J. Fluid Mech. (2007), *vol.* 584, *pp.* 373–396.
doi:10.1017/S002211200700643X Printed in the United Kingdom

Unbalanced instabilities of rapidly rotating stratified shear flows

J. VANNESTE[1] AND I. YAVNEH[2]

[1]School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh EH9 3JZ, UK

[2]Department of Computer Science, Technion, Haifa 32000, Israel

(Received 6 November 2006 and in revised form 6 March 2007)

The linear stability of a rotating stratified inviscid horizontal plane Couette flow in a channel is studied in the limit of strong rotation and stratification. Two dimensionless parameters characterize the flow: the Rossby number ϵ, defined as the ratio of the shear to the Coriolis frequency and assumed small, and the ratio s of the Coriolis frequency to the buoyancy frequency, assumed to satisfy $s \leqslant 1$. An energy argument is used to show that unstable perturbations must have large, $O(\epsilon^{-1})$, wavenumbers. This motivates the use of a WKB-approach which, in the first instance, provides an approximation for the dispersion relation of the various waves that can propagate in the flow. These are Kelvin waves, trapped near the channel walls, and inertia–gravity waves with or without turning points.

Although the waves have real phase speeds to all algebraic orders in ϵ, we establish that the flow is unconditionally unstable. This is the result of linear resonances between waves with oppositely signed wave momenta. Three modes of instabilities are identified, corresponding to the resonance between (i) a pair of Kelvin waves, (ii) a Kelvin wave and an inertia–gravity wave, and (iii) a pair of inertia–gravity waves. Whilst all three modes of instability are active when the Couette flow is anticyclonic, mode (iii) is the only possible instability mechanism when the flow is cyclonic.

We derive asymptotic estimates for the instability growth rates. These are exponentially small in ϵ, i.e. of the form $\mathrm{Im}\,\omega = a\exp(-\Psi/\epsilon)$ for some positive constants a and Ψ. For the Kelvin-wave instabilities (i), we obtain analytic expressions for a and Ψ; the maximum growth rate, in particular, corresponds to $\Psi = 2$. For the other types of instabilities, we make the simplifying assumption $s \ll 1$ and find that the maximum growth rates correspond to $\Psi = 2.80$ for (ii) and $\Psi = \pi$ for (iii). The asymptotic results are confirmed by numerical computations. These reveal, in particular, that the instabilities (iii) have much smaller growth rates in cyclonic flows than in anticyclonic flows, even though $\Psi = \pi$ in both cases.

Our results highlight the limitations of the so-called balanced models, widely used in geophysical fluid dynamics, which filter out Kelvin and inertia–gravity waves and hence predict the stability of Couette flow. They are also relevant to the stability of Taylor–Couette flows and of astrophysical accretion disks.

1. Introduction

Rapid rotation and strong density stratification characterize the dynamics of geophysical fluids, the atmosphere and the oceans in particular. Two dimensionless numbers are used to measure the importance of these two effects relative to nonlinear

advection: the Rossby number

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

and the Froude number

$$F = \frac{U}{ND}.$$

Here U is a typical horizontal velocity, $f > 0$ is the Coriolis parameter, N the Brunt–Väisälä frequency, and L and D are typical horizontal and vertical length scales. With $N > f$, as is typically the case, the Rossby number estimates the maximum ratio between the typical frequency of the (slow) advective motion (given by U/L), and the frequency of inertia–gravity waves (bounded from below by f). Its smallness, that is, $\epsilon \ll 1$, has an important dynamical consequence, namely the weakness of the interaction between advective motion and inertia–gravity waves. This, together with the observation that inertia–gravity waves have generally weak amplitudes in the atmosphere and oceans, has led to the development – and success – of the so-called balanced models, which filter out inertia–gravity waves completely. These models describe only the slow large-scale dynamics, termed balanced because of its closeness to hydrostatic and geostrophic balance. They can be derived asymptotically, using power-series expansions in ϵ, and, in principle, can achieve an arbitrary algebraic accuracy $O(\epsilon^n)$ (e.g. Warn *et al.* 1995; Warn 1997).

To understand balanced dynamics and its limitations more fully, it is important to identify and quantify the phenomena that balanced models fail to capture. Of particular interest are those unbalanced phenomena which occur in spite of the smallness of ϵ and cannot be suppressed by balancing the initial data. In the present paper, we consider one such mechanism, namely the instability of balanced flows to unbalanced, gravity-wave-like perturbations. Since this type of instability is absent from balanced models of arbitrary high accuracy (which all have qualitatively similar stability conditions; see Ren & Shepherd 1997),† the growth rates can be expected to be $o(\epsilon^n)$ for all $n \geqslant 1$ or, in other words, to be beyond all orders in ϵ, and typically exponentially small in ϵ. Our results confirm this scaling and show that the instability bands, i.e. the range of unstable wavenumbers, are exponentially narrow.

We note that unbalanced instabilities like the ones examined in this paper are distinct from the mechanism of spontaneous generation of inertia–gravity waves studied in Vanneste & Yavneh (2004). The two types of mechanisms are exponentially weak in a different sense: in the case of instabilities, the growth rates are found to be exponentially small in ϵ, whereas in the case of spontaneous generation it is the amplitude of the waves generated that is exponentially small. This difference may not be essential, however: if the unbalanced instabilities saturate at a level that decreases to zero with the growth rate, as is typical, then the inertia–gravity waves they excite will also have exponentially small amplitudes. Another difference is that the instabilities require an initial unbalanced perturbation, whereas spontaneous generation occurs from entirely balanced initial conditions. We emphasize that both mechanisms provide

† We refer here to the class of balanced models that filter out Kelvin waves as well as inertia–gravity waves. With suitable boundary conditions, other balanced models, such as the semi-geostrophic model and Salmon's L1 model, retain Kelvin waves in spite of their large frequencies for $O(1)$ wavenumbers (Allen, Barth & Newberger 1990; Ren & Shepherd 1997). These models therefore represent, in a distorted way, some of the instabilities which we term unbalanced and treat in this paper, namely those associated with Kelvin-wave interactions (see Kushner, McIntyre & Shepherd 1998, and §4.1 below).

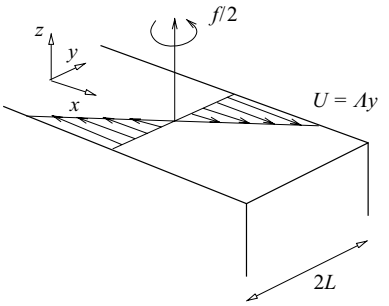

FIGURE 1. Schematic of the Couette flow, with velocity $(\Lambda y, 0, 0)$, of a three-dimensional Boussinesq fluid with constant Brunt–Väisälä frequency N. The domain rotates around the (vertical) z-axis at rate $f/2$; it is unbounded in the x- and z-directions, and bounded in the y-direction by two rigid walls separated by a distance $2L$.

potential sources of inertia–gravity waves in the atmosphere and oceans. What the exponential smallness indicates in both cases is that the effectiveness of these sources is highly sensitive to the Rossby number.

The specific flow whose stability we study is a horizontal Couette flow with velocity $(\Lambda y, 0, 0)$, modelled using the Boussinesq approximation with constant N. The geometry is that of a channel of width $2L$ on the f-plane (figure 1). A natural definition of a (signed) Rossby number for this flow is the ratio

$$\epsilon = \frac{\Lambda}{f}$$

of (minus) the basic-flow vorticity to the planetary vorticity. For $\epsilon > 0$ (< 0), the shear is anticyclonic (cyclonic). The other dimensionless parameter characterizing the flow can be taken to be the Prandtl ratio

$$s = \frac{f}{N}.$$

We restrict our attention to $s \leqslant 1$ and note that in the atmosphere and oceans $s \ll 1$ generally holds.

Because it is steady, the basic flow under consideration remains exactly balanced for all times, unlike generic time-dependent flows. Furthermore, it is stable in any balanced approximation, however accurate (assuming this filters out both inertia–gravity and Kelvin waves; see footnote on p. 374). This is because the shear is linear, and hence the potential vorticity is constant, whilst the only balanced instabilities possible in horizontal shears at small ϵ are inflectional instabilities, which require changes in the sign of the potential-vorticity gradient. Thus, with this flow, there are none of the difficulties in separating inertia–gravity waves from balanced motion that would appear for more complicated flows. The analysis reduces to a straightforward linear stability analysis, and the smallness of ϵ can be exploited to derive asymptotic results.

A number of workers have investigated shear-flow instabilities involving fast waves, either shallow-water surface waves or acoustic waves in compressible gas. These analyses were carried out mostly in the context of two-dimensional models, in either parallel or cylindrical geometry (Satomura 1981*a, b*, 1982; Narayan, Goldreich & Goodman 1987; Knessl & Keller 1992; Ford 1994; Balmforth 1996; Dritschel & Vanneste 2006), and of isentropic models (Papaloizou & Pringle 1987, and references

therein). The emphasis was not, however, on the small ϵ limit; indeed, in shallow water, flows with $\epsilon \ll 1$ and $F = O(\epsilon)$ are linearly stable as Ripa's theorem indicates (Ripa 1983). In contrast, the three-dimensional model examined here turns out to be always unstable, with growing modes whose horizontal and vertical wavenumbers scale as ϵ^{-1}. Our analysis has nevertheless many features in common with some of the works cited above, in particular the use of the WKB approximation. A common theme (in particular with Narayan *et al.* 1987) is also the interpretation of the instabilities in terms of (linear) resonances between modes with different signs of the conserved wave energy (or pseudoenergy) and wave momentum (or pseudomomentum) (see, e.g. Craik 1985; Ripa 1990, and references therein).

