differentiated with respect to T as in Kafiabad and Bartello (2018), which yields

$$\begin{aligned} \partial_T f_1^\pm &= \partial_T I^\pm(g^0, 0)/\omega^\pm = [I^\pm(g^0 + \partial_T g^0, 0) - I^\pm(g^0, 0) \\ &\quad - I^\pm(\partial_T g^0, 0)]/\omega^\pm \end{aligned} \quad (25)$$

using the relation $I^\pm(g^0 + \partial_T g^0, 0) = I^\pm(g^0, 0) + \partial_T I^\pm(g^0, 0) + I^\pm(\partial_T g^0, 0)$, which derives from the definition of the integral I^s , and using $\partial_T g^0 = -iI^0(g^0, 0)$. To find $\partial_T f_2^\pm$ and $\partial_T f_3^\pm$, the corresponding expressions get too complicated, and f_2^\pm and f_3^\pm are numerically differentiated in time by integrating the full model a few time steps as suggested by Baer and Tribbia (1977).

g. The effect of damping

In the numerical model it is necessary to include damping to suppress the gridscale noise generated by artificial numerical wave dispersion and the finite representation of the interaction integral I^s . The primitive equations are therefore equipped with biharmonic operators $-A\nabla^4\mathbf{u}$ and $-A\nabla^4b$, with a viscosity (or diffusivity) $A \sim O(\text{Ro})$ to damp the gridscale noise. The same parameter A is used for momentum and buoyancy since then the only effect of the numerical damping is that the matrix \mathbf{A} gets identical entries $iA|\mathbf{k}|^4$ on the diagonal [instead of zero for the case of vanishing damping as in Eq. (9)], the eigenvectors will stay identical as for the case $A = 0$, and the eigenvalues become complex with identical real part as for $A = 0$, but with imaginary part $A|\mathbf{k}|^4$.

There is however a complication: in Eq. (3), a term $-A\nabla^4p$ shows up—which enters the matrix \mathbf{A} on the diagonal in the spectral representation—but also a

damping term $A\nabla^4p|_{z=0}$ involving the surface pressure. By this term, the vertical modes get coupled not only by the nonlinear terms but also by the numerical damping. To avoid this coupling, the surface boundary condition is also equipped with damping and Eq. (2) is replaced by

$$\partial_t p|_{z=0} = -g \int_{-h}^0 \nabla \cdot \mathbf{u} dz - A\nabla^4 p|_{z=0}. \quad (26)$$

Vertical friction and mixing are not used here, since we find it difficult to avoid the coupling of the vertical modes in the discrete version of the equations. However, the lateral biharmonic damping is found to be sufficient to control gridscale noise in the numerical integrations.

Introducing the numerical damping terms, the eigenvalues ω^s obtain an imaginary part of $O(\text{Ro})$ which needs to be taken into account in the expansion. The first to fourth orders of the fast mode amplitude Eq. (10) then become

$$\begin{aligned} \Re(\omega^\pm) f_1^\pm &= I^\pm(g^0, 0) \\ \partial_t f_1^\pm - i\Re(\omega^\pm) f_2^\pm &= -iK^\pm(g^0, f_1^\pm) - f_1^\pm \Im(\omega^\pm)/\text{Ro} \\ \partial_t f_2^\pm - i\Re(\omega^\pm) f_3^\pm &= -iI^\pm(0, f_1^\pm) - iK^\pm(g^0, f_2^\pm) - f_2^\pm \Im(\omega^\pm)/\text{Ro} \\ \partial_t f_3^\pm - i\Re(\omega^\pm) f_4^\pm &= -iI^\pm(0, f_1^\pm + f_2^\pm) + iI^\pm(0, f_1^\pm) + iI^\pm(0, f_2^\pm) - iK^\pm(g^0, f_3^\pm) - f_3^\pm \Im(\omega^\pm)/\text{Ro}. \end{aligned} \quad (27)$$

Since the magnitude of $\Im(\omega^\pm) = A|\mathbf{k}|^4$ strongly depends on wavenumber, we find that the damping effect can in practice become zero order in Ro for large \mathbf{k} . In that case, Eq. (22) and Eq. (24) are used replacing Eq. (27) to determine f_n^\pm for large \mathbf{k} , but using the complex ω^\pm .

3. Numerical experiments

Numerical simulation of shear instabilities of balanced flow are diagnosed in terms of ageostrophic balanced modes given by Eq. (27). A single-layer model and a primitive equation model are used. Although less realistic than the primitive equation model, we use the single-layer model here because of its reduced complexity for which the method can be easier applied and tested. Further, it is possible to generate larger lateral shear in the initial conditions in the single-layer model, since in the primitive equation model too strong shear can generate convective or symmetric instabilities during the subsequent integration, even when the initial conditions are convectively stable.

The spatial discretization of the primitive equation model is outlined in the appendix, and it follows the example of standard C-grid ocean models such as the MITgcm (Marshall et al. 1997). The discretization of the single-layer model is similar and detailed in Eden et al.

(2019). The time-stepping scheme in both models is Euler forward for the first time step and afterward an Adam–Bashforth two-time level interpolation with adjusted weights to allow for a stable integration. In the primitive equation model, all vertical modes including the barotropic mode are integrated explicitly with sufficiently small time step, since we find that using an implicit method to solve for the barotropic mode to allow for a larger time step, the diagnostic of the higher-order ageostrophic balanced modes becomes inaccurate. The time step in the model is chosen according to the criteria given in the appendix.

a. The wave diagnostic

The following method to differentiate gravity waves from balanced flow in the model solution is applied: any given model state (\mathbf{u}, p) from the primitive equation model is projected on the vertical eigenfunctions using the discrete version of the eigenvalue problem appropriate to the numerical model given in the appendix. For the single-layer model, this first step is not necessary. Then an horizontal Fourier transform is applied to (\mathbf{u}_n, p_n) and the spectral amplitudes are projected on the left eigenvector \mathbf{p}^0 of the discrete version of the system matrix \mathbf{A} to obtain the slow mode amplitude g^0 . The inverse (vertical and horizontal) transform of $g^0 \mathbf{q}^0$ —where \mathbf{q}^0 is the right

eigenvector of \mathbf{A} —is then used to calculate the nonlinear tendency terms in the numerical model, which yields $I^\pm(g^0, 0)$ after forward transform and projection on \mathbf{p}^\pm , and thus finally f_1^\pm from Eq. (27). Corresponding steps yield the other ageostrophic balanced modes f_i^\pm . The inverse (vertical and horizontal) transform of $g^0\mathbf{q}^0 + \sum_{i=1}^K \text{Ro}^i f_i^\pm \mathbf{q}^\pm$ yields the balanced part of order K of the model state. The difference to the actual model state is interpreted as the residual gravity wave signal.

