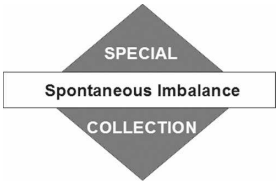

Exponential Smallness of Inertia–Gravity Wave Generation at Small Rossby Number

J. VANNESTE

School of Mathematics, and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, United Kingdom

(Manuscript received 11 April 2007, in final form 13 August 2007)

ABSTRACT

This paper discusses some of the mechanisms whereby fast inertia–gravity waves can be generated spontaneously by slow, balanced atmospheric and oceanic flows. In the small Rossby number regime relevant to midlatitude dynamics, high-accuracy balanced models, which filter out inertia–gravity waves completely, can in principle describe the evolution of suitably initialized flows up to terms that are exponentially small in the Rossby number ε, that is, of the form $\exp(-\alpha/\varepsilon)$ for some $\alpha > 0$. This suggests that the mechanisms of inertia–gravity wave generation, which are not captured by these balanced models, are also exponentially weak. This has been confirmed by explicit analytical results obtained for a few highly simplified models. These results are reviewed, and some of the exponential-asymptotic techniques that have been used in their derivation are presented. Two types of mechanisms are examined: spontaneous-generation mechanisms, which generate exponentially small waves from perfectly balanced initial conditions, and unbalanced instability mechanisms, which amplify unbalanced initial perturbations of steady flows. The relevance of the results to realistic flows is discussed.

1. Introduction

The atmosphere and oceans are typical examples of two-time-scale systems. As a result of fast rotation and strong stratification, their dynamics can be separated into a slow, or balanced, part evolving on an advective time scale L/U, and a fast part consisting of inertia–gravity waves (IGWs) evolving on time scales shorter than the inertial period f^{-1}. Thus, the Rossby number

$$\varepsilon = \frac{U}{fL},$$

which gives an estimate of the time-scale separation between the two types of motion, is generally small. A large time-scale separation leads to a weak coupling between the slow and fast components of the dynamics. For the atmosphere and oceans, which are forced mostly at low frequency, this has the consequence that the level of IGW activity often remains low; as a result, the dynamics can be represented accurately by balanced models, that is, models that filter out IGWs completely. This has been formalized by introducing the notion of slow manifolds, which are submanifolds of the state space on which the fast IGW motion is weak (Leith 1980; Lorenz 1980; Warn et al. 1995). A balance relation (such as geostrophic balance), which slaves fast dynamical variables to slow ones, is then regarded as the definition of a slow manifold, while the corresponding balanced model (such as the quasigeostrophic model) describes the dynamics constrained to this slow manifold.

By definition, slow manifolds are nearly invariant, in the sense that the full dynamics is nearly tangent to them. A natural question, therefore, is how close to completely invariant they can be. The answer to this question is clear in the case of finite-dimensional systems (MacKay 2004). In the absence of dissipative processes (the relevant assumption for large-scale atmospheric or oceanic dynamics), exactly invariant slow manifolds do not exist generically. Physically, this means that fast (or unbalanced) motion cannot be eliminated completely by initialization, that is, by projection of initial data onto a slow manifold; fast motion in the form of IGWs will emerge spontaneously in the flow evolution in spite of any attempt made at eliminating them from the initial conditions. This implies a fundamental limitation to the accuracy of balanced models as well as the existence of a physical mechanism for the generation of IGWs from balanced motion.

Corresponding author address: J. Vanneste, School of Mathematics, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom.
E-mail: j.vanneste@ed.ac.uk

DOI: 10.1175/2007JAS2494.1

The high accuracy that can be achieved by balanced models is remarkable, however. Indeed, the perturbative procedures used to define the slow manifolds (expansions in powers of ε or, equivalently, iterations) can in principle be carried out systematically to obtain a hierarchy of slow manifolds that are approximately invariant to higher and higher formal accuracy $O(\varepsilon^N)$ for $N = 1, 2, \ldots$. Restricting the dynamics to these slow manifolds then leads a hierarchy of balanced models that approximates to $O(\varepsilon^N)$ the evolution of balanced initial conditions. The nonexistence of a truly invariant slow manifold is reflected in the divergence of the perturbative procedures as $N \to \infty$, which limits the accuracy that can be achieved for fixed ε.

A maximum accuracy is attained by employing optimal truncation, that is, by stopping the perturbative procedures at an order $N(\varepsilon)$, which minimizes the error. General arguments then suggest that the most accurate balanced models obtained in this manner should have an error that is exponentially small in ε, namely of the type $\exp(-\beta/\varepsilon)$ for some $\beta > 0$. This can be proved for finite-dimensional toy models of geophysical fluids such as the Lorenz–Krishnamurthy (LK) model (Lorenz 1986; Lorenz and Krishnamurthy 1987) and others using standard methods (Gelfreich and Lerman 2002; Cotter and Reich 2006). For infinite-dimensional models of geophysical fluids, proofs are much more delicate and as yet only partial results have been obtained (Wirosoetisno 2004; Temam and Wirosoetisno 2007). One might nonetheless expect the exponential accuracy to hold true at least when the flows involved are sufficiently simple. Note that a key requirement for the exponential smallness in infinite-dimensional systems is that the frequency separation holds at all spatial scales; this is ensured in the small Rossby number regime on which this paper focuses.

The existence of exponentially accurate slow manifolds has an important physical implication: unbalanced, IGW phenomena that appear spontaneously from balanced motion are necessarily exponentially weak in the limit of small Rossby number. The qualifier "spontaneous" is crucial here, since it rules out all forms of IGW activity that can be eliminated by suitable initialization. This marks out a genuine physical mechanism for the generation of IGWs. This is in contrast to standard geostrophic adjustment (e.g., Reznik et al. 2001 and references therein), which describes the flow response to an initial imbalance that can have a variety of origins.

In the last few years, substantial efforts have been made to study the mechanism of spontaneous IGW generation and estimate its importance as a source of atmospheric IGW activity. In the small Rossby number regime, significant progress has been made along two lines of investigation. On the one hand, high-resolution numerical models of the three-dimensional fluid equations have been used to demonstrate that spontaneous IGW generation does occur in semirealistic flows such as baroclinic life cycles (e.g., O'Sullivan and Dunkerton 1995; Zhang 2004; Plougonven and Snyder 2005, 2007; Viúdez and Dritschel 2006; Viúdez 2006). To date, however, the amplitude dependence of the IGW on the Rossby number has not been established in these models. On the other hand, theoretical analysis of highly simplified models, typically described by ordinary differential equations, has established the exponential smallness of IGW generation explicitly (Vanneste and Yavneh 2004; Vanneste 2004; Ólafsdóttir et al. 2005). This has been achieved by obtaining asymptotic estimates for the IGW amplitude using the tools of exponential asymptotics (e.g., Olde Daalhuis 2003). The aim of the present paper is to review these analytical results while providing some background on the asymptotic methods employed. These methods shed some light on the nature of the IGW generation, showing, for instance, how a unique (noninvariant) optimal slow manifold can be defined, how the IGWs generated cannot be estimated from the knowledge of the leading-order balanced motion only, and how singularities of the balanced motion for complex values of time are crucial for the IGW amplitude. These conclusions are drawn on the basis of drastically simplified atmospheric models, but they likely hold true in more realistic situations, even though their explicit demonstration would be very challenging in this case.

This paper also discusses a set of recent results closely connected to the spontaneous generation of IGWs. These concern the instability of balanced flows to IGW-like or more generally unbalanced perturbations (McWilliams and Yavneh 1998; Yavneh et al. 2001; Molemaker et al. 2001, 2005; Plougonven et al. 2005; Vanneste and Yavneh 2007). In these papers, simple steady flows, which are stable under balanced approximations of arbitrary accuracy $O(\varepsilon^N)$, are shown to be in fact (spectrally) unstable for the full equations. The instabilities found involve unbalanced growing modes and hence are filtered out by balanced approximations. The exponential accuracy of the latter for small Rossby number indicates that the growth rates of the instabilities should be exponentially small. This has been confirmed, either numerically or analytically. Furthermore, the widths of the instability regions in parameter space (e.g., wavenumber space) are also shown to be exponentially small.

These steady-flow instabilities can be seen as simple, analytically tractable examples of a generic mechanism,

namely the instability of the slow manifold to infinitesimal perturbations off it. This mechanism differs from the spontaneous generation discussed above in that an initial imbalance is necessary. Unlike in the case of geostrophic adjustment, however, the amplitude of this initial imbalance does not affect the IGW-type behavior at later times, since this is controlled by saturation mechanisms that are likely independent of the initial perturbation. The effect of spontaneous generation and of unbalanced instabilities can therefore be expected to be similar, both providing mechanisms of IGW generation that may be difficult to distinguish in general time-dependent flows.

This paper is organized as follows: the construction of slow manifolds for two-time-scale systems is reviewed in section 2. This section briefly discusses an iteration scheme for the systematic improvement of the accuracy of slow manifolds, and it explains how this may be used to establish exponential accuracy (at least for finite-dimensional models). It also emphasizes the crucial difference between the small Rossby number regime considered in this paper, and the small Froude number regime. In the latter regime, time-scale separation holds only for a limited range of scales, leading to Lighthill-type wave generation, with power-law rather than exponential scaling of the wave amplitudes (see Ford et al. 2000, 2002; Saujani and Shepherd 2002). Section 3 is devoted to the asymptotic analysis of simple models of spontaneous generation of IGWs. It starts by a discussion of the forced harmonic oscillator, which provides an elementary example of the spontaneous generation of exponentially small fast oscillations analogous to IGWs. In this model, the slow part of the motion is represented by the forcing term, assumed to be determined a priori and independent of the evolution of the fast variables. This misses a central aspect of the dynamics of two-time-scale systems, namely that there is in general no exact split between slow and fast variables. Nevertheless, the forced harmonic oscillator provides an instructive example that serves to illustrate key asymptotic techniques (such as optimal truncation and Borel summation) that can be applied to somewhat more sophisticated applications. Two such applications, namely the Lorenz–Krishnamurthy model and sheared disturbances, are considered next. We review the explicit results obtained for these and comment on how they might generalize to more realistic models of the atmosphere and oceans. The unbalanced instabilities with exponentially small growth rates are examined in section 4. The instabilities considered there can be interpreted as resulting from the linear resonance between waves of a variety of types (Kelvin waves, IGWs, and edge waves); as a result they typically involve modes with asymptotically large wavenumbers, so that they can be analyzed in some detail using Wentzel–Kramers–Brillouin (WKB) approaches. The paper concludes in section 5 with a discussion.