In the presence of lateral boundaries, as is the case here, there are two types of unbalanced modes: inertia–gravity waves, which are oscillatory in the cross-stream direction, and Kelvin waves, which are trapped at each boundary. Instabilities involving the resonance of Kelvin waves have been studied by Kushner *et al.* (1998) for the model considered here, and by Yavneh, McWilliams & Molemaker (2001) and Molemaker, McWilliams & Yavneh (2001) in the annular geometry of the (stratified) Taylor–Couette flow (see also Rüdiger, Arlt & Shalybkov (2002) and Dubrulle *et al.* (2005) for astrophysical applications). These instabilities occur only for anticyclonic shears ($\Lambda > 0$). (Kelvin waves propagate with the boundary to their right and so can only be brought into resonance by anticyclonic flows.) Yavneh *et al.* (2001) and Molemaker *et al.* (2001) also identified other modes of instability in anticyclonic shears. These can be associated with the resonance between Kelvin and inertia–gravity waves, and between inertia–gravity waves. The first mechanism is analogous to the mixed-mode instabilities examined by Sakai (1989), McWilliams, Molemaker & Yavneh (2004), Molemaker, McWilliams & Yavneh (2005) and Plougonven, Muraki & Snyder (2005) in a variety of contexts. As we show, the second mechanism is also active in cyclonic shears ($\Lambda < 0$). Thus, we establish that the rotating stratified horizontal Couette flow is unconditionally unstable.

We provide both a qualitative description of the instabilities, which we interpret in terms of the resonance between modes with different momentum signatures, and quantitative results based on the WKB approximation and on numerical computations. For all the instabilities examined, the growth rates are exponentially small in ϵ because the resonant modes are localized exponentially on different sides of the channel.

The remainder of this paper is organized as follows. The linearized equations of motion governing the evolution of perturbations in Couette flow are introduced in § 2. The conservation laws for wave momentum and wave energy are also introduced there. The latter conservation law is used to show that the horizontal and vertical wavenumbers of growing perturbations must be $O(\epsilon^{-1})$ or larger. This motivates the WKB approach developed in §§ 3–4. In § 3, we formulate the eigenvalue problem for the normal modes of the system, then provide an approximate solution using a WKB expansion (§ 3.1). To all orders in ϵ, this leads to purely real eigenfrequencies or, in other words, to waves rather than growing modes. Instabilities with growth rates beyond all orders in ϵ are, however, possible, and we go on to show that they do occur. Focusing on the modes susceptible to instabilities, we give some details of the dispersion relation and structure of Kelvin waves (§ 3.2) and of inertia–gravity waves with turning points (§ 3.3). We then use arguments based on wave-momentum signature to show that the linear resonance between waves does lead to instabilities for both cyclonic and anticyclonic shears (§ 3.5). Section 4 is devoted to the estimation of the instability growth rates. A detailed asymptotic estimate for

Kelvin-wave instabilities, extending those of Kushner *et al.* (1998) and Yavneh *et al.* (2001), is derived in § 4.1. Rough estimates (focusing on the exponential dependence and ignoring order-one prefactors) are then obtained for the instabilities associated with the resonance between a Kelvin wave and an inertia–gravity wave (§ 4.2), and with the resonance between two inertia–gravity waves (§ 4.3). These estimates are confirmed by the numerical solutions of the eigenvalue problem presented in § 4.4. The paper concludes with a discussion in § 5.

2. Model

We consider small-amplitude perturbations to the Couette flow described in § 1 and in figure 1. The corresponding linearized equations of motion can be written as

$$\mathrm{D}_t u - (f - \Lambda) v = -\partial_x p, \tag{2.1}$$

$$\mathrm{D}_t v + f u = -\partial_y p, \tag{2.2}$$

$$\mathrm{D}_t w + \rho = -\partial_z p, \tag{2.3}$$

$$\mathrm{D}_t \rho - N^2 w = 0, \tag{2.4}$$

$$\partial_x u + \partial_y v + \partial_z w = 0, \tag{2.5}$$

where (u, v, w) are the components of the velocity perturbation, p is the pressure perturbation, ρ is the density perturbation scaled by ρ_0/g, with ρ_0 the constant reference density (that is, ρ is the negative of the buoyancy perturbation), and $\mathrm{D}_t = \partial_t + \Lambda y \partial_x$. The material conservation

$$\mathrm{D}_t q = 0$$

of the perturbation potential vorticity

$$q = (f - \Lambda)\partial_z \rho - N^2(\partial_x v - \partial_y u)$$

follows readily. We restrict our attention to perturbations with vanishing potential vorticity, $q = 0$, since this is a characteristic of unbalanced motion. (See Vanneste & Yavneh (2004) for a study of the generation of inertia–gravity waves from perturbations with $q \neq 0$.) With this restriction, the conservations of the wave energy (pseudoenergy)

$$\mathscr{E} = \iiint \left(\frac{|\boldsymbol{u}|^2}{2} + \frac{\rho^2}{2N^2} + \Lambda y \frac{u\partial_z \rho - w \partial_x \rho}{N^2} \right) \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \tag{2.6}$$

and of the wave momentum (pseudomomentum)

$$\mathscr{M} = \iiint \frac{u\partial_z \rho - w \partial_x \rho}{N^2} \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z \tag{2.7}$$

are readily derived, as detailed in Appendix A.

The conservation of $\mathscr{E}$ constrains the structure of unstable perturbations. Specifically, exponentially growing modes must have vanishing $\mathscr{E}$, for otherwise $\mathscr{E}$ would also grow exponentially (see, e.g. Ripa 1990). Completing the squares in (2.6), we rewrite $\mathscr{E}$ as

$$\mathscr{E} = \frac{1}{2} \iiint \left[\left(u + \frac{\Lambda y \partial_z \rho}{N^2} \right)^2 + v^2 + \left(w - \frac{\Lambda y \partial_x \rho}{N^2} \right)^2 + \frac{\rho^2}{N^2} - \frac{\Lambda^2 y^2}{N^4} \left((\partial_x \rho)^2 + (\partial_z \rho)^2 \right) \right] \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z.$$

Clearly, then, instability can occur only if the perturbation satisfies

$$\Lambda^2 y^2[(\partial_x \rho)^2 + (\partial_z \rho)^2] > \rho^2 N^2$$

somewhere in the channel. In terms of horizontal and vertical wavenumbers k and m, this gives the necessary condition for instability

$$\frac{N}{(k^2+m^2)^{1/2}} < \Lambda L. \tag{2.8}$$

This can be recognized as a supersonic condition: instability can only occur for modes whose phase speed is less than the maximum basic-flow velocity. With $s \leqslant 1$ as assumed, the supersonic condition implies that $L(k^2+m^2)^{1/2} \geqslant \epsilon^{-1}$, and therefore that modes involved in instabilities have asymptotically large wavenumbers. One interpretation of this result states that the Rossby number based on the wave scale, that is, $\Lambda L(k^2+m^2)^{1/2}/f$, is greater than unity for unstable modes.

We note that for the shallow-water model with depth H, the condition analogous to (2.8) is $(gH)^{1/2} < \Lambda L$ and is independent of the scale of the perturbation (Ripa 1983). It is never satisfied for sufficiently small Λ and thus, for order-one Burger number $(\epsilon/F)^2 = \epsilon^2 gH/(\Lambda^2 L^2)$, for sufficiently small ϵ. Thus the shallow-water analogue of our model is linearly stable in the limit $\epsilon \to 0$. This is an important difference between shallow-water and three-dimensional Boussinesq dynamics, related to the difference in dispersion relations, with phase speeds that are bounded from below (by $(gH)^{1/2}$) for the shallow-water model, but arbitrarily small (for $k, m \to \infty$ with k/m fixed) for the Boussinesq model.

3. Normal modes

Let us now consider normal-mode solutions of the linearized equations of motion (2.1)–(2.5). The supersonic condition (2.8) suggests that the wavenumbers k and m should be rescaled by ϵ. We therefore write the dependent variables in the form

$$u(x, y, z, t) = \hat{u}(y/L)\exp[\mathrm{i}\epsilon^{-1}L^{-1}(kx + mz/s) - f\omega t], \tag{3.1}$$

with similar expressions for v, w, p and ρ. Here we have redefined k, m and ω to be dimensionless wavenumbers and frequency, with their dimensional counterparts given by $k/(\epsilon L)$, $m/(\epsilon s L)$ and $f\omega$, respectively. Without loss of generality, we assume that $k > 0$. Note that the non-dimensionalization then implies that modes with $\omega > 0$ ($\omega < 0$) propagate to the right (left) in anticyclonic shear and to the left (right) in cyclonic shear.

In terms of the dimensionless k and m, the supersonic condition (2.8) is

$$r = (s^2k^2 + m^2)^{1/2} > 1. \tag{3.2}$$

Introducing the normal modes (3.1) into (2.1)–(2.5) leads to a system of ordinary differential equations for $\hat{u}$, $\hat{v}$, $\hat{w}$, $\hat{p}$ and $\hat{\rho}$. These independent variables can be eliminated in favour of $\hat{p}$, leading in particular to

$$\hat{u} = \frac{1}{\epsilon f L}\frac{\epsilon(1-\epsilon)\hat{p}' + k\hat{\omega}\hat{p}}{\hat{\omega}^2 - 1 + \epsilon}, \quad \hat{\rho} = -\frac{\mathrm{i}N}{\epsilon f L}\frac{m}{1 - s^2\hat{\omega}^2}\hat{p}, \quad \hat{w} = -\frac{1}{\epsilon N L}\frac{m\hat{\omega}}{1 - s^2\hat{\omega}^2}\hat{p}, \tag{3.3}$$

where a prime denotes differentiation with respect to the dimensionless variable y/L which we hereinafter denote simply by y. A second-order differential equation for $\hat{p}$, already obtained by Kushner *et al.* (1998), then follows. Denoting $\hat{p}$ by p for

simplicity, this equation is

$$\epsilon^2 p'' - \frac{2\epsilon^2 k\hat{\omega}}{1-\hat{\omega}^2-\epsilon} p' - \left(k^2 \frac{1-\hat{\omega}^2+\epsilon}{1-\hat{\omega}^2-\epsilon} + m^2 \frac{1-\hat{\omega}^2-\epsilon}{1-s^2\hat{\omega}^2} \right) p = 0, \tag{3.4}$$

where

$$\hat{\omega} = \omega - ky.$$

It is supplemented by the boundary conditions $\hat{v} = 0$, that is,

$$\epsilon \hat{c} p' + p = 0 \quad \text{at} \quad y = \pm 1, \tag{3.5}$$

where $\hat{c} = c - y = \omega/k - y$. Note that the singularities of (3.4) for $\hat{\omega}^2 = 1 - \epsilon$ are removable: in particular, they are absent from the equation for $\hat{u}$ equivalent to (3.4) and given in Appendix B.