b. The single-layer model

The single-layer model on a double periodic domain size of 10×5 is initialized with an unstable zonal flow with an initial profile of $\mathbf{u}(y) = (2 \exp[-(y - 1.25)^2/0.1^2 - (y - 3.75)^2/0.1^2], 0)$ to which a small random perturbation is added. The initial conditions are balanced to first order, higher-order balances do not change the results discussed here. Since we use scaled equations as detailed after Eq. (1), the scaled Coriolis parameter is set to one and we also choose $c = 1$, such that the single layer corresponds to the first baroclinic mode using the scaled depth $h = \pi$. The domain is resolved with 400×200 grid points and we choose biharmonic friction and mixing operators as detailed in section 2g with diffusivity $A = (\text{Ro}/2)\Delta x^4$. The zonal jet begins to meander and dissolves finally into eddies in the simulations. The length scale of the eddies is close to the scaled (first baroclinic) Rossby radius $c = 1$, which determines our choice of the model domain. Increasing the resolution of the model domain therefore does not change much the results presented here since the lateral shear instability is well resolved.

We present simulations with the single layer using different parameters, from $\text{Ro} = 0.02$ to $\text{Ro} = 0.3$, that is, ranging from mesoscale conditions with small Ro to the so-called submesoscale conditions with finite Ro . Table 1 presents statistics of the local Rossby number $\text{Ro}_l = \text{Ro}|\partial_x v - \partial_y u|$ after the onset of the instability in the simulations, showing that locally much larger Ro_l can be reached in the simulations than specified by the parameter Ro .

Figure 1 shows a time series of the ageostrophic balanced zonal velocity and the residual wave signal up to fourth order and the full thickness p at $x = 0$ in a simulation with $\text{Ro} = 0.1$. With the onset of the instability at $t \approx 10$, the ageostrophic balanced mode increases in amplitude in all orders. Figure 1a shows the zonal velocity of the first-order ageostrophic mode u_1 and Fig. 1b shows $u - u_0$. We call u_0 and u_i the zonal velocity corresponding to the inverse transform of $g^0\mathbf{q}^0$ and $\text{Ro}^i f_i^\pm \mathbf{q}^\pm$, respectively, while u denotes the actual zonal velocity in the model. Note that u_1 is the first-order ageostrophic

TABLE 1. Spatial mean $\overline{\text{Ro}_l}$, spatial maximum $\text{Ro}_m = \max(\text{Ro}_l)$, and mean of the largest 5% $\text{Ro}_{5\%}$ of the actual local Rossby number $\text{Ro}_l = \text{Ro}|\partial_x v - \partial_y u|$ in the simulations with the single-layer model for different parameters Ro after the onset of the instability.

Ro	$\overline{\text{Ro}_l}$	Ro_m	$\text{Ro}_{5\%}$
0.02	0.04	0.30	0.21
0.05	0.10	0.75	0.52
0.10	0.20	1.49	1.04
0.15	0.28	2.23	1.49
0.20	0.32	2.80	1.70
0.25	0.34	3.45	1.49
0.30	0.41	4.00	1.77

velocity in quasigeostrophic approximation. Since the difference between u_1 and $u - u_0$ is very small, there is no residual wave signal in the first order. The same is true for the second order, shown in Figs. 1c and 1d, and also for the third order (Figs. 1e,f).

Only in the fourth order (Figs. 1g,h) do small differences between u_4 and $u - u_0 - u_1 - u_2 - u_3$ begin to show up, which points to a small residual gravity wave signal in the fourth order. This residual wave signal increases going to larger Ro . Figure 2 shows that the wave signal begins to emerge for $\text{Ro} = 0.15$ already in the third order, and that at fourth order a clear large-scale near-inertial wave signal is present.

Figure 3a shows the residual wave energy $E_K = |\mathbf{u} - \sum_{n=0}^K \mathbf{u}_n|^2/2 + (p - \sum_{n=0}^K p_n)^2/2$ as a function of Ro normalized with the total energy $|\mathbf{u}|^2/2 + p^2/2$ (which does not vary much for different Ro) during the peak of the instability in the simulations. The first-order residual energy E_1 scales as Ro^4 such that all the energy can be attributed to the second-order ageostrophic balanced mode for all Ro , since the amplitude of the second-order ageostrophic balanced mode scales as Ro^2 . Correspondingly, E_2 scales roughly as Ro^6 and E_4 as Ro^8 for small Ro , but this scaling behavior gets distorted for larger Ro due to the emergence of the residual gravity wave signal seen in Fig. 2. An exponential fit $\sim \exp(-2/\text{Ro})$ seems to fit the residual gravity wave signal. For small Ro , however, the exponential fit deviates, and also E_4 does not scale as Ro^{10} for small Ro . This points toward a problem in the diagnosis of the fourth-order ageostrophic balanced mode with our method, most likely due to numerical precision errors. In any case, the diagnostics reveals a finite gravity wave signal for $\text{Ro} > 0.1$. This wave signal could be related to spontaneous gravity wave emission generated by shear instability of the balanced flow, but note that local Rossby numbers Ro_l in Table 1 exceed unity such that symmetric instability due to lateral shear becomes possible. Therefore, the wave signal could also relate to this instability process instead