2. Slow manifolds

The dynamical decoupling between balanced flow on the one hand, and IGWs on the other, relies on the existence of gap in the frequency spectrum of the equations of motion for a rotating stratified fluid on the f plane. When these are linearized and wave solutions of the form $\exp(i\mathbf{k}\cdot\mathbf{x})$ are introduced, the dispersion relation is found to have the three branches

$$\omega_{\rm bal} = 0 \quad \text{and} \quad \omega_{\rm igw} = \pm(f^2 \sin^2\vartheta + N^2 \cos^2\vartheta)^{1/2}, \tag{2.1}$$

where ϑ is the angle between the wavevector $\mathbf{k}$ and the horizontal. These branches are identified as the balanced or vortical mode, and upward and downward propagating inertia–gravity waves. The frequency spectrum (2.1) implies two time scales for the nonlinear equations of motion: the inertia–gravity time scale, bounded from above by f^{-1} (assuming a Brunt–Väisälä frequency $N > f$), and the nonlinear, advective time scale L/U. Choosing the latter as the reference time scale, the equations of motion can then be written in the form

$$\frac{\partial \mathbf{s}}{\partial t} = \mathcal{N}_{\mathbf{s}}(\mathbf{s}, \mathbf{f}), \tag{2.2}$$

$$\frac{\partial \mathbf{f}}{\partial t} + \frac{1}{\varepsilon}\mathcal{L}\mathbf{f} = \mathcal{N}_{\mathbf{f}}(\mathbf{s}, \mathbf{f}) \tag{2.3}$$

when suitable dependent variables are used (Warn et al. 1995). These distinguish the single slow field $\mathbf{s}$ (potential vorticity or a linearized version thereof) from the two fast fields $\mathbf{f} = (\mathbf{f}_1, \mathbf{f}_2)$ (e.g., divergence and ageostrophic vorticity). In (2.2)–(2.3), $\mathcal{L}$ is a linear operator, with a purely imaginary spectrum satisfying

$$\text{spec}\ \mathcal{L} \subset \{i\omega : \omega \in \mathbf{R}, |\omega| \geq 1\}, \tag{2.4}$$

and $\varepsilon = U/(fL) \ll 1$ is the Rossby number ($\mathbf{R}$ represents the set of real numbers). The other expressions, $\mathcal{N}_{\mathbf{s}}$ and $\mathcal{N}_{\mathbf{f}}$, group the nonlinear terms whose dependence on ε has been ignored for convenience. Note that the toy models for balance such as the LK model are also of the form (2.2)–(2.3).

Since $\varepsilon \ll 1$, the dynamics can be simplified using asymptotic methods developed for two-time-scale systems. The most standard of these is averaging, whereby (2.2)–(2.3) are averaged over the fast $O(\varepsilon^{-1})$ time scale

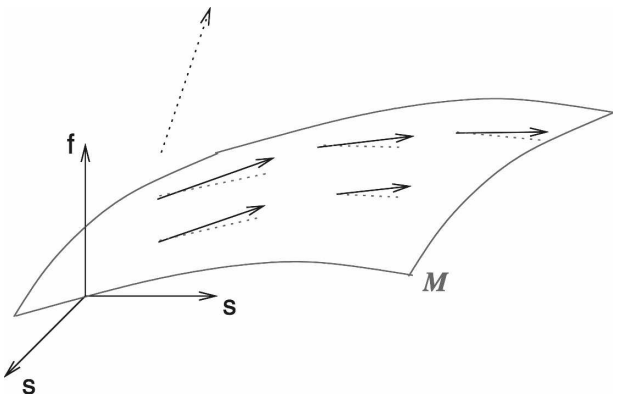

FIG. 1. Schematic of a slow manifold for a 3D system. The slow manifold $\mathcal{M}$ has the dimension of the set of slow variables $\mathbf{s}$, and is nearly invariant so that the vector fields governing the evolution (arrows) are nearly tangent to it. The slowness of the evolution on $\mathcal{M}$ is indicated by the fact that some distance away from it the vector field has a large component in the direction of the fast variable $\mathbf{f}$.

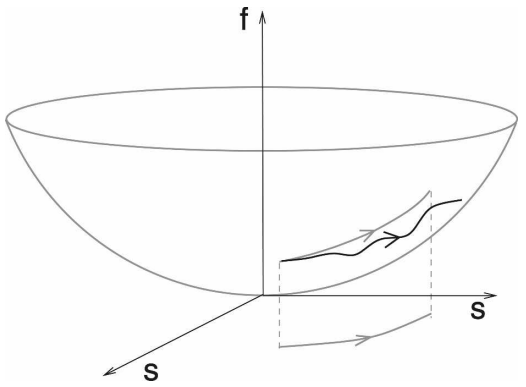

FIG. 2. Schematic illustrating the evolution of an initially balanced solution: the solution (wiggly curve) starts on a slow manifold but leaves it when fast oscillations appear. A projection of this solution on the slow manifold and the trace of this projection on the slow plane (where it could be tracked by a balanced model) are also shown.

to leave only slow equations for $\mathbf{s}$ and for slowly varying amplitudes of the fast variables $\mathbf{f}$. Averaging has been applied only recently in the geophysical context, however (e.g., Majda and Embid 1998; Babin et al. 2000; Wirosoetisno et al. 2002 and references therein), and with rather theoretical motivations. Instead, what has been extensively employed is the idea of balance, which seeks to filter out fast inertia–gravity waves completely, mainly on the grounds that these are weak in many parts of the atmosphere and oceans as well as poorly constrained by observations.

a. Construction of slow manifolds

At a basic level, balance starts by noting that $\mathbf{f} = O(\varepsilon)$ is a solution of (2.3). So one can define a first slow manifold by the balance condition $\mathbf{f} = 0$, which corresponds to geostrophic balance, and project the dynamics onto it, leading to the simplest balanced model $\partial_t \mathbf{s} = \mathcal{N}_{\mathbf{s}}(\mathbf{s}, 0)$, namely the quasigeostrophic model. The idea can be refined to obtain more and more accurate slow manifolds. By slow manifold, we mean, following MacKay (2004), a submanifold $\mathcal{M}$ of the state space ($\mathbf{s}$, $\mathbf{f}$) with the same dimension as $\mathbf{s}$, which is approximately invariant and on which the motion is slow. The accuracy of a slow manifold can be defined by the size of the component of the vector field ($\partial_t \mathbf{s}$, $\partial_t \mathbf{f}$) transverse to it. See Fig. 1 for an illustration.

Several procedures can be devised to derive slow manifolds with systematically improved accuracy. They seek to determine a balance relation of the form

$$\mathbf{f} = \mathbf{F}(\mathbf{s}; \varepsilon), \tag{2.5}$$

such that the superbalance equation

$$\varepsilon \mathcal{N}_{\mathbf{s}}(\mathbf{s}, \mathbf{F}) \cdot \partial_{\mathbf{s}} \mathbf{F} + \mathcal{L}\mathbf{F} = \varepsilon \mathcal{N}_{\mathbf{f}}(\mathbf{s}, \mathbf{F}), \tag{2.6}$$

obtained by introducing (2.5) into (2.3) and using (2.2) to eliminate $\partial_t \mathbf{s}$, can be satisfied with as small a remainder as possible. Here $\cdot$ denotes summation over the components of $\mathbf{s}$. Once an expression for $\mathbf{F}$ has been found, the corresponding balanced model is derived by constraining the dynamics on the slow manifold to obtain the closed evolution equation

$$\partial_t \mathbf{s} = \mathcal{N}_{\mathbf{s}}[\mathbf{s}, \mathbf{F}(\mathbf{s}; \varepsilon)]$$

for $\mathbf{s}$. See Fig. 2 for an illustration.

Approximate solutions to (2.6) can be obtained systematically by introducing power-series expansion for $\mathbf{F}$ or by devising iterative procedures. Both types of methods provide successively improved approximations $\mathbf{F}^{(n)}$, $n = 1, 2, \ldots$, with $\mathbf{F}^{(0)} = 0$. The simplest iterative procedure is

$$\varepsilon \mathcal{N}_{\mathbf{s}}[\mathbf{s}, \mathbf{F}^{(n)}] \cdot \partial_{\mathbf{s}} \mathbf{F}^{(n)} + \mathcal{L}\mathbf{F}^{(n+1)} = \varepsilon \mathcal{N}_{\mathbf{f}}[\mathbf{s}, \mathbf{F}^{(n)}], \tag{2.7}$$

which requires at each iterate the inversion of the linear operator $\mathcal{L}$. An alternative is

$$\varepsilon \mathcal{N}_{\mathbf{s}}[\mathbf{s}, \mathbf{F}^{(n+1)}] \cdot \partial_{\mathbf{s}} \mathbf{F}^{(n)} + \mathcal{L}\mathbf{F}^{(n+1)} = \varepsilon \mathcal{N}_{\mathbf{f}}[\mathbf{s}, \mathbf{F}^{(n+1)}]. \tag{2.8}$$

This construction is more costly, because a nonlinear equation needs to be solved at each iteration, but it has an attractive feature: the slow manifolds $\mathbf{f} = \mathbf{F}^{(n)}(\mathbf{s}; \varepsilon)$

at each n exactly contain the steady solutions of (2.2)–(2.3). This is seen from the fact that steady solutions, with $\mathcal{N}_{\mathbf{s}}(\mathbf{s}, \mathbf{f}) = 0$ and $\mathcal{L}\mathbf{f} - \varepsilon\mathcal{N}_{\mathbf{f}}(\mathbf{s}, \mathbf{f}) = 0$, evidently satisfy (2.8). As argued by MacKay (2004), it is natural to impose this property, since steady solutions are unambiguously slow. However, constructions of slow manifolds, which differ from that given in (2.8) by negligible errors, may not satisfy this property. This is, in particular, the case of the iteration (2.7). On the other hand, the slow manifolds obtained by setting the nth derivatives of the divergence and ageostrophic vorticity to 0 (e.g., McIntyre and Norton 2000; Mohebalhojeh and Dritschel 2001 and references therein) do contain all steady solutions.

