Exponentially small inertia-gravity waves and the breakdown of quasi-geostrophic balance

J. Vanneste[*]

School of Mathematics,

University of Edinburgh,

Edinburgh, United Kingdom

I. Yavneh

Department of Computer Science,

Technion,

Haifa, Israel

3rd revision, July 2003

[*]Corresponding author address: J. Vanneste, School of Mathematics, University of Edinburgh, King's Buildings, Mayfield Road, Edinburgh EH9 3JZ, United Kingdom. Email: J.Vanneste@ed.ac.uk

The spontaneous generation of inertia-gravity waves by balanced motion is investigated in the limit of small Rossby number $\epsilon \ll 1$. Particular (sheared-disturbance) solutions of the three-dimensional Boussinesq equations are considered. For these solutions, there is a strict separation between balanced motion and inertia-gravity waves for large times. This makes it possible to estimate the amplitude of the inertia-gravity waves that are generated spontaneously from perfectly balanced initial conditions. It is shown analytically using exponential asymptotics, and confirmed numerically, that this amplitude is proportional to $\epsilon^{-1/2} \exp(-\alpha/\epsilon)$, with a constant $\alpha > 0$ and a proportionality constant that are given in closed form. This result demonstrates the inevitability of inertia-gravity-wave generation and hence the non-existence of an invariant slow manifold; it also exemplifies the remarkable, exponential, smallness of the wave generation for $\epsilon \ll 1$. The importance of the singularity structure of the balanced motion for complex values of time is emphasised, and some general implications of the results are discussed.

1   Introduction

It has long been recognised that the large-scale dynamics of the atmosphere and oceans is dominated by its low-frequency, nearly geostrophic, balanced component, with comparatively little energy in the fast, inertia-gravity waves. The concept of slow manifold (Leith 1980, Lorenz 1980) formalises this observation: in the state space of the primitive equations (or of some more general equations), it introduces a manifold of reduced dimensionality in the neighbourhood of which the dynamics is thought to be confined. The initialisation procedures used for the integration of the primitive equations can then be regarded as projections of initial conditions on a slow manifold; and balanced models, which filter out inertia-gravity waves, can be regarded as constrained systems obtained by projecting the dynamics of the primitive equations on a slow manifold (see, e.g., Warn, Bokhove, Shepherd & Vallis (1995)).

There has been considerable debate about the slow manifold. Formally, this manifold can be defined as the solution to Lorenz' superbalance equation (Lorenz 1980), obtained by assuming the existence of unknown functional relations slaving all the dynamical variables in the primitive equations to a single variable. An exact solution for these functional relations would imply the existence of an exactly invariant slow manifold; as a consequence, a suitable initialisation would lead to a dynamics that is entirely free of inertia-gravity waves and so may be described exactly by a balanced model of reduced dimensionality.

Approximate solutions to the superbalance equation (geostrophic balance among others), corresponding to approximately invariant slow manifolds, have been derived perturbatively using the smallness of the Rossby number ϵ and Froude number F which reflect the fast rotation and strong stratification of the atmosphere and oceans. But, it has become gradually accepted that the superbalance equation does not admit

an exact solution (Warn 1997, Lorenz & Krishnamurthy 1987). Thus, while solutions of the primitive equations evolve in the neighbourhood of approximate slow manifolds, they have full dimensionality and cannot be described exactly by any balanced model. This is Warn's picture of a fuzzy slow manifold (Warn 1997, Warn & Menard 1986) or quasi-manifold (Ford, McIntyre & Norton 2000) in which the fuzziness results from a fundamental physical process, the spontaneous generation of free inertia-gravity waves by the balanced motion, and cannot be eliminated by initialisation.

This picture has emerged mostly from the study of low-order models of the primitive equations (e.g. Lorenz (1986)): on the one hand, the divergence of perturbative procedures for solving the superbalance equation suggested the non-existence of solutions (Vautard & Legras 1986, Warn & Menard 1986); on the other hand, some explicit solutions starting on a putative slow manifold were shown to develop inertia-gravity waves (Lorenz & Krishnamurthy 1987, Camassa 1995, Bokhove & Shepherd 1996, Camassa & Tin 1996). There are fewer explicit results for partial-differential equations. The recent ones due to Ford et al. (2000) concern the shallow-water equations in small-Froude-number limit $F \ll 1$, with a Rossby number $\epsilon = O(1)$. In this limit, there is no frequency separation between the balanced motion and long inertia-gravity waves (which are slow), and balance is found to break down at an algebraic order in F, with the generation of $O(F^2)$ inertia-gravity waves. In contrast, in the rotation-dominated, quasi-geostrophic limit $\epsilon \ll 1$, $F = O(\epsilon) \ll 1$, there is a formal frequency separation between balanced motion and inertia-gravity waves, suggesting a much weaker gravity-wave generation and hence a greater accuracy of balance. (See Saujani & Shepherd (2002) and Ford, McIntyre & Norton (2002) for a discussion contrasting the two regimes.)

A number of arguments indicate that the inertia-gravity waves generated in the quasi-geostrophic regime are, in fact, exponentially small in ϵ (e.g. proportional to

4

$\exp(-\alpha/\epsilon)$ for some constant $\alpha > 0$). Firstly, the superbalance equation can be solved formally, order by order in ϵ, using series expansions or iterations, but the procedures employed turn out to be asymptotic rather than convergent. Such lack of convergence is very often the fingerprint of non-analytic, typically exponential, terms which the formal procedures fail to capture. In fact, optimally truncated asymptotic series can be shown to provide approximate solutions to the superbalance equations with exponentially small residuals (see Gelfreich & Lerman (2002) for a treatment of general finite-dimensional Hamiltonian systems with a single fast frequency, and Wirosoetisno, manuscript in preparation, for a treatment of geophysical models). Secondly, a number of steady flows in geostrophic balance have been shown to be unstable to unbalanced instabilities involving gravity-wave-like perturbations, with growth rates that are exponentially small in ϵ (Ford 1994, Nore & Shepherd 1997, Kushner, McIntyre & Shepherd 1998, Yavneh, McWilliams & Molemaker 2001, Molemaker, McWilliams & Yavneh 2001). Finally, numerical experiments using low-order models directly indicate an exponential scaling for the amplitude of the inertia-gravity waves that emerge from well balanced initial conditions (Lorenz & Krishnamurthy 1987, Warn & Menard 1986, Bokhove & Shepherd 1996). This numerical evidence is supported to some extent by analytic results, both formal (Lorenz & Krishnamurthy 1987) and rigorous (Camassa 1995, Camassa & Tin 1996), which provide asymptotic estimates for particular solutions to Lorenz's five-component model (Lorenz 1986) starting on a local slow manifold. Although these estimates exhibit an exponential dependence in the Rossby number ϵ, they are only suggestive because the assumptions on which they rely break down in the limit $\epsilon \to 0$.[1]

---

[1]The asymptotic results of Lorenz & Krishnamurthy (1987), Camassa (1995) and Camassa & Tin (1996) assume a small rotational Froude number F/ϵ and an order-one frequency separation between inertia-gravity waves and balanced motion (i.e., in the shallow-water context of Lorenz's model, $\epsilon F/\sqrt{\epsilon^2 + F^2} = O(1)$; see Bokhove & Shepherd (1996)).

The aims of the present paper are (i) to present an example of generation of exponentially small inertia-gravity waves for solutions of the full three-dimensional primitive equations (or more exactly of the non-hydrostatic Boussinesq equations); and (ii) to provide an asymptotic estimate for the amplitude of the waves that is valid in the quasi-geostrophic limit $\epsilon \to 0$ (and $F = O(\epsilon) \ll 1$). The solutions that we consider are important because they allow to definitely rule out the possibility of an exactly invariant slow manifold for the primitive equations. This cannot be inferred from the corresponding result for the five-component model of Lorenz & Krishnamurthy (1987): exponentially small phenomena are notoriously sensitive to perturbations of the models and are most probably affected by the drastic truncation that leads to the five-component model. Our solutions, and the asymptotic technique used to derive them, are also interesting because they illustrate how the generation of exponentially weak inertia-gravity waves is associated with the non-analyticity of the balanced motion for complex values of the time t. The role of singularities of the balanced motion in the complex t-plane can be understood heuristically. These singularities determine the exponential decay rate of the Fourier transform of the balanced signal at large frequencies; they therefore determine the amplitude of the balanced signal at the (large) frequencies of the inertia-gravity waves and hence control the excitation of these waves (Warn 1997). However, the precise amplitude of the excited inertia-gravity waves cannot be determined in an immediate manner; as the asymptotic technique we employ demonstrates, it depends on the structure the inertia-gravity waves when analytically continued for complex values of t.

The particular solutions of the three-dimensional non-hydrostatic Boussinesq equations that we consider have been previously examined by McWilliams & Yavneh (1998). They consist of the so-called sheared disturbances, obtained by superposing a single plane wave to a horizontal Couette flow. For these solutions, which generalise

those originally found by Kelvin (Thomson 1887), the various dynamical fields have the spatial structure of a plane wave, with constant streamwise and vertical wavenumbers, and with a cross-stream wavenumber which changes linearly with time. There is then an exact nonlinear reduction of the primitive partial-differential equations to a single second-order inhomogeneous ordinary differential equation for the time-dependent amplitude of one of the dynamical variables. It is this ordinary differential equation, previously solved numerically by McWilliams & Yavneh (1998) for Rossby number $\epsilon = O(1)$, that we treat asymptotically for $\epsilon \ll 1$ using the techniques of exponential asymptotics (see, e.g., Segur, Tanveer & Levine (1991) and Hakim (1998) for an introduction to these techniques). Such techniques are essentially formal, although they have been proved to provide rigorous results in some applications (e.g., Byatt-Smith & Davie (1990)); we therefore confirm the validity of our asymptotic results by a comparison with numerical results.