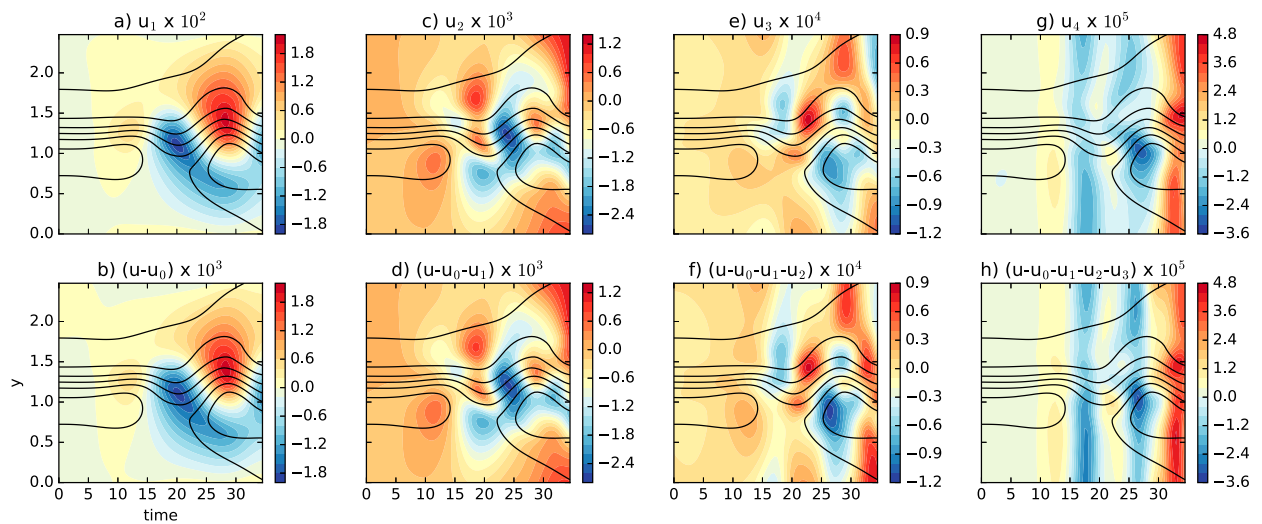


FIG. 1. (top) The ageostrophic balanced zonal velocity and (bottom) residual wave signal up to fourth order at $x = 0$ for $Ro = 0.1$ in the single-layer model. The terms u_0 and u_i denote the zonal velocity corresponding to the inverse transform of $g^0 \mathbf{q}^0$ and $Ro^i f_i^\pm \mathbf{q}^\pm$, respectively. Also shown is the full thickness p as contour lines. Note the differences in the color scales.

of spontaneous gravity wave emission in the classical Lighthill sense.

c. The primitive equation model

The primitive equation model is configured similar to the single-layer model on a double periodic domain of size $5 \times 5 \times 1$ resolved by $255 \times 255 \times 40$ grid points, in the zonal, meridional, and vertical direction, respectively. The model is initialized with a zonal velocity shaped in the horizontal identical to the single-layer model, but here with the first vertical mode as vertical structure. The initial velocity, pressure, and buoyancy fields are balanced to first order in Ro ; higher-order

balancing does not change the results. Since we use scaled equations, the (scaled) background stratification (or Brunt-Väisälä frequency) is one and the Coriolis parameter is also one. Compared to the single-layer model, the amplitude of the initial zonal velocity u is reduced to 0.2 to exclude the occurrence of static instabilities with $\partial_z b \leq -1$ that arise at higher amplitudes of u and show up locally in the simulations. The scaled first baroclinic Rossby radius is $L_r = c_1 = 1/\pi$, which is as in the single-layer model close to the length scale of the instabilities showing up after finite time in the integrations, and also determines the choice of the model domain. Therefore, increasing the horizontal and

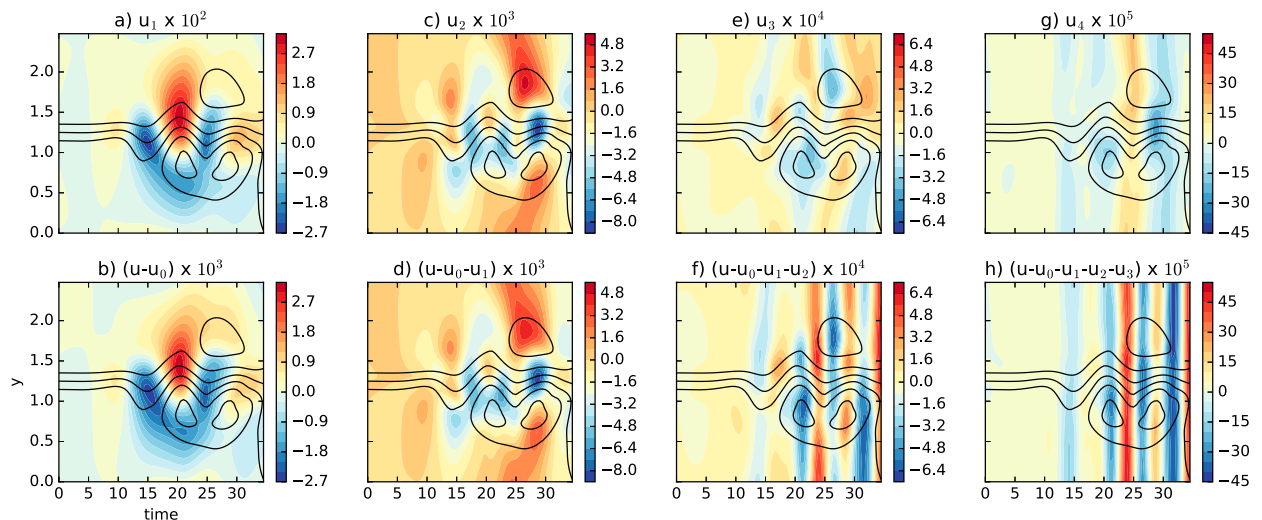


FIG. 2. As in Fig. 1, but for $Ro = 0.15$.