b. Exponential accuracy

Regardless of the procedure followed for the derivation of balanced relations, whether iterations are as described here or power-series expansions, two key points can be made. First, there is in principle no obstacle to carrying out the procedure formally to an arbitrary order n and so define slow manifolds with arbitrary accuracy ε^n. Second, the procedure is asymptotic and not convergent: the error

$$\mathcal{E}_n(\mathbf{s};\varepsilon) = \varepsilon^{-1}\mathcal{L}\mathbf{F}^{(n)}(\mathbf{s};\varepsilon) + \mathcal{N}_{\mathbf{s}}[\mathbf{s}, \mathbf{F}^{(n)}(\mathbf{s};\varepsilon)] \cdot \partial_{\mathbf{s}}\mathbf{F}^{(n)}(\mathbf{s};\varepsilon) - \mathcal{N}_{\mathbf{f}}[\mathbf{s}, \mathbf{F}^{(n)}(\mathbf{s};\varepsilon)],$$

which estimates the component of the motion transverse to the slow manifold $\mathbf{f} = \mathbf{F}^{(n)}$, is formally $O(\varepsilon^n)$, but for fixed ε it increases rapidly as $n \to \infty$. One cannot expect, in fact, to find an exactly invariant slow manifold that would satisfy (2.6) exactly. The best that can be done then is to choose the value of $n = N(\varepsilon)$ that minimizes the error for a fixed ε. This is the standard procedure in asymptotics (e.g., Bender and Orszag 1999): divergent series are truncated optimally, with a number of terms $N(\varepsilon)$ that minimizes the remainder. The remainder typically grows with n like some factorial, so that optimal truncation can be expected to lead to an exponentially small remainder. For instance, when the remainder is proportional to $n!\varepsilon^n$, the choice $N \sim \varepsilon^{-1}$ gives a remainder of the form $\varepsilon^{-1/2}\exp(-1/\varepsilon)$, which is exponentially small in ε.

The exponential smallness of the optimal remainder $\mathcal{E}_{\mathrm{n}}(\mathbf{s};\varepsilon)$ can be proved directly for the iteration scheme (2.7) for finite-dimensional models (Gelfreich and Lerman 2002; Cotter 2004; Cotter and Reich 2006) and for the primitive equation under the assumption that the viscosity is $O(1)$ (Temam and Wirosoetisno 2007). It is likely that the proofs can be extended to the more complicated scheme (2.8), which has the advantage of better accuracy near steady flows. The general conclusion of this type of analysis is that slow manifolds can be defined that are invariant up to exponentially small errors, in the sense that motion transverse to the slow manifold is exponentially smaller than motion tangent to it. From the physical viewpoint, this means that unbalanced phenomena that are genuinely spontaneous are exponentially small.

c. Lighthill radiation

It is probably worth emphasizing the difference between the small Rossby number regime considered here, and the small Froude number regime treated in great detail by Ford (1994) and Ford et al. (2000) in the context of the shallow-water model [see Saujani and Shepherd (2002) and Ford et al. (2002) for a discussion]. In the latter regime, the spectrum of the linear operator in (2.3) is not bounded from below by a large parameter. To see this in the context of the Boussinesq model, we write the dispersion relation (2.1) scaled by U/L as

$$\omega_{\mathrm{igw}} = \varepsilon^{-1}(\sin^2\vartheta + s^2\cos^2\vartheta)^{1/2},$$

where $s = N/f$. In terms of the Froude number $F = U/(NH)$, where H is a typical vertical scale, and of the aspect ratio $\delta = H/L$,

$$s = \frac{\varepsilon}{\delta F}.$$

Now, the small Froude number regime corresponds to the assumptions $F \ll 1$ and $\varepsilon = O(1)$. It is then clear that the frequency is not necessarily large: waves with $\cos\vartheta = O(\delta F)$ have order-one frequencies. Thus, even though there is a small parameter F such that most inertia–gravity modes have large $O(F^{-1})$ frequencies, this does not hold for all wavenumbers, and the frequency spectrum is not bounded from below by (a multiple of) F^{-1}. As a result, the slow manifold theory does not apply.

If one assumes that the balanced motion is localized in space, waves with $kL = O(F) \ll 1$ and $mH = O(1)$ are slow and resonate with the balanced motion. This type of resonance between slow motion and long waves with similarly slow frequencies is typical of Lighthill radiation. This type of radiation is generic for nondispersive waves, acoustic waves (Lighthill 1952; Crow 1970), and gravitational waves (Einstein 1918; Landau and Lifschitz 1975) in particular; in the geophysical context it has been examined for gravity waves for shallow-water (Ford et al. 2000 and references therein) and continuously stratified flows [Plougonven and Zeitlin

(2002); also see Vanneste (2006) for Lighthill radiation in a model extending the LK model].

The asymptotic treatment of Lighthill radiation takes advantage of the spatial separation between waves and slow motion, and shows that wave generation by balanced motion is small in F, although in an algebraic rather than exponential manner. In this treatment, the balanced motion acts as a point source for the waves that can be computed consistently from the leading-order approximation to the balanced motion. When basic invariants are taken into account, the source appears as a quadrupole.

Note that in the case of internal gravity waves, the dispersion relation indicates that waves with large horizontal wavelengths are not the only ones that can resonate with the balanced motion. The condition $\cos\vartheta = O(\delta F)$ can be also met, for instance, for $kL = O(1)$ and $mH = O(F^{-1}) \gg 1$. Radiation of waves with short vertical wavelengths $mH \gg 1$ is therefore possible, but unless the slow motion has itself some fine vertical scales, this radiation can be expected to be exponentially weak in F (because the amplitude of the slow motion decays exponentially with m for $m \gg 1$ if the slow motion has a smooth large-scale vertical structure).

3. Spontaneous generation

We now return to the small Rossby number regime on which this paper concentrates. The conclusion of the previous section is that slow manifolds of exponential accuracy can be defined, at least formally in the PDE case. This indicates that truly spontaneous unbalanced phenomena, that is, those that cannot be completely eliminated by initialization, are at most exponentially small. This has now been demonstrated in several model flows, and in toy models such as the Lorenz–Krishnamurthy model.

a. Forced harmonic oscillator

Perhaps the simplest example of generation of exponentially small fast oscillations is the forced oscillator, governed by

$$\varepsilon^2\ddot{x} + x = g(t), \tag{3.1}$$

where the overdot denotes differentiation with respect to t, for some $g(t)$, say with $g \to 0$ as $t \to \pm\infty$. The solutions to this equation are the sum of the rapidly oscillating homogeneous solutions $\exp(\pm it/\varepsilon)$, which might be thought of as representing IGWs, and a particular inhomogeneous solution. If defined properly, the latter can be free of fast oscillations, at least for some time, and so be identified as the slow or balanced solution. It can be found asymptotically, starting with $x^{(0)}(t) = g(t)$. Higher corrections can be found using a series expansion

$$x_{\mathbf{s}}(t) = \sum_{n=0}^{N} \varepsilon^{2n} x^{(n)}(t), \tag{3.2}$$

with

$$x^{(n)}(t) = (-1)^n \frac{d^{2n}g}{dt^{2n}}. \tag{3.3}$$

No fast oscillations appear at any order, and optimal truncation shows that an exponential accuracy can be achieved. However, the exact solution

$$x(t) = \frac{1}{\varepsilon}\int_{-\infty}^{t} g(t') \sin\frac{t-t'}{\varepsilon}\, dt', \tag{3.4}$$

obtained assuming that $x, \dot{x} \to 0$ as $t \to -\infty$, immediately shows that there may be oscillations if $g(t)$ has poles in complex time. In the asymptotic evaluation of this integral, the balanced solution can be identified with the contribution of the endpoint $t' = t$, which is obtained by successive integration by parts to recover (3.2)–(3.3). The poles of $g(t)$ do contribute to exponentially small terms. The poles closest to the real axis, located at t_* and $\bar{t}_*$ say, where the overbar denotes the complex conjugate, will give the dominant contribution. If

$$g(t) \sim \frac{a}{t - t_*} \quad \text{as} \quad t \to t_*, \tag{3.5}$$

with $t_* = \alpha - i\beta$, $\beta > 0$, for instance, a computation based on closing the integration contour in the complex plane indicates that the fast oscillation

$$x_{\mathbf{f}}(t) \sim -2\pi\varepsilon^{-1} e^{-\beta/\varepsilon}\, \mathrm{Im}[iae^{i(t-\alpha)/\varepsilon}], \tag{3.6}$$

which is clearly exponentially small in ε, appears for $t > \alpha$. The appearance of the oscillatory term is an instance of the Stokes phenomenon in which an exponentially small term is switched on at the tail of an optimally truncated asymptotics series when a parameter (here t) crosses a Stokes line (here $\mathrm{Re}\ t = \alpha$) in the complex plane (e.g., Ablowitz and Fokas 1997).