Besides the reduction from partial to ordinary differential equations, two properties of the sheared-disturbance solutions are crucial to our analysis. First, because of the phase mixing associated with the reduction in horizontal scale that the disturbance undergoes for large $|t|$, the dynamics becomes extremely simple in the limit $t \to \pm\infty$; so simple in fact that there is, in this limit, an unambiguous separation between the balanced motion (which has a constant amplitude) and the inertia-gravity waves (which correspond to a rapidly oscillating amplitude). This makes it possible to consider a completely balanced "initial" condition for $t \to -\infty$ and evaluate the amplitude of the (spontaneously generated) inertia-gravity waves for $t \to +\infty$. This property, similar to that of the perturbed homoclinic orbit considered by Lorenz & Krishnamurthy (1987) in their five-component model, eliminates the conceptual difficulties that exist for most solutions for which a completely balanced state cannot be defined. The second crucial property of the sheared-disturbance solutions is the

7

simple but non-trivial time dependence of the balanced motion, with simple poles in the complex t-plane. This allows the amplitude of the inertia-gravity waves to be evaluated analytically for $\epsilon \ll 1$ using a straightforward exponential-asymptotic technique.

The paper is structured as follows. Section 2 is devoted to a brief description of the second-order ordinary differential equation which governs the evolution of sheared-disturbance solutions in a Boussinesq fluid. In section 3 we consider the small-Rossby-number limit and give the asymptotic expression for the amplitude of the exponentially small inertia-gravity waves that are generated spontaneously from a balanced state. This expression is confirmed by numerical solutions of the ordinary differential equation for a few different values of the parameters. A simplified asymptotic expression is also given for the geophysically relevant situation of a strongly stratified fluid. The paper concludes with a Discussion in section 4. The analytic results are presented with only brief indications about their derivation. The interested reader is referred to three Appendices for the details. Appendix B, in particular, provides the details of the exponential-asymptotics calculation of the amplitude of the inertia-gravity waves.

2 Model

Following McWilliams & Yavneh (1998) we consider the evolution of a perturbed Couette flow $(\Sigma y, 0, 0)$ in a tri-dimensional Boussinesq fluid on an f-plane. An exact, nonlinear reduction from the nonlinear partial-differential equations to linear ordinary differential equations is possible for perturbations of the form

$$a(t)\mathrm{e}^{\mathrm{i}[k(x-\Sigma yt)+ly+mz]} + \mathrm{c.c.},\tag{2.1}$$

8

with time-dependent amplitudes $a(t)$ distinct for each dynamical field. Such perturbations are the sheared disturbances first discussed by Kelvin (Thomson 1887). They consist of plane waves with a wavevector $(k, l - k\Sigma t, m)$ which depends explicitly on time through its second component.

The amplitude of the vertical component of the vorticity field, denoted by $\zeta(t)$, can be shown to satisfy the linear equation

$$\ddot{\zeta} + b(t)\dot{\zeta} + c(t)\zeta = N^2 \frac{k^2 + (l - \Sigma k t)^2}{k^2 + (l - \Sigma k t)^2 + m^2}, \qquad (2.2)$$

with

$$
\begin{aligned}
b(t) &= \frac{2\Sigma k m^2 (l - \Sigma k t)}{[k^2 + (l - \Sigma k t)^2][k^2 + (l - \Sigma k t)^2 + m^2]}, \\
c(t) &= \frac{(f - \Sigma)m^2}{k^2 + (l - \Sigma k t)^2 + m^2}\left[f - \frac{2\Sigma k^2}{k^2 + (l - \Sigma k t)^2}\right] + \frac{N^2[k^2 + (l - \Sigma k t)^2]}{k^2 + (l - \Sigma k t)^2 + m^2},
\end{aligned}
$$

where $f > 0$ is the Coriolis frequency and N the constant Brunt–Väisälä frequency. See Appendix A for a derivation.

Rescaling time according to $t \mapsto t/|\Sigma|$, choosing the origin of time so that $l = 0$, we rewrite (2.2) in dimensionless form. Introducing the Rossby number

$$\epsilon := \frac{|\Sigma|}{f} \ll 1,$$

and the Prandtl and aspect ratios

$$S := \frac{N^2}{f^2} \quad \text{and} \quad \delta := \frac{m}{k},$$

we obtain

$$\epsilon^2 \left[\ddot{\zeta} + b(t)\dot{\zeta}\right] + c(t)\zeta = S \frac{1 + t^2}{1 + \delta^2 + t^2}, \qquad (2.3)$$

where

$$
\begin{aligned}
b(t) &= \frac{-2\delta^2 t}{(1 + t^2)(1 + \delta^2 + t^2)}, \\
c(t) &= \frac{\delta^2(1 \mp \epsilon)}{1 + \delta^2 + t^2}\left(1 \mp \frac{2\epsilon}{1 + t^2}\right) + S \frac{1 + t^2}{1 + \delta^2 + t^2}.
\end{aligned}
$$

9

Here and in what follows, the $\mp$ signs distinguish the anticyclonic shear ($\Sigma > 0$, upper sign) from the cyclonic shear ($\Sigma < 0$, lower sign). Henceforth, we assume that $S > 1$.

We are interested in the limit of a small Rossby number $\epsilon \ll 1$, with the other parameters δ and S formally of $O(1)$. Correspondingly, the Froude number is small:

$$F := \frac{|\Sigma|\delta}{N} = \epsilon \frac{\delta}{S^{1/2}} \ll 1.$$

The factor $\delta/S^{1/2}$ can be identified as the rotational Froude number, i.e., the ratio of the horizontal scale k^{-1} to the radius of deformation $N/(fm)$ (its square is the inverse of the Burger number). In the standard quasi-geostrophic scaling, $\delta/S^{1/2} = O(1)$ while both S and δ are large, corresponding to a strong stratification and to a disturbance with small vertical scale; we will consider this regime as a particular case of our more general results.

3 Inertia-gravity-wave amplitude

We first obtain a formal solution of (2.3) using a regular perturbation expansion in powers of ϵ. A straightforward calculation gives

$$\zeta = \zeta_{\mathrm{bal}} := \zeta^{(0)} + \epsilon \zeta^{(1)} + \cdots, \tag{3.1}$$

where

$$\zeta^{(0)} = \frac{1 + t^2}{1 + \delta^2/S + t^2}, \quad \zeta^{(1)} = \frac{\pm \delta^2 (3 + t^2)}{S(1 + \delta^2/S + t^2)^2}, \quad \cdots.$$

This expansion provides the balanced approximation to the dynamics, i.e., the dynamics on a slow manifold, with $\zeta^{(0)}$ corresponding to the quasi-geostrophic approximation and higher-order terms to more accurate balance relations. The expansion (3.1) is expected to be asymptotic rather than convergent as a result of exponentially small terms which it cannot capture. These terms, which represent inertia-gravity waves, could only vanish if an exactly invariant slow manifold existed; we show that

it is not the case by computing explicitly the amplitude of the inertia-gravity waves generated for a solution that is initially gravity-wave free.

In general it is difficult to separate balanced motion and inertia-gravity waves unambiguously, because there is not a single definition for a (non-invariant) slow manifold. However, for the model under consideration, an unambiguous separation is possible in the limit $|t| \to \infty$. This is because in this limit the solution to (2.3) simply becomes

$$\zeta \sim 1 + C^{\gtrless} \cos(S^{1/2}t/\epsilon + \varphi^{\gtrless}), \tag{3.2}$$

where $C^{\gtrless}$ and $\varphi^{\gtrless}$ are real constants, and the $\gtrless$ signs refer to the limits $t \to \pm\infty$, respectively. Clearly, the slow, balanced motion corresponds to the constant term in (3.2) while the inertia-gravity waves correspond to the rapidly oscillating term. In writing (3.2), we have anticipated that $\zeta(t)$ is purely real for all t; this is obviously the case if, as will be assumed, $\zeta(t)$ is real for $t \to -\infty$. Physically, the rapidly oscillating term in (3.2) then describes a one-dimensional standing inertia-gravity wave whose orientation is being slowly rotated by the background shear (recall (2.1)). In fact, it is a pure buoyancy wave: the influence of rotation vanishes asymptotically for $|t| \to \infty$.

A perfectly balanced condition can be imposed for $t \to -\infty$ by choosing $C^< = 0$. We show below that this choice leads to a non-zero, albeit exponentially small, $C^>$, i.e., non-vanishing inertia-gravity waves for $t \to \infty$. This establishes the inevitable breakdown of balance or, equivalently, the non-existence of an exactly invariant slow manifold.