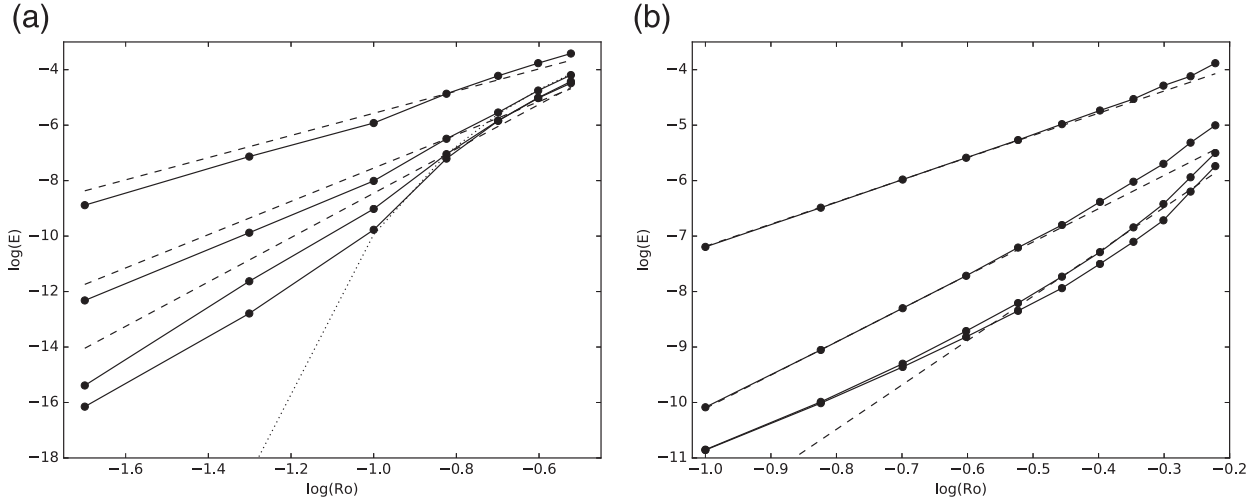


FIG. 3. (a) Total residual wave energy $E_K = |\mathbf{u} - \sum_{n=0}^K \mathbf{u}_n|^2/2 + p - (\sum_{n=0}^K p_n)^2/2$ for $K = 1, \dots, 4$ normalized with the total energy integrated over the model domain at the peak of the instability in the single-layer model. The term E_1 corresponds to the uppermost dots and E_4 to the lowest dots. (b) As in (a), but for the primitive equation model and with total residual wave energy $E_K = |\mathbf{u} - \sum_{n=0}^K \mathbf{u}_n|^2/2 + (b - \sum_{n=0}^K b_n)^2/2$ also normalized with the total energy. Dashed lines show different power laws Ro^4 , Ro^6 , and Ro^8 , and the dotted line in (a) an exponential scaling law $\exp(-2/\text{Ro})$.

vertical resolution does not change the integrations significantly and also does not affect the decomposition into balanced and unbalanced flow.

The scaled barotropic Rossby radius L_0 is set to 5 by adjusting the parameter $g = L_0^2/L_r^2$ accordingly. Larger values of L_0 do not change the simulations much, but require a much smaller time step, therefore this rather small ratio between L_0 and L_r is chosen. The lateral damping is identical to the single-layer model. While we time step the model only once in the single-layer model to calculate $\partial_T f_n^\pm$ in Eq. (27), we find it necessary to calculate 50 time steps in the primitive equation model since otherwise large gridscale noise shows up in the higher-order f_n^\pm . Furthermore, it is necessary to exchange Eq. (27) with Eq. (22) and Eq. (24) but with complex ω^\pm for the 40 largest wavenumbers, as discussed in section 2g.

As for the single-layer model, we present simulations for different Ro with a range from $\text{Ro} = 0.10$ to $\text{Ro} = 0.60$, that is, from mesoscale conditions with small Ro to the so-called submesoscale conditions with finite Ro . Table 2 presents statistics of the local Rossby number $\text{Ro}_l = \text{Ro} |\partial_x v - \partial_y u|$ after the onset of the instability in the simulations. Although smaller in general than in the single-layer model because of the reduced initial zonal velocity amplitude, also here larger Ro_l are reached locally in the simulations than specified by the parameter Ro .

Figure 4 shows the ageostrophic balanced vertical velocity up to fourth order, that is, from w_1 to w_4 and the corresponding residual wave signals in the primitive

equation model after the onset of the instability for $\text{Ro} = 0.5$. Again, the first-order ageostrophic vertical velocity w_1 is very similar to the actual vertical velocity, and the second-order balanced vertical velocity w_2 is very similar to $w - w_1$, which is the second-order residual signal. This similarity between w_2 and $w - w_1$ indicates that most of the residual signal seen in $w - w_1$ is in fact related to the balanced mode (and is the so-called ageostrophic balanced mode or slaved mode), and not to the unbalanced wave signal it appears to be. Note that w_1 corresponds to the quasigeostrophic vertical velocity, which can also be diagnosed from the omega equation (not shown), and that w_2 is an order of magnitude smaller but still entirely related

TABLE 2. As in Table 1, but at the surface of the primitive equation model. The last line shows the same quantities but for the simulation with a larger amplitude of 1 in the zonal velocity of the initial condition instead of 0.2.

Ro	$\overline{\text{Ro}}_l$	Ro_m	$\text{Ro}_{5\%}$
0.10	0.01	0.19	0.10
0.15	0.02	0.28	0.15
0.20	0.03	0.37	0.19
0.25	0.03	0.45	0.25
0.30	0.04	0.53	0.31
0.35	0.05	0.62	0.34
0.40	0.06	0.70	0.40
0.45	0.07	0.77	0.47
0.50	0.07	0.88	0.46
0.55	0.08	0.92	0.58
0.60	0.08	1.01	0.57
0.10	0.08	1.10	0.53