The phenomenon is illustrated by Fig. 3 obtained for the forcing function $g(t) = \mathrm{sech}\ t$, which has its poles nearest to the real t axis at $\pm i\pi/2$. The left panel shows numerical solutions of (3.1) for several values of ε. The fast oscillations generated across the Stokes line $\mathrm{Re}\ t = 0$ (since $\alpha = 0$) appear clearly, with amplitudes that vary very rapidly with ε and can be verified to agree with Eq. (3.6). The right panel of the figure shows the difference $x_{\mathbf{r}}$ between the numerical solution of (3.1) and the balanced approximation (3.2) for $\varepsilon = 0.15$ and several

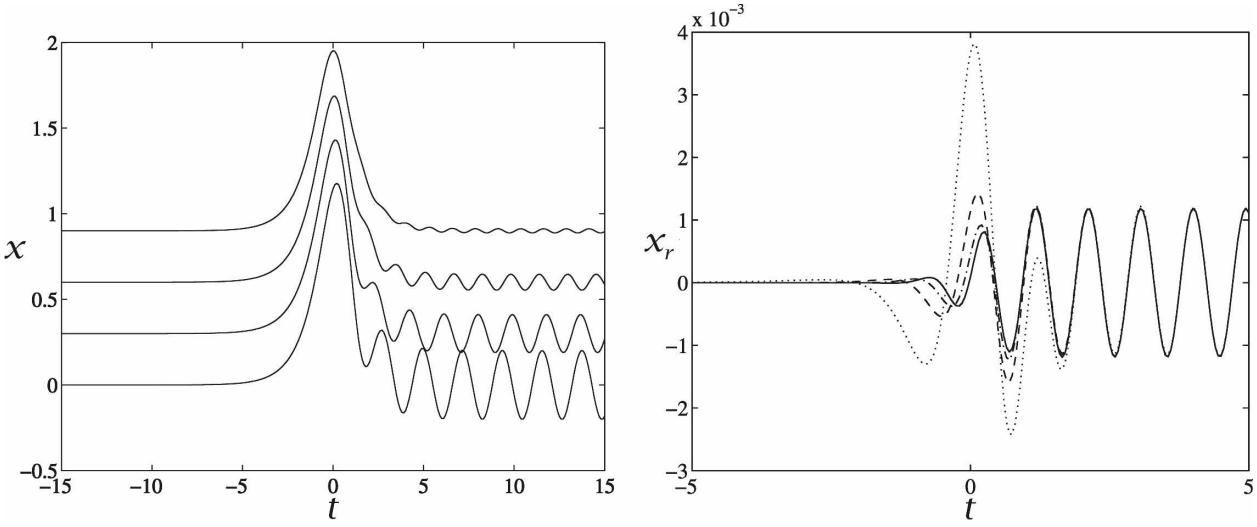

FIG. 3. Solutions $x(t)$ of the harmonic-oscillator Eq. (3.1) with forcing $g(t) = \text{sech } t$ balanced for $t \to -\infty$. (left) The solutions obtained for $\varepsilon = 0.2, 0.25, 0.3, 0.35$ (with decreasing vertical offsets) illustrate the switching on of fast oscillations at $t = 0$ and the sharp decrease of their amplitude with ε. (right) The difference $x_{\mathbf{r}}(t) = x(t) - x_{\mathbf{s}}(t)$ between the numerical solution obtained for $\varepsilon = 0.15$ and the balanced approximation (3.2) truncated at $N = 1$ (dotted line), 2 (dashed line), 3 (dashed–dotted line), and 4 (solid line).

truncations N. It demonstrates how optimal truncation can be used to isolate the fast oscillations from the slow balanced motion: when an optimal truncation $N = 4$ (or $N = 5$, not shown) is used, the dominant part of $x_{\mathbf{r}}$ consists of the fast oscillations switched on at $t = 0$, and the balanced part has been virtually eliminated.

By examining balanced series near optimal truncation, it is possible to capture the manner in which the fast oscillations appear as the Stokes line is crossed (Berry 1989). It is instructive to perform this analysis for the solution (3.4) of the linear equation (3.1). This is carried out in the appendix, where we derive the result

$$x_{\mathbf{f}}(t) \sim -2\pi\varepsilon^{-1}e^{-\beta/\varepsilon}\,\mathrm{Erf}\left[\frac{t-\alpha}{(2\beta\varepsilon)^{1/2}}\right]\mathrm{Im}[iae^{i(t-\alpha)/\varepsilon}], \tag{3.7}$$

where Erf is a scaled and shifted error function [see (A.5)] making the transition between 0 and 1 as its argument changes from $-\infty$ to $+\infty$. This formula shows precisely how the switching on of the oscillations occurs over a time interval of order $O(\varepsilon^{1/2})$, intermediate between the slow and fast time scales of system. It is confirmed by Fig. 4, which compares the prediction (3.7) with the estimate $\mathrm{x}_{\mathbf{r}}$ obtained by subtracting from the numerical solution the optimally truncated balanced approximation (3.2). The forcing is $g(t) = \text{sech } t$, as in Fig. 3, $\varepsilon = 0.15$, and the truncation of the balanced approximation is $N = 5$. The asymptotic prediction and numerical approximation of the fast oscillations are found to match very accurately.

The asymptotic analysis of the appendix separates the slow balanced motion from the fast oscillations to a higher-than-exponential accuracy. However, because there remains a small error, the two components of the motion are not defined unambiguously. It is interesting to note that, in principle, it is possible to remove any ambiguity and separate balanced motion from fast oscillations completely. One approach consists in defining the balanced part of the motion by its asymptotic ex-

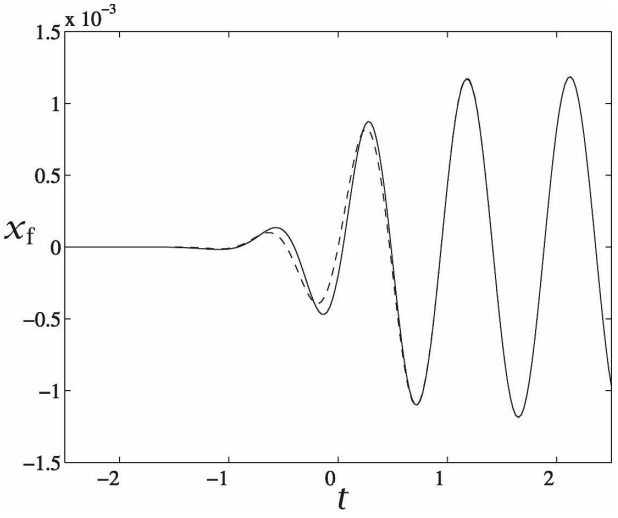

FIG. 4. Comparison between the fast oscillations for $g(t) = \text{sech } t$ and $\varepsilon = 0.15$ predicted from (3.6) (solid line) and their estimate obtained by subtracting from the numerical solution the balanced approximation (3.2) truncated at $N = 5$.

pansion (3.2), requiring that it holds in a large sector of the complex t plane. The sector needs to be large enough to include regions where the fast oscillations have switched dominance with the balanced solution. This enforces the exact vanishing of the oscillations, for if these had a nonzero amplitude, they would be larger than the balanced solution in these regions and the asymptotics (3.2) would not hold. In our example, for instance, one should require that (3.2) continues to hold for $|\mathrm{Im}\ t| > \beta$, since the oscillations switch dominance with the balanced solution for $|\mathrm{Im}\ t| = \beta$. Another approach uses Borel summation. This provides an interpretation of the divergent series (3.2), which we now demonstrate.

Borel summation relies on the observation that the terms in (3.2) grow like $(2n)!$ [see (A.3)]. Thus the function $X(t; \xi)$ defined by

$$X(t; \xi) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \frac{d^{2n} g(t)}{dt^{2n}} \xi^n$$

converges for $|\xi| < 1$ and can be continued analytically for $|\xi| > 1$. Formally, the series (3.2) can be recovered from $X(t; \xi)$ by the Laplace-like transform

$$x(t) = \int_0^{\infty} X(t; \varepsilon^2 s^2) e^{-s}\, ds. \tag{3.8}$$

This transform defines $x(t)$ uniquely once the integration contour in the complex s plane is specified. Now, $X(t; \varepsilon^2 s^2)$ has singularities in the complex s plane, and these correspond to fast oscillations. Thus, choosing the contour in (3.8) to exclude the singularities defines a perfectly balanced, oscillation-free solution. However, if such a balanced solution is chosen initially, it does not remain balanced for all times because the singularities move with t and can subsequently cross the integration contour and contribute an oscillatory term to the integral (3.8): this is how the Stokes phenomenon appears in the Borel summation.

The singularities of $X(t; \xi)$ in the ξ plane are controlled by the poles of $g(t)$. Considering the behavior of $g(t)$ near the pole t_*, we see from the asymptotics (A.3) that X has the behavior

$$X(t; \varepsilon^2 s^2) \sim \frac{a}{t - t_*} \sum_{n=0}^{\infty} (-1)^n \left(\frac{\varepsilon s}{t - t_*} \right)^{2n} = \frac{a(t - t_*)}{(t - t_*)^2 + (\varepsilon s)^2},$$

when $\varepsilon s \approx \pm i(t - t_*)$. The pole at $s = -i(t - t_*)/\varepsilon$ crosses the integration contour in (3.8) when $t = \alpha$. For larger t, it yields a contribution to x, which together with the complex-conjugate contribution associated with the second pole $\bar{t}_*$ of $g(t)$, corresponds precisely to the fast oscillation (3.6).

Several important conclusions, with implications for the process of wave generation, can be drawn from the elementary example (3.1). First, the exponential smallness of the oscillations depends crucially on the smoothness of $g(t)$: the smoother it is, the farther its singularities will be from the real axis, and the smaller the wave excitation. If $g(t)$ is not analytic in a strip around the real t axis, then the fast oscillations are not exponentially small in ε but power-law-like, with the power depending on the degree of regularity of $g(t)$.

Second, the amplitude of the oscillations can also can also be understood in terms of the Fourier transform $\hat{g}(\omega)$ of $g(t)$. For large frequency ω, this is governed, again, by the singularities of $g(t)$ nearest the real axis. With (3.5), the asymptotics is in fact

$$\hat{g} \sim -iae^{-\beta\omega} e^{-i\alpha\omega}. \tag{3.9}$$

The fast oscillations can be thought of as being resonantly excited by the forcing. The result (3.6) can then be explained by the fact that, according to (3.9), the contribution of the forcing with resonant frequency $\omega = \varepsilon^{-1}$ is precisely proportional to $\exp(-\beta/\varepsilon)$. This crude argument only gives the exponential dependence in (3.6); the prefactor and the details of the switching on can however be determined from a careful analysis in the frequency domain similar to that carried out in the appendix in the time domain. However, the large-ω behavior of the frequency spectrum of the balanced motion gives a first diagnostic for the possible spontaneous generation of IGWs, one that can be applied to more complicated systems, including solutions of the partial-differential equations of geophysical fluid dynamics.