Figure 1 shows a typical solution, with $S = \delta^2 = 10$ and $\epsilon = 0.25$ (anticyclonic shear). The dashed line is an approximate balanced solution,

$$\tilde{\zeta}_{\text{bal}} = \zeta^{(0)} + \epsilon\zeta^{(1)}. \tag{3.3}$$

The solid line is obtained by integrating (2.3) from $-T$ to T for some large T (100 in

11

this example), with initial conditions

$$\zeta(-T) = \tilde{\zeta}_{\text{bal}}(-T), \quad \dot{\zeta}(-T) = \dot{\tilde{\zeta}}_{\text{bal}}(-T). \tag{3.4}$$

(Details of the numerical method are given in section 3.) It is easily verified that $\tilde{\zeta}_{\text{bal}}(t) = \zeta_{\text{bal}}(t) + O(\epsilon^2 t^{-4})$ for $\epsilon \to 0, |t| \to \infty$, so we can make the initial conditions arbitrarily well balanced by choosing T sufficiently large. We find that the solution remains exceptionally well balanced, even though ϵ is moderate, until t is close to zero. At this point inertia-gravity waves are generated, and the solution eventually takes the expected form of a regular inertia-gravity-wave oscillation about the balanced solution.

The inertia-gravity waves are evidently absent from the perturbation expansion (3.1) but they can be captured by a WKB solution: introducing a solution of the form

$$\zeta = \zeta_{\text{igw}} = (g_0 + \epsilon g_1 + \cdots) \, e^{i\epsilon^{-1} \int_0^t \omega(t') \mathrm{d}t'}$$

into the homogeneous version of (2.3) gives

$$\omega^2 g_0 - S \frac{1 + \delta^2/S + t^2}{1 + \delta^2 + t^2} g_0 = 0, \tag{3.5}$$

at leading order, and, taking this into account,

$$2i\omega \dot{g}_0 + i\dot{\omega} g_0 - \frac{2i\delta^2 t}{(1+t^2)(1+\delta^2+t^2)} \omega g_0 \mp \frac{\delta^2(3+t^2)}{(1+t^2)(1+\delta^2+t^2)} g_0 = 0 \tag{3.6}$$

at the next order. Solving (3.5) yields two roots,

$$\omega = S^{1/2} \left(\frac{1 + \delta^2/S + t^2}{1 + \delta^2 + t^2} \right)^{1/2}, \tag{3.7}$$

and its negative. These can be recognised as the dimensionless frequencies of inertia-gravity waves with wavevector $(k, -kt, m)$. Solving (3.6) gives

$$g_0 = A \frac{1}{\omega^{1/2}} \left(\frac{1+t^2}{1+\delta^2+t^2} \right)^{1/2} e^{ih(t)},$$

12

where A is a complex constant, and

$$h(t) = \frac{\mp \delta^2}{2} \int_0^t \frac{3 + t'^2}{(1 + t'^2)(1 + \delta^2 + t'^2)\omega(t')} \mathrm{d}t'.$$ (3.8)

Note that the lower limit of integration in this integral can be chosen freely, because of the presence of the arbitrary constant A in g_0; we have chosen 0 for convenience.

The generation of exponentially small oscillations is associated with the breakdown of the regular perturbation expansion (3.1) near singularities of its terms in the complex t-plane. Typically, if $t_0 \in \mathbb{C}$ is such a singularity, one expects the appearance of oscillations whose amplitude scales roughly like

$$\exp\left[-\frac{1}{\epsilon} \left| \mathrm{Im} \int_0^{t_0} \omega(t)\, \mathrm{d}t \right| \right].$$

The largest contribution therefore is expected to come from the singularity t_0 leading to the smallest value of the absolute-value term. In the case of the expansion (3.1), there is a pair of singularities at

$$t = \pm t_\star := \pm \mathrm{i} \left(1 + \delta^2/S \right)^{1/2}.$$

At first sight it might be expected from (2.3) that singularities in the balanced solution ζ_{bal} also appear for $t = \pm \mathrm{i}$; however, this is not the case. This is first hinted at by the first few terms in ζ_{bal}, which are all regular at $\pm \mathrm{i}$, and can be confirmed by examining the Frobenius solutions of (2.3) about $\pm \mathrm{i}$.[2]

To capture the exponentially small oscillations associated with $\pm t_\star$, one proceeds as follows. (For a clear introduction to the treatment of exponentially small effects, see Hakim (1998).) Equation (2.3) is integrated not for real t between $\pm\infty$, but on a path in the complex plane that comes close to $t_\star$ (the contribution of $-t_\star$ is simply

---

[2]The Frobenius solutions have indices 0 and 2 so that generically a logarithmic contribution is expected to appear, with $t = \pm \mathrm{i}$ as branch points; however, it turns out that this logarithmic contribution vanishes for (2.3).

complex conjugate to that of $t_\star$ and is added at the end). Away from $t_\star$, the regular perturbation technique can be employed and the solution is given by $\zeta = \zeta_{\text{bal}} + \zeta_{\text{igw}}$. The amplitude of the inertia-gravity waves is zero to the left of $t_\star$ and to be determined to the right. Near $t_\star$ the regular perturbation breaks down and one needs to rescale the equation in an inner region. There, the oscillations, which are generated through a Stokes phenomenon (Olver 1974, Ablowitz & Fokas 1997, Berry 1989), have the same order as the balanced solution, so they can be determined, in fact by solving a standard connection problem. Matching the inner solution with the outer solution gives the connection between the amplitudes of the inertia-gravity waves on both sides of $t_\star$; on the real axis this amplitude is exponentially small. This programme is carried out in Appendix B. We remark that the linearity of (2.3) suggests employing a WKB approximation to estimate $\zeta(t)$ (including its balanced component); although this approximation does not directly capture the exponentially small terms, it could be used as a basis for an alternative approach for their estimation (cf. Hinch (1991, §7.5.6)).

In Appendix B the vorticity of the inertia-gravity waves to the right of the singularities $\pm t_\star$ is found to be

$$\zeta_{\text{igw}} \sim \frac{2|K| e^{-\alpha/\epsilon}}{\epsilon^{1/2}} \frac{1}{\omega^{1/2}} \left(\frac{1+t^2}{1+\delta^2+t^2} \right)^{1/2} \cos \left[\frac{1}{\epsilon} \int_0^t \omega(t') \, dt' + h(t) \pm \frac{\pi}{2} \right], \qquad (3.9)$$

where

$$|K| = \frac{\pi^{1/2} \delta \, e^{\pm\beta}}{2^{1/2}(1+\delta^2/S)^{1/4}}, \qquad (3.10)$$

with the $\pm$ signs corresponding to the anticyclonic and cyclonic shears, respectively. The parameters α and β are derived in Appendix B and given explicitly by (B.2) and (B.4). Taking the limit of (3.9) for $t \to \infty$ and comparing with (3.2) we find that the amplitude of the inertia-gravity waves that are generated spontaneously by the

balanced motion is

$$C^> \sim \frac{2|K|}{S^{1/4}} \frac{\mathrm{e}^{-\alpha/\epsilon}}{\epsilon^{1/2}}. \tag{3.11}$$

This exponentially small, non-zero value establishes unambiguously the breakdown of quasi-geostrophic balance for the sheared-disturbance solutions of the three-dimensional Boussinesq equations.

Although we focus on the inertia-gravity-wave amplitude in the limit $t \to \pm\infty$, it is worth recalling from the theory of the Stokes phenomenon (see, e.g., Berry (1989) or Ablowitz & Fokas (1997)) that the exponentially small terms are in fact switched on abruptly in a small $(O(\epsilon^{1/2}))$ neighbourhood of the intersection between the real t-axis and a Stokes line. Here the relevant Stokes line is defined by $\mathrm{Re} \int_{t_\star}^{t} \omega(t')\mathrm{d}t' = 0$ and is a segment of the imaginary axis, so that the inertia-gravity waves can be predicted to appear suddenly in the neighbourhood of $t = 0$. This explains what is observed in Figure 1.

The parameters α and β appearing in (3.9) both depend on δ and S, and both are defined by elliptic integrals; α controls the exponential smallness of the inertia-gravity waves, β the order-one prefactor $|K|$. Contour plots of α and β are shown in Figure 2. The logs are in base 10, and the values are shown for $S \in (1, 100)$, $\delta^2 \in (0.1, 100)$. Note, in particular, that while α is always positive, β can have either sign. This is interesting because the only asymmetry between anticyclonic and cyclonic shears stems from the factor $\mathrm{e}^{\pm\beta}$ in $|K|$. When $\beta > 0$, as is the case for realistic (large) values of δ and S, the amplitude of the inertia-gravity waves is larger for the anticyclonic shear than for the cyclonic shear; the opposite is true for $\beta < 0$.

Simple expressions for α and β in the limit of large S and δ are derived in Appendix C. In the standard quasi-geostrophic scaling, with $\delta, S \gg 1$ and a ratio of the horizontal scale to the radius of deformation

$$b := \frac{\delta}{S^{1/2}} = O(1),$$

15

(i.e., an $O(1)$ Burger number) we find

$$\alpha \sim \frac{\pi(1 + b^2)}{4b} \quad \text{and} \quad \beta \sim \frac{\pi b}{4}.$$

Note that α is minimised for $b = 1$, hence the amplitude of the waves is maximised for $\delta^2 \sim S$, i.e., for the aspect ratio $m/k \sim N/f$. Another limit, mainly of academic interest, namely $S \gg 1$, $\delta = O(1)$, is also considered in Appendix C. In this limit α and β can be expressed as simple elliptic integrals, and it is possible to show that the change in the sign of β occurs for $\delta = 0.953$, consistent with Figure 2.