3.1. *WKB approximation*

Together, (3.4) and (3.5) constitute an eigenvalue problem from which the dispersion relation giving ω as a function of k and m can be derived. Taking advantage of the small parameter ϵ, this eigenvalue problem can be solved approximately using the WKB method. To this end, we first expand (3.4) in powers of ϵ, introducing the expansion

$$\omega = \omega_0 + \epsilon\omega_1 + \cdots$$

of the frequency. All the coefficients ω_n, $n = 0, 1, \ldots$ turn out to be real. Taking into account that $p' = O(\epsilon^{-1})$ and $p'' = O(\epsilon^{-2})$, we rewrite (3.4) as

$$\epsilon^2 p'' - \lambda^2 p - \frac{2\epsilon^2 k\hat{\omega}_0}{1-\hat{\omega}_0^2} p' + \epsilon h p = O(\epsilon^2), \tag{3.6}$$

where

$$\lambda^2 = k^2 + m^2 \frac{1-\hat{\omega}_0^2}{1-s^2\hat{\omega}_0^2}, \tag{3.7}$$

and

$$h = \frac{2k^2}{\hat{\omega}_0^2 - 1} + \frac{m^2}{1-s^2\hat{\omega}_0^2} + 2\omega_1 \frac{m^2(1-s^2)\hat{\omega}_0}{(1-s^2\hat{\omega}_0^2)^2}.$$

We introduce solutions of the form

$$p = (g_\pm + \epsilon g_{1\pm} + \cdots) \exp\left[\pm|\epsilon|^{-1} \int^y \lambda(y')\, \mathrm{d}y' \right] \tag{3.8}$$

into (3.6) and find that $g_\pm$ satisfies

$$\frac{g_\pm'}{g_\pm} = -\frac{\lambda'}{2\lambda} + \frac{k\hat{\omega}_0}{1-\hat{\omega}_0^2} \mp \sigma \frac{h}{2\lambda}, \tag{3.9}$$

where $\sigma = \operatorname{sgn}\epsilon$ is equal to $+1$ for an anticyclonic shear and -1 for a cyclonic shear. The solution can be written as

$$g_\pm = A \left(\frac{1-\hat{\omega}_0^2}{\lambda} \right)^{1/2} \exp\left[\mp\sigma \int^y \frac{h(y')}{2\lambda(y')} \mathrm{d}y' \right], \tag{3.10}$$

where A is an arbitrary complex constant. Note that this solution is single-valued near $\hat{\omega}_0 = \pm 1$, consistent with the observation that the singularities of (3.4) for $\hat{\omega}^2 = 1 - \epsilon$ are removable: the multi-valuedness caused by the square-root factor in (3.10) is cancelled by that of the integral in the argument of the exponential.

We can classify the solutions (3.8) according to the sign of λ^2 in the channel and distinguish:

Kelvin waves (KWs) for which $\lambda^2 > 0$ for $-1 \leqslant y \leqslant 1$. These modes are trapped exponentially near one of the boundaries, with $O(\epsilon)$ trapping scale.

inertia–gravity waves (IGWs) which satisfy $\lambda^2 < 0$ in at least part of the channel. There they have an oscillatory structure with $O(\epsilon)$ wavelength.

We now derive approximate dispersion relations for both types of waves. Together with information on the signature of their wave momentum discussed in § 3.5, these allow the prediction of instabilities associated with KW–KW, KW–IGW and IGW–IGW resonances. Asymptotic estimates for the growth rates of these instabilities are derived in § 4, where they are compared with numerical results.

3.2. *Kelvin waves*

We first consider WKB solutions to (3.4) for which $\lambda^2 > 0$. Two independent solutions can be written as

$$p_- = \left(\frac{1-\hat{\omega}_0^2}{\lambda}\right)^{1/2} \exp\left\{-|\epsilon|^{-1}\int_{-1}^{y}\left[\lambda(y') - \epsilon\frac{h(y')}{2\lambda(y')}\right]\mathrm{d}y'\right\}[1+O(\epsilon)], \quad (3.11)$$

$$p_+ = \left(\frac{1-\hat{\omega}_0^2}{\lambda}\right)^{1/2} \exp\left\{-|\epsilon|^{-1}\int_{y}^{1}\left[\lambda(y') - \epsilon\frac{h(y')}{2\lambda(y')}\right]\mathrm{d}y'\right\}[1+O(\epsilon)]. \quad (3.12)$$

The dispersion relation is found from the boundary conditions in the form

$$\begin{vmatrix} \epsilon\hat{c}(-1)p'_-(-1) + p_-(-1) & \epsilon\hat{c}(-1)p'_+(-1) + p_+(-1) \\ \epsilon\hat{c}(1)p'_-(1) + p_-(1) & \epsilon\hat{c}(1)p'_+(1) + p_+(1) \end{vmatrix} = 0. \quad (3.13)$$

Since the off-diagonal terms are exponentially small, the dispersion relation factorizes to all orders into two branches corresponding to KWs trapped at each boundary. We denote by $\mathrm{KW}_\pm$ the branch trapped at $y = \pm 1$, respectively; the corresponding frequency satisfies

$$\epsilon\hat{c}(1)p'_+(1) + p_+(1) = 0, \quad \epsilon\hat{c}(-1)p'_-(-1) + p_-(-1) = 0. \quad (3.14)$$

At leading order in ϵ, these two relations reduce to

$$\sigma\hat{c}_0(1)\lambda(1) + 1 = 0, \quad -\sigma\hat{c}_0(-1)\lambda(-1) + 1 = 0,$$

with solutions

$$\hat{c}_0(1) = -\frac{\sigma}{r}, \quad \hat{c}_0(-1) = \frac{\sigma}{r}. \quad (3.15)$$

(In addition, there are spurious solutions $k\hat{c}_0 = \pm\sigma$.) Thus, the KWs localized near $y = \pm 1$ have the leading-order dispersion relation

$$c_0 = 1 - \frac{\sigma}{r}, \quad c_0 = -1 + \frac{\sigma}{r}. \quad (3.16)$$

Higher-order approximations for the KW dispersion relation can be obtained by pursuing the expansion in powers of ϵ, each leading to a purely real correction to (3.15).

3.3. *Inertia–gravity waves*

In the region where the IGW is oscillatory, two independent WKB solutions (3.8) can be written as

$$p = \left(\frac{1-\hat{\omega}_0^2}{\ell}\right)^{1/2} \exp\left\{\pm\mathrm{i}|\epsilon|^{-1}\int^{y}\left[\ell(y') + \epsilon\frac{h(y')}{2\ell(y')}\,\mathrm{d}y'\right]\right\}, \quad (3.17)$$

where $\ell > 0$ is defined by

$$\ell^2 = -\lambda^2.$$

Depending on the value of ω, IGWs can have at most two turning points, i.e. points where $\lambda = \ell = 0$, in the channel. These are located at

$$y_\pm = c_0 \pm \frac{1}{r}(1+\delta^2)^{1/2} \quad \text{where} \quad \delta = m/k, \tag{3.18}$$

on either side of the 'critical level' $y = c_0$ where $\hat{\omega}_0 = 0$. The mode structure is then oscillatory for $y < y_-$ and $y > y_+$, and exponential for $y_- < y < y_+$. Here, we concentrate on modes with at least one turning point since, as argued in § 3.5 below, the presence of a turning point is necessary for instability. These IGWs are localized on one side of the channel and exponentially small on the opposite boundary.

Let us consider one such IGW that is decaying exponentially with y in $[y_-, y_+]$ and denote the corresponding solution by p_-. (Its counterpart, growing exponentially in $[y_-, y_+]$ and denoted by p_+, is readily deduced using the symmetry $(y, c) \mapsto (-y, -c)$.) In $[y_-, y_+]$, the solution p_- can be written as

$$p_- \sim A\left(\frac{1-\hat{\omega}_0^2}{\lambda}\right)^{1/2} \exp\left\{-|\epsilon|^{-1}\int_{y_-}^{y}\left[\lambda(y') - \epsilon\frac{h(y')}{2\lambda(y')}\right]\mathrm{d}y'\right\}. \tag{3.19}$$

The boundary condition (3.5) at $y = 1$ is satisfied automatically to all orders in ϵ. The form (3.19) breaks down in an $\epsilon^{2/3}$ neighbourhood of y_-, where it is replaced by an Airy function Ai. In $[-1, y_-]$, the solution is given by a linear combination of the two solutions in (3.17). The connection formula, which relates the two arbitrary constants to A and is found by matching with the Airy function, gives (cf. Bender & Orszag 1999, equation (10.4.16))

$$p_- \sim 2A\left(\frac{1-\hat{\omega}_0^2}{\ell}\right)^{1/2} \sin\left\{|\epsilon|^{-1}\int_{y}^{y_-}\left[\ell(y') + \epsilon\frac{h(y')}{2\ell(y')}\right]\mathrm{d}y' + \frac{\pi}{4}\right\}. \tag{3.20}$$

The dispersion relation is then found by applying the boundary condition (3.14) at $y = -1$, leading to

$$-\sigma\hat{c}(-1)\ell(-1)\cos S(y_-) + \sin S(y_-) = O(\epsilon), \tag{3.21}$$

where

$$S(y_-) = |\epsilon|^{-1}\int_{-1}^{y_-}\left[\ell(y') + \epsilon\frac{h(y')}{2\ell(y')}\right]\mathrm{d}y' + \frac{\pi}{4}.$$

Solving for $S(y_-)$, we find

$$S(y_-) = n\pi + \tan^{-1}[\sigma\hat{c}(-1)\ell(-1)] + O(\epsilon), \tag{3.22}$$

where n is an integer. At leading order this gives

$$\int_{-1}^{y_-}\ell(y)\,\mathrm{d}y = (n - 1/4)\pi|\epsilon|, \tag{3.23}$$

which determines c_0 implicitly. The next-order relation determines c_1. To obtain (3.23), we have assumed that $n|\epsilon| = O(1)$ so that the term in $\tan^{-1}$ can be neglected against the other two terms in (3.22). Equation (3.23) may, however, be argued to remain useful for $n|\epsilon| \ll 1$ since the balance of terms in (3.22) remains the same in this regime as in the regime $n|\epsilon| = O(1)$. To see this, note that the boundary condition for $\epsilon \to 0$ implies that $y_- \to -1$. Therefore, $\ell(-1) \to \ell(y_-) = 0$ and the $\tan^{-1}$ term

in (3.22) remains negligible. The 1/4 correction to n, negligible for $n|\epsilon| = O(1)$, is of course important when $n|\epsilon| \ll 1$.