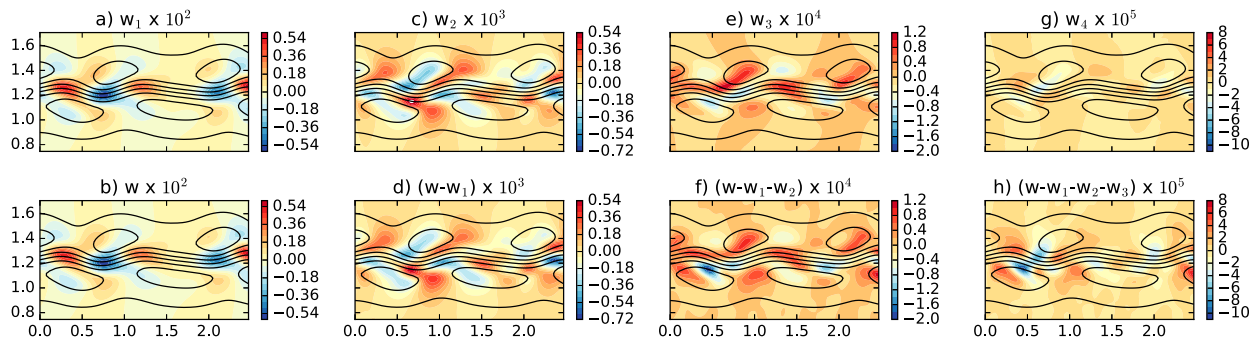


FIG. 4. (top) Ageostrophic balanced vertical velocity (color; from w_1 to w_4) and (bottom) residual wave signal (from $w - w_1$ to $w - \sum_{i=1}^3 w_i$) up to fourth order at $z = 0.2$ in the primitive equation model for $Ro = 0.5$. The term w_i denotes the vertical velocity corresponding to the inverse transform of $Ro^i f_i^\pm \mathbf{q}^\pm$. Shown is a time step after the onset of the instability. Also shown is the full pressure as contour lines. Note the different color scales.

to the balanced mode. There is also good agreement between w_3 and $w - w_1 - w_2$ but small differences already show up, while w_4 deviates more from the corresponding residual wave signal. This behavior is also seen for smaller Ro , that is, agreement up to w_3 but deviation for w_4 .

The apparent wave signal at fourth order (i.e., in $w - w_1 - w_2 - w_3$ in Fig. 4) for small Ro is most likely related to numerical precision errors and not to a wave signal. This can be seen in Fig. 3b in which the total residual energy at the fourth order (E_4) shows a very similar scaling to that of the third order (E_3). Since E_3 contains mostly ageostrophic balanced modes (as seen in Fig. 4 for $Ro = 0.5$ which also holds for different Ro), this indicates that the fourth-order residual energy E_4 is also largely composed of the ageostrophic balanced modes. Thus, the wavelike signals seen in vertical velocity in Fig. 4h are only an apparent signal, arising most likely from a numerical artifact, which we have not been able to remove², although we use the discretized equations consistent on the grid level to eliminate numerical inaccuracy (we found that small inconsistencies on the discrete level can lead to first-order inaccuracies in the flow decomposition). Similar to Fig. 3a, however, for $Ro > 0.5$ a deviation from the corresponding power laws is seen in all E_n in Fig. 3b, but here for much larger Ro . These results lead us to conclude that the residual wave signal is very small in this configuration of the primitive equation model and begins to be detectable with our

decomposition method only at very large Ro . Then it appears to follow also an exponential power law as in Fig. 3a. We therefore conclude that there is hardly any spontaneous wave emission by the shear instability of the balanced flow in this model simulation for $Ro < 0.5$.

The situation changes when allowing for amplitudes of the zonal velocity in the initial conditions larger than 0.2. In this case, however, static instabilities with $\partial_z b \leq -1$ are generated by the growing shear instability, even when the initial conditions are stably stratified. This is not the case in the previous model experiment, where the amplitude was chosen as large as possible but without static instabilities in the simulations. For the experiment with larger velocity amplitudes, we thus apply a convective adjustment to the model with a vertical diffusivity, which takes a value of $K = 1$ if $\partial_z b < -1$ and $K = 0$ elsewhere, but we do not account for this convective adjustment in the diagnostic of the ageostrophic balanced mode.

Figure 5 shows the ageostrophic balanced vertical velocity up to third order, w_1 , w_2 , and w_3 , and the corresponding residual wave signals (i.e., w , $w - w_1$, $w - w_1 - w_2$) for an initial zonal velocity amplitude of 1 and for $Ro = 0.1$. The actual local Rossby numbers Ro_l for this simulation are also listed in Table 2. They are larger than in the previous simulations with the primitive equation model avoiding static instabilities. Already in the vertical velocity w a wave signal is present which dominates in the second and higher orders over the balanced mode which is shown in Figs. 5a, 5c, and 5e. The wave crests in Figs. 5b, 5d, and 5f are oriented roughly perpendicular with a wavelength of about 0.2–0.3 with a larger vertical mode than the initial conditions. The emergence of the wave signal shows up for all Ro we have tested (0.02–0.3) and can be related to the regions of static instabilities with (or close to) $\partial_z b \leq -1$ in space and time. We therefore conclude that the waves

²Note that tests with balanced initializations of the primitive equation model without any subsequent instability also show the absence of a fast adjustment by gravity waves up to third order, but a fast wave signal due to an adjustment from the initial conditions in the fourth order, indicating an imperfect balanced initialization on this order. These integrations are, however, not presented here.

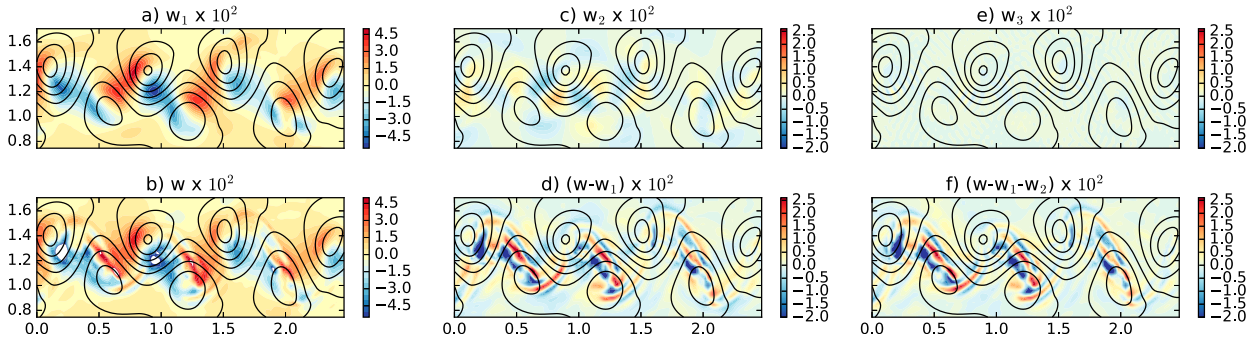


FIG. 5. As in Fig. 4, but with larger amplitude of 1 in the zonal velocity of the initial condition and for $Ro = 0.1$.

we do detect in our simulations are actually generated by convective or symmetric instabilities occurring in regions where $\partial_z b \sim -1$ —or where potential vorticity becomes negative as criterion for symmetric instability (Olbers et al. 2012)—but not by spontaneous wave emission.