We have confirmed the analytic result (3.11) for the inertia-gravity-wave amplitude by solving (2.3) numerically for several values of the parameters. We choose a sufficiently large time T, and integrate (2.3) from $-T$ to T, with initial conditions (3.4). We estimate the final inertia-gravity-wave amplitude at time T using equation (3.2):

$$C^> \approx \left\{ \left[\zeta(T) - \tilde{\zeta}_{\text{bal}}(T) \right]^2 + \frac{\epsilon^2}{S} \left[\dot{\zeta}(T) - \dot{\tilde{\zeta}}_{\text{bal}}(T) \right]^2 \right\}^{1/2}.$$

We use a fourth-order Runge–Kutta integration scheme. T is chosen large enough, and the resolution fine enough, such that the estimated $C^>$ remains unchanged to five significant digits when the resolution and/or T are doubled.

Figure 3 shows $C^>$ as a function of $1/\epsilon$ for $\delta^2 = S = 10$, both for an anticyclonic and for a cyclonic shear. The figure compares the analytic and numerical estimates for $C^>$, revealing an excellent agreement even for moderate values of ϵ. Note that with the logarithmic scale chosen for $C^>$, a straight line corresponds to an exponential dependence on $1/\epsilon$; the slight departure from a straight line noticeable for moderate ϵ results from the factor $\epsilon^{-1/2}$ in (3.11). The values of $C^>$ in Figure 3 also emphasise the extreme smallness of the inertia-gravity waves that are emitted spontaneously. Similar results were obtained for different values of δ and S.

Figure 4 shows the relative error in the analytic estimate for $C^>$, i.e., the difference

16

between the analytically computed and the numerically estimated values, normalised by the numerical value, again for $\delta^2 = S = 10$. Evidently, the error in the analytical values tends to zero approximately linearly with ϵ, suggesting that $C^>$ is in fact given by

$$C^> = \frac{2|K|}{S^{1/4}} \frac{\mathrm{e}^{-\alpha/\epsilon}}{\epsilon^{1/2}} (1 + c_1 \epsilon + \ldots).$$

Carrying out the asymptotic calculation of Appendix B to higher order would no doubt confirm this numerical finding.

4 Discussion

This paper presents an example of the generation of inertia-gravity waves by balanced motion in the small-Rossby-number regime $\epsilon \ll 1$. Exponential-asymptotic calculations and numerical solutions provide a firm basis to the qualitative picture that has emerged in recent years: the spontaneous generation of waves is unavoidable and no exactly invariant slow manifold can be defined — in other words, the slow manifold is inherently fuzzy — but the amplitude of the waves generated is extremely small, in fact exponentially small in ϵ. Our results fall short of a genuine proof of this fact, however, because the exponential-asymptotic method that we use is a formal one and does not provide strict error bounds; a proof providing such error bounds would be desirable.

The model studied takes advantage of an exact reduction of the full three-dimensional (non-hydrostatic Boussinesq) partial-differential equations to a linear ordinary-differential equation. It is, however, not particularly realistic, notably because it involves unbounded, infinite-energy solutions for which there is no meaningful energy conservation. Furthermore, the combination of unbounded domain and uniform background shear introduces an apparent difficulty of interpretation in that the slow

17

time dependence of the fields, which involves the factor $\exp(-ik\Sigma yt)$ (see (2.3)), may not seem truly slow for large $|y|$. That this is a simple Doppler-shift effect, with no dynamical consequences for the interactions between balanced motion and inertia-gravity waves, should however be clear: the factor $\exp(-ik\Sigma yt)$ being common to all fields, it drops from the problem completely. To remove any ambiguity, it would nevertheless be of interest to examine the issue of inertia-gravity-wave generation using solutions of the Boussinesq (or shallow-water) that are less contrived that the sheared distrubances considered here.

In spite of the limitations of our model, we believe that many of the features of the inertia-gravity-wave generation that it reveals are generic. Among these features, the role played by the singularities of the balanced motion in the complex t-plane is noteworthy. In a sense, it is at these singularities that the wave generation takes place (although for the usual real values of t they appear at the intersection with the Stokes lines emanating from the singularities). When deriving simplified models of the interactions between balanced and inertia-gravity dynamics, it is therefore important that the approximations made do not alter the singularity structure of the balanced motion.[3]

Our estimation of the inertia-gravity-wave amplitude is of course much simplified by the special spatio-temporal structure (2.1) of the sheared-disturbance solutions that we consider. For more general solutions of the Boussinesq equations, the location of singularities of the balanced motion in the complex t-plane will depend on space. One can therefore expect the inertia-gravity waves to be generated across time-dependent front-like structures. An asymptotic theory providing the exponen-

---

[3]As an illustration, we note that neglecting the term $\mp 2\epsilon/(1+t^2)$ of $c(t)$ in (2.3) — a reasonable approximation for t real but not for t near $\pm i$ — introduces spurious singularities in ζ_{bal} at $t = \pm i$. As a result, the amplitude of the inertia-gravity waves is grossly overestimated: it remains proportional to $\exp(-\alpha/\epsilon)$ but with the upper limit of integration in the definition (B.2) of α replaced by 1.

tially small amplitude of the gravity waves in this general situation would be useful; its derivation is likely to be challenging, however, since the study of the Stokes phenomenon for partial-differential equations is still in its infancy.

In our analysis, we exploit the fact that dynamics of the model is asymptotically simple for $t \to \pm\infty$. This makes it possible to separate balanced motion and inertia-gravity waves unambiguously, and hence to estimate the amplitude of the waves generated spontaneously. We note, however, that it is not crucial to consider, as we do, initial conditions for $t \to -\infty$ and diagnose the wave amplitude for $t \to \infty$. Since the waves are switched on abruptly in a $O(\epsilon^{1/2})$ neighbourhood of $t = 0$ (where the Stokes line is crossed, see Berry (1989)), there is a natural unique definition of the balanced part of the motion, and hence of the inertia-gravity waves, for any finite t away from 0: for $t < 0$, the balanced motion is obtained by forward integration of (2.3) from balanced initial conditions $\zeta = 1$, $\dot{\zeta} = 0$ for $t \to -\infty$; for $t > 0$ it is obtained by backward integration from the same balanced initial conditions for $t \to \infty$. What this defines can be viewed as two disconnected parts of a manifold which is not completely slow (although the oscillations are minuscule for $\epsilon^{-1/2}t \gg 1$) and not globally invariant.

This construction does not generalise straightforwardly to more complicated situations in which there may be many Stokes lines intersecting the real t-axis. However, the study of the Stokes phenomenon (Berry 1989) suggests an alternative separation between balanced motion and inertia-gravity waves which, albeit not completely free of arbitrariness, may prove useful. For a certain class of problems, Berry (1989) notes that the subdominant term in an asymptotic expansion can be computed with an asymptotically small error by subtracting from the full (typically numerical) solution the series expansion of the dominant term, provided that this series be optimally truncated. (The truncation error is then $\epsilon^{1/2}$ smaller than the subdominant term.)

19

Transposed to the context of balanced dynamics, this conclusion implies that the inertia-gravity waves can be defined uniquely up to asymptotically small terms as the difference between a full (unbalanced) solution and an optimally truncated (balanced) series solution (cf. Warn & Menard (1986)). Whether this conclusion holds needs to be verified for specific models; if it does, it provides a practical way of removing the ambiguities in the separation between balanced and wave motions which have marred the debate about the slow manifold.

Our analysis addresses the issue of spontaneous generation of inertia-gravity waves in an initialised flow. A related issue concerns the coupling between balanced motion and inertia-gravity waves in flows with a significant level of wave activity. Recent work on the subject using averaging techniques (Babin, Mahalov & Nicolaenko 2002, Majda & Embid 1998, Reznik, Zeitlin & Jelloul 2001, Wirosoetisno, Shepherd & Temam 2002, and references therein) show that, to first order in ϵ, this coupling is of catalytic type, with the balanced motion influencing the propagation of inertia-gravity waves without changing their energy. It is possible that the energy transfer from balanced motion to inertia-gravity waves is exponentially small, even in the presence of excited waves. To capture this transfer would thus require to perform some form of averaging beyond all orders in ϵ.

To conclude, let us return to the important distinction mentioned in the Introduction between the small-Rossby-number regime $\epsilon \ll 1$ which is the focus of this paper, and the small-Froude-number regime $F \ll 1$ with $\epsilon = O(1)$ studied by Ford et al. (2000). As recently emphasised by Saujani & Shepherd (2002) and Ford et al. (2002), the inertia-gravity waves generated by the balanced motion are exponentially small in ϵ in the first regime — because of the frequency separation between waves and balanced motion, while they scale like some power of F in the second regime — because of the absence of frequency separation for sufficiently long waves. It is

in this second regime, in which the balanced motion generates (slow) inertia-gravity waves with a much larger spatial scale (Ford et al. 2000), that the mechanism of wave emission is closely analogous to the Lighthill radiation of acoustic waves (Lighthill 1952).