In principle, however, our derivation of (3.23) is not valid in the limit $n|\epsilon| \to 0$. Indeed, as the turning point $y_- \to -1$, the surrounding $O(\epsilon^{2/3})$ region in which the WKB solution is not valid overlaps with the boundary. The boundary condition should therefore be applied not to (3.20), but to the Airy-function continuation of (3.19). We have not attempted this cumbersome procedure since the numerical results of § 3.4 show that (3.23) provides a satisfactory approximation for $n|\epsilon| \ll 1$.

Let us write the dispersion relation (3.23) for c_0 in a convenient form. Define μ by

$$c_0 = -1 + \frac{(1+\delta^2)^{1/2}}{r}(\mu + 1),$$

where $\delta = m/k$, so that

$$y_- = -1 + \frac{(1+\delta^2)^{1/2}}{r}\mu.$$

The assumption that this turning point is inside the channel imposes the restriction $0 < \mu < 2r(1+\delta^2)^{-1/2}$. Introducing the integration variable Y, with $y = -1+(1+\delta^2)^{1/2}Y/r$, reduces (3.23) to the expression

$$\int_0^\mu \left[\frac{(Y-\mu-1)^2 - 1}{1-\nu^2(Y-\mu-1)^2}\right]^{1/2} \mathrm{d}Y = \frac{(s^2+\delta^2)^{1/2}}{1+\delta^2}(n-1/4)\pi|\epsilon|, \tag{3.24}$$

where

$$\nu^2 = \frac{s^2(1+\delta^2)}{s^2+\delta^2}.$$

This defines implicitly a function $\mu(\delta, s, n, |\epsilon|)$ with values in $[0, \nu^{-1} - 1]$, from which c_0 is deduced. Taking both the solution p_- and its symmetric p_+ into account, we find the two families of branches

$$c_0 = \pm 1 \mp \frac{(1+\delta^2)^{1/2}}{r}[\mu(\delta, s, n, |\epsilon|) + 1], \tag{3.25}$$

corresponding to modes exponentially small near $y = \mp 1$ and denoted by $\mathrm{IGW}_\pm$, respectively. Again, higher-order approximations to the phase velocity can, in principle, be computed, leading to real corrections to c_0 in powers of ϵ. Note that, at leading order in ϵ, the dispersion relation is the same for both signs of ϵ, that is, for both cyclonic and anticyclonic flows. An asymmetry only appears at higher order.

For $n = O(1)$, the form of (3.25) remains valid, even though μ may depend on the parameters in a more complicated manner (involving Airy functions) than given by (3.24). As $\epsilon \to 0$, $\mu \to 0$, and the leading-order dispersion relation reduces to

$$c_0 \sim \pm 1 \mp \frac{(1+\delta^2)^{1/2}}{r}, \tag{3.26}$$

corresponding to $y_\pm \to \pm 1$. This is a robust result, insensitive to the details of the boundary condition. It implies that the branches of the dispersion relation come closer together as $|\epsilon|$ decreases, accumulating on the curve (3.26). To estimate the distance between successive branches, we can examine the the small-μ behaviour of the integral on the left-hand side of (3.24). This leads to the approximation

$$\mu \sim \frac{1}{2}\left[\frac{3\delta(1-s^2)^{1/2}}{1+\delta^2}(n-1/4)\pi|\epsilon|\right]^{2/3},$$

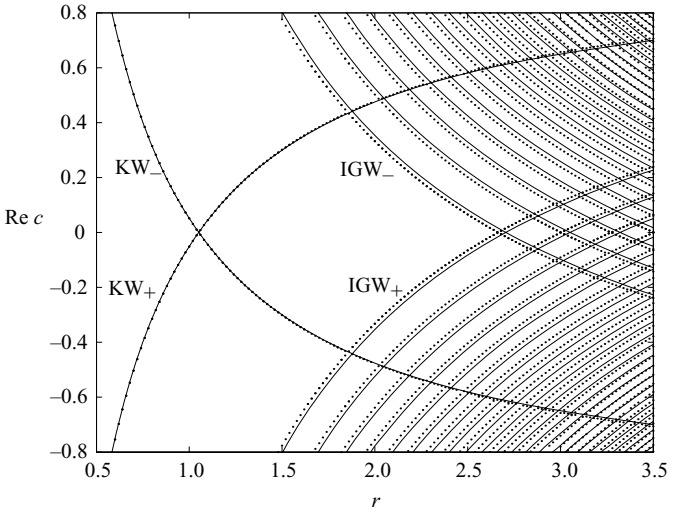

FIGURE 2. Dispersion relation for an anticyclonic flow, with $\epsilon = 0.1$, $\delta = 2$ and $s = 0.1$. The two Kelvin waves (KW) are shown along with many inertia–gravity waves (IGW). The asymptotic estimates for $|\epsilon| \ll 1$ (solid curves) are compared with numerical solutions (dotted curves).

which indicates that the branches $n = 1, 2, \ldots$ of the dispersion relation (3.25) are $O(\epsilon^{2/3})$ apart. This is consistent with the Airy function scaling which, as discussed above, is relevant to the boundary condition at $y = -1$ when $\epsilon \to 0$ for fixed n.

The asymptotic results (3.16) and (3.25) provide a first approximation to the dispersion relation of KWs and IGWs. We have extended this by solving the eigenvalue problem (3.4)–(3.5) (or rather the equivalent formulation (B 1)–(B 2) in terms of $\hat{u}$) numerically. Our numerical solver is the same as that used in Yavneh *et al.* (2001), employing a second-order finite-volume discretization of (B 1)–(B 2). For given physical parameters and wavenumbers, m and k, we search for eigenfrequencies for which the matrix representing the discretized system is singular. The codes are implemented in MATLAB, with the search performed using the *fminsearch* function that employs the Simplex algorithm.

3.4. *Dispersion relation*

Figures 2 and 3 show the dispersion relation for anticyclonic and cyclonic flows, respectively. The parameters have been chosen as $\epsilon = 0.1$, $\delta = 2$ and $s = 0.1$, but the qualitative features remain the same for a wide range of values. The numerical results (dotted curves) are compared with the asymptotic estimates (solid curves) to confirm the validity of the latter. For KWs, we have used an $O(\epsilon)$-accurate estimate which improves on (3.16) by adding the term $\epsilon c_1 = \mp\epsilon\sigma(2r)$ derived in Appendix C. For IGWs, we have used the estimate (3.25). The agreement between asymptotic and numerical results is good, including for the first few IGWs modes for which, as explained above, the asymptotics is not strictly valid. In the anticyclonic case, it is difficult to relate each asymptotic curve to its numerical counterpart visually for higher modes (say $n = 7, 8, \ldots$). This is because the distance between branches decreases rapidly with the mode number n, and becomes quickly of the same order as, or even larger than, the error in the asymptotic dispersion relation. This error remains nonetheless small, and consistent with the $O(\epsilon)$ terms neglected in the derivation. In

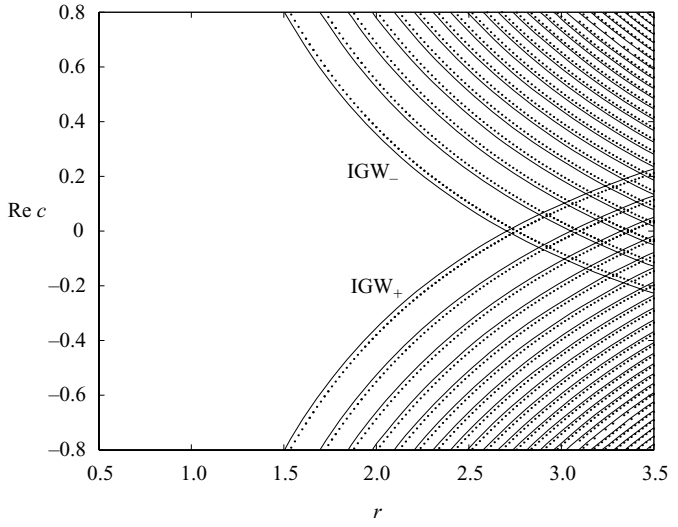

FIGURE 3. Dispersion relation for a cyclonic flow, with $\epsilon = -0.1$. The other parameters and the notation are the same as in figure 2.

the cyclonic case, the error of the asymptotic estimates remains smaller than the distance between branches even for higher modes. To understand the reason for this asymmetric asymmetric behaviour between anticyclonic and cyclonic shear would require us to continue the asymptotic developments to the next order in ϵ; this is a daunting task which we have not attempted for IGWs.

Figures 2 and 3 show multiple intersections between the branches $\mathrm{IGW}_\pm$ of the dispersion relation and hence demonstrate the existence of numerous linear resonances between these modes. In the anticyclonic case, there are additional intersections (and hence resonances) between the branches $\mathrm{KW}_\pm$, and between $\mathrm{KW}_\pm$ and $\mathrm{IGW}_\mp$. (The KWs do not appear in figure 3 for the cyclonic case because they have $|c| > 1$. They are not involved in linear resonances.) The simple intersections between branches of the dispersion relation observed in the figures are in fact spurious: they result from the finite resolution of the plot for the numerical results, and from the limited accuracy for the asymptotic ones. General arguments about linear resonances indicate that simple intersections should not be expected; rather, there are two possible generic behaviours: (i) mode conversion, when the phase velocities remain real and the two curves, rather than intersecting, locally form the two branches of a hyperbola; or (ii) instability, when the phase velocities on the two branches become complex conjugates with non-zero imaginary parts, forming an 'instability bubble' (cf. Howard & MacKay 1987). The two situations are distinguished by the signs of quadratic invariants, such as the wave momentum, along the colliding branches: generically, (i) mode conversion occurs when both signs are the same; and (ii) instability occurs when the signs differ (e.g. Cairns 1979). We now show that the latter situation is the relevant one in our problem by examining the sign of the wave momentum for KWs and IGWs in the WKB approximation.