Experiment with the primitive equation model without strong lateral but vertical shear, where predominantly baroclinic instability shows up, behaves the same as the above discussed experiments. For amplitudes in vertical shear that do not generate static instabilities, hardly any wave signal is detected, only with increased shear and occasional static instabilities a clear wave signal similar to Fig. 5 is generated.

4. Summary and discussion

This study is an elaborate attempt to find spontaneous gravity wave emission during lateral and vertical shear instability of balanced flow. We have designed a novel numerical tool to uniquely differentiate the slow balanced flow from the fast gravity waves up to fourth order based on an asymptotic expansion in the Rossby number Ro as suggested in, for example, Warn et al. (1995) and Kafiabad and Bartello (2018). Here, we apply the method the first time up to fourth order and in a model of growing instabilities to answer the specific question of spontaneous wave emission and its dependency on Ro , whereas the method was applied to a model of decaying isotropic turbulence by Kafiabad and Bartello (2018) only up to second order in Ro . The concept is applied here on the discrete level of the numerical models since we find that this high accuracy is necessary to allow for orders of the expansion larger than one. In principle, our tool allows for arbitrary precision, in practice this limit is hampered by numerical (lack of precision) and mathematical (lack of asymptotic limit) issues. Unlike the available approximate methods to differentiate between waves and balanced flow, such as the Lagrangian frequency method (Shakespeare and Taylor 2015) or the

optimal potential vorticity balance method (Viúdez and Dritschel 2004), a Fourier transform of the model state is necessary in Warn et al.'s approach. The application of our decomposition tool is therefore more straightforward to idealized settings, whereas the approximate methods can be more readily applied to realistic scenarios but at the cost of a less accurate decomposition and perhaps a misinterpretation of the wavelike signals. We plan a comparison of the performance of the different methods in idealized models to transfer the results and to use the methods in the realistic ocean models as well.

In a numerical simulation of lateral shear instability with a single-layer model our novel tool can be successfully applied up to fourth order, which allows us to diagnose the gravity wave signal to this order of accuracy. Only in the case of a strong lateral shear, a near inertial wave signal is detected at higher orders of Ro in the model simulations. This result of weak wave generation is in agreement to McIntyre and Norton (2000) who also reported a minor role of wave generation in a single layer appropriate to the global atmosphere. The small but finite wave generation at large Ro we found could be related to spontaneous gravity wave emission in the sense of Lighthill radiation, but since local Rossby numbers well exceed unity, symmetric instability due to the lateral shear could also be the cause of the wave signal. The amplitude of the wave signal, however, shows an exponential scaling with respect to Ro in agreement to previous suggestions of classical spontaneous wave generation (Vanneste and Yavneh 2004; Vanneste 2013).

As compared to the decomposition up to fourth order in Ro in the single-layer model, our novel tool can robustly identify wave signals at least up to third order in Ro in the vertically resolved model, where the reduction in accuracy is most likely related to the error in numerical precision. In the vertically resolved model with lateral and vertical shear small enough to avoid static instabilities, the diagnosis and the scaling behavior with respect to Ro suggest that there is a negligibly small

wave signal. Only if the shear becomes strong enough to generate static instabilities the emission of waves is detected with lateral wavelengths of a fraction of the Rossby radius and a higher vertical mode than the balanced flow. This emission is localized in space and time to the occurrence of static instabilities and the waves closely resemble the wavelike signals often seen before in numerical simulations (e.g., [Plougonven and Snyder 2007](#); [Hien et al. 2018](#); [Chouksey et al. 2018](#)).

Relatively larger amplitudes of waves generated by convective or symmetric instabilities compared to the ones generated by spontaneous emission are also found and discussed in [Chouksey \(2018\)](#) in a suite of numerical experiments with different configurations and for a range of different Ro. There, however, the method to differentiate waves from balanced flow follows the one by [Machenbauer \(1977\)](#) or uses the omega equation, and is thus only accurate to first order in Ro, such that the amplitude of spontaneous wave emission is most likely overestimated.

We cannot rule out the possibility that in other model configurations than what we have discussed here (and in [Chouksey 2018](#)) spontaneous wave emission in the sense of Lighthill radiation plays a more important role. We have tested several different configurations including also ones with vertical shear as before but only weak lateral shear and always found the same result: hardly any wave generation in general, but enhanced wave generation when static instabilities are generated by the balanced flow. In all such cases, the wave signal is localized in space and time to the sites of static instability. We also cannot rule out the possibility that waves at much larger or smaller wavelengths than allowed by our model grid can be generated. On the other hand, the model and laboratory experiments in the literature refer to spontaneously generated waves at similar wavelength as the dynamical process under investigation, in agreement to the (true) wave signal seen in our experiments with convective and/or symmetric instabilities.

These results suggest a relatively small role of spontaneous wave emission in the classical sense of Lighthill radiation, and emphasize the role of convective or symmetric instabilities during frontogenesis for the generation of internal waves. We speculate that the internal wave signals seen and discussed in previous studies based on numerical simulations and laboratory experiments (e.g., [Plougonven and Snyder 2007](#); [Hien et al. 2018](#); [Chouksey et al. 2018](#)) may also be generated by convective or symmetric instabilities instead of the classical spontaneous emission mechanism.

Acknowledgments. This paper is a contribution to the Collaborative Research Centre TRR 181 “Energy Transfer in Atmosphere and Ocean” funded by the

Deutsche Forschungsgemeinschaft (DFG, German Research Foundation), Projektnummer 274762653.

APPENDIX

Discrete Relations

The discretization of the primitive equations is given by

$$\begin{aligned}\partial_t u - \overline{v^{i+1}}^- + \delta_i^+ p &= n^u - A(\delta_i^- \delta_i^+ + \delta_j^- \delta_j^+)^2 u, \\ \partial_t v + \overline{u^{i+1}}^+ + \delta_j^+ p &= n^v - A(\delta_i^- \delta_i^+ + \delta_j^- \delta_j^+)^2 v, \\ \partial_t b + \overline{w_k}^- &= n^b - A(\delta_i^- \delta_i^+ + \delta_j^- \delta_j^+)^2 b, \quad (\text{A1})\end{aligned}$$

and follows a standard C-grid arrangement. The indices i, j, k denote discretization in the $x, y,$ and z directions, but shown only when they are relevant in the individual terms. Finite differencing operators $\delta_i^+ p = (p_{i+1} - p_i)/\Delta x = \delta_{i+1}^- p$, where Δx is the grid spacing in x direction, and averaging operators $\overline{p_i}^+ = (p_i + p_{i+1})/2 = \overline{p_{i+1}}^-$ and similar for the other directions with grid spacing Δy and Δz are introduced. The vertical grid has grid points $k = 1, \dots, N$ with pressure and horizontal velocity at depth level $z_k = \Delta z(N - k - 1/2)$, vertical velocity at depth level $z_k + \Delta z/2$, and surface pressure p_s at $z_N + \Delta z/2 = 0$. The surface boundary condition is then given by

$$\partial_t p_s = -g\Delta z \sum_{k'=1}^N \nabla \cdot \mathbf{u}_{k'} - A(\delta_i^- \delta_i^+ + \delta_j^- \delta_j^+)^2 p_s. \quad (\text{A2})$$

As discussed above, the damping term is included here since otherwise the vertical modes get coupled by the damping. The nonlinear terms $n^u, n^v,$ and n^b are given for the primitive equations by

$$\begin{aligned}n_u &= -\delta_i^- (\overline{u_i^+} \overline{u_i^+}) - \delta_j^- (\overline{v_j^+} \overline{u_j^+}) - \delta_k^- (\overline{w_k^+} \overline{u_k^+}), \\ n_v &= -\delta_i^- (\overline{u_j^+} \overline{v_j^+}) - \delta_j^- (\overline{v_j^+} \overline{v_j^+}) - \delta_k^- (\overline{w_k^+} \overline{v_k^+}), \\ n_b &= -\delta_i^- (\overline{u b_i^+}) - \delta_j^- (\overline{v b_j^+}) - \delta_k^- (\overline{w b_k^+}). \quad (\text{A3})\end{aligned}$$

For the single-layer model, the discretization of the nonlinear terms in the momentum equation follows the energy-conserving scheme by [Sadourny \(1975\)](#) and similar to n_b in Eq. (A3) for the thickness equation.

Integrating the discrete buoyancy equation vertically and using the hydrostatic relation $\delta_k^+ p = \overline{b_k^+}$, the surface boundary condition, and the integrated continuity equation $w_k = -\Delta z \sum_{k'=1}^k \nabla \cdot \mathbf{u}_{k'}$ yields

$$\begin{aligned}\partial_t p_k &= -M(\nabla \cdot \mathbf{u}_k) - \Delta z \left(n_k^b/2 + \sum_{k'=k+1}^{k'=N} n_{k'}^b \right) \\ &\quad - A(\delta_i^- \delta_i^+ + \delta_j^- \delta_j^+)^2 p_k \quad (\text{A4})\end{aligned}$$

with the discrete vertical mode operator

$$M(\phi_k) = \Delta z \left(g - \frac{\Delta z}{2} \right) \sum_{k'=1}^N \phi_{k'} + \Delta z^2 \sum_{k'=k}^{k'=N} \sum_{k''=1}^{k'} \phi_{k''} - \frac{\Delta z^2}{4} \phi_k. \tag{A5}$$

The discrete operator M has eigenvalues c_n^2 and eigenfunctions ϕ_k^n with $M(\phi_k^n) = c_n^2 \phi_k^n$. Taking two finite differences of the eigenvalue equation yields

$$\left(c_n^2 + \frac{\Delta z^2}{4} \right) \delta_k^- \delta_k^+ \phi^n = -\phi_k^n, \quad \left(g + \frac{\Delta z}{2} \right) \delta_N^+ \phi^n + \phi_N^n = 0, \\ \delta_1^- \phi^n = 0. \tag{A6}$$

Both the integral form and the solution to the differential form yield identical eigenfunctions and eigenvalues for the discrete grid. Expressing $\mathbf{u}_{i,j,k} = \sum_n \mathbf{u}_{ij}^n \phi_k^n$ and $p_{i,j,k} = \sum_n p_{ij}^n \phi_k^n$ and applying a horizontal Fourier transform yields the discrete version of the system matrix \mathbf{A} of Eq. (9) given in Eden et al. [2019, their Eq. (50)], although here equipped with damping terms on the diagonal and for vertical mode number n . The eigenvectors and eigenvalues of \mathbf{A} are therefore also identical to the ones in Eden et al. [2019, their Eqs. (51)–(53)] but with $c = c_n$, and the eigenvalues also get an imaginary part related to the damping. The corresponding vertical mode expressions for the auxiliary variables b and w are $w_{i,j,k} = \sum_n (c_n^2 + \Delta z^2/4) \nabla \cdot \mathbf{u}_n \delta_k^+ \phi^n$ and $b_{i,j,k} = \sum_n p_n [\delta_k^- \phi^n - \Delta z / (4c_n^2) (\phi_k^n + \phi_{k-1}^n)]$.

The time stepping scheme solving Eq. (9) (i.e., its inverse Fourier transform) is

$$\mathbf{z}_{n+1} - \mathbf{z}_n = \Delta_t [i\mathbf{A} \cdot (\alpha \mathbf{z}_n + \beta \mathbf{z}_{n-1}) + \alpha \hat{\mathbf{n}}(\mathbf{z}_n) + \beta \hat{\mathbf{n}}(\mathbf{z}_{n+1})] \tag{A7}$$

with time levels n , the time step Δt , and with the interpolation weights $\alpha = 1.5 + \epsilon$, $\beta = -(0.5 + \epsilon)$. The parameter $\epsilon > 0$ allows for a stable integration if Δt is chosen appropriately small. The value of Δt can be estimated from the eigenvalues of the linear discrete system $\bar{\omega}^s$, which differ in general from the physical eigenvalues ω^s and are given by

$$\bar{\omega}_{1,2} = i \ln(z_{1,2}) / \Delta t, \quad z_{1,2} = \frac{1}{2} - i\alpha \frac{\Delta t}{2} \omega^s \\ \pm \frac{1}{2} \sqrt{1 - 4i\Delta t \omega^s \left(\beta + \frac{\alpha}{2} \right) - (\Delta t \omega^s \alpha)^2}. \tag{A8}$$

The complex frequency $\bar{\omega}_1^s$ corresponds to a spurious numerical mode, which is always strongly damped and cannot be seen in the integrations. The other frequency $\bar{\omega}_2^s$ converges to ω^s for $\Delta t \rightarrow 0$, but for finite Δt it differs from ω^s and becomes complex. For $\epsilon = 0$, $\Im(\bar{\omega}_2^s) > 0$ for all

wavenumbers such that the scheme becomes unstable. But the parameter $\epsilon > 0$ can be chosen for given Δt , such that $\Im(\bar{\omega}_2^s) < 0$ for all vertical modes and lateral wavenumbers possible on the model grid. We choose $\epsilon = 0.1$ and adjust Δt accordingly. The integrations are then always stable.

REFERENCES

Baer, F., and J. J. Tribbia, 1977: On complete filtering of gravity modes through nonlinear initialization. *Mon. Wea. Rev.*, **105**, 1536–1539, [https://doi.org/10.1175/1520-0493\(1977\)105<1536:OCFOGM>2.0.CO;2](https://doi.org/10.1175/1520-0493(1977)105<1536:OCFOGM>2.0.CO;2).

Chouksey, M., 2018: Disentangling gravity waves from balanced flow. Ph.D. thesis, Universität Hamburg, <https://10.17617/2.3000108>.

—, C. Eden, and N. Brüggemann, 2018: 2018: Internal gravity wave emission in different dynamical regimes. *J. Phys. Oceanogr.*, **48**, 1709–1730, <https://doi.org/10.1175/JPO-D-17-0158.1>.

Eden, C., M. Chouksey, and D. Olbers, 2019: Mixed Rossby–gravity wave–wave interactions. *J. Phys. Oceanogr.*, **49**, 291–308, <https://doi.org/10.1175/JPO-D-18-0074.1>.

Ford, R., M. McIntyre, and W. Norton, 2000: Balance and the slow quasimanifold: Some explicit results. *J. Atmos. Sci.*, **57**, 1236–1254, [https://doi.org/10.1175/1520-0469\(2000\)057<1236:BATSQS>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<1236:BATSQS>2.0.CO;2).

Hien, S., J. Rolland, S. Borchert, L. Schoon, C. Zülicke, and U. Achatz, 2018: Spontaneous inertia–gravity wave emission in the differentially heated rotating annulus experiment. *J. Fluid Mech.*, **838**, 5–41, <https://doi.org/10.1017/jfm.2017.883>.

Hoskins, B., I. Draghici, and H. Davies, 1978: A new look at the ω -equation. *Quart. J. Roy. Meteor. Soc.*, **104**, 31–38, <https://doi.org/10.1002/qj.49710443903>.

Kafiabad, H. A., and P. Bartello, 2018: Spontaneous imbalance in the non-hydrostatic Boussinesq equations. *J. Fluid Mech.*, **847**, 614–643, <https://doi.org/10.1017/jfm.2018.338>.

Leith, C. E., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.*, **37**, 958–968, [https://doi.org/10.1175/1520-0469\(1980\)037<0958:NNMIAQ>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<0958:NNMIAQ>2.0.CO;2).

Lighthill, J., 1978: *Waves in Fluids*. Cambridge University Press, 516 pp.

Lorenz, E. N., 1980: Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.*, **37**, 1685–1699, [https://doi.org/10.1175/1520-0469\(1980\)037<1685:ASAQGE>2.0.CO;2](https://doi.org/10.1175/1520-0469(1980)037<1685:ASAQGE>2.0.CO;2).

—, 1986: On the existence of a slow manifold. *J. Atmos. Sci.*, **43**, 1547–1558, [https://doi.org/10.1175/1520-0469\(1986\)043<1547:OTEOAS>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<1547:OTEOAS>2.0.CO;2).

Machenhauer, B., 1977: On the dynamics of gravity oscillations in a shallow water model with applications to normal mode initialization. *Beitr. Phys. Atmos.*, **50**, 253–271.

Marshall, J., A. Adcroft, C. Hill, L. Perelman, and C. Heisey, 1997: A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *J. Geophys. Res.*, **102**, 5753–5766, <https://doi.org/10.1029/96JC02775>.

McIntyre, M. E., and W. A. Norton, 2000: Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.*, **57**, 1214–1235, [https://doi.org/10.1175/1520-0469\(2000\)057<1214:PVIOAH>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<1214:PVIOAH>2.0.CO;2).

Olbers, D., J. Willebrand, and C. Eden, 2012: *Ocean Dynamics*. Springer, 703 pp.

Plougonven, R., 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, <https://doi.org/10.1175/JAS3953.1>.

- Sadourny, R., 1975: The dynamics of finite-difference models of the shallow-water equations. *J. Atmos. Sci.*, **32**, 680–689, [https://doi.org/10.1175/1520-0469\(1975\)032<0680:TDOFDM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1975)032<0680:TDOFDM>2.0.CO;2).
- Shakespeare, C. J., and J. Taylor, 2015: The spontaneous generation of inertia–gravity waves during frontogenesis forced by large strain: Numerical solutions. *J. Fluid Mech.*, **772**, 508–534, <https://doi.org/10.1017/jfm.2015.197>.
- Vanneste, J., 2013: Balance and spontaneous wave generation in geophysical flows. *Annu. Rev. Fluid Mech.*, **45**, 147–172, <https://doi.org/10.1146/annurev-fluid-011212-140730>.
- , and I. Yavneh, 2004: Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.*, **61**, 211–223, [https://doi.org/10.1175/1520-0469\(2004\)061<0211:ESIWAT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2004)061<0211:ESIWAT>2.0.CO;2).
- Viúdez, Á., and D. G. Dritschel, 2004: Optimal potential vorticity balance of geophysical flows. *J. Fluid Mech.*, **521**, 343–352, <https://doi.org/10.1017/S0022112004002058>.
- Warn, T., O. Bokhove, T. Shepherd, and G. Vallis, 1995: Rossby number expansions, slaving principles, and balance dynamics. *Quart. J. Roy. Meteor. Soc.*, **121**, 723–739, <https://doi.org/10.1002/qj.49712152313>.

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.