In our model, the spatial scale of the inertia-gravity waves is by construction that of the balanced motion; there is, therefore, no direct analogue of the Lighthill-like radiation examined by Ford et al. (2000). There is, however, an asymptotic regime in which the frequency separation between inertia-gravity waves and balanced motion is not complete, so that a power-law dependence of the wave amplitude on the asymptotic parameter might be expected. This regime corresponds to $\epsilon = O(1)$, $\delta, S \gg 1$ and $\delta/S^{1/2} = O(1)$. With this scaling, the inertia-gravity-wave frequency found by a WKB analysis with $S^{1/2}$ or $S^{1/2}/\epsilon = N/|\Sigma|$ (to encompass the non-rotating case $\epsilon = f = 0$) as the large parameter is again $\omega(t)/\epsilon$ with $\omega(t)$ given in (3.7). It is large, specifically $O(S^{1/2}/\epsilon)$, for $t \gtrsim O(\delta)$ but of $O(1)$, and hence matching the balanced motion frequency, for $t \ll \delta$. A boundary-layer analysis may be carried out to relate the outer behaviour of the solution for $t = O(\delta)$ to the inner behaviour for $t = O(1)$ and deduce the amplitude of inertia-gravity waves that are generated spontaneously from a balanced state at $t \to -\infty$. This analysis does not appear to yield a closed form for the inertia-gravity-wave amplitude because the inner dynamics is governed by an ordinary differential equation not significantly simpler than (2.3), but it indicates that the amplitude scales like $\delta^{1/2}$ or, equivalently, $S^{1/4}/\epsilon^{1/2}$. (Note that the positive power appears because the time interval during which the inertia-gravity-wave frequency is $O(1)$ increases with δ.) This scaling is confirmed by numerical solutions of (2.3) whose results are shown in Figure 5.

A   Derivation of (2.2)

The equations of motion, employing the Boussinesq approximation, for an inviscid stratified fluid in an environment rotating with a constant Coriolis frequency f are

$$D_t U - fV = -\frac{P_x}{\bar{\rho}} , \tag{A.1}$$

$$D_t V + fU = -\frac{P_y}{\bar{\rho}} , \tag{A.2}$$

$$D_t W + \frac{g\rho}{\bar{\rho}} = -\frac{P_z}{\bar{\rho}} , \tag{A.3}$$

$$D_t \rho + W \rho_{0z} = 0 , \tag{A.4}$$

$$U_x + V_y + W_z = 0 , \tag{A.5}$$

where (U, V, W) are the usual Cartesian velocity components, P is the reduced pressure, the fluid density has been written as $\rho_0(z) + \rho(x, y, z, t)$, with $\bar{\rho}$ denoting its (constant) mean value, and $D_t := \partial_t + U\partial_x + V\partial_y + W\partial_z$ is the material derivative. We denote the buoyancy by $B := -g\rho/\bar{\rho}$, and the geopotential by $\Phi := P/\bar{\rho}$.

Motivated by McWilliams and Yavneh (1998), we consider unbounded solutions which are a sum of a uniform shear flow and a finite-amplitude perturbation:

$$
\begin{aligned}
(U, V, W, \Phi, B) &= \left[\Sigma y, 0, 0, -\frac{1}{2}f\Sigma y^2, 0\right] \\
&+ [u(t), v(t), w(t), \phi(t), b(t)]\, \mathrm{e}^{i[k(x-\Sigma yt)+ly+mz]} , \tag{A.6}
\end{aligned}
$$

where $-\Sigma$ is the constant vorticity of the uniform shear flow, and k, l, m are constant wavenumbers. This form, together with the incompressibility condition (A.5), ensures

that the material derivatives of (U, V, W, Φ, B) in (A.1)–(A.5) simply transform into the time derivatives of the corresponding amplitudes (u, v, w, ϕ, b). Substituting (A.6) into (A.1)–(A.5) therefore produces the set of linear ordinary differential equations

$$\dot{u} + (\Sigma - f)v = -\mathrm{i}k\phi, \tag{A.7}$$

$$\dot{v} + fu = -\mathrm{i}(l - \Sigma kt)\phi, \tag{A.8}$$

$$\dot{w} - b = -\mathrm{i}m\phi, \tag{A.9}$$

$$\dot{b} + N^2 w = 0, \tag{A.10}$$

$$ku + (l - \Sigma kt)v + mw = 0, \tag{A.11}$$

where $N^2 := -(g/\bar{\rho})\mathrm{d}\rho_0/\mathrm{d}z$ is the buoyancy frequency squared, which is assumed to be constant, and the overdot denotes the time derivative. Taking $\mathrm{i}k(A.8) - \mathrm{i}l(A.7)$ yields

$$\dot{\zeta} + (f - \Sigma)\mathcal{D} = 0, \tag{A.12}$$

where $\zeta := \mathrm{i}kv - \mathrm{i}(l - \Sigma kt)u$ is the amplitude of the vertical component of the vorticity of the finite perturbation, and $\mathcal{D} := \mathrm{i}ku + \mathrm{i}(l - \Sigma kt)v$ is the amplitude of the horizontal velocity divergence. From (A.10), (A.11) and (A.12) we obtain

$$\dot{q} = 0, \tag{A.13}$$

where $q = N^2\zeta + \mathrm{i}m(f - \Sigma)b$ is the amplitude of the Ertel potential vorticity associated with the perturbation. Next, we take $\mathrm{i}k(A.7) + \mathrm{i}l(A.8)$ and obtain

$$\dot{\mathcal{D}} + \frac{2\Sigma k(l - \Sigma kt)}{\lambda^2}\mathcal{D} = \lambda^2\phi + \left(f - \frac{2k^2\Sigma}{\lambda^2}\right)\zeta, \tag{A.14}$$

where

$$\lambda^2 := k^2 + (l - \Sigma kt)^2. \tag{A.15}$$

Elimination of ϕ from (A.14), using (A.9), (A.11) and the definitions of $\mathcal{D}$ and q, yields after multiplication by m^2,

$$(\lambda^2 + m^2)\dot{\mathcal{D}} + \frac{2\Sigma km^2(l - \Sigma kt)\mathcal{D}}{\lambda^2} = \left[m^2\left(f - \frac{2k^2\Sigma}{\lambda^2}\right) + \frac{N^2\lambda^2}{f - \Sigma}\right]\zeta - \frac{q\lambda^2}{f - \Sigma}. \tag{A.16}$$

Finally, we combine (A.12) and (A.16) to obtain a single equation for ζ:

$$\ddot{\zeta} + b(t)\dot{\zeta} + c(t)\zeta = \frac{q\lambda^2}{\lambda^2 + m^2}, \tag{A.17}$$

where

$$
\begin{aligned}
b(t) &= \frac{2(l - \Sigma k t)\Sigma k m^2}{\lambda^2(\lambda^2 + m^2)}, \\
c(t) &= \frac{m^2(f - \Sigma)(f - 2k^2\Sigma/\lambda^2) + N^2\lambda^2}{\lambda^2 + m^2}.
\end{aligned}
$$

Since this equation is linear, $\zeta(t)$ is simply proportional to the constant potential-vorticity amplitude q which can therefore be chosen arbitrarily. The choice $q = N^2$ is convenient. Introducing this into (A.17) and using the explicit form (A.15) of λ we finally obtain (2.2).

B   Exponential asymptotics

We start by considering the solution of (2.3) away from the singular point at $t = t_\star$, in the outer region where $|t - t_\star| = O(1)$. In this region the balanced part of $\zeta(t)$ is captured by a regular perturbation expansion, while the inertia-gravity wave part is captured by the WKB approximation. We then examine the solution of (2.3) in an inner region near $t_\star$ and carry out the necessary matching between inner and outer solutions.

B.1   Outer expansion $|t - t_\star| = O(1)$

From (3.1) and (3.7)–(3.8), we have the outer solution

$$
\begin{aligned}
\zeta &= \frac{1 + t^2}{1 + \delta^2/S + t^2} + \cdots \\
&\quad + \frac{1}{\omega^{1/2}}\left(\frac{1 + t^2}{1 + \delta^2 + t^2}\right)^{1/2}\left\{A^{\gtrless}\mathrm{e}^{\mathrm{i}\left[\epsilon^{-1}\int_0^t \omega(t')\mathrm{d}t' + h(t)\right]} + B^{\gtrless}\mathrm{e}^{-\mathrm{i}\left[\epsilon^{-1}\int_0^t \omega(t')\mathrm{d}t' + h(t)\right]}\right\},
\end{aligned} \tag{B.1}
$$

where the $\gtrless$ of the complex constants $A^{\gtrless}, B^{\gtrless}$ refer to the right and left of $t_\star$, respectively. We assume $A^< = B^< = 0$ and want to compute $A^>$ and $B^>$. (In fact we know that $A^> = 0$ since a non-zero value would correspond to an exponentially large solution on the real axis — only subdominant, exponentially small terms may be switched on through a Stokes phenomenon.) To match the outer solution (B.1) to the inner solution to be derived, we need the asymptotic form of ζ when t is on a path approaching $t_\star$. It is convenient to take this path to follow anti-Stokes lines where

$$\mathrm{Im} \int_0^t \omega(t')\,\mathrm{d}t' = \mathrm{const} = \frac{1}{\mathrm{i}} \int_0^{t_*} \omega(t')\,\mathrm{d}t' =: \alpha,$$

so that the WKB terms have a constant amplitude at leading order. Here, we have defined the (positive) constant α which can also be rewritten explicitly using (3.7) as the real integral

$$\alpha = S^{1/2} \int_0^{(1+\delta^2/S)^{1/2}} \left(\frac{1 + \delta^2/S - x^2}{1 + \delta^2 - x^2} \right)^{1/2} \mathrm{d}x \tag{B.2}$$

and can be evaluated explicitly in terms of elliptic functions. It is easy to check that the anti-Stokes lines, $\mathcal{L}^{\gtrless}$ say for the line tending to $\pm\infty$, are tangent to

$$\mathcal{L}^< : \arg(t - t_\star) = -\frac{5\pi}{6} \quad \text{and} \quad \mathcal{L}^> : \arg(t - t_\star) = -\frac{\pi}{6}$$

at $t_\star$. See Figure 6. On $\mathcal{L}^<$, the solution ζ has the following behaviour for $\epsilon^{2/3} \ll |t - t_\star| \ll 1$:

$$\zeta = \frac{\mathrm{i}\delta^2}{2S(1 + \delta^2/S)^{1/2}} \frac{1}{t - t_\star} + \cdots \tag{B.3}$$