3.5. *Wave-momentum signature*

IGWs and KWs have different leading-order approximations to their wave momentum. To see this, we introduce (3.1) and (3.3) into (2.7) and assume that

ω is real. This gives

$$\mathscr{M} = \frac{2N^2m^2}{f^2L^2\epsilon^3}\int\left[\frac{\epsilon(1-\epsilon)}{2(\hat{\omega}^2-1+\epsilon)(1-s^2\hat{\omega}^2)}\frac{\mathrm{d}|p|^2}{\mathrm{d}y}+\frac{k\hat{\omega}(1-s^2+s^2\epsilon)}{(\hat{\omega}^2-1+\epsilon)(1-s^2\hat{\omega}^2)^2}|p|^2\right]\mathrm{d}y$$
$$= \frac{2N^2m^2}{f^2L^2\epsilon^3}\hat{\mathscr{M}}, \tag{3.27}$$

where again we denote $\hat{p}$ simply by p. The last line defines the dimensionless wave momentum $\hat{\mathscr{M}}$ which we will use hereinafter. For IGWs, the first term is negligible: indeed, in the regions where p oscillates rapidly, $\mathrm{d}|p|^2/\mathrm{d}y = O(1)$, while in the possible regions where p decays exponentially, $\mathrm{d}|p|^2/\mathrm{d}y = O(\epsilon^{-1})$ only for a range of y of size ϵ; both types of region thus contribute at $O(\epsilon)$ to $\hat{\mathscr{M}}$. This leads to the leading-order approximation

$$\hat{\mathscr{M}} \sim \int \frac{k\hat{\omega}(1-s^2)}{(\hat{\omega}^2-1)(1-s^2\hat{\omega}^2)^2}|p|^2\,\mathrm{d}y \quad \text{for IGWs.} \tag{3.28}$$

Given that the denominator $(\hat{\omega}^2-1)$ cancels with the same factor in $|p|^2$ (see (3.8)–(3.10)), it is clear that instability involving IGWs implies that $\hat{\omega}$ changes sign. It follows that there is at least one turning point in the channel, as announced, since the absence of turning points ($\ell^2 > 0$) implies that $|c| > 1$. Assuming there are turning points, the sign of $\hat{\mathscr{M}}$ for the two types of IGW considered in § 3.3 is then

$$\hat{\mathscr{M}} < 0 \ \text{ for IGW}_+, \quad \hat{\mathscr{M}} > 0 \ \text{ for IGW}_-.$$

For KWs, the two terms in (3.27) have a similar, $O(\epsilon)$, order of magnitude. Using (3.8), we find that

$$\hat{\mathscr{M}}_\pm \sim \frac{|\epsilon|}{2[\hat{\omega}^2(\pm1)-1][1-s^2\hat{\omega}^2(\pm1)]}\left[\pm\sigma+\frac{k(1-s^2)}{\lambda(\pm1)[1-s^2\hat{\omega}^2(\pm1)]}\right]|p(\pm1)|^2.$$

Using the dispersion relation for Kelvin waves, $\hat{\omega}(\pm1) = \mp\sigma k/r + O(\epsilon)$ in our non-dimensionalization, and its consequence $\lambda(\pm1) = r$ (see (3.15) and (C 2*a*)), this reduces to

$$\hat{\mathscr{M}}_\pm \sim \mp\frac{\epsilon r^4}{2m^4}|p(\pm1)|^2 \quad \text{for KWs,}$$

leading to the following signs:

$$\hat{\mathscr{M}} \lessgtr 0 \ \text{ for KW}_+, \quad \hat{\mathscr{M}} \gtrless 0 \ \text{ for KW}_- \quad \text{when} \quad \epsilon \gtrless 0,$$

differing in the anticyclonic and cyclonic cases.

With the wave-momentum signatures just obtained, it is clear from figures 2 and 3 that the numerous intersections of branches correspond to waves with oppositely signed $\hat{\mathscr{M}}$. This establishes the existence of many modes of instability, both for anticyclonic and cyclonic shears. The main difference between the two configurations is that instabilities involving only KWs are possible only for anticyclonic shear.

All the instabilities are associated with the interactions of modes which are exponentially localized on different sides of the channel. Therefore, their interaction is exponentially weak and, as a consequence, the growth rates of the instabilities and range of unstable wavenumbers are exponentially small in ϵ, as anticipated in § 1. As the asymptotic calculations of the next section show, such small growth rates are somewhat delicate to capture analytically. However, the interpretation in terms of interactions of waves with oppositely signed $\hat{\mathscr{M}}$ makes it possible to predict instability

robustly, without detailed calculations. This prediction only makes an assumption of genericity for the linear resonances: there could be no instability in the unlikely event that the modes involved are completely uncoupled.

4. Instabilities

4.1. *KW–KW instabilities*

We start our study of the weak instabilities associated with mode interactions by deriving an estimate for the growth rate of the instability that arises through the resonance of KWs in anticyclonic shear. This instability has been examined in some detail by Kushner *et al.* (1998) and by Yavneh *et al.* (2001). Because it is the strongest instability, with physical relevance in Taylor–Couette and accretion disks (see Dubrulle *et al.* 2005), we present here a complete asymptotic derivation of the growth rate. For the KW–IGW and IGW–IGW instabilities considered in §§ 4.2–4.3, we limit the derivation to the exponential behaviour of the growth rate as $\epsilon \to 0$. The method we now describe could, however, be applied to these instabilities as well, should a more accurate estimate be required.

To obtain the growth of the instability, we must reconsider the dispersion relation (3.13) in the vicinity of the resonance point, taking into account exponentially small terms. Let $c_\star$ and $r_\star$ be the values of r and c at resonance. By symmetry, $c_\star = 0$. According to (3.16) (with $\sigma = 1$ corresponding to the anticyclonic shear),

$$r = r_\star = 1 + O(\epsilon).$$

Thus, resonance occurs on an ellipse with semi-axes $1/s$ and 1 in the (k, m)-plane, and the instability region is an exponentially small annulus around this ellipse. It is best parameterized using the polar coordinates (r, θ), with

$$sk = r\cos\theta, \qquad m = r\sin\theta.$$

Now, take

$$c = C, \quad r = r_\star + R,$$

where C and R are exponentially small. This can be introduced into the dispersion relation (3.13); using the fact that $(c = 0, r = r_\star)$ satisfy (3.14), a Taylor expansion leaves only terms that are exponentially small. In the coefficients of C and R in these terms, we can approximate $r_\star$ by its leading-order estimate 1. Noting that, in this approximation,

$$\lambda(\pm 1) \approx \lambda_\star(\pm 1) + \frac{(s^2 - \cos^2\theta)R \pm \cos^2\theta(1 - s^2)C}{s^2\sin^2\theta},$$

we find the dispersion relation in the form

$$\left(\frac{s^2 - \cos^2\theta}{s^2\sin^2\theta}\right)^2 (R^2 - C^2) = 4\mathrm{e}^{-2\Psi/\epsilon}, \tag{4.1}$$

where

$$\Psi = \int_{-1}^{1} \left[\lambda_\star(y) - \epsilon\frac{h_\star(y)}{2\lambda_\star(y)}\right] \mathrm{d}y,$$

and the subscript $\star$ indicates evaluation at the resonance point. A consistent approximation of Ψ requires us to include the $O(\epsilon)$ contribution to $\lambda_\star$ in the first term of the integrand. To this end, we compute the KW dispersion relation to $O(\epsilon)$

in Appendix C and find that $r_\star = 1 + \epsilon/2 + O(\epsilon^2)$. This leads to

$$\Psi = \Psi_0 + \epsilon\Psi_1,$$

where

$$\Psi_0 = \int_{-1}^{1} \lambda_0(y)\,\mathrm{d}y \tag{4.2}$$

with

$$\lambda_0 = \left[\frac{\cos^2\theta(1-y^2) + s^2\sin^2\theta}{s^2(1-\cos^2\theta y^2)}\right]^{1/2},$$

and

$$\Psi_1 = \frac{1}{2}\int_{-1}^{1}\left[\frac{\lambda_0(y)}{1-\cos^2\theta y^2} - \frac{\cos^2\theta y^2}{\lambda_0(y)s^2(1-\cos^2\theta y^2)} - \frac{h_0(y)}{\lambda_0(y)}\right]\mathrm{d}y, \tag{4.3}$$

with

$$h_0(y) = \frac{2\cos^2\theta}{\cos^2\theta y^2 - s^2} + \frac{\sin^2\theta}{1-\cos^2\theta y^2}.$$

The second integral has to be interpreted as a Cauchy principal value at the singularities $y = \pm s/\cos\theta$ of $h_0(y)$ when these are in $[-1, 1]$. With this result, the dispersion relation (4.1) can be rewritten as

$$C = [R^2 - \alpha^2 \mathrm{e}^{-2\Psi_0/\epsilon}]^{1/2}, \tag{4.4}$$

where

$$\alpha = \frac{2s^2\sin^2\theta\,\mathrm{e}^{-\Psi_1}}{s^2 - \cos^2\theta}.$$

Equation (4.4) is the first main result of this paper. It provides the leading-order asymptotics for the growth rate of the KW–KW instability (after multiplication by k) as $\epsilon \to 0$ and for arbitrary $s \leqslant 1$). It also makes evident the exponential smallness of the growth rate and of the instability-bandwidth. Its validity is confirmed in § 4.4 where it is compared with numerical results.

The minimum of Ψ_0, and hence the maximum growth rate, is attained for $\theta = \pi/2$, for which $\Psi_0 \sim 2$. Thus, at the crude level of exponential dependence on ϵ, we obtain the estimate

$$\log \mathrm{Im}\,\omega \sim -\frac{2}{\epsilon} \quad \text{as} \quad \epsilon \to 0, \tag{4.5}$$

for the largest growth rate $\mathrm{Im}\,\omega = k\mathrm{Im}\,C$. Note that because $\theta = \pi/2$ implies that $k = 0$ and hence $\omega = 0$, the maximum growth rate is, in fact, achieved for θ slightly less than $\pi/2$; however, this does not affect the exponential dependence in (4.5) (see below).