A third conclusion is that detailed knowledge of $g(t)$ is crucial if one is to accurately determine the oscillations that are switched on. With the amplitude of the oscillations taking the general form $c\varepsilon^b \exp(-a/\varepsilon)$ for some constants a, b and c, the position of the poles of $g(t)$ determines a, the order of the poles determines b, and the coefficients of the poles determine c. This has the important consequence that if $g(t)$ is also a function of ε, given perturbatively as

$$g(t; \varepsilon) = g_0(t) + \varepsilon g_1(t) + \cdots, \tag{3.10}$$

all the terms $g_n(t)$ in the expansion for $n \le N$, where N is the optimal number of terms, need to be known in order to estimate $x_{\mathbf{f}}$ correctly. If some $g_n(t)$ for $n \ge 1$ has a pole nearer the real t axis than $g_0(t)$, then an estimate of $x_{\mathbf{f}}$ based on the leading-order approximation $g_0(t)$ gives an incorrect exponential dependence. More

likely, all the $g_n(t)$ have the same pole, but with an increasing order:

$$g_n(t) \sim \frac{a_n}{(t - t_*)^{n+1}}, \tag{3.11}$$

for instance. In this case, each of the terms $\varepsilon^n g_n(t)$ of $g(t)$ contributes to $x_\mathbf{f}$ with the same ε dependence, so that an estimate of the oscillations switched on based on $g_0(t)$ gives a value for $x_\mathbf{f}$ that differs from the correct one by an $O(1)$ factor. Thinking of $g(t)$ in the forced oscillator as a caricature for the balanced motion in more sophisticated models, we conclude that knowledge of the balanced solutions to all algebraic orders, at least near the singularities, is in general necessary in order to estimate correctly the amplitude of the fast oscillations generated spontaneously.

b. Lorenz–Krishnamurthy model

The Lorenz–Krishnamurthy model (Lorenz 1986; Lorenz and Krishnamurthy 1987) has been used extensively to study slow manifolds and spontaneous wave generation. It is a simple set of five ODEs of the form (2.2)–(2.3) for three slow variables $\mathbf{s} = (u, v, w)$ and two fast variables $\mathbf{f} = (x, y)$ derived heuristically from the f-plane shallow-water equations. It reads

$$\begin{aligned} \dot{u} &= -vw + \varepsilon b v y, \\ \dot{v} &= uw - \varepsilon b u y, \\ \dot{w} &= -uv, \\ \varepsilon \dot{x} &= -y, \\ \varepsilon \dot{y} &= x + buv, \end{aligned} \tag{3.12}$$

with ε (a variant of) the Rossby number and b a rotational Froude number (inverse square root of the Burger number). Some intuition about the behavior of the LK model is gained if the constancy of $u^2 + v^2$ is exploited to reduce the system to four equations (Camassa 1995; Bokhove and Shepherd 1996). Scaling time, we can take $u = \cos(\theta/2)$ and $v = \sin(\theta/2)$; with the rescaling $(w, x, y) \mapsto (w, x, y)/2$, (3.12) reduces to

$$\dot{\theta} = w - \varepsilon b y, \tag{3.13}$$

$$\dot{w} = -\sin\theta, \tag{3.14}$$

$$\varepsilon \dot{x} = -y, \tag{3.15}$$

$$\varepsilon \dot{y} = x + b \sin\theta. \tag{3.16}$$

In this form, the LK system can be recognized as describing the dynamics of a slow nonlinear pendulum with angle θ from the vertical coupled in some manner with a stiff spring (with constant ε^{-1}) of extension x; see Fig. 5. [A similar set of equations governs the dynamics of the elastic pendulum or stiff spring, also used as a toy model for studying slow manifolds (Lynch 2002).] The problem of finding a slow manifold has then the following interpretation: for a given pendulum angle and angular velocity, find the initial spring extension and speed that minimize the spring's subsequent oscillations. The oscillations that are generated in spite of this optimal initialization are the analog of waves generated spontaneously in fluids.

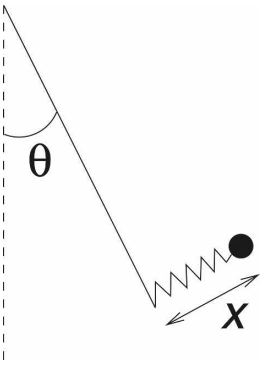

FIG. 5. Schematic of the mechanical equivalent of the Lorenz–Krishnamurthy model: a slow pendulum, described by the angle θ, is coupled to a stiff spring with extension x.

For $\varepsilon = 0$, the pendulum decouples from the spring, and the corresponding equations of motion can be solved in terms of elliptic functions. The solutions are periodic, except for the homoclinic solution, which asymptotically approaches $\theta = \pm\pi$ as $t \to \pm\infty$. For $\varepsilon \neq 0$, balanced solutions can be found order by order. These have poles in the complex t plane whose order increases with the order in ε in the manner indicated by (3.10)–(3.11). Thus, to estimate the fast oscillations generated spontaneously by the balanced motion requires one to analyze the dynamics to all orders in the neighborhood of the singularities. This analysis is carried out in Vanneste (2004) first for the homoclinic solution, with a single pair of dominant poles at $t_* = -i\pi/2$ and $\bar{t}_* = i\pi/2$, then for the periodic solutions with

$$u^{(0)} = \mathrm{cn}(t/k, k), \quad v^{(0)} = -\mathrm{sn}(t/k, k) \quad \text{and}$$
$$w^{(0)} = -\mathrm{dn}(t/k, k),$$

where $0 \leq k < 1$ is a fixed parameter, and cn, sn, and dn are the Jacobian elliptic functions. These solutions have a periodic array of contributing poles at $t_* = 2nkK(k) \pm ikK'(k)$, $n = 0, \pm 1, \pm 2, \ldots$, where $K(k)$ is the complete elliptic integral of the first kind and

$K'(k) = K(\sqrt{1 - k^2})$ (e.g., Abramowitz and Stegun 1965, chapter 17). A Borel summation technique is used to obtain the fast oscillations switched on in the form

$$x_{\mathbf{f}} = -2\pi\varepsilon^{-2}\kappa e^{-\beta/\varepsilon}\cos\frac{t-\alpha}{\varepsilon}, \qquad (3.17)$$

where, as before, β is the imaginary part of the poles nearest to the real axis: $\beta = kK'(k)$. In (3.17), κ is a function of b that is deduced from the late asymptotics of $x_{\mathbf{s}}$ (i.e., from the large-n form of the coefficients x_n in the expansion of $x_{\mathbf{s}}$ in powers of ε) near the poles. This is found by solving numerically (nonlinear) algebraic recurrence relations. This function is shown in Fig. 6. It increases with b from $b = 0$ to $b \approx 1$, then oscillates with a decreasing amplitude. The zeros for discrete values of $b > 0$ indicate that for these particular values, the amplitude of the inertia–gravity oscillations is even smaller than $\exp(-\beta/\varepsilon)$.

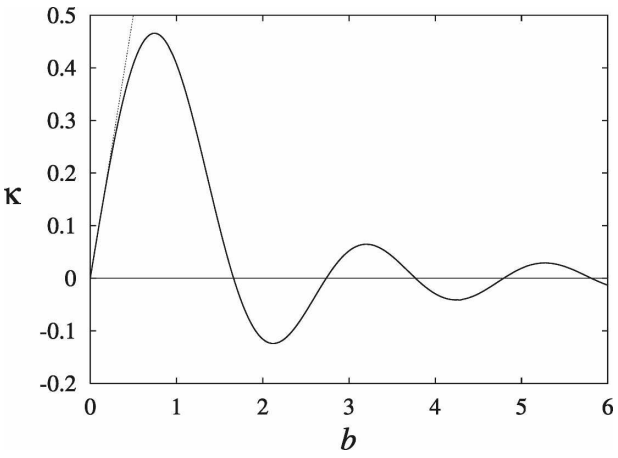

FIG. 6. Dependence on b of the prefactor κ appearing in the asymptotic Eq. (3.17) for the amplitude of the spontaneous fast oscillations in the Lorenz–Krishnamurthy model.

In the limit of small b, corresponding to weak coupling between the pendulum and the spring, $\kappa \sim b$. This result can be obtained directly by substituting the leading-order approximation for θ into the right-hand side of (3.16) and solving the resulting forced-oscillator equation (see Lorenz and Krishnamurthy 1987). This can be interpreted as the use of a Lighthill-like approach for the LK model; a key point is that this approach gives an incorrect estimate for the fast-oscillation amplitude when b is finite.

It is interesting to note that the generation of the fast oscillations is a linear process in the sense that the exponentially small oscillations that are switched on every time a Stokes line is crossed (i.e., periodically in time in the LK model) simply superpose linearly. This means that the amplitude of the fast oscillations typically grows over time. Note, however, that the initial conditions can be tuned in such a manner that initial oscillations combine with those generated by the Stokes phenomenon to yield periodic solutions. This is how the existence of exactly periodic solutions, noted by Lorenz (1986) and Bokhove and Shepherd (1996) as exact (but not slow) solutions of the superbalance equation (2.6) for the LK model, can be interpreted from the exponential-asymptotic viewpoint. (The existence of these periodic solutions is also guaranteed by general persistence arguments; see MacKay 2004; Gelfreich and Lerman 2003.)

c. Sheared disturbances

There is little doubt that spontaneous generation of exponentially small IGWs takes place in rapidly rotating fluids in a manner that is analogous to the generation of fast oscillations in the LK model. Yet this is delicate to demonstrate explicitly because the methods of exponential asymptotics are difficult to adapt to PDEs. They require a substantial amount of explicit information on the balanced solution; this is rarely available for solutions that are nontrivial enough for spontaneous generation to take place.