To derive a similar result for $\mathcal{L}^>$ we first note the following asymptotic behaviours as $t \to t_\star$:

$$\omega \sim a^{1/2} \mathrm{e}^{\pi \mathrm{i}/4} (t - t_\star)^{1/2},$$

where

$$a := \frac{2S^2(1 + \delta^2/S)^{1/2}}{\delta^2(S - 1)},$$

25

and

$$\int_0^t \omega(t')\,\mathrm{d}t' = \mathrm{i}\alpha + \frac{2}{3}a^{1/2}\mathrm{e}^{\pi\mathrm{i}/4}(t-t_\star)^{3/2} + \cdots,$$

$$\left(\frac{1+t^2}{1+\delta^2+t^2}\right)^{1/2} \sim \frac{\mathrm{e}^{\pi\mathrm{i}/2}}{(S-1)^{1/2}}.$$

Note that the factor $\mathrm{e}^{\pi\mathrm{i}/2}$ (as opposed to $\mathrm{e}^{-\pi\mathrm{i}/2}$) in $(1+t^2)^{1/2}$ appears because we are considering the limit $t \to t_\star$ on $\mathcal{L}^>$, and the continuation of the square root is made to the right of the branch point at i. See Figure 6. We also need the leading-order approximation to $h(t)$ defined in (3.8), which we write

$$h(t) \sim h(t_\star) =: \mp\mathrm{i}(\beta + \mathrm{i}\gamma),$$

thus defining the real numbers β and γ. Some care is necessary to evaluate β and γ since $h(t)$ is not analytic between the real axis and $t_\star$ because of the pole of the integrand at $t = \mathrm{i}$. In the definition of $h(t)$, t describes $\mathcal{L}^>$. Taking, as before, a path to the right of the singularity at $t = \mathrm{i}$ which we then reduce to a segment of the imaginary axis indented to the right of $t = \mathrm{i}$, we obtain

$$\beta = \frac{\delta^2}{2S^{1/2}}\mathcal{P}\int_0^{(1+\delta^2/S)^{1/2}} \frac{3-x^2}{(1-x^2)(1+\delta^2-x^2)^{1/2}(1+\delta^2/S-x^2)^{1/2}}\mathrm{d}x, \qquad \text{(B.4)}$$

where $\mathcal{P}$ denotes Cauchy-principal-value integral, and

$$\gamma = -\frac{\pi\mathrm{i}}{2}.$$

Note that one could also choose a path to the left of $t = \mathrm{i}$: if this is done consistently our final results below are recovered unchanged; this reflects the single-valuedness of the solution near $t = \mathrm{i}$.

Introducing the asymptotic behaviours just deduced into (B.1), we find the outer solution along $\mathcal{L}^>$ when approaching $t_\star$ in the form

$$\zeta = \frac{\mathrm{i}\delta^2}{2S(1+\delta^2/S)^{1/2}}\frac{1}{t-t_\star} + \cdots$$

26

$$+\frac{\mathrm{e}^{3\pi\mathrm{i}/8}}{a^{1/4}(S-1)^{1/2}}\frac{1}{(t-t_\star)^{1/4}}\left[A^>\mathrm{e}^{-\alpha/\epsilon\pm\beta\mp\pi\mathrm{i}/2}\,\mathrm{e}^{\frac{2}{3}a^{1/2}\mathrm{e}^{3\pi\mathrm{i}/4}(t-t_\star)^{3/2}/\epsilon}\right. \tag{B.5}$$

$$\left.+\,B^>\mathrm{e}^{\alpha/\epsilon\mp\beta\pm\pi\mathrm{i}/2}\,\mathrm{e}^{\frac{2}{3}a^{1/2}\mathrm{e}^{-\pi\mathrm{i}/4}(t-t_\star)^{3/2}/\epsilon}\right]+\cdots.$$

The constants $A^>$ and $B^>$ are determined by matching (B.3) and (B.5) with the inner solution which we now derive.

B.2   Inner expansion $|t - t_\star| \ll 1$ and matching

There is a turning point at $t_\star$ where the regular-perturbation and WKB expansions break down. The usual turning-point scaling holds, with an inner region of size $\epsilon^{2/3}$. Consequently, we introduce the new independent variable τ defined by

$$t = t_\star + \epsilon^{2/3}a^{-1/3}\tau. \tag{B.6}$$

We also expand ζ according to

$$\zeta(t) = (\epsilon a)^{-2/3}\phi(\tau) + \cdots. \tag{B.7}$$

Introducing into (2.3), we find at leading-order the inhomogeneous Airy equation

$$\phi'' + \mathrm{i}\tau\phi = \frac{-S}{S-1},$$

whose solution can be written in terms of Airy functions and, for the inhomogeneous part, Scorer functions, with argument $\mathrm{e}^{-5\pi\mathrm{i}/6}\tau$, $\mathrm{e}^{-\pi\mathrm{i}/6}\tau$ or $\mathrm{e}^{\pi\mathrm{i}/2}\tau$ (Abramowitz & Stegun 1965, Olver 1974). The solution that matches the outer solution (B.3) on $\mathcal{L}^<$ can be written as

$$\phi = \frac{\pi S\mathrm{e}^{-2\pi\mathrm{i}/3}}{S-1}\mathrm{Hi}(\mathrm{e}^{-\pi\mathrm{i}/6}\tau), \tag{B.8}$$

where Hi is a Scorer function (Olver 1974). To verify the matching, we use the large-z asymptotic of $\mathrm{Hi}(z)$, valid when the argument of z is π as is the case for (B.8) on $\mathcal{L}^<$; we find

$$\mathrm{Hi}(z) \sim \frac{-1}{\pi z}. \tag{B.9}$$

27

This gives

$$\phi(\tau) \sim \frac{iS}{S-1}\frac{1}{\tau},$$

for $|\tau| \gg 1$ on $\mathcal{L}^<$ which is seen to match (B.3) when (B.6)–(B.7) are taken into account. To find the asymptotic behaviour of (B.8) on $\mathcal{L}^>$ and match with (B.5), we use the connection formula

$$\mathrm{Hi}(z) = \mathrm{e}^{-2\pi i/3}\mathrm{Hi}(\mathrm{e}^{-2\pi i/3}z) + 2\mathrm{e}^{\pi i/6}\mathrm{Ai}(\mathrm{e}^{2\pi i/3}z)$$

(see F. W. J. Olver, http://dlmf.nist.gov/Contents/AI/) and obtain the alternative representation for (B.8),

$$\phi = \frac{\pi S}{S-1}\left[\mathrm{e}^{2\pi i/3}\mathrm{Hi}(\mathrm{e}^{-5\pi i/6}\tau) + 2\mathrm{e}^{\pi i/2}\mathrm{Ai}(\mathrm{e}^{\pi i/2}\tau)\right]. \tag{B.10}$$

The behaviour for $|\tau| \gg 1$ on $\mathcal{L}^>$ is derived from (B.9) (again the argument of z is π) and from the large-z asymptotic formula

$$\mathrm{Ai}(z) \sim \frac{\mathrm{e}^{-\frac{2}{3}z^{3/2}}}{2\pi^{1/2}z^{1/4}}.$$

Introducing into (B.10) gives the asymptotic expression

$$\phi = \frac{iS}{S-1}\frac{1}{\tau} + \frac{\pi^{1/2}S\mathrm{e}^{-5\pi i/8}}{S-1}\frac{1}{\tau^{1/4}}\mathrm{e}^{\frac{2}{3}\mathrm{e}^{-\pi i/4}\tau^{3/2}} + \cdots.$$

Comparison with (B.5) taking (B.6)–(B.7) into account yields $A^> = 0$, as expected, and

$$\frac{\mathrm{e}^{\alpha/\epsilon\mp\beta\pm\pi i/2}\mathrm{e}^{3\pi i/8}}{a^{1/4}(S-1)^{1/2}}B^> = \frac{\pi^{1/2}S\mathrm{e}^{-5\pi i/8}}{\epsilon^{1/2}a^{3/4}(S-1)}.$$

Using the definition of a we finally find the exponentially small gravity-wave amplitude to be

$$B^> = K\epsilon^{-1/2}\,\mathrm{e}^{-\alpha/\epsilon},$$

with the order-one prefactor K given by

$$K = \frac{\pi^{1/2}\delta\,\mathrm{e}^{\pm\beta\pm\pi i/2}}{2^{1/2}(1+\delta^2/S)^{1/4}}. \tag{B.11}$$

This completes the asymptotic calculation of the inertia-gravity-wave amplitude. Taking into account the (complex conjugate) contributions of both turning points at $t = \pm t_\star$, we find from (B.1) the inertia-gravity-wave part of $\zeta(t)$ for positive t in the form (3.9).