Estimates more precise than (4.5) can, of course, be inferred from (4.4). Focusing on the limit $\theta \to \pi/2$, we note that C depends on the relationship between s and θ. A distinguished limit is found for $s = O(\cos\theta) \ll 1$. This corresponds to the regime with $s \ll 1$ and $\delta = k/m = O(1)$, which we term the quasi-geostrophic regime, since it corresponds to the quasi-geostrophic scaling implying, in particular, the hydrostatic approximation (k/m can be recognized as the square root of the Burger number based on the wave scale). Taking the limit $\theta \to \pi/2$ of (4.2)–(4.4) with $k = s/\cos\theta$ fixed then yields

$$\Psi_0 \sim 1 + \frac{1+k^2}{k}\tan^{-1}k, \quad \alpha \sim 2\frac{|1-k|^{k-1}}{|1+k|^{k+1}}.$$

The maximum of the imaginary part of the phase speed is then obtained for $k \to 0$ and given by $\mathrm{Im}\, c \sim 2\exp(-2/\epsilon)$, consistent with Yavneh *et al.*'s (2001) equation (35). The maximum of the growth rate $\mathrm{Im}\,\omega$ is easily seen to be attained for $k = O(\epsilon^{1/2})$ and to be a factor $\epsilon^{1/2}$ smaller than the maximum of $\mathrm{Im}\, c$. In dimensional terms, this means that the horizontal and vertical scales are both small, but have different orders of magnitudes, scaling as $\epsilon^{1/2}$ and ϵ^1, respectively.

4.2. *KW–IGW instabilities*

The KW–IGW instabilities occur for anticyclonic flows through the resonance of an IGW, which has one turning point and is localized on one side of the channel, with a KW localized on the other side. To estimate their growth rates, we can consider a solution consisting of a linear combination of the IGW_- given by (3.19)–(3.20) which is oscillatory near $y = -1$, and the KW_+ given by (3.12). (The other combination, of IWG_+ with KW_-, has the same growth rate, by symmetry.) A calculation similar to that carried out for KW–KW instabilities could, in principle, be performed to obtain the leading-order behaviour of the growth rate. However, this requires the derivation of the IGW dispersion relation accurate to $O(\epsilon)$, involving an inordinate amount of calculation. We shall therefore limit ourselves to the determination of the exponential behaviour of $\mathrm{Im}\,\omega$ (that is, to the determination of the constant Ψ_0 such that $\log \mathrm{Im}\,\omega \sim -\Psi_0/\epsilon$ as $\epsilon \to 0$) in the instability regions, and ignore the order-one prefactor in the expression of $\mathrm{Im}\,\omega$.

As in the case of KW–KW instabilities, the two colliding modes are exponentially small on opposite boundaries, where their amplitude is of a similar order of magnitude. We determine the exponential behaviour of $\mathrm{Im}\,\omega$ by determining this amplitude, given explicitly by $\exp(-\Psi_0/\epsilon)$. Note that Ψ_0 controls not only the exponential smallness of the growth rate, but also that of the width of the instability bands.

For simplicity, we restrict our analysis to the quasi-geostrophic scaling $s \ll 1$, $\delta = O(1)$. For $s \ll 1$ and $\sigma = 1$, the phase speeds of colliding KW_+ and IGW_- branches given in (3.16) and (3.25) reduce at leading order to

$$c = 1 - \frac{1}{m}, \quad c = -1 + \left(\frac{1}{k^2} + \frac{1}{m^2}\right)^{1/2},$$

respectively. The corresponding resonance condition

$$\frac{1}{m} + \left(\frac{1}{k^2} + \frac{1}{m^2}\right)^{1/2} = 2,$$

that is,

$$k = \frac{1}{2}\left(\frac{m}{m-1}\right)^{1/2} \quad \text{with} \quad m > 1 \tag{4.6}$$

defines a curve in the (k, m)-plane in the vicinity of which instabilities are concentrated. For KW–IGW instabilities, since there is a single turning point y_- in the channel, Ψ_0 is given as

$$\Psi_0 = \int_{y_-}^{1} \lambda(y)\, \mathrm{d}y. \tag{4.7}$$

For $s \ll 1$, the integrand $\lambda(y)$, given in (3.7), can be approximated by

$$\lambda(y) = km[(y_+ - y)(y - y_-)]^{1/2}, \tag{4.8}$$

with $y_\pm$ reducing to

$$y_\pm = c \pm \left(\frac{1}{k^2} + \frac{1}{m^2}\right)^{1/2} = \begin{cases} 3 - 2/m, \\ -1. \end{cases} . \tag{4.9}$$

Introducing (4.6) and (4.8)–(4.9) into (4.7) gives the expression

$$\Psi_0 = \frac{1}{8[m(m-1)]^{1/2}} \left[(2m-1)^2 \left(\pi + 2\sin^{-1}\frac{1}{2m-1}\right) + 4(m(m-1))^{1/2}\right]. \tag{4.10}$$

The maximum growth rate of the KW–IGW instability is given by the minimum value of Ψ_0, found numerically from (4.10) to be

$$\Psi_0 = 2.80\ldots \quad \text{for} \quad k = 1.04\ldots, \quad m = 1.30\ldots. \tag{4.11}$$

Thus, we obtain the asymptotics

$$\log \operatorname{Im} \omega \sim -\frac{2.80}{\epsilon} \quad \text{as} \quad \epsilon \to 0, \tag{4.12}$$

for the growth rate of KW–IGW instabilities. Comparison with (4.5) then indicates that these are considerably weaker than the KW–KW instabilities.

4.3. *IGW–IGW instabilities*

We now consider the instabilities that result from the resonance between IGWs. These are particularly important for cyclonic flows since they provide the only mode of instability in this case. In fact, as can be expected from the leading-order dispersion relation (3.25), the dominant behaviour of these instabilities is unaffected by the sign of the shear, so that the exponential dependence on $1/\epsilon$ is identical for anticyclonic and cyclonic shears. What differs between the two cases, however, is the order-one prefactor which we do not estimate analytically.

IGW–IGW instabilities occur when a solution p_- of the form (3.19)–(3.20) is resonant with its counterpart p_+. The modes have then two turning points $y_\pm$ in the channel, leading to the necessary condition $r \geqslant (1+\delta^2)^{1/2}$ for the instability. We now estimate the factor Ψ_0 controlling the exponential smallness of the instability growth rates. As in the previous section, we restrict our attention to the quasi-geostrophic scaling $s \ll 1$ and $\delta = O(1)$. We furthermore consider only the instability with the largest growth rate, associated with the (symmetric) resonance of the gravest ($n = 1$) IGW modes, and for which $c = 0$ to all orders in ϵ. The resonance condition is therefore

$$\frac{1}{k^2} + \frac{1}{m^2} = 1.$$

Since for $n = O(1)$, the two turning points are $y_\pm = \pm 1$ at leading order in ϵ, Ψ_0 is computed as

$$\Psi_0 = km \int_{y_-}^{y_+} [(y_+ - y)(y - y_-)]^{1/2}\, \mathrm{d}y = \frac{\pi k m}{2}$$

using (4.8). The minimum value is therefore

$$\Psi_0 = \pi \quad \text{for} \quad k = m = \sqrt{2}, \tag{4.13}$$

and the exponential scaling of the growth rate given by

$$\log \operatorname{Im} \omega \sim -\frac{\pi}{\epsilon} \quad \text{as} \quad \epsilon \to 0, \tag{4.14}$$

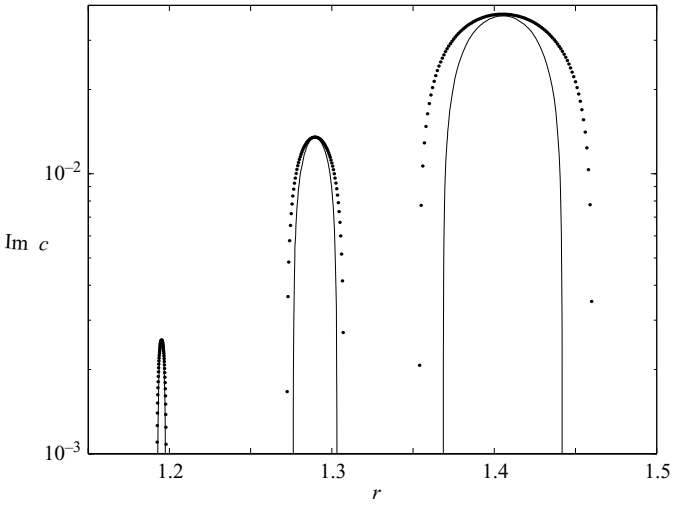

FIGURE 4. Imaginary part of the phase speed $\mathrm{Im}\, c$ as a function of r (in linear–logarithmic scale) for the KW–KW instability in anticyclonic flows with $s = 0.1$. Numerical (dots) and asymptotic (solid curves) results are compared for $\epsilon = 0.3, 0.4, 0.5$ (with increasing $\mathrm{Im}\, c$), and $\theta = \pi/2$ (i.e. $k = 0$ and $m = r$).

for both anticyclonic and cyclonic flows. This is smaller than the growth rates for the KW–KW and the KW–IGW instabilities (4.5) and (4.12) by the exponentially small factors $\exp(-1.14/\epsilon)$ and $\exp(-0.34/\epsilon)$, respectively.

4.4. *Numerical computation of growth rates*

We now present comparisons of the growth rate, or rather $\mathrm{Im}\, c$, computed numerically with the asymptotic results of §§4.1–4.3. The numerical method employed is that described in §3.4 where $\mathrm{Re}\, c$ was considered. For the small values of ϵ examined here, $\mathrm{Im}\, c$ is small and the bands of unstable wavenumbers are narrow, so that fine resolution in y is required to capture $\mathrm{Im}\, c$ accurately. In order to ensure high accuracy, we successively double the grid resolution until results are unchanged to at least four significant digits. This required grids of sizes ranging from about 250 mesh points for strong or moderate instabilities, to as many as 16 000 mesh points for very weak instabilities. This may be improved upon by using non-uniform grids with high resolution only in regions where the solution changes fast. The search for the bands of instabilities in (k, m) is delicate, but made possible by the excellent approximations afforded by the asymptotic results.

We start by considering the KW–KW instability of anticyclonic flows. Figure 4 shows $\mathrm{Im}\, c$ as a function of r for $\theta = \pi/2$ and $\epsilon = 0.3, 0.4, 0.5$ in the instability bands. The dots represent numerically computed values; the solid line are computed analytically using (4.4). Note that we only have an estimate for $r_\star$ to the algebraic accuracy $O(\epsilon)$, while the bands are exponentially narrow. Hence, we use the numerical results for determining $r_\star$, taking it as the value of r for which $\mathrm{Im}\, c$ is maximized. The narrowing of the instability bands with ϵ is clearly exhibited in the figure, and the small-ϵ analytical approximation quickly converges to the numerical results as ϵ becomes small. The dependence of $\mathrm{Im}\, c$ on θ is illustrated by figure 5 which compares numerical and asymptotic estimates for the maximum value of $\mathrm{Im}\, c$ as a function of θ for $s = 0.1$ and $\epsilon = 0.3$, 0.4 and 0.5. The value of $\mathrm{Im}\, c$ in the quasi-geostrophic scaling $s \ll 1$, $\delta = O(1)$, that is, the limit $\theta \to \pi/2$, is also indicated. The figure

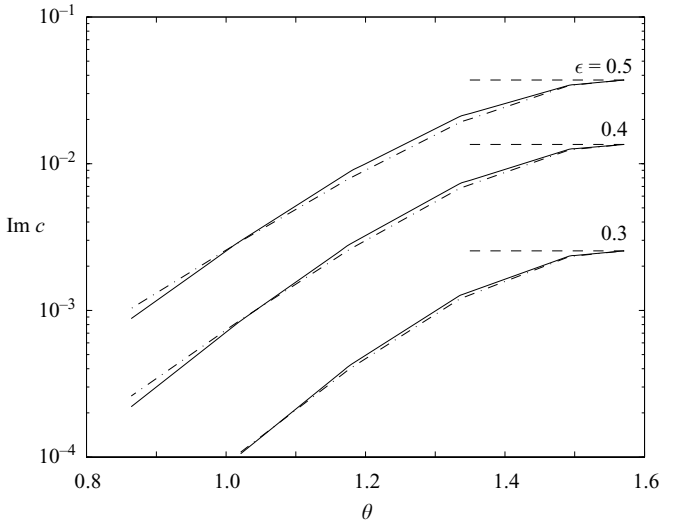

FIGURE 5. Maximum of Im c as a function of θ (in linear–logarithmic scale) for the KW–KW instability of anticyclonic flows with $s = 0.1$. Numerical (dash-dotted curves) and asymptotic (solid curves) results are compared for $\epsilon = 0.3, 0.4, 0.5$. The limits of Im c as $\theta \to 0$, corresponding to the quasi-geostrophic (dashed) scaling $s \ll 1$ and $\delta = O(1)$ are also indicated.

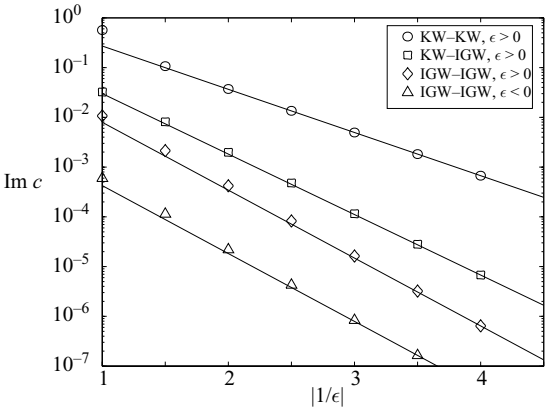

FIGURE 6. Maximum of Im c as a function of $1/|\epsilon|$ (in linear–logarithmic coordinates) for all the instability mechanisms examined in this paper, both for anticyclonic ($\epsilon > 0$) and cyclonic ($\epsilon < 0$) flows. The asymptotic estimates (solid lines) are compared with numerical results (symbols) for $s = 0.1$.

confirms the accuracy of the asymptotic estimate and shows the rapid decrease of Im c as θ decreases from $\pi/2$.

Our results for all the types of instability are summarized in figure 6. This compares asymptotic and numerically computed values of Im c as a function of $1/|\epsilon|$ for KW–KW, KW–IGW and IGW–IGW instabilities in anticyclonic flows, and IGW–IGW instabilities in cyclonic flows. The values of Im c displayed correspond to the maximum over m and k for fixed $s = 0.1$. For KW–KW instabilities, the asymptotic estimates are obtained from (4.4). For KW–GW and IGW–IGW instabilities, we use (4.12) and (4.14), respectively. These give Im c only up to a multiplicative constant which we

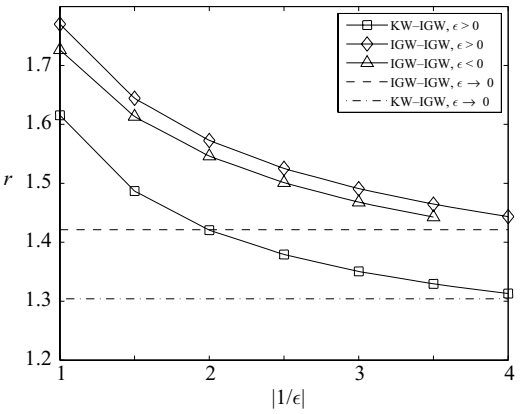

FIGURE 7. Numerical estimates of the value of $r = (k^2 + m^2)^{1/2}$ which maximizes $\mathrm{Im}\, c$ for KW–IGW and IGW–IGW instabilities as a function of $1/|\epsilon|$ for $s = 0.1$. The asymptotic limit for $\epsilon \to 0$ obtained from (4.11) and (4.13) are also shown.

fix by matching the asymptotic and numerical results for the smallest values of $|\epsilon|$ shown in the figure. In the linear–logarithmic coordinates used, the numerical points line up with the predicted straight lines for larger $|\epsilon|$, thus confirming the validity of the asymptotic analysis. Further support is provided by the fact that the values of k and m for which $\mathrm{Im}\, c$ is maximized are close to the estimates (4.11) and (4.13). This is illustrated by figure 7 which compares the value of r maximizing $\mathrm{Im}\, c$ obtained numerically with the asymptotic prediction. This figure was computed by fixing the ratio $\delta = m/k$ at the predicted optimal values (1.25 for KW–IGW instabilities and 1 for IGW–IGW instabilities). With the small value $s = 0.1$ chosen, it can be interpreted as a confirmation that the value of m leading to the maximum $\mathrm{Im}\, c$ is predicted well by the asymptotics. It is evident from both figures 6 and 7 that the match between the numerical and analytical results is quite good even for ϵ moderately small.

Examining figure 6 more closely, we see that the instabilities become substantial for $\epsilon \approx 1$, especially KW–KW instabilities. Observe that, as predicted by the analysis, the decay of the growth rate in IGW–IGW instability as ϵ becomes small is the same for cyclonic and anticyclonic flows, and yet the growth rates of cyclonic flow are smaller by a factor of about 20. Thus, the $O(1)$ prefactor in the asymptotics of $\mathrm{Im}\, c$ for IGW–IGW instabilities, ignored in (4.14), turns out to be numerically very different for anticyclonic and cyclonic flows. The smallness of this prefactor in the cyclonic case means that the instability remains exceedingly weak even for $\epsilon \approx 1$, and probably irrelevant in many physical situations.

5. Discussion

This paper examines the linear stability of a horizontal Couette flow of a rapidly rotating strongly stratified inviscid fluid. The main conclusion is that the flow is unconditionally unstable: unbalanced instabilities, associated with linear resonances between Kelvin and inertia–gravity waves, occur for arbitrarily small Rossby numbers $\epsilon = \Lambda/f$. The growing perturbations have small horizontal and vertical scales, with typical wavenumbers or spatial-decay rates of the order of ϵ^{-1}. Physically, it is easy to understand why asymptotically small scales are a key ingredient of the instabilities. The phase locking between different waves which underlies the instability mechanisms requires the wave phase speed to be comparable with the basic flow velocity, and this

only occurs for small-scale waves. The need for small vertical scales also explains why the instabilities examined in this paper have no direct counterparts in shallow-water flows; these are stable for small enough $|\epsilon|$ because of the inherent restriction in vertical structure imposed by the shallow-water approximation. This illustrates some of the limitations of using the shallow-water model as a testbed for the study of the interactions between balanced motion and fast waves in the atmosphere and ocean.

Our conclusion that the rotating stratified Couette flow is always unstable is of course in sharp contrast with the one that may be drawn from balanced models which filter out Kelvin and inertia–gravity waves. Regardless of their accuracy, which can be any power ϵ^n, they predict the stability of flows without inflection points such as the Couette flow. There is no contradiction, however, since the growth rates found for the unbalanced instabilities are exponentially small in $|\epsilon|$. Note that balanced models which do not filter out Kelvin waves can represent the instability mode associated with Kelvin-wave resonance. As demonstrated by Kushner *et al.* (1998) with the semi-geostrophic model, the representation is only accurate when the Kelvin waves involved have low frequencies (that is, have aspect ratio $\delta \gg 1$ or, in dimensional terms, $k/m \ll f/N$).

In practice, the exponential dependence of the growth rate with the Rossby number ϵ means that the instabilities are exceedingly weak when $|\epsilon|$ is small, but can become important suddenly as $|\epsilon|$ increases towards 1 and beyond. If the instabilities are to play a significant role in the breakdown of balance in geophysical flows, this will therefore be in a manner that is extremely sensitive to the Rossby number. To illustrate this, a rough estimate gives e-folding times of the Kelvin-wave instability of the order of 1000 days for $\epsilon = 0.25$, 10 days for $\epsilon = 0.5$ and 1 day $\epsilon = 0.75$ (assuming that the asymptotic formula provides a correct order of magnitude for $\epsilon = O(1)$, as the numerics suggest). These estimates indicates that Rossby numbers $\epsilon \gtrsim 1/2$ are necessary for the instabilities discussed in this paper to be relevant to atmospheric and oceanic dynamics. In astrophysical applications, where typical time scales are much longer, the instabilities may well play a role for significantly smaller values of $|\epsilon|$.