The sheared disturbances in a rotating Boussinesq fluid studied in McWilliams and Yavneh (1998) and Vanneste and Yavneh (2004) provide an example of a flow for which the IGWs generated spontaneously can be described completely explicitly. This is possible because the spatial structure of this type of solutions is so simple that the evolution is governed by ODEs that can be treated using standard exponential-asymptotic methods. Sheared disturbances date back to Kelvin (Thomson 1887) and consist of plane waves superposed to a horizontal Couette flow $\mathbf{u} = (\Sigma y, 0, 0)$ in an unbounded domain. The Rossby number is $\varepsilon = |\Sigma|/f$, where Σ is the shear, and if one makes the hydrostatic approximation (not made in Vanneste and Yavneh 2004), there is only one additional parameter, namely $b = mf/(kN) = O(1)$. The effect of the shear means that plane waves have a cross-stream wavenumber that changes linearly in time. All the perturbation fields can in fact be written in the form

$$u = \hat{u}(t)\exp\{i[kx + (l - k\Sigma t)y + mz]\}, \qquad (3.18)$$

involving time-dependent amplitudes $\hat{u}(t)$, $\hat{v}(t)$, etc. The ODEs for these amplitudes can be reduced to a single second-order inhomogeneous ODE for a single amplitude, for instance, the vertical-vorticity amplitude $\hat{\zeta}$, once the constancy of the potential-vorticity amplitude $\hat{q}$ is taken into account. This ODE reads

$$\varepsilon^2\left(\frac{d^2\hat{\zeta}}{dt^2}-\frac{2t}{1+t^2}\frac{d\hat{\zeta}}{dt}\right)+\left[(1\mp\varepsilon)\left(1\mp\frac{2\varepsilon}{1+t^2}\right)+\frac{1+t^2}{b^2}\right]\hat{\zeta}=\frac{1+t^2}{b^2}\hat{q}, \tag{3.19}$$

when time is nondimensionalized using $|\Sigma|^{-1}$ as a reference time, and shifted so as to have the cross-stream wavenumber $l = 0$. In (3.19), we have taken advantage of the linearity to scale ζ by N^2; similarly, there is no loss of generality in taking the potential-vorticity amplitude $\hat{q} = 1$. The $\mp$ sign corresponds to $\Sigma \gtrless 0$, that is, to anticyclonic and cyclonic flows, respectively. Note that (3.19) can be rewritten as a system of three equations of the form (2.2)–(2.3), for example with $\mathbf{f} = (\hat{\zeta}, d\hat{\zeta}/dt)$ and $\mathbf{s} = \hat{q}$, but with a time-dependent $\mathcal{L}$.

A balanced solution of (3.19) is readily constructed by straightforward perturbation expansion, starting as

$$\hat{\zeta}_{\mathbf{s}} = \frac{1+t^2}{1+b^2+t^2}+\cdots. \tag{3.20}$$

The inertia–gravity oscillations are homogeneous solutions of (3.19) and can be captured using a WKB expansion

$$\hat{\zeta}_{\mathbf{f}} = e^{i\int_0^t \omega(t')dt'/\varepsilon}[g_0(t)+\varepsilon g_1(t)+\cdots], \tag{3.21}$$

where the frequency

$$\omega = \pm\frac{1}{b}(1+b^2+t^2)^{1/2} \tag{3.22}$$

can be recognized as the nondimensionalization of the inertia–gravity frequency (2.1) in the hydrostatic approximation and for the wavevector $(k, -kt, m)$ [cf. (3.18) with $l = 0$]. Now, the balanced motion (3.20) has poles in complex time, at $t_* = i(1+b^2)^{1/2}$ and $\bar{t}_*$. These are precisely the turning points of (3.21), that is, the locations in the complex time plane where the frequency (3.22) vanishes. In the vicinity of these points there is, of course, no frequency separation between (3.20) and (3.21), and so it is not surprising that oscillations (3.21) can be generated. A more detailed analysis confirms that there is indeed a Stokes phenomenon for the Stokes line Re $t = 0$ joining t_* to $\bar{t}_*$ across which inertia–gravity oscillations in the form (3.22) are switched on. The amplitude of these oscillations is easily estimated. This is done by noting that the amplitude is roughly $O(1)$ near the turning points, and decreases by a factor

$$e^{-\beta/\varepsilon} = \exp\left[\frac{i}{\varepsilon}\int_0^{t_*}\omega(t')\,dt\right] \tag{3.23}$$

for $t = 0$, the point at the Stokes line that is crossed for real t. The factor β evaluates explicitly as

$$\beta = \frac{\pi(1+b^2)}{4b},$$

and controls the exponential smallness of the fast oscillations switched on, as in sections 3a and 3b. The analysis of Vanneste and Yavneh (2004) and Ólafsdóttir et al. (2005) gives the full leading-order asymptotics of the oscillation amplitude, which in addition to the exponential dependence (3.23), contains a factor $\varepsilon^{-1/2}$ as well as an $O(1)$ prefactor. Note that it is only through the latter that the sign of the shear Σ influences the oscillation amplitude. It turns out that this amplitude is larger for anticyclonic shear than for cyclonic shear.

The sheared-disturbance example just described is useful because it demonstrates the spontaneous generation of exponentially small inertia–gravity oscillations in a case where explicit asymptotic computations can be carried out. It concerns a rather unrealistic flow, however, if only because, in the unbounded domain assumed, both the basic shear and the disturbance have infinite energy. It is nonetheless possible to envision that the mechanism at work leads to spontaneous generation in more realistic situations. One such situation concerns the evolution of a localized perturbation, such as a localized vortex, in a horizontal shear. If the localization is sufficient, the shear can be approximated by a Couette flow. The results obtained for sheared disturbances can then be exploited directly by writing all the dynamical fields of localized perturbations as superpositions of sheared modes. For instance, the zonal velocity reads

$$u = \int\int\int \hat{u}(k,l,m,t)\exp\{i[kx+(l-k\Sigma t)y+mz]\}\,dk\,dl\,dm,$$

where $\hat{u}(k, l, m, t)$ is an amplitude to be determined. The corresponding vertical-vorticity amplitude $\hat{\zeta}(k, l, m, t)$ then satisfies an equation of the form (3.19), with t replaced by $t - l/(\Sigma k)$: thus its balanced and IGW components can be deduced directly from (3.20)–(3.21). An initial condition can be imposed by specifying the (localized) potential vorticity and requiring that the solution be fully balanced. In the subsequent evolution, the sheared modes undergo a Stokes phenomenon at the time $t = l/(\Sigma k)$ that depends on their wavenumber. The superposition of the inertia–gravity wave oscillations generated at each of these times represents an inertia–gravity wavepacket that is emitted spontaneously as a result of the shear and subsequently propa-

gates freely. A consequence of the superposition is that the Stokes phenomenon is smeared out in the sense that it does not occur abruptly at a fixed time, but takes place continuously. The explicit computation of the form of the inertia–gravity wavepacket is a quite involved task, which is carried out in Ólafsdóttir (2006) and Ólafsdóttir et al. (2008).

The examples of this section illustrate how the generation of IGWs can be associated with the nontrivial dependence of the balanced flow on time. The mountain-wave example examined by Muraki (2003) illustrates a complementary situation, where it is the spatial dependence of a steady balanced solution that is associated with the wave generation. In general, one can of course expect that the complete spatiotemporal structure of the balanced flow controls the wave emission. However, it remains unclear how that can be captured by analytical methods.

4. Unbalanced instabilities

So far we have been concerned by the evolution of flows that are initialized in such a way that there are no IGWs at $t = 0$. This can be conceived in the models described because the asymptotic treatments allow for the definition of a unique, perfectly balanced state, through the use of Borel summation, for instance. The wave generation that takes place subsequently is then truly spontaneous.

A different process whereby significant levels of IGW activity can be attained is the unbalanced instability of balanced flows. In this case, gravity wave–like perturbations initially present in the flow are amplified by an instability mechanism leading to finite-amplitude gravity waves. If the instability mechanism is genuinely unbalanced, then it should be absent from any balanced model of the flow. The exponential accuracy of balanced models then implies that the growth rate of the instability is exponentially small in the Rossby number.

This expectation is borne out by the stability analysis of several simple steady flows. For these, there is no ambiguity in the initial state of balance, and perturbations can be restricted to be unbalanced by taking their initial potential vorticity to vanish. In most of the flows studied, the instabilities are associated with the linear resonance between different types of wave modes. In the case of the vertically sheared flows studied by Plougonven et al. (2005) and Molemaker et al. (2005), it is the resonance between an edge wave and an inertia–gravity wave that leads to an instability; for the horizontally sheared flows considered by Yavneh et al. (2001), Molemaker et al. (2001), McWilliams et al. (2004), and Vanneste and Yavneh (2007), it is the resonance between either two Kelvin waves, two IGWs, or a Kelvin wave and an IGW. For the resonance to be achieved, unbalanced modes need to be phase locked by the background shear, which is only possible if their phase speed is of the same order as the background-flow speed. For IGWs with horizontal wavenumber k, this gives the condition

$$\omega_{\mathrm{igw}} = (f^2 \sin^2\vartheta + N^2 \cos^2\vartheta)^{1/2} \approx kU,$$

or, in nondimensional terms,

$$(\sin^2\vartheta + s^2 \cos^2\vartheta)^{1/2} \approx \varepsilon kL. \tag{4.1}$$

If U is bounded, this is possible in the regime $s \geq 1$ of interest to us, for disturbances with a large streamwise wavenumber

$$k = O(\varepsilon^{-1}L^{-1}) \gg 1. \tag{4.2}$$

The disturbances remain hydrostatic provided that the vertical wavenumber also is large:

$$m = O(\varepsilon^{-1}H^{-1}) \gg 1. \tag{4.3}$$

The first constraint on the unstable model can be rephrased by stating that the Rossby number that is based on the unstable-mode scale, namely εkL, must be of order one. Note that a similar reasoning applied to shallow-water or multilayer systems leads to the conclusion that phase locking is impossible at small ε, because the gravity wave frequencies increase with k; unbalanced instabilities can then only occur for ε exceeding some threshold (see Ford 1994; Dritschel and Vanneste 2006). Note also that the condition means that the instability processes are restricted to the models governed by PDEs, since ODE models typically involve motion at fixed spatial scales that are independent of the Rossby number. This is a caution against taking the formal representation (2.2)–(2.3) too literally when dealing with PDE models: because the nonlinear term on the right-hand side of (2.3) involves unbounded operators (spatial derivatives), it may become as large as (or even larger than) the linear term $\mathcal{L}\mathbf{f}/\varepsilon$ if the variables $\mathbf{f}$ have small spatial scales.

The condition (4.2) is useful because it gives the possibility of using a WKB approach to solve the eigenvalue problem giving the instability growth rates. In spite of the fact that there is some form of frequency matching between the flow and inertia–gravity frequency, the growth rates found in this manner turn out to be exponentially small. This is because of the spatial structure of the two modes whose resonance leads to instability: they are localized exponentially in different

parts of the flow, with a localization scale proportional to ε, so that their coupling is exponentially weak. The WKB analysis has been carried out explicitly for a horizontal Couette flow (Vanneste and Yavneh 2004) but could also be applied to other types of flows such as those studied by Molemaker et al. (2005) and McWilliams et al. (2004). In the latter situation, the effect of the critical layer would need to be taken into account.

We finally note that another type of instability, also with an exponentially small growth rate, has been examined by McWilliams and Yavneh (1998). This is the rotating, stratified version of the standard elliptical instability (e.g., Kerswell 2002) of flows with elliptic streamlines. Taking advantage of the linear dependence of the velocity field in the spatial coordinates, perturbations can be written as plane waves with time-dependent wavevector and amplitude. The wavevector is periodic in time, and as a result the single amplitude equation that can be derived has periodic coefficients. It is in fact analogous to (3.19), with periodic coefficients and a zero right-hand side. McWilliams and Yavneh (1998) found numerically that the solution to this equation can be exponentially growing for certain wavenumbers, and that the growth rate scales exponentially with ε. We note that asymptotic approximations for the growth rate could be obtained by applying the theory of the Hill equation with the large parameter developed by Weinstein and Keller (1987).

5. Discussion

This paper discusses how some simple mechanisms of IGW generation at small Rossby number can be captured asymptotically. A key point is that, in the simple models studied, suitably defined balanced models (which filter out IGWs completely) can describe the motion up to errors that are exponentially small in the Rossby number. IGW generation and, more generally, spontaneous unbalanced phenomena, do nonetheless take place, with an amplitude that is by necessity exponentially small. Analyzing these phenomena asymptotically thus requires one to use the so-called exponential-asymptotic techniques that go beyond the standard order-by-order perturbation methods. These techniques can be applied with relative ease to highly truncated models governed by ODEs such as the LK model but remain difficult to implement for PDEs [see Vanneste (2006), however, for the exponential-asymptotic analysis of a mixed ODE/PDE model]. As a result, the exponential smallness of IGW amplitudes has been demonstrated explicitly for toy models only.

If this exponential smallness holds for the real atmosphere and oceans, it implies that significant IGW generation can only take place in regions where the Rossby number is not small but larger than about 1. Furthermore, in the regions where it is significant, the generation is likely to be highly heterogeneous in space and time since the exponential dependence sharply amplifies the local variations in the Rossby number.

So far, rigorous results guaranteeing the exponential smallness of IGW-like oscillations have been obtained in the contexts of ODE models [with the exception of Temam and Wirosoetisno (2007), who assume large dissipation]. In the absence of general results for PDEs, it is of interest to speculate about which mechanisms may lead to IGW amplitudes that are significantly larger. A crucial ingredient for exponential smallness is the smoothness in time of the balanced motion or, equivalently, the exponential decay of the frequency spectrum of the balanced motion. One can expect that the relationship between the frequency content of the balanced motion at the IGW frequency and the amplitude of the IGWs generated spontaneously, demonstrated explicitly for ODE models, also holds for PDEs. If the frequency spectrum does not decay exponentially, IGW amplitudes will not be exponentially small in the Rossby number, but much larger, scaling for instance like a power of the Rossby number for power-law frequency spectra. For laminar flows, the exponential decay of the spectrum is not in doubt. For more realistic turbulent flows, however, the situation is much more complex.

A simple scaling argument suggests that quasigeostrophic turbulence and, by extension, turbulence in more accurate balanced models, should be characterized by exponential frequency spectra. The argument notes that the frequency ω of eddies is related to their inverse scale k by $\omega^2 \propto \int_{k_0}^{k} k'^2 E(k')\, dk'$ (Kraichnan 1971). This is essentially independent of k if the energy spectrum $E(k)$ is steeper than k^{-3}, as is thought to be the case for quasigeostrophic turbulence in the enstrophy cascade. When this is the case, the frequency spectrum is controlled by the large-scale flow, which is presumably smooth in time, and so the spectrum can be expected to decay exponentially. This is a crude argument, of course, which would need to be assessed by a more detailed analysis [see Babiano et al. (1987) for some results showing very steep—though perhaps not exponentially decaying—frequency spectra]. In particular, it is not yet clear whether the relevant frequency spectrum should be Eulerian or Lagrangian [see Tennekes (1975) for different predictions for both], although one intuitively expects the Lagrangian frequency spectrum to be the relevant one. In addition, it

should be kept in mind that the energy spectrum can be shallower than k^{-3} even for small Rossby number dynamics: in the presence of a horizontal boundary, for instance, surface effects can lead to kinetic energy spectra in $k^{-5/3}$ (e.g., Pierrehumbert et al. 1994; Tulloch and Smith 2006).

Some aspects of the unbalanced instabilities discussed in section 4 are worth emphasizing. The resonance of IGWs through which the instabilities arise is made possible by two facts: (i) order-one nonlinear frequencies kU are achieved for asymptotically large horizontal wavenumbers k; and (ii) the IGW dispersion relation allows for $O(1)$ frequencies with large k if the vertical wavenumbers m are also large [see (4.1)–(4.3)] or if nonhydrostatic effects come into play. The first point highlights that the infinite-dimensional nature of fluids is a crucial ingredient of the instabilities. The second shows how the details of the dispersion relation matter. This points, in particular, to the limitation of the shallow-water model for the study of unbalanced phenomena in the atmosphere and oceans: the instabilities discussed in this paper have no immediate shallow-water analogs because shallow-water surface waves with large k necessarily have large frequencies.

Note that the resonance mechanisms underlying the instabilities result from effects that are ignored in the formal representation (2.2)–(2.3) of geophysical fluid models. In spite of this, the instabilities turn out to be exponentially weak. The explanation for the exponential smallness in this case appears to be the spatial localization of the unstable modes rather than frequency separation. The two effects are, however, closely related since it can be argued that localization is necessary for modes to be unstable (see Vanneste and Yavneh 2007).

There remains a number of outstanding issues that need to be addressed if the generation of unbalanced motion at small Rossby number is to be fully understood. As mentioned, one of these is the question of whether there are any mechanisms that lead to IGW amplitudes that are larger than exponentially small. Another one is the development of asymptotic techniques providing predictions for wave amplitudes in realistic flows. Yet another one is the potential role of unbalanced motion as a dissipative mechanism for balanced flows (e.g., McWilliams et al. 2001). Related to this are questions about the maintenance of balance and the part played by wave radiation in this maintenance. Future work will examine some of these issues.

Acknowledgments. The author thanks D. Wirosoetisno and A. Olde Daalhuis for useful discussions and comments.

APPENDIX

Stokes Phenomenon

In this appendix we derive Eq. (3.6) describing the switching on of the fast oscillations for the forced oscillator (3.1). We start by writing

$$\varepsilon x = \mathrm{Im}\, e^{it/\varepsilon} I, \quad \text{where} \quad I = \int_{-\infty}^{t} g(t') e^{-it'/\varepsilon}\, dt'. \tag{A.1}$$

To all algebraic orders in ε, the integral I is approximated by the contribution of its endpoint $t' = t$. Successive integrations by parts give

$$I = ie^{-it/\varepsilon} \sum_{n=0}^{N-1} (-i\varepsilon)^n \frac{d^n g(t)}{dt^n} + R_N,$$

consistent with (3.2)–(3.3), with the remainder

$$R_N = (-i\varepsilon)^N \int_{-\infty}^{t} \frac{d^N g(t')}{dt'^N} e^{-it'/\varepsilon}\, dt'. \tag{A.2}$$

A key point is that for large N, the Nth derivative of $g(t')$ is controlled by the singularities t_* and $\bar{t}_*$ nearest to the real axis. Thus we have

$$\frac{d^N g(t')}{dt'^N} \sim (-1)^N N! \left[\frac{a}{(t' - t_*)^{N+1}} + \frac{\bar{a}}{(t' - \bar{t}_*)^{N+1}} \right]. \tag{A.3}$$

With this result, we can estimate the remainder (after an additional integration by parts) as

$$|R_N| \sim \left\{ \frac{\varepsilon}{[(t-\alpha)^2 + \beta^2]^{1/2}} \right\}^{N+1} N!$$

This indicates that the optimal truncation should have $N \sim [(t-\alpha)^2 + \beta^2]^{1/2}/\varepsilon$, leading to an exponentially small error.

We now focus on the neighborhood of the Stokes line crossing at $t = \alpha$. The optimal truncation then dictates that

$$N = \left\lfloor \frac{\beta}{\varepsilon} \right\rfloor + \gamma - 1, \tag{A.4}$$

where $\lfloor\,\rfloor$ denotes the integral part, and $\gamma = O(1)$ (the -1 is introduced for convenience). Using (A.3), we write the remainder (A.2) as

$$R_N = (i\varepsilon)^N N! \int_{-\infty}^{t} \left[\frac{a}{(t' - t_*)^{N+1}} + \frac{\bar{a}}{(t' + \bar{t}_*)^{N+1}} \right] e^{-it'/\varepsilon}\, dt'.$$

With our choice Im $t_* = -\beta < 0$, only the first of the two terms in the brackets contributes significantly to this integral. Taking (A.4) into account, we therefore have

$$R_N \sim (i\varepsilon)^N N! a e^{i\alpha/\varepsilon} \int_{-\infty}^{t-\alpha} e^{-[it' + \beta \log(t' + i\beta)]/\varepsilon} e^{-\gamma \log(t' + i\beta)}\, dt'$$

after shifting t' by α. This integral can then be estimated by expansion near the saddle point $t' = 0$ of $f(t') = it' + \beta \log(t' + i\beta)$. This leads to

$$R_N \sim (i\varepsilon)^N N! a e^{-i\alpha/\varepsilon} e^{-(N+1)\log(i\beta)} \int_{-\infty}^{t-\alpha} e^{-t'^2/(2\beta\varepsilon)}\, dt'.$$

Using Stirling's formula and the definition of the shifted error function

$$\mathrm{Erf}\,(\cdot) = \frac{1}{2}[1 + \mathrm{erf}(\cdot)], \tag{A.5}$$

this simplifies after a number of cancellations into

$$R_N \sim -2\pi i a e^{-\beta/\varepsilon} \mathrm{Erf}\left[\frac{t-\alpha}{(2\beta\varepsilon)^{1/2}}\right]. \tag{A.6}$$

Returning to (A.1), we see that, to leading order, the remainder of the optimally truncated expansion corresponds to the rapidly oscillating term

$$x_{\mathbf{f}}(t) \sim -2\pi\varepsilon^{-1} e^{-\beta/\varepsilon} \mathrm{Erf}\left[\frac{t-\alpha}{(2\beta\varepsilon)^{1/2}}\right] \mathrm{Im}[i a e^{i(t-\alpha)/\varepsilon}]$$

switched on as t goes through α. Note that this term is larger, by a factor proportional to $\varepsilon^{-1/2}$, than the last term in the series x_s truncated at $n = N$.

REFERENCES

Ablowitz, M. J., and A. S. Fokas, 1997: *Complex Variables: Introduction and Applications.* Cambridge University Press, 647 pp.

Abramowitz, M., and I. A. Stegun, 1965: *Handbook of Mathematical Functions.* Dover, 1046 pp.

Babiano, A., C. Basdevant, P. LeRoy, and R. Sadourny, 1987: Single particle dispersion, Lagrangian structure function, and Lagrangian energy spectrum in two-dimensional incompressible turbulence. *J. Mar. Res.,* **45,** 107–131.

Babin, A., A. Mahalov, and B. Nicolaenko, 2000: Fast singular oscillating limits and global regularity for the 3D primitive equations of geophysics. *Modél. Math. Anal. Num.,* **34,** 201–222.

Bender, C. M., and S. A. Orszag, 1999: *Advanced Mathematical Methods for Scientists and Engineers.* Springer, 593 pp.

Berry, M. V., 1989: Uniform asymptotic smoothing of Stokes's discontinuities. *Proc. Roy. Soc. London,* **422A,** 7–21.

Bokhove, O., and T. G. Shepherd, 1996: On Hamiltonian balanced dynamics and the slowest invariant manifold. *J. Atmos. Sci.,* **53,** 276–297.

Camassa, R., 1995: On the geometry of an atmospheric slow manifold. *Physica D,* **84,** 357–397.

Cotter, C. J., 2004: Model reduction for shallow water dynamics: Balance, adiabatic invariance, and subgrid modeling. Ph.D. thesis, Imperial College, 153 pp.

——, and S. Reich, 2006: Semigeostrophic particle motion and exponentially accurate normal forms. *Multiscale Model. Simul.,* **5,** 476–496.

Crow, S. C., 1970: Aerodynamic sound emission as a singular perturbation problem. *Stud. Appl. Math.,* **49,** 21–44.

Dritschel, D. G., and J. Vanneste, 2006: Instability of a shallow-water potential-vorticity front. *J. Fluid Mech.,* **561,** 237–254.

Einstein, A., 1918: Über Gravitationswellen. *Sitzungsberichte der Königlich Preußische Akademie der Wissenschaften zu Berlin,* 154–167.

Ford, R., 1994: The instability of an axisymmetric vortex with monotonic potential vorticity in rotating shallow water. *J. Fluid Mech.,* **280,** 303–334.

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

——, ——, and ——, 2002: Reply. *J. Atmos. Sci.,* **59,** 2878–2882.

Gelfreich, V., and L. Lerman, 2002: Almost invariant elliptic manifold in a singularly perturbed Hamiltonian system. *Nonlinearity,* **15,** 447–457.

——, and ——, 2003: Long periodic orbits and invariant tori in a singularly perturbed Hamiltonian system. *Physica D,* **176,** 125–146.

Kerswell, R. R., 2002: Elliptical instability. *Annu. Rev. Fluid Mech.,* **34,** 83–113.

Kraichnan, R. H., 1971: Inertial-range transfer in two- and three-dimensional turbulence. *J. Fluid Mech.,* **47,** 525–535.

Landau, L. D., and E. M. Lifshitz, 1975: *The Classical Theory of Fields.* 4th ed. Pergamon, Vol. 2, 402 pp.

Leith, C. E., 1980: Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.,* **37,** 958–968.

Lighthill, M. J., 1952: On sound generated aerodynamically. I. General theory. *Proc. Roy. Soc. London,* **211A,** 564–587.

Lorenz, E. N., 1980: Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.,* **37,** 1685–1699.

——, 1986: On the existence of a slow manifold. *J. Atmos. Sci.,* **43,** 1547–1557.

——, and V. Krishnamurthy, 1987: On the nonexistence of a slow manifold. *J. Atmos. Sci.,* **44,** 2940–2950.

Lynch, P., 2002: The swinging spring: A simple model for atmospheric balance. *Geometric Methods and Models,* I. Roulstone and J. Norbury, Eds., Vol. II, *Large-Scale Atmosphere–Ocean Dynamics,* Cambridge University Press, 64–108.

MacKay, R. S., 2004: Slow manifolds. *Energy Localisation and Transfer,* T. Dauxois et al., Eds., World Scientific, 149–192.

Majda, A. J., and P. Embid, 1998: Averaging over fast gravity waves for geophysical flows with unbalanced initial data. *Theor. Comput. Fluid Dyn.,* **11,** 155–169.

McIntyre, M. E., and W. A. Norton, 2000: Potential vorticity inversion on a hemisphere. *J. Atmos. Sci.,* **57,** 1214–1235; Corrigendum, **58,** 949.

McWilliams, J. C., and I. Yavneh, 1998: Fluctuation growth and instability associated with a singularity of the balance equations. *Phys. Fluids,* **10,** 2587–2596.

——, J. M. Molemaker, and I. Yavneh, 2001: From stirring to mixing of momentum: Cascades from balanced flows to dissipation in the oceanic interior. *From Stirring to Mixing in a Stratified Ocean: Proc. 'Aha Huliko'a Hawaiian Winter Workshop,* Honolulu, HI, University of Hawaii at Manoa, 59–66.

——, M. J. Molemaker, and I. Yavneh, 2004: Ageostrophic, anticyclonic instability of a geostrophic, barotropic boundary current. *Phys. Fluids,* **16,** 3720–3725.

Mohebalhojeh, A. R., and D. G. Dritschel, 2001: Hierarchies of balance conditions for the *f*-plane shallow water equations. *J. Atmos. Sci.,* **58,** 2411–2426.

Molemaker, M. J., J. C. McWilliams, and I. Yavneh, 2001: Instability and equilibration of centrifugally stable stratified Taylor–Couette flow. *Phys. Rev. Lett.,* **86,** 5270–5273.

——, ——, and ——, 2005: Baroclinic instability and loss of balance. *J. Phys. Oceanogr.,* **35,** 1505–1517.

Muraki, D. J., 2003: Revisiting Queney's flow over a mesoscale ridge. Preprints, *Ninth Conf. on Mountain Meteorology,* Aspen, CO, Amer. Meteor. Soc., 5.3.

Ólafsdóttir, E. I., 2006: Atmospheric-wave generation: An exponential-asymptotic analysis. Ph.D. thesis, University of Edinburgh, 121 pp.

——, A. B. Olde Daalhuis, and J. Vanneste, 2005: Stokes-multiplier expansion in an inhomogeneous differential equation with a small parameter. *Proc. Roy. Soc. London,* **461A,** 2243–2256.

——, ——, and ——, 2008: Inertia–gravity wave radiation by a sheared vortex. *J. Fluid Mech.,* **596,** 169–189.

Olde Daalhuis, A. B., 2003: Exponential asymptotics. *Orthogonal Polynomials and Special Functions: Leuven 2002,* E. Koelink and W. van Assche, Eds., Springer, 211–244.

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.

Pierrehumbert, R. T., I. M. Held, and K. L. Swanson, 1994: Spectra of local and non-local two-dimensional turbulence. *Chaos Solitons Fractals,* **4,** 1111–1116.

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 C. Snyder, 2005: Gravity waves excited by jets: Propagation versus generation. *Geophys. Res. Lett.,* **32,** L18802, doi:10.1029/2005GL023730.

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

——, D. J. Muraki, and C. Snyder, 2005: A baroclinic instability that couples balanced motions and gravity waves. *J. Atmos. Sci.,* **62,** 1545–1559.

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

Saujani, S., and T. G. Shepherd, 2002: Comments on "Balance and the slow quasimanifold: Some explicit results." *J. Atmos. Sci.,* **59,** 2874–2877.

Temam, R., and D. Wirosoetisno, 2007: Exponentially accurate approximations for the primitive equations of the ocean. *Discrete Contin. Dyn. Syst.,* **7B,** 425–440.

Tennekes, H., 1975: Eulerian and Lagrangian time microscales in isotropic turbulence. *J. Fluid Mech.,* **67,** 561–567.

Thomson, W., 1887: Stability of fluid motion—Rectilineal motion of viscous fluid between two parallel planes. *Philos. Mag.,* **24,** 188–196.

Tulloch, R., and K. S. Smith, 2006: A theory for the atmospheric energy spectrum: Depth-limited temperature anomalies at the tropopause. *Proc. Natl. Acad. Sci. USA,* **103,** 14 690–14 694.

Vanneste, J., 2004: Inertia–gravity wave generation by balanced motion: Revisiting the Lorenz–Krishnamurthy model. *J. Atmos. Sci.,* **61,** 224–234.

——, 2006: Wave radiation by balanced motion in a simple model. *SIAM J. Appl. Dyn. Syst.,* **5,** 783–807.

——, and I. Yavneh, 2004: Exponentially small inertia–gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.,* **61,** 211–223.

——, and ——, 2007: Unbalanced instabilities of rapidly rotating stratified shear flows. *J. Fluid Mech.,* **584,** 373–396.

Viúdez, Á., 2006: Spiral patterns of inertia–gravity waves in geophysical flows. *J. Fluid Mech.,* **562,** 73–82.

——, and D. G. Dritschel, 2006: Spontaneous emission of inertia–gravity wave packets by balanced geophysical flows. *J. Fluid Mech.,* **553,** 107–117.

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

Weinstein, M. I., and J. B. Keller, 1987: Asymptotic behavior of stability regions for Hill's equation. *SIAM J. Appl. Math.,* **47,** 941–958.

Wirosoetisno, D., 2004: Exponentially accurate balance dynamics. *Adv. Differ. Equations,* **9,** 177–196.

——, T. G. Shepherd, and R. M. Temam, 2002: Free gravity waves and balanced dynamics. *J. Atmos. Sci.,* **59,** 3382–3398.

Yavneh, I., J. C. McWilliams, and M. J. Molemaker, 2001: Non-axisymmetric instability of centrifugally-stable stratified Taylor–Couette flow. *J. Fluid Mech.,* **448,** 1–21.

Zhang, F., 2004: Generation of mesoscale gravity waves in upper-tropospheric jet–front systems. *J. Atmos. Sci.,* **61,** 440–457.