C   Asymptotics of α and β

The regime most relevant to atmospheric and oceanic dynamics is characterised by $\delta, S \gg 1$ and $b = \delta/S^{1/2} = O(1)$. Substituting $bS^{1/2}$ for δ and taking the limit $S \to \infty$ in (B.2) we find

$$\alpha \sim \frac{1}{b} \int_0^{(1+b^2)^{1/2}} (1 + b^2 - x^2)^{1/2}\, \mathrm{d}x = \frac{\pi(1 + b^2)}{4b}.$$

Proceeding similarly in (B.4) gives

$$\beta \sim \frac{b}{2} \mathcal{P} \int_0^{(1+b^2)^{1/2}} \frac{3 - x^2}{(1 - x^2)(1 + b^2 - x^2)^{1/2}} \mathrm{d}x.$$

This integral can be evaluated analytically by first writing it as the sum of two integrals using $3 - x^2 = 1 - x^2 + 2$, then changing the integration variable from x to θ with $x = (1 + b^2)^{1/2} \sin \theta$ to obtain

$$
\begin{aligned}
\beta &\sim \frac{b}{2} \left[\int_0^{\pi/2} \mathrm{d}\theta + 2\mathcal{P} \int_0^{\pi/2} \frac{\mathrm{d}\theta}{1 - (1 + b^2)\sin^2 \theta} \right] \\
&\sim \frac{\pi b}{4},
\end{aligned}
$$

since the second integral vanishes. Note that β is positive in this regime, and that α is minimised for $b = 1$, whereupon it tends to $\pi/2$.

Another asymptotic regime of interest, in particular regarding the change of sign of β, corresponds to $\delta = O(1)$ and $S \gg 1$. Letting $S \to \infty$ in (B.2) yields

$$\alpha \sim S^{1/2} \int_0^1 \left(\frac{1 - x^2}{1 + \delta^2 - x^2} \right)^{1/2} \mathrm{d}x = S^{1/2} \int_0^{\pi/2} \frac{\cos^2 \theta\, \mathrm{d}\theta}{(1 + \delta^2 - \sin^2 \theta)^{1/2}},$$

which can be expressed in terms of complete elliptic integrals. To derive the corresponding result for β care needs to be exercised, because the limit $S \to \infty$ leads the coalescence of the singularities at $x = 1$ and $x = (1 + \delta^2/S)^{1/2}$ of the integrand in (B.4). As before, we write β as the sum of two integrals:

$$\beta = \frac{\delta^2}{2S} \left\{ \int_0^{(1+\delta^2/S)^{1/2}} \frac{\mathrm{d}x}{(1 + \delta^2 - x^2)^{1/2}(1 + \delta^2/S - x^2)^{1/2}} \right.$$
$$\left. + 2 \int_0^{(1+\delta^2/S)^{1/2}} \left[\frac{1}{(1 + \delta^2 - x^2)^{1/2}} - \frac{1}{\delta} \right] \frac{\mathrm{d}x}{(1 - x^2)(1 + \delta^2/S - x^2)^{1/2}} \right\}.$$

In the second integral, the term proportional to $1/\delta$ whose contribution can be shown to vanish has been added to remove the singularity of the integrand at $x = 1$. The limit $S \to \infty$ can now be taken. With $x = \sin\theta$, we find after some manipulations

$$\beta \sim \frac{\delta^2}{2S} \left\{ \int_0^{\pi/2} \frac{\mathrm{d}\theta}{(1 + \delta^2 - \sin^2\theta)^{1/2}} \right.$$
$$\left. - \frac{2}{\delta} \int_0^{\pi/2} \frac{\mathrm{d}\theta}{(1 + \delta^2 - \sin^2\theta)^{1/2}[(1 + \delta^2 - \sin^2\theta)^{1/2} + \delta]} \right\}$$

which, again, can be expressed in terms of elliptic integrals. A numerical calculation then shows that $\beta > 0$ for $\delta > 0.953$ and $\beta < 0$ for $\delta < 0.953$.

References

Ablowitz, M. J. & Fokas, A. S. 1997, *Complex variables: introduction and applications*, Cambridge University Press.

Abramowitz, M. & Stegun, I. A. 1965, *Handbook of mathematical functions*, Dover.

Babin, A., Mahalov, A. & Nicolaenko, B. 2002, Fast singular oscillating limits of stably-stratified 3D Euler and Navier–Stokes equations, and ageostrophic wave fronts, *in* J. Norbury & I. Roulstone, eds, Large-scale atmosphere–ocean dynamics I: analytical methods and numerical models, Cambridge University Press, pp. 126–201.

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

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

Byatt-Smith, J. G. & Davie, A. M. 1990, A rigorous proof of an exponentially small estimate for a boundary value arising from an ordinary differential equation, *Proc. R. Soc. Edinburgh* **114A**, 243–268.

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

Camassa, R. & Tin, S.-K. 1996, The global geometry of the slow manifold in the Lorenz–Krishnamurthy model, *J. Atmos. Sci.* **53**, 3251–3264.

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

Ford, R., McIntyre, M. E. & Norton, W. A. 2000, Balance and the slow quasi-manifold: some explicit results, *J. Atmos. Sci.* **57**, 1236–1254.

Ford, R., McIntyre, M. E. & Norton, W. A. 2002, Reply to comments by S. Saujani and T. G. Shepherd on "Balance and the slow quasi-manifold: some explicit results", *J. Atmos. Sci.* **59**, 2878–2882.

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

Hakim, V. 1998, Asymptotic techniques in nonlinear problems: some illustrative examples, *in* C. Godrèche & P. Manneville, eds, Hydrodynamics and nonlinear instabilities, Cambridge University Press, chapter 3, pp. 295–386.

Hinch, E. J. 1991, *Perturbation methods*, Cambridge University Press.

Kushner, P. J., McIntyre, M. E. & Shepherd, T. G. 1998, Coupled Kelvin wave and mirage-wave instabilities in semi-geostrophic dynamics, *J. Phys. Oceanogr.* **28**, 737–762.

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. R. Soc. Lond. A* **211**, 564–587.

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

Lorenz, E. N. 1986, On the existence of a slow manifold, *J. Atmos. Sci.* **43**, 1547–1557.

Lorenz, E. N. & Krishnamurthy, V. 1987, On the nonexistence of a slow manifold, *J. Atmos. Sci.* **44**, 2940–2950.

Majda, A. J. & Embid, P. 1998, Averaging over fast gravity waves for geophysical flows with unbalanced initial data, *Theoret. Comput. Fluid. Dynamics* **11**, 155–169.

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

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

Nore, C. & Shepherd, T. G. 1997, A Hamiltonian weak-wave model for shallow-water flow, *Proc. R. Soc. Lond.* **453**, 563–580.

Olver, F. W. J. 1974, *Asymptotics and special functions*, Academic Press, New York. Computer Science and Applied Mathematics.

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

Saujani, S. & Shepherd, T. G. 2002, Comments on "Balance and the slow quasi-manifold: some explicit results", *J. Atmos. Sci.* **59**, 2874–2877.

Segur, H., Tanveer, S. & Levine, H., eds 1991, *Asymptotics beyond all orders*, Plenum Press, New York.

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

Vautard, R. & Legras, B. 1986, Invariant manifolds, quasi-geostrophy and initialization, *J. Atmos. Sci.* **43**, 565–584.

Warn, T. 1997, Nonlinear balance and quasi-geostrophic sets, *Atmos.-Ocean* **35**, 135–145.

Warn, T. & Menard, R. 1986, Nonlinear balance and gravity-inertial wave saturation in a simple atmospheric model, *Tellus* **38A**, 285–294.

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

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

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

Figure captions

Figure 1: The evolution of $\zeta(t)$ (solid line) is compared to $\tilde{\zeta}_{\mathrm{bal}}$ (dashed) for $\epsilon = 0.25$, $\delta^2 = S = 10$.

Figure 2: Contour plots of α and β as functions of S and δ^2.

Figure 3: Inertia-gravity-wave amplitude $C^>$ as a function of $1/\epsilon$ for $\delta^2 = S = 10$, and for an anticyclonic and a cyclonic shear.

Figure 4: Relative error of the analytic estimate for $C^>$, assuming the numerical results to be exact. The parameters are the same as in Figure 3.

Figure 5: Inertia-gravity-wave amplitude as a function of S for $\epsilon = 0.25$ and $\delta = S$.