In the literature, most attention has been paid to anticyclonic flows, and in particular to the coupled Kelvin-wave instability occurring in these flows. Our results clarify that cyclonic flows are also unstable, through an instability mechanism involving coupled inertia–gravity waves. This mechanism is also active in anticyclonic flows where, along with the instability mode mixing Kelvin and inertia–gravity waves, it provides an alternative to the well-studied instability due to Kelvin-wave resonance (see Molemaker *et al.* 2001; Yavneh *et al.* 2001). The focus on anticyclonic flows and Kelvin-wave instabilities is justified in practice by the associated growth rate being much larger than those of the other instability mechanisms, exponentially larger, in fact, in the limit $\epsilon \to 0$. The instability of the cyclonic flows is especially weak. This weakness is not completely accounted for by the exponential dependence on $1/\epsilon$, since this is the same for both anticyclonic and cyclonic flows whereas the growth rates obtained numerically are very different. We conclude, then, that the exponential dependence and the $O(1)$ prefactor conspire to make the instability of cyclonic flows extremely weak, even for moderate $|\epsilon|$.

The WKB approach used in this paper could be extended to examine the instability in more general rotating stratified shear flows. Obvious applications are the stratified Taylor–Couette flow (Molemaker *et al.* 2001; Yavneh *et al.* 2001), which differs from the problem studied here by the presence of curvature terms, and the stability of accretion disks (Rüdiger *et al.* 2002; Dubrulle *et al.* 2005). Additional physical effects that it would be of interest to study include different boundary conditions (in

particular the case of infinite domains for which no Kelvin waves exist), viscous and thermal damping, and non-zero potential-vorticity gradients, leading to the existence of critical levels for neutral modes (cf. Balmforth 1996).

J. V. was funded by a NERC Advanced Research Fellowship.

Appendix A. Conservation laws

Let

$$M = (u\partial_z\rho - w\partial_x\rho)/N^2.$$

Denoting integration over the periodic domain in x and z by

$$\langle\cdot\rangle = \iint \cdot \,\mathrm{d}x\,\mathrm{d}z,$$

we compute

$$\begin{aligned} N^2\partial_t\langle M\rangle &= \langle \partial_t u\partial_z\rho - \partial_z u\partial_t\rho - \partial_t w\partial_x\rho + \partial_x w\partial_t\rho\rangle \\ &= \langle (f-\Lambda)v\partial_z\rho - \partial_x p\partial_z\rho - N^2\partial_z uw + \partial_z p\partial_x\rho\rangle \\ &= -N^2\partial_y\langle uv\rangle, \end{aligned} \tag{A 1}$$

where we have used integration by parts and periodicity extensively, and, for the last line, $q = 0$ and the incompressibility equation. The conservation for the quadratic wave momentum (or pseudomomentum),

$$\mathscr{M} = \iiint (u\partial_z\rho - w\partial_x\rho)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z/N^2,$$

follows by integration in y, using the boundary conditions $v = 0$.

The perturbation energy $\mathscr{E}'$, with density $|\boldsymbol{u}|^2/2 + \rho^2/(2N^2)$, is not conserved but satisfies

$$\frac{\mathrm{d}\mathscr{E}'}{\mathrm{d}t} = -\iiint \Lambda uv\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Integrating by parts the right-hand side and using (A 1) gives a conservation law for the wave energy (or pseudoenergy)

$$\mathscr{E} = \iiint \left(\frac{|\boldsymbol{u}|^2}{2} + \frac{\rho^2}{2N^2} + \Lambda y\frac{u\rho_z - w\rho_x}{N^2}\right)\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Note that the conservation of both $\mathscr{M}$ and $\mathscr{E}$ can also be derived from the exact conservation laws for momentum, energy and potential vorticity for the full system, that is, basic flow plus perturbation.

Appendix B. Equation for $\hat{u}$

In § 2, the eigenvalue problem satisfied by normal-mode solutions is formulated as the second-order differential equation (3.4) for $\hat{p}$ and its associated boundary condition (3.5) (cf. Kushner *et al.* 1998). An alternative formulation, employed by Yavneh *et al.* (2001), uses $\hat{u}$ instead of $\hat{p}$ as the dependent variable. It has the advantage that the removable singularities that appear in (3.4) are absent. For completeness, we record this alternative formulation as

$$\epsilon^2\left(\frac{1-s^2\hat{\omega}^2}{K}\hat{u}'\right)' - \left(\frac{k^2(1-s^2\hat{\omega}^2) + m^2(1-\epsilon-\hat{\omega}^2)}{K} + \frac{2\epsilon(1-\epsilon)s^2k^2m^2\hat{\omega}^2}{K^2}\right)\hat{u} = 0, \tag{B 1}$$

where

$$K = (1-s^2\hat{\omega}^2)k^2 + (1-\epsilon)^2m^2.$$

The associated boundary conditions are

$$\epsilon\hat{u}' + \frac{(1-\epsilon)m^2\hat{\omega}}{k(1-s^2\hat{\omega}^2)}\hat{u} = 0 \quad \text{at} \quad y = \pm 1. \tag{B 2}$$

This is the formulation used for the numerical computation of the normal modes.

Appendix C. Kelvin-wave dispersion relation

In this Appendix, we derive the dispersion relation for KWs accurate to $O(\epsilon)$, as is necessary to obtain the leading-order asymptotics of the KW-instability growth rate.

The dispersion relation for $\mathrm{KW}_\pm$ valid to all orders in ϵ are given in (3.14). It is solved at leading order in § 3.2 to give (3.16). At the next order, we find the two equations

$$\pm\sigma c_1\lambda(\pm 1)g_\pm(\pm 1) + \hat{c}_0(\pm 1)g'_\pm(\pm 1) = 0, \tag{C 1}$$

which allow the determination of the $O(\epsilon)$ contribution to the frequency ω_1. Note that the contributions of the $O(\epsilon)$ terms neglected in (3.11)–(3.12) cancel in these two equations when (3.15) is taken into account. Equation (3.9) can be used to express the derivatives of $g_\pm$; the following results are therefore useful:

$$\lambda(\pm 1) = r, \tag{C 2a}$$

$$-\frac{\lambda'(\pm 1)}{2\lambda} = \frac{\pm\sigma(1-s^2)k^2 r}{m^2}, \tag{C 2b}$$

$$\frac{k\hat{\omega}_0(\pm 1)}{1-\hat{\omega}_0^2(\pm 1)} = \frac{\mp\sigma k^2 r}{r^2-k^2}, \tag{C 2c}$$

$$\frac{\mp\sigma h(\pm 1)}{2\lambda(\pm 1)} = \mp\sigma r\left(\tfrac{1}{2} - \frac{k^2}{r^2-k^2}\right) + c_1\frac{(1-s^2)k^2r^2}{m^2}. \tag{C 2d}$$

Using these and (3.9), (C 1) gives the first-order correction to the frequencies (3.15),

$$c_1 = \frac{\mp\sigma}{2r}. \tag{C 3}$$

REFERENCES

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

Balmforth, N. J. 1996 Shear instability in shallow water. *J. Fluid Mech.* **387**, 97–127.

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

Cairns, R. A. 1979 The role of negative energy waves in some instabilities of parallel flow. *J. Fluid Mech.* **92**, 1–14.

Craik, A. D. D. 1985 *Wave Interactions and Fluid Flows.* Cambridge University Press.

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

Dubrulle, B., Marié, L., Normand, C., Richards, D., Hersant, F. & Zahn, J. 2005 A hydrodynamic shear instability in stratified disks. *Astron. Astrophys.* **429**, 1–13.

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

Howard, J. E. & MacKay, R. S. 1987 Linear stability of symplectic maps. *J. Math. Phys.* **28**, 1036–1051.

Knessl, C. & Keller, J. B. 1992 Stability of rotating shear flow in shallow water. *J. Fluid Mech.* **244**, 605–614.

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**, 513–518.

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

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.

Molemaker, M. J., McWilliams, J. C. & Yavneh, I. 2005 Baroclinic instability and loss of balance. *J. Phys. Oceanogr.* **35**, 1505–1517.

Narayan, R., Goldreich, P. & Goodman, J. 1987 Physics of modes in a differentially rotating system – analysis of shearing sheet. *Mon. Not. R. Astr. Soc.* **228**, 1–41.

Papaloizou, J. C. B. & Pringle, J. E. 1987 The dynamical stability of differentially rotating discs – III. *Mon. Not. R. Astr. Soc.* **225**, 267–283.

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

Ren, S. & Shepherd, T. G. 1997 Lateral boundary contributions to wave-activity invariants and nonlinear stability theorems for balanced dynamics. *J. Fluid Mech.* **345**, 287–305.

Ripa, P. 1983 General stability conditions for zonal flow in a one-layer model on the β-plane or the sphere. *J. Fluid Mech.* **126**, 463–489.

Ripa, P. 1990 Positive, negative and zero wave energy and the flow stability problem in the Eulerian and Lagrangian–Eulerian descriptions, *Pure Appl. Geophys.* **133**, 713–732.

Rüdiger, G., Arlt, R. & Shalybkov, D. 2002 Hydrodynamic stability in accretion disks under the combined influence of shear and density stratification. *Astron. Astrophys.* **391**, 781–787.

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

Satomura, T. 1981*a* An investigation of shear instability in a shallow water. *J. Met. Soc. Japan* **59**, 148–167.

Satomura, T. 1981*b* Supplementary notes on shear instability in a shallow water. *J. Met. Soc. Japan* **59**, 168–171.

Satomura, T. 1982 An investigation of shear instability in a shallow water, part II: numerical experiment. *J. Met. Soc. Japan* **60**, 227–244.

Vanneste, J. & Yavneh, I. 2004 Exponentially small inertia–gravity waves and the breakdown of quasi-geostrophic balance. *J. Atmos. Sci.* **61**, 211–223.

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

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

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.