Figure 6: Integration in the complex t-plane: regular perturbation and WKB techniques are used to solve (2.3) along the anti-Stokes lines $\mathcal{L}^<$ and $\mathcal{L}^>$; the outer solutions so obtained are matched with an inner solution found in a neighbourhood of the singular point $t_\star$. The segment of the imaginary axis $\mathrm{Re}\, t = 0$ joining $t_\star$ to its conjugate is a Stokes line.
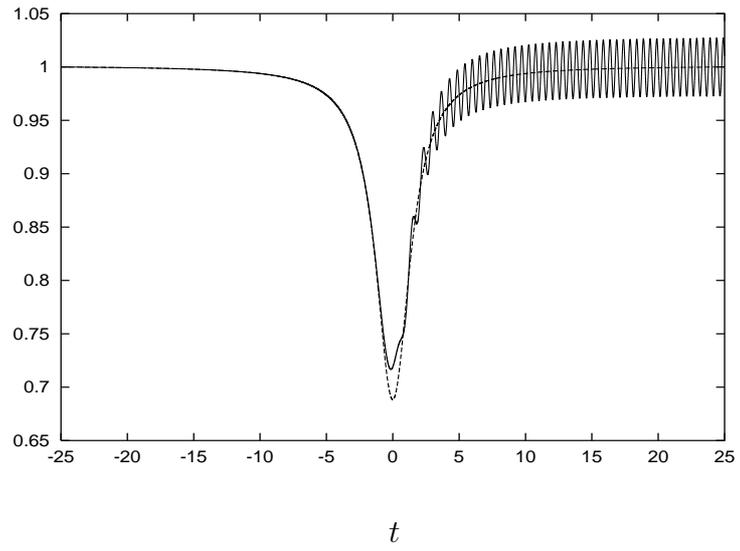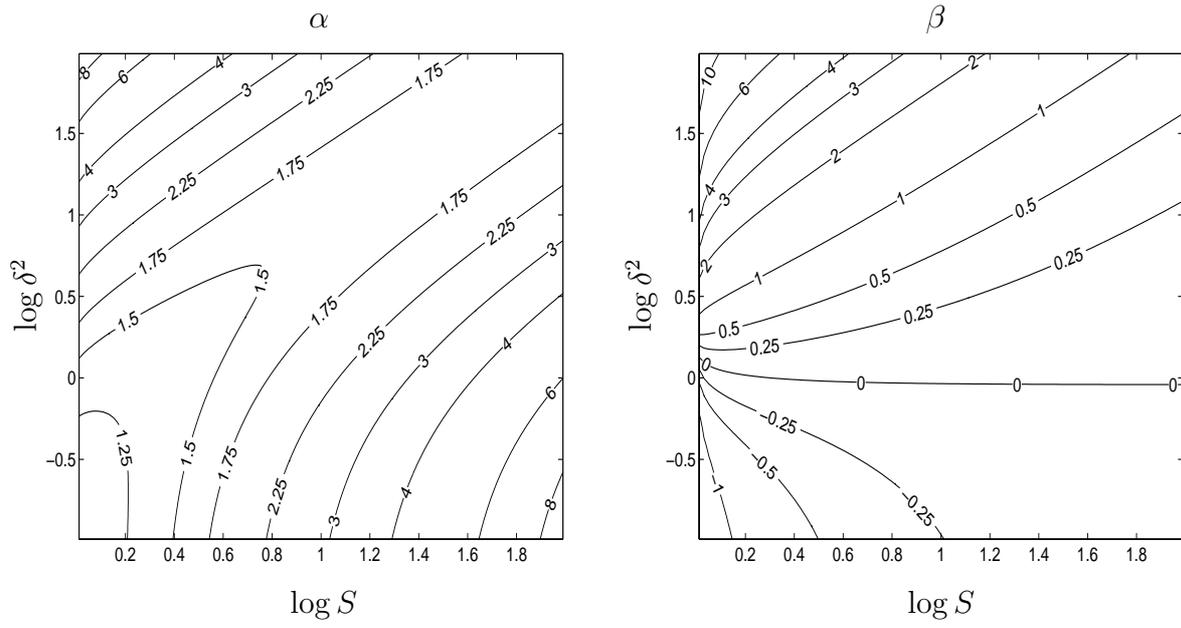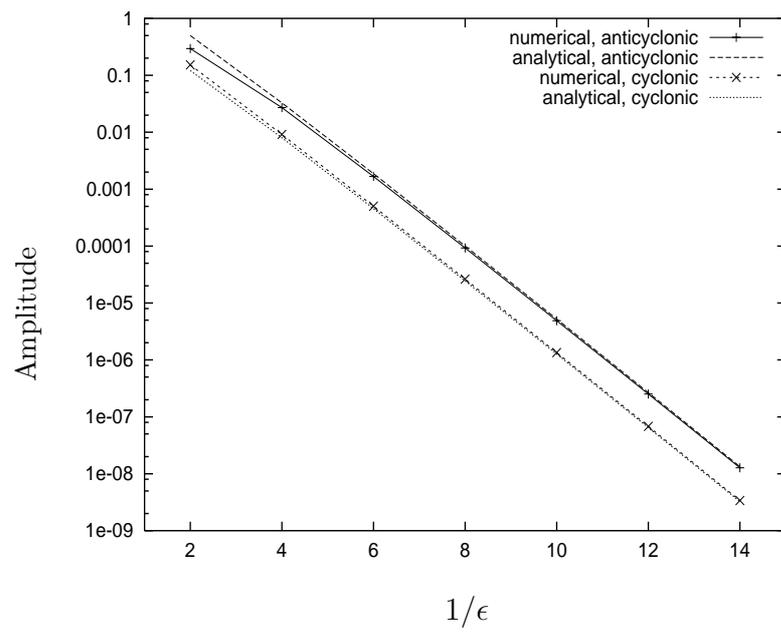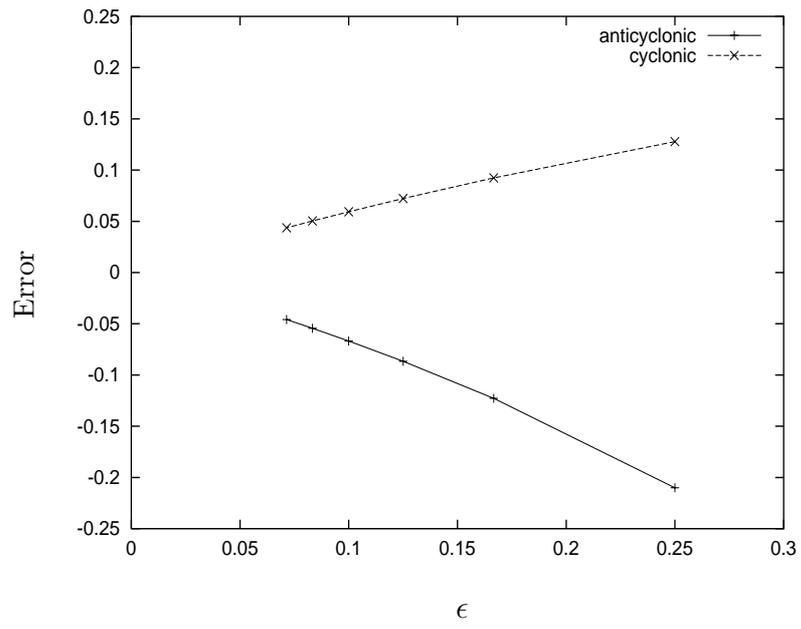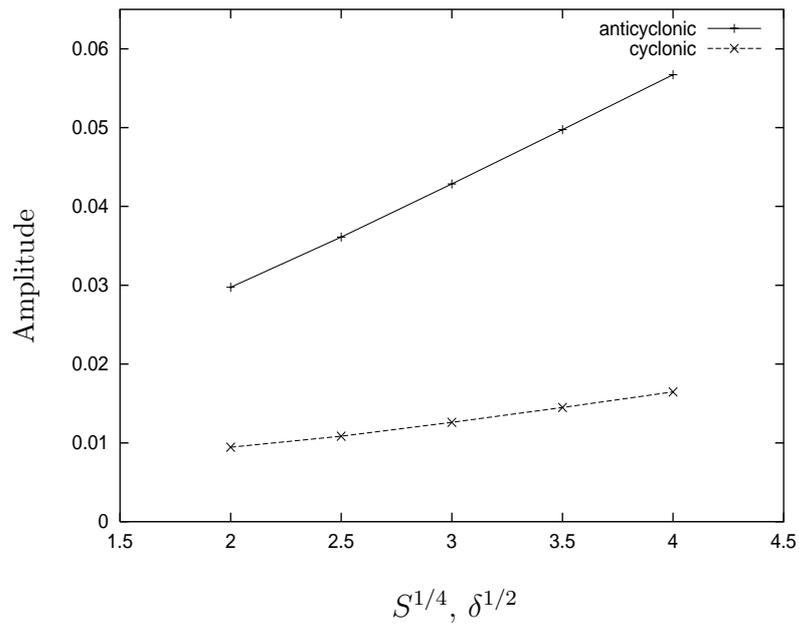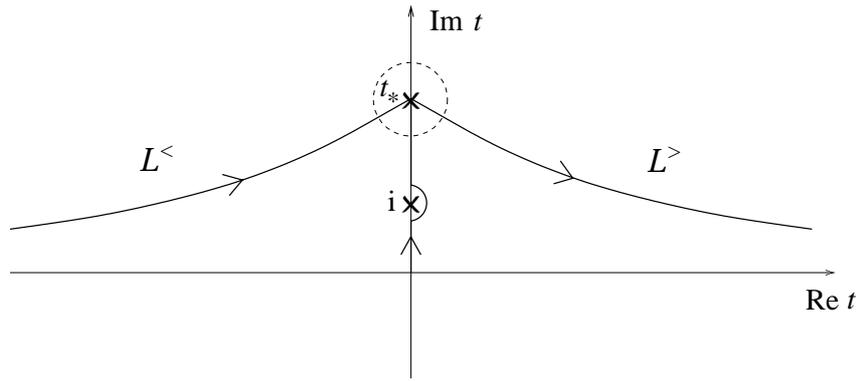
Figure 1:

Figure 2:

Figure 3:

Figure 4:

Figure 5:

Figure 6: