Vertical scale selection in inertial instability†

R. C. KLOOSTERZIEL[1] AND G. F. CARNEVALE[2]

[1]School of Ocean & Earth Science & Technology, University of Hawaii, Honolulu,
HI 96822, USA

[2]Scripps Institution of Oceanography, University of California, San Diego, La Jolla,
CA 92093, USA

The linear instability of a barotropic flow with uniform horizontal shear in a stratified rotating fluid is investigated with respect to perturbations invariant in the alongflow direction. The flow can be inertially unstable if there is sufficiently strong anticyclonic shear, but only for sufficiently high Reynolds numbers Re. We determine the critical Reynolds numbers required for amplification of the instability as a function of Prandtl number, strength of the stratification and magnitude of the shear. The vertical scales at the onset of the instability are calculated. For Prandtl number $P < 1.44$ instability always sets in through stationary overturning motions, for $P > 1.44$ it may also commence through overstable (oscillatory) motions. For Re exceeding the critical value, we determine the vertical scale of the most rapidly amplifying modes and the corresponding growth rates and how they vary with Re, P, the shear and the strength of stratification.

1. Introduction

Inertial instability was first primarily studied in the context of the stability of zonally symmetric atmospheric flows, encircling the Earth's rotation axis. Whereas a wide variety of instabilities may arise, inertial instability is generally understood to refer to instability due to zonally invariant perturbations. Because of this presumed alongflow invariance of the perturbations, it is also sometimes called symmetric instability. An example of the instability is the appearance of the well-known toroidal Taylor–Couette vortices in flows between concentric rotating cylinders. For such circular flows it is called centrifugal stability. Rayleigh's (1916) criterion for centrifugal instability is that the square of the angular momentum decreases with radius, or, equivalently, that the Rayleigh discriminant $\Phi < 0$ (Drazin & Reid 1981). For a parallel shear flow in a system rotating with angular velocity Ω, the criterion for instability is (Charney 1973)

$$\Phi = fQ < 0, \qquad \text{with} \qquad Q = f + \omega_z \qquad (1.1)$$

the absolute vorticity, $f = 2\Omega$ the Coriolis parameter, and ω_z the vertical (z) component of the relative vorticity (along the axis of rotation, see figure 1 for symbol definitions). This is for barotropic flows, i.e. flows with no vertical shear. In geophysical flows, (1.1) is the inviscid criterion for instability of barotropic parallel flows on a mid-latitude f-plane. For baroclinic flows in a stably stratified fluid, the criterion for instability remains $fQ < 0$, with Q replaced by the Ertel potential

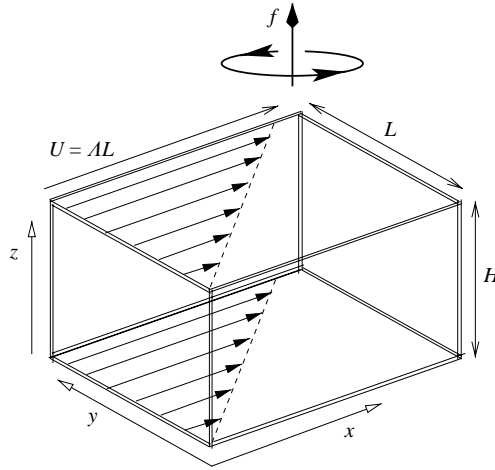† With an appendix by Stephen D. Griffiths

FIGURE 1. Diagram of the anticyclonically sheared barotropic flow in a rotating system. Inertial instability is only possible when the Rossby number for the anticyclonic shear $Ro = -\Lambda/f < -1$, so that the Rayleigh discriminant $\Phi = fQ = f(f - \Lambda) < 0$.

vorticity (Hoskins 1974). On the equatorial β-plane, the criterion for instability follows by setting $f = \beta y$, with y latitude ($y = 0$ at the equator) and β the latitudinal gradient of the Coriolis parameter at the equator.

Inviscid linear theory predicts that the inertial instability is manifested as over-turning motions in the meridional plane, largely confined to the region where $fQ < 0$ (Ooyama 1966). This has been confirmed in numerous studies. In ideal fluids, the overturning motions associated with inertial instability amplify most rapidly if their scales are vanishingly small (e.g. Stone 1966; Dunkerton 1981; Bayly 1988; Smyth & Peltier 1994; Smyth & McWilliams 1998; Griffiths 2003*a*; Billant & Gallaire 2005). For barotropic flows generally the growth rate s of normal-modes perturbations (proportional to $\exp(st)$, with t time), is bounded from above according to $s < |\overline{fQ}|^{1/2}$, with $\overline{fQ}$ the negative minimum of fQ (see Griffiths 2003*b*; Kloosterziel, Carnevale & Orlandi 2007). This inviscid maximum growth rate is attained for infinitely shallow overturning motions, i.e. when vertical scales become vanishingly small.

For barotropic flows in a stratified environment, there is generally a threshold *vertical* scale above which no growth occurs. This vertical cutoff scale decreases with increasing buoyancy frequency N. Since diffusion preferentially damps smaller scales, for increasingly stronger stratification the cutoff scale may become so small that no growth will occur at all, unless the Reynolds number Re is sufficiently high. In that case, only perturbations with vertical scales within a finite range will amplify. Within this range a maximum growth rate is found at a specific vertical scale. Hence, if a flow is subjected to small (symmetric) perturbations and the 'fastest' mode is excited, one expects meridional motions with this scale to emerge, before the evolution becomes truly fully nonlinear (see e.g. Griffiths 2003*a*; Kloosterziel *et al.* 2007).

Despite the extensive research on inertial instability, few studies have addressed the scale selection in any detail. It appears that only Griffiths' (2003*a*) study of a uniform shear flow on the equatorial β-plane contains precise predictions regarding the scale selection and maximal growth rates. The marginal stability boundary was first discussed by Dunkerton (1981). In both studies (with constant buoyancy frequency N), the Prandtl number was $P = 1$. When $P \neq 1$, Dunkerton (1982) found that instability

can set in through overstability, but few specific results regarding the scale selection are mentioned.

Whereas any finite shear at the equator implies (inviscid) inertial instability, on the f-plane it requires a flow to have sufficiently strong anticyclonic shear. The scale selection in such flows has not been investigated theoretically before in any detail. From the point of view of mid-latitude (ocean) numerical modelling, it is important to know what Reynolds number is minimally required to allow the instability to unfold. Also it is important to know what length scales are associated with the instability, since the instability will not be observed if these scales fall below the resolution in a numerical model. In particular, the present study was motivated by the question of what Reynolds numbers and resolutions are required if values of N/f are increased to large oceanically relevant values, and whether the scale selection depends sensitively on the value of P.

In this paper we analyse the linear (inertial) instability of a barotropic flow on the f-plane with uniform anticyclonic horizontal shear, which implies constant $fQ < 0$. The environment is linearly stratified, so that N is constant. The flow is artificially terminated at lateral vertical walls, where we prescribe free-slip boundary conditions. The normal-modes analysis is straightforward because the dispersion relation is algebraic (a cubic in the normal-modes growth rate). The problem is briefly discussed by Dunkerton (1982), who stated that the marginal stability criteria are qualitatively similar to that for the uniform shear flow on the β-plane. But, details were not provided.

In §2 we discuss the linear perturbation equations and the boundary conditions. The exact marginal stability boundaries are calculated in §3. We formulate it in terms of critical Reynolds numbers as a function of $|fQ|/f^2$, N/f and P. The vertical scales of the normal modes that appear at the onset of instability are also determined.

In §4 we determine the fastest growing modes and the growth rates as we move into the unstable regime, with emphasis on the limit of large Re. In §4.1 we first consider Prandtl number $P = 1$, which is a special case because the maximal growth rates and corresponding vertical scales can be determined in closed form. But, expansions in inverse fractional powers of Re are found to be remarkably accurate, in that the first few terms give excellent approximations. In §4.2 we compare the results obtained thus far with those reported in the equatorial β-plane studies by Dunkerton (1981, 1982) and Griffiths (2003a). Finally, in §4.3 we discuss the scale selection when $P \neq 1$.

In §5 we conclude with a summary of the main results and remarks on how this work may be extended to more realistic unbounded flows with, for example, uniform shear in a transition region between two counterflowing uniform currents. For such a flow one would have constant $fQ < 0$ in a strip of finite width, while $fQ = f^2$ outside this region. We expect that the structure of the marginal stability boundaries, and the Reynolds number dependence in the scale selection process, will qualitatively survive if the analysis is extended to such unbounded flows, but this has not yet been fully investigated.

A substantial clarification of the differences between our results on the one hand, and the β-plane flows and barotropic vortices on the other, is provided in Appendix B (by S. D. Griffiths).

2. The model and simplifying assumptions

We consider a steady barotropic parallel flow $u = U(y)$ in the x-direction, with uniform shear $dU/dy = \Lambda$, i.e. $U = U_0 + \Lambda y$. The fluid is assumed to be Boussinesq and

has a stable, linear density stratification $\rho(z) = \rho_0(1 - (N^2/g)z)$; g is the gravitational constant, ρ the density and ρ_0 a reference density; $N^2 = -(g/\rho_0)\mathrm{d}\rho/\mathrm{d}z \geqslant 0$ is the square of the buoyancy frequency, which is constant. The density is assumed to depend linearly on temperature. The flow is in geostrophic and hydrostatic balance in the rotating system. The stability with respect to perturbations invariant in the alongflow (x)-direction is as usual investigated by introducing velocity perturbations $u', v', w'(y, z, t)$, density perturbation $\rho'(y, z, t)$ and pressure perturbation $p'(y, z, t)$, and subsequent linearization about the steady state. The linearized equations are (see e.g. Emanuel 1979)

$$\left(\frac{\partial}{\partial t} - \nu\nabla^2\right)u' + [\Lambda - f]v' = 0, \tag{2.1}$$

$$\left(\frac{\partial}{\partial t} - \nu\nabla^2\right)v' + fu' = -\frac{1}{\rho_0}\frac{\partial p'}{\partial y}, \tag{2.2}$$

$$\left(\frac{\partial}{\partial t} - \nu\nabla^2\right)w' + g\frac{\rho'}{\rho_0} = -\frac{1}{\rho_0}\frac{\partial p'}{\partial z}, \tag{2.3}$$

$$\left(\frac{\partial}{\partial t} - \kappa\nabla^2\right)\rho' + w'\frac{\mathrm{d}\rho}{\mathrm{d}z} = 0, \tag{2.4}$$

$$\frac{\partial v'}{\partial y} + \frac{\partial w'}{\partial z} = 0; \tag{2.5}$$

ν, κ are viscosity and diffusivity of heat, respectively, and ∇^2 is the Laplace operator. The flow is laterally terminated by rigid free-slip walls at $y = 0, L$. That is, we assume

$$v' = 0, \quad \frac{\partial w'}{\partial y} = 0 \quad \text{at} \quad y = 0, L. \tag{2.6}$$

Additional boundary conditions for ρ' and u' are required. These have to be consistent with the governing equations (2.1)–(2.4) and (2.6). For example, in the inviscid limit, (2.1) implies with (2.6) that $\partial_t u' = 0$ at $y = 0, L$. Hence, if at some initial time $u' = 0$ at the boundaries, it will remain so. The vorticity equation

$$\frac{\partial}{\partial t}\left[\frac{\partial w'}{\partial y} - \frac{\partial v'}{\partial z}\right] - f\frac{\partial u'}{\partial z} + \frac{g}{\rho_0}\frac{\partial \rho'}{\partial y} = \nu\nabla^2\left[\frac{\partial w'}{\partial y} - \frac{\partial v'}{\partial z}\right] \tag{2.7}$$

shows that in the inviscid limit $-f\partial_z u' + (g/\rho_0)\partial_y\rho' = 0$ at the boundaries. If we assume that $u' = 0$ at $y = 0, L$, then in the inviscid limit it follows that $\partial_y\rho' = 0$ at $y = 0, L$. Hence the condition

$$u' = 0, \quad \frac{\partial \rho'}{\partial y} = 0 \quad \text{at} \quad y = 0, L \tag{2.8}$$

is consistent with (2.1)–(2.4) and (2.6). If $\nu \neq 0$ and $\kappa \neq 0$, the boundary conditions (2.6) and (2.8) are also consistent with (2.1)–(2.4), and they are used in what follows.

A streamfunction ψ for the meridional velocity perturbations is introduced as

$$v' = -\frac{\partial \psi}{\partial z}, \qquad w' = \frac{\partial \psi}{\partial y}, \tag{2.9}$$

so that the vorticity $\partial_y w' - \partial_z v' = \nabla^2\psi$, with $\nabla^2 = \partial_y^2 + \partial_z^2$. Through cross-differentiation a single equation for ψ follows:

$$\mathrm{D}_\kappa\mathrm{D}_\nu^2(\nabla^2\psi) + \mathrm{D}_\kappa\partial_z(fQ\partial_z\psi) + \mathrm{D}_\nu\partial_y(N^2\partial_y\psi) = 0, \qquad \text{where} \quad fQ = f(f - \Lambda), \tag{2.10}$$

and

$$D_\nu = \partial_t - \nu\nabla^2, \qquad D_\kappa = \partial_t - \kappa\nabla^2. \tag{2.11}$$

In this study $N^2 =$ constant and $fQ =$ constant. With the boundary conditions (2.6) and (2.8), (2.10) allows simple normal-modes solutions

$$\psi = \exp(st)\sin(l\pi y/L)\exp(im\pi z/H), \qquad l = 1, 2, 3, \ldots, \tag{2.12}$$

with the real part understood. The streamlines in the meridional plane will resemble cells of overturning motions of alternating sign. Note that we treat the domain in the vertical as unbounded. The vertical (z) coordinate has been non-dimensionalized with an arbitrary length scale H. Without loss of generality we put $H = L$ in the analysis that follows. In §5 we discuss how things change when $H \neq L$. While l can only take discrete values, the non-dimensional vertical wavenumber m can take any real positive value.

If $\kappa = \nu = 0$, the dispersion relation is

$$s^2 = -\frac{m^2 fQ + l^2 N^2}{l^2 + m^2}. \tag{2.13}$$

Instability is clearly only possible if $fQ < 0$. When $fQ < 0$, growth occurs for all m exceeding the buoyancy cutoff wavenumber, i.e. for

$$m > \frac{N}{|fQ|^{1/2}} l. \tag{2.14}$$

For any given l, the growth rate approaches the inviscid maximum $s = s_\star = |fQ|^{1/2}$ in the limit $m \to \infty$, i.e. for cells that become infinitely 'flat' in the vertical direction.

3. Linear stability boundaries

When $\nu \neq 0$ and $\kappa \neq 0$, instability can also only occur when $fQ < 0$. This is not necessarily true for baroclinic flows (vertically sheared flows), which can be unstable when $fQ > 0$ and the Prandtl number $P \neq 1$ (McIntyre 1970; Kloosterziel & Carnevale 2007). We replace fQ in (2.10) with $-|fQ|$, with the understanding that $fQ < 0$. Substitution of (2.12) in (2.10) then yields a cubic in s. Non-dimensionally it can be written as

$$\tilde{s}^3 + B\tilde{s}^2 + C\tilde{s} + D = 0, \qquad \text{with} \qquad B = \left(2 + \frac{1}{P}\right)a, \tag{3.1}$$

$$\left.\begin{array}{l} C = \left(1 + \dfrac{2}{P}\right)a^2 - |\widetilde{fQ}| + \dfrac{l^2\pi^2 E}{a}[|\widetilde{fQ}| + \tilde{N}^2], \\[2mm] D = \dfrac{1}{P}(a^3 - |\widetilde{fQ}|a + l^2\pi^2 E[|\widetilde{fQ}| + P\tilde{N}^2]). \end{array}\right\} \tag{3.2}$$

$\tilde{s} = s/f$ is the non-dimensional growth rate, $\widetilde{fQ} = fQ/f^2$, $\tilde{N}^2 = N^2/f^2$, $P = \nu/\kappa$ is the Prandtl number and

$$a = \left(l^2 + m^2\right)\pi^2 E, \qquad \text{where} \qquad E = \frac{\nu}{fL^2} \tag{3.3}$$

is the Ekman number (the ratio of viscous to Coriolis forces). We will use the relations

$$E = \frac{|Ro|}{Re}, \qquad |Ro| = \Lambda/f = |\widetilde{fQ}| + 1, \qquad Re = \frac{\Delta U L}{\nu}, \tag{3.4}$$

	$D < 0$	$D = 0$	$D > 0$
$BC - D < 0$	$s > 0$	$s > 0$	$s < 0$
	$\mathrm{Re}\,s < 0$	$s = 0$	$\mathrm{Re}\,s > 0$
	$\mathrm{Re}\,s < 0$	$s < 0$	$\mathrm{Re}\,s > 0$
	unstable	unstable	unstable
$BC - D = 0$	$s > 0$	$s = 0$	$s < 0$
	$\mathrm{Re}\,s < 0$	$s = 0$	$s = +\mathrm{i}\omega$
	$\mathrm{Re}\,s < 0$	$s < 0$	$s = -\mathrm{i}\omega$
	unstable	convection	overstable
$BC - D > 0$	$s > 0$	$s = 0$	$s < 0$
	$\mathrm{Re}\,s < 0$	$\mathrm{Re}\,s < 0$	$\mathrm{Re}\,s < 0$
	$\mathrm{Re}\,s < 0$	$\mathrm{Re}\,s < 0$	$\mathrm{Re}\,s < 0$
	unstable	convection	stable

TABLE 1. A summary of the properties of the three roots of the cubic (3.1) when B, C, D are real and $B > 0$. Whenever the sign of the real part ($\mathrm{Re}\,s$) is given, s can either be real or is one of a complex-conjugate pair. In the overstable case the frequency ω is determined by $\omega^2 = C$.

with $Ro < -1$ the Rossby number for the unstable anticyclonic shear flow ($\widetilde{fQ} < 0$), Re the Reynolds number and $\triangle U = \Lambda L$ the velocity change across the width of the domain. The dispersion relation (3.1) is the same as for the stratified Taylor–Couette problem studied by Thorpe (1966). The dispersion relation for when the hydrostatic approximation is made and horizontal diffusion is neglected can be found in Dunkerton (1982). For convenience, we drop the tildes on $\tilde{s}$, $\widetilde{fQ}$ and $\widetilde{N^2}$.

Since B, C and D are real, either the three roots of the cubic are real or one root is real and the other two are complex conjugates. The flow will be unstable if there is at least one positive real root $s > 0$ or when there is a pair of complex-conjugate roots with a real part $\mathrm{Re}\,s > 0$. Instability associated with a real root $s > 0$ is called stationary instability. Instability associated with complex-conjugate roots with $\mathrm{Re}\,s > 0$, is called oscillatory instability.

In Appendix A we show that if there is a real root $s > 0$, then $s < |fQ|^{1/2}$. If there is a pair of complex conjugate roots $s = b \pm \mathrm{i}\omega$ with $b > 0$, then $\mathrm{Re}\,s = b < |fQ|^{1/2}$. Thus, the growth rates for both stationary instability and oscillatory instability are bounded from above by $|fQ|^{1/2}$.

Because B is always positive, the marginal stability boundaries can be determined exactly by examination of the signs of D and $BC - D$. This was shown in Kloosterziel & Carnevale (2003). Table 1 contains all the required information. For sufficiently low Reynolds numbers, both $D > 0$ and $BC - D > 0$ for all $a > 0$. In that case the flow is stable. In §3.1 we discuss the transition from $D > 0$ to instability associated with $D < 0$, whilst in §3.2 we discuss the transition from $BC - D > 0$ to instability implied by $BC - D < 0$. In §3.3, we establish the exact marginal stability boundary by combining the results from §3.1 and §3.2.

3.1. The convection boundary

Consider the coefficient D given by (3.2) and recall that $a > 0$ for any real l, m. For given $P, E, |fQ|, l$ and N^2, there may be a range of positive a-values for which $D < 0$. This range shrinks to zero width on increasing E, or lowering the Reynolds number Re. At that point, $D = 0$ for one particular $a = a_c$ with a corresponding root $s = 0$,
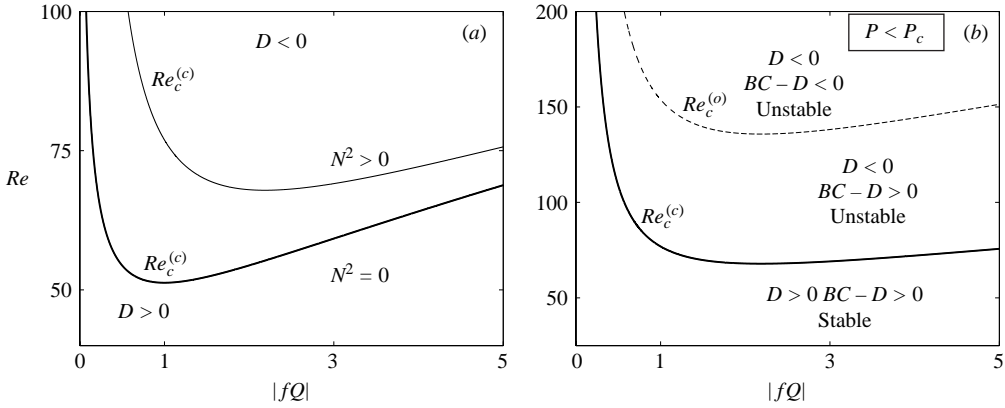
FIGURE 2. (*a*) The convection boundary $Re_c^{(c)}$ (3.5) as a function of $|fQ|$ for $N^2 = 0$ (thick line) and $N^2 = 0.5$, $P = 1$ (thin line). (*b*) The convection boundary (solid line) and the overstability boundary (dashed line) $Re_c^{(o)}$ (3.7) for $N^2 = 0.5$ and $P = 1$. For $P < P_c \approx 1.44$, we have $Re_c^{(c)} < Re_c^{(o)}$ for all $|fQ|$ and all $N^2 > 0$. Panel (*b*) is representative of all cases with $P < P_c$ when the marginal stability boundary is $Re_c^{(c)}$ (thick line). Stability/instability as indicated follows from table 1.

but $D > 0$ for all other a. For those a, one root will be real and negative. For small enough Re, D will become positive for all $a > 0$.

The variable a is continuous since the vertical wavenumber m is continuous. Critical values are found by simultaneously solving $D = 0$ and $\partial D / \partial a = 0$. The critical Reynolds number is

$$Re_c^{(c)} = R_c^{1/2} \frac{(|fQ| + 1)[|fQ| + PN^2]}{|fQ|^{3/2}}, \qquad \text{where} \qquad R_c = \frac{27\pi^4}{4} \qquad (3.5)$$

is the critical Rayleigh number for the onset of convection in the classical Rayleigh–Bénard problem with free-slip boundary conditions at the top and bottom (Chandrasekhar 1961). If $Re < Re_c^{(c)}$, then $D > 0$ for any normal-modes perturbation. When $Re > Re_c^{(c)}$, we have $D < 0$ for some perturbations and the system will be unstable. When $Re = Re_c^{(c)}$, we have $D = 0$ when $l = 1$ and $m = m_c$, with

$$m_c^2 = \left(\frac{|fQ|}{3} \right)^{1/2} \frac{Re_c^{(c)}}{\pi^2 (|fQ| + 1)} - 1 = \frac{1}{2} + \frac{3P}{2} \frac{N^2}{|fQ|}. \qquad (3.6)$$

But, $D > 0$ for all other $\{l, m\}$. An example of $Re_c^{(c)}$ as a function of $|fQ|$ is shown in figure 2(*a*). We call $Re_c^{(c)}$ the 'convection' boundary, for reasons explained below. With increasing N^2 or P (when $N^2 \neq 0$), the curves shift to higher Re-values. For $l \neq 1$, the critical Reynolds number follows by multiplying the right-hand side of (3.5) by l^2. Hence, $l = 1$ is the 'most dangerous' horizontal wavenumber, in the sense that the transition from $D > 0$ to $D < 0$ occurs for the smallest critical Reynolds number.

3.2. *The overstability boundary*

When $Re < Re_c^{(c)}$, stability is not guaranteed: although $D > 0$ for any perturbation, roots of the cubic can have $\text{Re}s > 0$ if $BC - D < 0$ for some $\{l, m\}$ (see table 1). Solving simultaneously for $BC - D = 0$ and $\partial (BC - D) / \partial a = 0$, we find that if

$$Re < Re_c^{(o)} = R_c^{1/2} \frac{(|fQ| + 1)}{|fQ|^{3/2}} \left[\frac{1 + P}{P} |fQ| + \frac{(1 + P)^2}{2P^2} N^2 \right], \qquad (3.7)$$

then $BC - D > 0$ for any perturbation. We call $Re_c^{(o)}$ the 'overstability' boundary. For $Re > Re_c^{(o)}$, we have $BC - D < 0$ for some perturbations, which implies instability. When $Re = Re_c^{(o)}$, we have $BC - D = 0$ for $l = 1$ and $m = m_c$, with

$$m_c^2 = \frac{P}{P+1} \left(\frac{|fQ|}{3} \right)^{1/2} \frac{Re_c^{(o)}}{\pi^2(|fQ| + 1)} - 1 = \frac{1}{2} + \frac{3(P+1)}{4P} \frac{N^2}{|fQ|}. \tag{3.8}$$

For all other $\{l, m\}$, we have $BC - D > 0$ when $Re = Re_c^{(o)}$.

3.3. *The marginal stability boundary*

If $Re < Re_c^{(c)}$ and $Re < Re_c^{(o)}$, the flow is stable since then for any perturbation $D > 0$ and $BC - D > 0$. If $Re_c^{(c)} < Re_c^{(o)}$, instability will first set in when the Reynolds number is increased to $Re = Re_c^{(c)}$. On the other hand, if $Re_c^{(c)} > Re_c^{(o)}$, instability will commence when Re reaches $Re = Re_c^{(o)}$. To determine which critical Reynolds number is greater, we equate (3.5) to the right-hand side of (3.7). We find that $Re_c^{(c)} = Re_c^{(o)}$ if

$$|fQ| = |fQ|_i = \frac{N^2}{P} \left[P^3 - \tfrac{1}{2}(1 + P)^2 \right], \tag{3.9}$$

where $|fQ|_i$ is the intersection of the convection and overstability surface. Since $|fQ| > 0$, they intersect only if $P^3 > (1/2)(1 + P)^2$, i.e. for $P > P_c \approx 1.44$, with P_c the positive real root of $P^3 - (1/2)(1 + P)^2 = 0$. This critical Prandtl number was first found by Thorpe (1966) in his study of stratified Taylor–Couette flow.

When $P < P_c$, we have $Re_c^{(c)} < Re_c^{(o)}$ for all $|fQ|$ and N^2 so that the flow is stable for all $Re < Re_c^{(c)}$. An example is shown in figure 2(b). When Re exceeds $Re_c^{(c)}$, there are $\{l, m\}$ for which $D < 0$, which implies instability. Hence, the marginal stability boundary is $Re_c^{(c)}$. At the boundary all modes decay ($D > 0$ and $BC - D > 0$), except for $\{l, m\} = \{1, m_c\}$, with m_c determined by (3.6). Perturbations with that vertical wavenumber render $D = 0$ and $BC - D > 0$, and can therefore excite a non-decaying neutral mode with $s = 0$, i.e. steadily overturning motions or stationary convection. Note that according to (3.6), if $P \ll 1$, instability can set in for far lower wavenumbers than the buoyancy cutoff wavenumber, given in (2.14).

If $P > P_c$, the marginal stability boundary is $Re_c^{(o)}$ for $0 < |fQ| < |fQ|_i$ and $Re_c^{(c)}$ for $|fQ| > |fQ|_i$. An example is shown in figure 3. When $|fQ| < |fQ|_i$ instability sets in through overstability at the marginal stability boundary. That is, if the Reynolds number reaches $Re = Re_c^{(o)}$, perturbations with vertical wavenumber m_c given by (3.8) can excite non-decaying, oscillatory modes with $s = \pm i\omega$ because then $D > 0$ and $BC - D = 0$. For all other wavenumbers, the modes decay. In either case ($P > P_c$ or $P < P_c$), for small $|fQ|$ and corresponding large critical Re, $m_c \propto Re^{1/3}$ when $N \neq 0$. The smallest critical Reynolds number is found in the limit $N^2 \to 0$ at $|fQ| = 1$, i.e. $Re = (27\pi^4)^{1/2} \approx 51.3$.

4. Vertical scale selection in unstable flows

4.1. *Special case: Prandtl number $P = 1$*

When $\nu = \kappa$, equation (3.3) is over-differentiated. The analysis should be based on the quadratic:

$$(s + a)^2 = |fQ| - \frac{l^2\pi^2 E}{a} [|fQ| + N^2]. \tag{4.1}$$
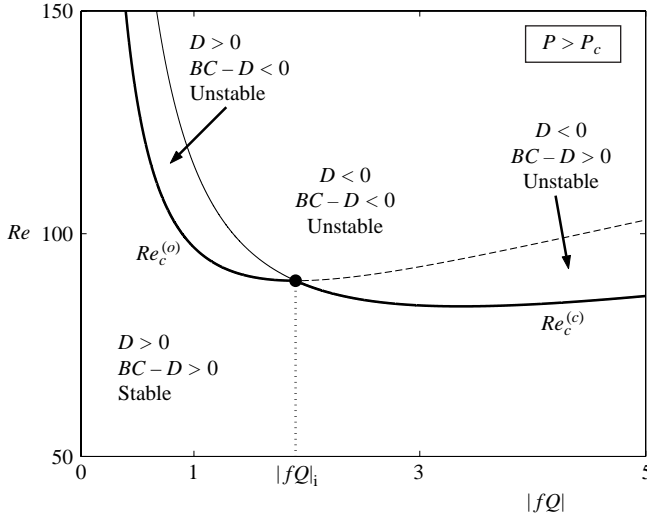
FIGURE 3. An example of the marginal stability boundary representative of when $P > P_c$ ($N^2 = 0.5$, $P = 2.5$). The convection boundary $Re_c^{(c)}$ (thin solid) and the overstability boundary $Re_c^{(o)}$ (thin dashed line) intersect at $\{Re, |fQ|\} = \{Re_i, |fQ|_i\}$ (indicated by ●). The marginal stability boundary is the thick solid line. To the left of $|fQ|_i$ it is the overstability boundary $Re_c^{(o)}$, to the right the convection boundary $Re_c^{(c)}$. $|fQ|_i$ is determined by (3.9).

Overstability is impossible, as is seen by substituting $s = i\omega$. We simplify (4.1) by writing

$$s = |fQ|^{1/2}\hat{s}, \qquad a = \epsilon|fQ|^{1/2}\hat{a}, \qquad \text{with} \qquad \epsilon = \frac{l\mathscr{C}}{Re^{1/2}} \tag{4.2}$$

and

$$\mathscr{C} = \left(\frac{\pi^2}{2}\right)^{1/2} \frac{(|fQ| + 1)^{1/2}(|fQ| + N^2)^{1/2}}{|fQ|^{3/4}}. \tag{4.3}$$

This transforms (4.1) into

$$(\hat{s} + \epsilon\hat{a})^2 = 1 - \frac{2\epsilon}{\hat{a}}. \tag{4.4}$$

Differentiating (4.4) with respect to $\hat{a}$, we find $(\hat{s} + \epsilon\hat{a})(\epsilon + \partial\hat{s}/\partial\hat{a}) = \epsilon/\hat{a}^2$. Maximal $\hat{s} = \hat{s}_\star$ are found by setting $\partial\hat{s}/\partial\hat{a} = 0$. Substitution in (4.4) shows that the maximum occurs for $\hat{a} = \hat{a}_\star(\epsilon)$, where $\hat{a}_\star$ is the positive real root of

$$\hat{a}_\star^4 - 2\epsilon\hat{a}_\star^3 = 1, \qquad \text{and} \qquad \hat{s}_\star = \hat{a}_\star(\hat{a}_\star - 3\epsilon). \tag{4.5}$$

When $\epsilon = 0$, $\hat{s}_\star$ takes the largest possible value $\hat{s}_\star = 1$, and $\hat{a}_\star = 1$. For $\epsilon = \epsilon_c = (1/3)^{3/4}$, $\hat{a}_\star = 3\epsilon$ and $\hat{s}_\star = 0$. For $\epsilon > \epsilon_c$ all modes are damped. Hence, there will be instability for $0 \leqslant \epsilon < \epsilon_c$. With the definition of ϵ it follows again that $l = 1$ is the most dangerous horizontal wavenumber and that the marginal stability boundary is $Re = Re_c^{(c)}$, as in (3.5), but with $P = 1$.

The positive real root $\hat{a}_\star$ of the quartic in (4.5) is

$$\hat{a}_\star(\epsilon) = \frac{1}{2}\left[\epsilon + (\epsilon^2 + h)^{1/2}\right] + \frac{1}{2}\left[\left(\epsilon + (\epsilon^2 + h)^{1/2}\right)^2 - 2\left(h + \frac{\epsilon h}{(\epsilon^2 + h)^{1/2}}\right)\right]^{1/2}, \tag{4.6}$$

with $h = h(\epsilon) = [-2\epsilon^2 + (4\epsilon^4 + (4/3)^3)^{1/2}]^{1/3} - [2\epsilon^2 + (4\epsilon^4 + (4/3)^3)^{1/2}]^{1/3}$ the negative real root of $h^3 + 4h + 4\epsilon^2 = 0$. These solutions are found using Ferrari's and Vieta's

method dating back to the 16th century (see Kline 1972). Substitution of $\hat{a}_\star$ in the equation for $\hat{s}_\star$ in (4.5) yields the exact expression for $s_\star = |fQ|^{1/2}\hat{s}_\star(\epsilon)$. For small ϵ

$$\hat{a}_\star(\epsilon) = 1 + \tfrac{4}{8}\epsilon + \tfrac{3}{8}\epsilon^2 + O(\epsilon^3), \qquad \hat{s}_\star(\epsilon) = 1 - 2\epsilon - \tfrac{1}{2}\epsilon^2 + O(\epsilon^3). \qquad (4.7)$$

Because the exact solutions have little curvature for $0 \leqslant \epsilon \leqslant \epsilon_c$, the first two or three terms in (4.7) give excellent approximations, except near $\epsilon \approx \epsilon_c$. For $l = 1$

$$s_\star = |fQ|^{1/2} \left[\underbrace{1 - \overbrace{\frac{2\mathscr{C}}{Re^{1/2}}}^{\text{first-order}} - \frac{\mathscr{C}^2}{2Re}}_{\text{second-order}} - O\left(Re^{-3/2}\right) \right]. \qquad (4.8)$$

Thus, for large Re

$$s_\star \approx |fQ|^{1/2} \left[1 - (2\pi^2)^{1/2} \frac{(|fQ| + N^2)^{1/2}}{|fQ|^{3/4}} \frac{(|fQ| + 1)^{1/2}}{Re^{1/2}} \right]. \qquad (4.9)$$

We call $m = m_\star$ the vertical wavenumber of the fastest growing modes. Since $a = \epsilon|\Phi|^{1/2}\hat{a}$ and $a = (l^2 + m^2)\pi^2 E$, we find that for $l = 1$

$$m_\star^2 = \frac{Re|fQ|^{1/2}\epsilon\hat{a}_\star(\epsilon)}{\pi^2(|fQ| + 1)} - 1 = \frac{|fQ|^{1/2}}{\pi^2(|fQ| + 1)} \left[\underbrace{\overbrace{\mathscr{C}Re^{1/2} + \frac{1}{2}\mathscr{C}^2}^{\text{first-order}} + \frac{3\mathscr{C}^3}{8Re^{1/2}}}_{\text{second-order}} + O(Re^{-1}) \right] - 1. \qquad (4.10)$$

For large Re therefore

$$m_\star^2 \approx \frac{1}{4}\frac{N^2}{|fQ|} + \frac{(|fQ| + N^2)^{1/2}}{(2\pi^2)^{1/2}|fQ|^{1/4}} \frac{Re^{1/2}}{(|fQ| + 1)^{1/2}}. \qquad (4.11)$$

In figure 4(a) we show the exact $m_\star$ and the first-order approximation (4.11). This example is representative of all possible combinations of $|fQ|$ and N^2 in that for Re close to but larger than the critical value $Re_c^{(c)}$, the approximation is already very close to the exact value. In figure 4(b) we compare the exact $s_\star$ with the first-order approximation (4.9). When $Re \approx Re_c^{(c)}$, the error is substantial but rapidly diminishes as Re increases.

4.2. Comparison with Dunkerton's and Griffiths' equatorial β-plane results

For a barotropic flow with uniform horizontal shear $\mathrm{d}U/\mathrm{d}y = \Lambda$ in an environment with constant buoyancy frequency N, Griffiths (2003a) has described to some detail the linear vertical scale selection in inertial instability on the equatorial β-plane, assuming $P = 1$. The marginal stability boundary was previously determined by Dunkerton (1981). As is common in theoretical equatorial β-plane studies, horizontal diffusion was ignored and the hydrostatic approximation was made.

Since $fQ = \beta y(\beta y - \Lambda)$, it varies with latitude and $fQ \leqslant 0$ for $0 \leqslant y \leqslant \Lambda/\beta$. A negative minimum $fQ = \overline{fQ} = -(\Lambda/2)^2$ is found at $y = \bar{y} = \Lambda/2\beta$. This implies that the maximum inviscid normal-modes growth rate is $s_\star = |\overline{fQ}|^{1/2} = \Lambda/2$, taking $\Lambda > 0$. A set of equations similar to (4.5) follows (Griffiths 2003a, equation 8a, b):

$$\hat{m}_\star^6 - 4\epsilon\hat{m}_\star^5 = 1, \qquad \hat{s}_\star = \hat{m}_\star^2(\hat{m}_\star - 5\epsilon), \qquad \text{where} \qquad \epsilon = \left(\frac{2\nu N^2 \beta^2}{\Lambda^5}\right)^{1/3} \qquad (4.12)$$

is a non-dimensional parameter, $\hat{s}_\star = (2/\Lambda)s_\star$ the non-dimensional maximum growth rate and $\hat{m}_\star = (\epsilon\Lambda^2/N\beta)m_\star$ the non-dimensional vertical wavenumber associated with
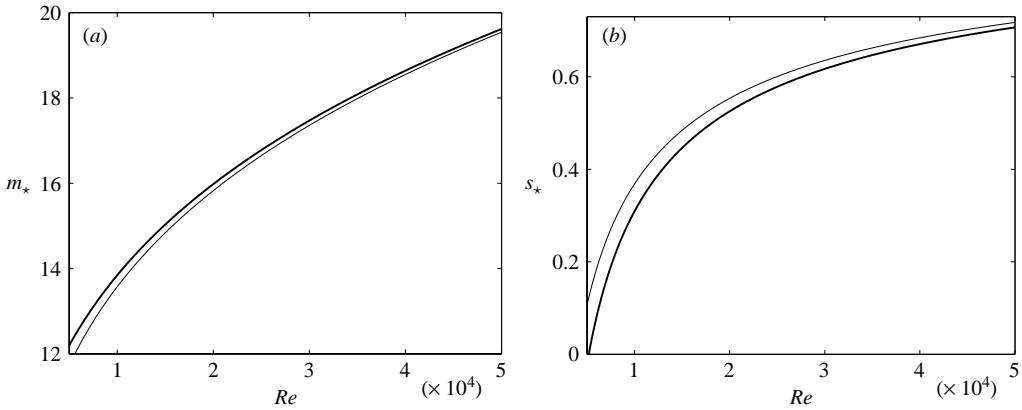
FIGURE 4. (*a*) The wavenumber $m_\star$ of the fastest growing mode determined by the exact solution (4.10) as a function of $Re \in [5000, 50{,}000]$ for $|fQ|=1$, $N^2=100$ and $P=1$ (thick solid line). The thin line is the first-order approximation of $m_\star$ (4.11). (*b*) The exact maximum growth rate $s_\star = |fQ|^{1/2}\hat{s}_m$ as a function of Re (thick solid line) and the first-order approximation of $s_\star$ (thin line) given in (4.9). Growth commences for $Re = Re_c^{(c)} \approx 5180$ with $m_c \approx 12$. The second-order approximations (not shown) virtually coincide with the exact solutions for all Re.

maximal growth. Whereas we use a non-dimensional vertical wavenumber, in Griffiths' study $m_\star$ has the dimension of length^{-1}.

At the marginal stability boundary instability sets in through stationary convection ($s=0$) when $\epsilon = \epsilon_c = 5^{-5/6}$, with a corresponding $\hat{m}_\star = \hat{m}_c = 5\epsilon_c = 5^{1/6}$. Defining the Reynolds number as in (3.4) by $Re = \Delta U L / \nu$, with $\Delta U = \Lambda L$ and $L = \Lambda/\beta$ the width of the unstable domain, Griffiths' and Dunkerton's results imply

$$Re_c^{(c)} = 2 \times 5^{5/2} \frac{N^2}{\Lambda^2}, \qquad m_c = 5\frac{N}{\Lambda}, \qquad \text{where} \qquad Re = \frac{\Lambda^3}{\nu\beta^2}. \qquad (4.13)$$

The critical wavenumber m_c along the stability boundary has been non-dimensionalized by multiplying with the length scale $L = \Lambda/\beta$.

If we also make the hydrostatic approximation and ignore horizontal diffusion, the first term in (2.10) containing $\nabla^2\psi$ must be replaced by $\partial_z^2\psi$ (hydrostatic approximation), and ∇^2 by ∂_z^2 in D_ν and D_κ, defined in (2.11) (no horizontal diffusion). The consequence is that we must replace $|fQ| + PN^2$ with PN^2 in (3.5), and the constant $1/2$ in (3.6) must be dropped. For $P=1$, we find that

$$Re_c^{(c)} = R_c^{1/2} \frac{N^2\Lambda}{|fQ|^{3/2}}, \qquad m_c = \left(\frac{3}{2}\right)^{1/2} \frac{N}{|fQ|^{1/2}}. \qquad (4.14)$$

Identifying $|fQ| \leftrightarrow (\Lambda/2)^2$, we see that (4.13) and (4.14) are essentially the same.

In Dunkerton (1982), the stability analysis of the equatorial shear flow is extended to $P \neq 1$. The marginal stability criterion for the existence of a neutrally stable mode ($s=0$) is given in terms of the critical shear and corresponding vertical wavenumber (equation 1.2a,b in Dunkerton 1982). His criterion can again be expressed as a critical Reynolds number $Re_c^{(c)}$, with PN^2 replacing N^2 in (4.13). The critical vertical wavenumber m_c follows by replacing N with $\sqrt{PN^2}$. On the f-plane, $Re_c^{(c)}$ and m_c follow in the same fashion by substituting PN^2 for N^2 in (4.14). Further, Dunkerton concluded that for $P \gtrsim 3/2$ instability sets in through overstability, but no explicit

expression equivalent to our critical Reynolds number $Re_c^{(o)}$ (3.7) is given. It is not clear whether this critical $P_c \approx 3/2$ should actually be Thorpe's (1966) $P_c \approx 1.44$, which is a root of an explicit cubic, given above after equation (3.9). Given the cumbersome nature of the dispersion relation, it appears that Dunkerton found P_c numerically. Dunkerton also mentions that for a constant shear flow on the f-plane, bounded by rigid walls, the marginal stability criteria closely resemble those for the β-plane flow. Our explicit expressions confirm this.

Whereas the marginal stability criteria for the β-plane flow and f-plane flow are qualitatively the same, we find significant differences with regard to the scale selection at large Re. For small ϵ or large Re, Griffiths (2003a) found using (4.12) that for the β-plane flow with $P = 1$

$$\hat{m}_\star = 1 + \frac{2\epsilon}{3} + \frac{10\epsilon^2}{9} + \cdots, \qquad \text{so that} \qquad \hat{s}_\star = 1 - 3\epsilon - \frac{28\epsilon^2}{9} + \cdots. \qquad (4.15)$$

With Re as defined in (4.13), it follows that for large Re

$$s_\star = \frac{\Lambda}{2}\hat{s}_\star \approx \frac{\Lambda}{2}\left[1 - 3\left(\frac{2N^2}{\Lambda^2 Re}\right)^{1/3}\right], \qquad m_\star \approx \left(\frac{N Re}{2\Lambda}\right)^{1/3}. \qquad (4.16)$$

If again we also make the hydrostatic approximation and ignore horizontal diffusion, we must replace $|f Q| + N^2$ with N^2 in (4.9) and (4.11). The correction to the maximum growth rate in $s_\star$ is then proportional to $(N^2/Re)^{1/2}$. But, for the β-plane flow the correction scales with $(N^2/Re)^{1/3}$. For our f-plane flow we find $m_\star \propto (N^2 Re)^{1/4}$, whereas on the β-plane $m_\star \propto (N Re)^{1/3}$. Not surprisingly, in either case the correction to the maximum inviscid growth rate is proportional to $\nu m_\star^2$.

4.3. *General case: arbitrary Prandtl number*

The question of how the vertical scale selection is affected by varying P is now considered. For our uniform shear flow on the f-plane we assume that when $P \neq 1$, for large Re the fastest growing mode is associated with a positive real root of (3.1), i.e. with stationary instability. Maximal $s = s_\star$ and corresponding $a = a_\star$ (or $m_\star$) can be determined by simultaneously solving the cubic (3.1) and the equation that follows by taking the $\partial/\partial a$-derivative of (3.1) and equating the result to zero after setting $\partial s/\partial a = 0$. We seek power series solutions for $s = s_\star$ and $a = a_\star$ of the form

$$s_\star = |f Q|^{1/2}[\hat{s}_0 + \hat{s}_1\epsilon + \hat{s}_2\epsilon^2 + \cdots], \quad a_\star = \epsilon^\alpha|f Q|[\hat{a}_0 + \hat{a}_1\epsilon + \hat{a}_2\epsilon^2 + \cdots], \qquad (4.17)$$

with $\epsilon = Re^{-1/2}$. Setting $l = 1$, we find $\alpha = 1$, $\hat{s}_0 = 1$, $\hat{a}_0 = \mathscr{C}$, $\hat{s}_1 = -2\mathscr{C}$, $\hat{a}_1 = \frac{1}{2}\mathscr{C}^2$ and

$$\hat{s}_2 = -\frac{1}{2}\mathscr{C}^2 + \left(\frac{\mathscr{C}^2}{P} - \mathscr{D}^2\right), \qquad \hat{a}_2 = \frac{3\mathscr{C}^3}{8} + \frac{1+P}{2P^2}[P\mathscr{C}\mathscr{D}^2 - \mathscr{C}^3]. \qquad (4.18)$$

$\mathscr{C}$ is defined in (4.3) and

$$\mathscr{D} = \left(\frac{\pi^2}{2}\right)^{1/2}\frac{(|f Q| + 1)^{1/2}(|f Q|/P + N^2)^{1/2}}{|f Q|^{3/4}}. \qquad (4.19)$$

For large Re we obtain

$$s_\star = |f Q|^{1/2}\left[\underbrace{1 - \overbrace{\frac{2\mathscr{C}}{Re^{1/2}}}^{\text{first-order}} + \frac{\hat{s}_2}{Re}}_{\text{second-order}} + O(Re^{-3/2})\right], \qquad (4.20)$$

and

$$m_\star^2 = \frac{|f Q|^{1/2}}{\pi^2(|f Q| + 1)} \left[\overbrace{\mathscr{C} Re^{1/2} + \underbrace{\frac{1}{2} \mathscr{C}^2 + \frac{\hat{a}_2}{Re^{1/2}}}} {}^{\text{first}-\text{order}} + O(Re^{-1}) \right] - 1. \quad (4.21)$$

If we assume that the fastest growth is associated with the real part $\mathrm{Re}\,s = b$ of complex conjugate roots $s = b \pm i\omega$ (oscillatory instability), we can proceed in the same fashion with the cubic for b, which is given in Appendix A (equation (A4)). This yields a maximal $b = b_\star$, with

$$b_\star \approx |f Q|^{1/2} \left[\frac{1}{2} - \left(\frac{1 + P}{P} \right)^{1/2} \frac{\mathscr{C}}{Re^{1/2}} \right], \quad (4.22)$$

and corresponding

$$m_\star^2 \approx \frac{|f Q|^{1/2}}{\pi^2(|f Q| + 1)} \left[\left(\frac{P}{1 + P} \right)^{1/2} \mathscr{C} Re^{1/2} + \frac{(P^2 - 4P - 1)\mathscr{C}^2}{(1 + P)^2} \right] - 1. \quad (4.23)$$

This would imply that for large Re, $s_\star$ exceeds $b_\star$. We therefore expect that the fastest growing modes are associated with stationary instability when Re becomes sufficiently large. Further, comparing (4.20) with (4.8), we see that the maximum growth rates start to differ with the $\hat{s}_2/Re$ term when $P \neq 1$. Similarly, $m_\star$ differs from (4.10) beginning with the $\hat{a}_2/Re^{1/2}$ term when $P \neq 1$. Thus, $s_\star$ and $m_\star$ converge for large Re to the values found for $P = 1$, i.e. (4.9) and (4.11), respectively. Since the specific value of P does not affect the leading-order behaviour, for large Re the scale selection and maximum growth rate are not sensitive to the value of the Prandtl number.

These predictions can easily be tested by numerically searching for the maximal growth rates as a function of m. An example is shown in figure 5. This is representative of all cases with $P > P_c$ and $|f Q| < |f Q|_i$. The marginal stability boundary is as in figure 3. When Re crosses the overstability boundary $Re_c^{(o)}$, oscillatory instability commences, associated with complex-conjugate roots of the cubic for which $BC - D < 0$ and $D > 0$ (table 1). The fastest growth is for the vertical wavenumber $m_\star$ (thick dashed line in figure 5a) with the fastest growth rate $s_\star$ being the real part $b_\star$ of a pair of complex-conjugate roots (thick dashed line in figure 5b). As Re is further increased, the maximum growth rate becomes associated with a positive real root (stationary instability). There is a jump in $m_\star$ to a higher value. For ever increasing Re, the fastest growth remains associated with a positive real root (thick solid lines in figure 5a, b). The discontinuous behaviour at $Re \approx 1.8 \times 10^4$ is not related to crossing the convection boundary $Re_c^{(c)}$ as we move vertically in figure 3 to the left of $|f Q|_i$. In this example $Re_c^{(c)} \approx 3.6 \times 10^4$, and the positive real root remains associated with $BC - D < 0$ and $D > 0$. We are not able to predict analytically at what Reynolds number the jump will occur.

Also shown in figure 5 are the first- and second-order approximations to $m_\star$ and $s_\star$, given in (4.21) and (4.20), respectively. These approximations are not good at the lower Reynolds numbers when the fastest growth is associated with oscillatory instability. For larger Re, when the maximum growth is associated with stationary instability, they become excellent.

Since instability always sets in through overstability when $P > P_c$ and $|f Q| < |f Q|_i$, the fastest growing modes for Reynolds numbers slightly above $Re_c^{(o)}$ must be associated with a pair of complex-conjugate roots. In that case we expect that
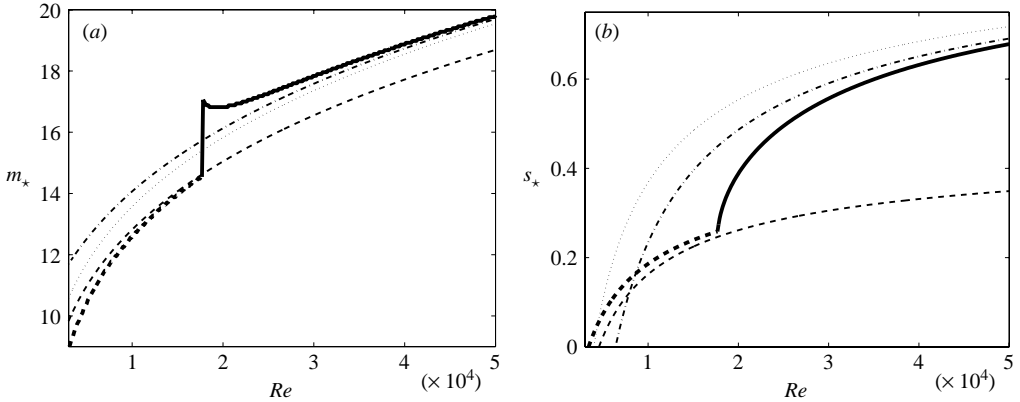
FIGURE 5. (*a*) Numerically determined $m_\star$ as a function of $Re \in [3000, \ 50,000]$ for $|fQ| = 1$, $N^2 = 100$ and $P = 7$ (thick dashed/solid) and (*b*) the corresponding maximum growth rate $s_\star$ (thick dashed/solid). In both panels thick solid lines indicate stationary instability and thick dashed lines oscillatory instability. Growth commences at $Re = Re_c^{(o)} \approx 3400$ with $m_c \approx 9.3$. This example is representative of all cases with $P > P_c$ and $|fQ| < |fQ|_i$. Also shown in (*a*) is the the first-order approximation (thin dotted) and the second-order approximation (dash-dotted) according to (4.21). In (*b*) the thin dotted line is the first-order approximation to $s_\star$ and the dash-dotted line the second-order approximation, given by (4.20). In (*a*) the thinner dashed line is the first-order approximation (4.23) and in (*b*) the first-order approximation to $b_\star$ given by (4.22).

$m_\star$ will, to leading order, be determined by (4.23), and the fastest growth rate by $b_\star$ as in (4.22). This is also verified in figure 5.

When $P > P_c$, but $|fQ| > |fQ|_i$, instability will set in through stationary convection when $Re = Re_c^{(c)}$ (see figure 3). For slightly larger Re, the fastest growth is associated with a positive real root and wavenumber which renders $D < 0$. For further increasing Re, stationary instability persists. In all cases we looked at, including extremely large and small $|fQ|$, N^2, we find essentially the situation as in figure 4: the first-order approximations are good except near $Re \approx Re_c^{(c)}$, while the second-order approximations are for all Re virtually indistinguishable from the exact, numerically determined values.

What remains are cases with $P < P_c$, when instability will always set in through stationary convection (see figure *b*). For $P \approx 1$, again the first- and second-order approximations are good and excellent, respectively. For $P \ll 1$, we find generally that for Re near the critical value, the first-order approximation is quite good but the second-order approximation not good at all. An example is shown in figure 6. The first-order approximations to $m_\star$ and $s_\star$ are close to the calculated values, but near $Re = Re_c^{(c)} \approx 100$, the second-order approximations are dismal. This is because for $P \ll 1$, the coefficients $\hat{a}_2$ and $\hat{s}_2$ (4.18) become large compared to $\hat{a}_1$ and $\hat{s}_1$, respectively.

5. Summary and discussion

We have studied a simple model of a barotropic shear flow that is inertially unstable in a region of uniform anticyclonic shear. This region of constant $fQ < 0$ is artificially terminated at free-slip sidewalls. The analytical simplicity of the problem studied in this paper allowed us in §3 to determine the exact critical Reynolds number required
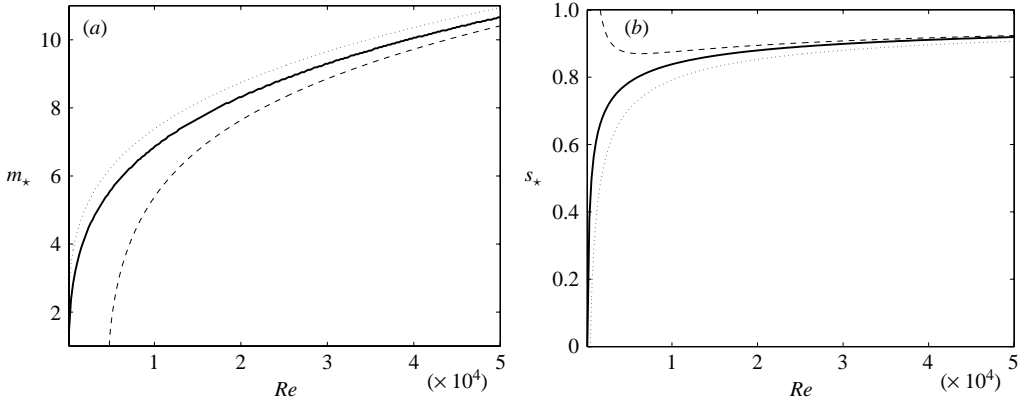
FIGURE 6. (a) $m_\star$ as a function of $Re \in [100, 50,000]$ for $|fQ| = 1$, $N^2 = 10$ and $P = 0.1$ (thick line), the first-order approximation (dotted line) and the second-order approximation (dashed line). (b) The corresponding maximum growth rate $s_\star$ (solid), the first-order approximation (dotted line) and the second-order approximation (dashed line). Growth commences for $Re = Re_c^{(c)} \approx 100$ with $m = m_c \approx 1.4$.

for the onset of instability as a function of the ambient buoyancy frequency N, the Prandtl number P and the magnitude of the anticyclonic shear, measured by $|fQ|$. Since the dispersion relation (3.1) is the same as for stratified Taylor–Couette flow in the narrow-gap approximation and free-slip boundary conditions, these results can be seen as a precise description of the marginal stability criteria for that problem. In the pioneering study of Thorpe (1966) and in, for example, Caton, Janiaud & Hopfinger (2000), only asymptotic results are reported, i.e. for when viscosity becomes very small.

We can now estimate that, for example, if $P = 1$ the critical Reynolds number $Re = Re_c^{(c)} \approx 560$ and $m_c \approx 4$, when $N^2/f^2 = 10$ and $|fQ| = 1$ (Rossby number $Ro = -\Lambda/f = -2$). If $N^2 = 10^4 f^2$ (a value often quoted as typical for the oceans), this increases to $Re_c^{(c)} \approx 5 \times 10^5$ and $m_c \approx 120$. With $P = 7$, instability for $|fQ| = 1$ sets in through overstability (because $|fQ| < |fQ|_i \approx 4 \times 10^4$) at $Re = Re_c^{(o)} \approx 3 \times 10^5$ with $m = m_c \approx 90$. For smaller $|fQ|$ (weaker shear), these critical numbers become even larger. Thus, for such flows extremely high Reynolds numbers would be required to observe any incipient amplification of the inertial instability. Also in such cases with $m_c \approx 10^2$, in numerical simulations rather high resolutions in the vertical would be required to resolve the overturning motions. Certainly two or three grid points per cell with height $\Delta z = L/m_c$ would not be sufficient and one would have to use 10^3 grid points in the vertical. But, m_c gives the total number of cells of overturning motions found over a total depth H equal to the width L of the unstable region, because we used a vertical length scale $H = L$ from the outset. If the barotropic shear is confined to, say, a total depth $H = L/10$, and $m_c = 100$, the number of cells would be $H/L \times m_c = 10$. Thus with an aspect ratio $H/L = 1/10$, the number of grid points in the vertical can be reduced by a factor of 10.

In §4.1 we calculated the vertical wavenumber $m_\star$ of the fastest growing modes and the growth rate $s_\star$, when Re exceeds the critical value for instability and $P = 1$. For large Re, we predict that $m_\star \propto Re^{1/4}$, while the correction to the maximum inviscid growth rate $s_\star$ is predicted to be proportional to $Re^{-1/2}$. This disagrees with the equatorial β-plane predictions of Griffiths (2003a), as noted in §4.2. But, the marginal

stability criteria for our f-plane flow and the β-plane flow are, apart from constants of proportionality, the same.

The correction to the maximum inviscid growth rate in $s_\star$ for the β-plane flow scales with $Re^{-1/3}$, while $m_\star \propto Re^{1/3}$. These β-plane results compare well with the findings of Kloosterziel *et al.* (2007). They studied the scale selection in inertially unstable barotropic vortices in a rotating and uniformly stratified environment, taking $P = 1$. Through numerical simulations, the vertical scale and growth rate of the fastest growing perturbations were measured in the early, linear stage of the unfolding of the instability. Both N/f and Re were varied, but owing to numerical constraints both parameters were confined to modest ranges. To a fair degree of precision, it was found that $m_\star \propto Re^{1/3}$, with a viscous correction to the inviscid maximum growth rate proportional to $Re^{-1/3}$. This agreement with the, at first sight, seemingly unrelated predictions of Griffiths (2003*a*) is not a coincidence, as has become clear while preparing the revision of this manuscript. In Appendix B (by S. D. Griffiths), it is shown that for parallel shear flows with *variable* fQ, generally for large Re and corresponding large $|m|$, the same scalings as for the uniform shear flow on the β-plane are expected. The reason is that in the limit of large Re, the eigenvalue problems converge to the canonical form of the quantum harmonic oscillator equation. A similar asymptotic analysis for circular vortices has as yet not been performed. But, it seems unlikely that the scaling laws will be different.

The asymptotic analysis in Appendix B further suggests that, when $P \neq 1$, for sufficiently large Re the maximal growth rates and the vertical scale selection become weakly dependent on the value of P. This was explicitly shown in §4.3 for our f-plane flow with $fQ = $ constant. But, the scaling laws we have found, such as $m_\star \propto Re^{1/4}$, cannot be inferred from an analysis similar to that in Appendix B, which predicts $m_\star \propto Re^{1/3}$. The reason is that for flows with variable fQ and a negative minimum $\overline{fQ}$ at $y = \bar{y}$, the fastest growing normal modes have a structure that varies with y as $\exp(-c|m|(\overline{fQ''})^{1/2}(y - \bar{y})^2/2)$, with c a positive function of N and $\overline{fQ}$, and $\overline{fQ''} = \mathrm{d}^2(fQ)/\mathrm{d}y^2$ at $y = \bar{y}$, which is of course positive if there is a simple negative minimum. Hence, for increasing $|m|$, the modes become ever more confined to a region about $y = \bar{y}$. This 'shrinking' of the modes in the cross-stream (y) direction does not occur if fQ has a constant region where its value is minimal and negative, since then $\overline{fQ''} = 0$.

If more realistic unbounded flows with, say, uniform shear in a transition region between two counterflowing currents are considered, an analysis is still feasible, in particular when the hydrostatic approximation is made and horizontal diffusion is ignored. In the simplest possible model, with constant $fQ < 0$ in a strip of finite width, and $fQ = f^2$ outside this region, the eigenvalue problem which determines the relation between the cross-stream (y) structure of the normal modes and the vertical wavenumber, growth rate, etc., is essentially that encountered in quantum mechanics when solutions of the Schrödinger equation for a finite-depth square potential well are sought. The equivalent eigenvalue problem corresponding to the bounded flow studied in this paper is that of the Schrödinger equation for an infinitely deep square potential well. For the simple unbounded flow, the eigenmodes have a y-dependent part $\tilde{\psi}(y)$, which is non-zero outside the unstable strip, i.e. $\tilde{\psi}(y) \neq 0$ where $fQ = f^2$ (the stable region). But $|\tilde{\psi}(y)|$ decays exponentially in the stable region, and ever more rapidly with increasing Re and vertical wavenumber m. Hence, $\tilde{\psi}(y)$ becomes more and more confined to the unstable region where there is constant $fQ < 0$, but never shrinks to zero width. For large $|m|$, the modes differ little from the sinusoidal mode with $l = 1$ that we find for the bounded flow. A detailed comparison of the results

presented in this paper with the marginal stability criteria and scale selection for such an unbounded flow is left for future work. If the hydrostatic approximation is *not* made and horizontal diffusion is *not* ignored, instead of a Schrödinger-like equation $d_y^2 \tilde{\psi}(y) - F(\cdots)\tilde{\psi}(y) = 0$, equations follow containing the sixth-order derivative (when $P = 1$), or the eighth-order derivative (when $P \neq 1$). In view of the discussion in §4.2, we expect little to be gained from a study of these far more complicated eigenvalue problems, unless $N^2 \ll |fQ|$ (when $P = 1$) or $PN^2 \ll |fQ|$ (when $P \neq 1$).

Appendix A. Bounds on growth rates

The cubic (3.1) can be written as

$$\left(s + \frac{a}{P}\right)[(s+a)^2 - |fQ|] + \left(s + \frac{a}{P}\right)\frac{\pi^2 l^2 E}{a}|fQ| + (s+a)\frac{\pi^2 l^2 E}{a}N^2 = 0. \quad \text{(A 1)}$$

If there is a positive real root $s > 0$, then $(s+a)^2 - |fQ|$ has to be negative, since E, P, l^2, a and $|fQ|$ are all positive and $N^2 \geqslant 0$. Hence we must have $s < |fQ|^{1/2}$. There can also be complex-conjugate pairs of roots. Setting $s = b \pm i\omega$ in (A 1), the real part is

$$\left(b + \frac{a}{P}\right)\left[(b+a)^2 - \omega^2 - |fQ| + \frac{l^2\pi^2 E}{a}|fQ|\right] + (b+a)\left(\frac{l^2\pi^2 E}{a}N^2 - 2\omega^2\right) = 0, \quad \text{(A 2)}$$

and the imaginary part is

$$\pm i\omega\left[(b+a)^2 - \omega^2 - |fQ| + \frac{l^2\pi^2 E}{a}|fQ| + 2\left(b + \frac{a}{P}\right)(b+a) + \frac{l^2\pi^2 E}{a}N^2\right] = 0. \quad \text{(A 3)}$$

Since by assumption $\omega \neq 0$, an equation $\omega^2 = \cdots$ follows. Substitution in (A 2) yields

$$(b+a)[(b+a)^2 - |fQ|] + (b+a)\left[\frac{|fQ|}{a} + \frac{N^2}{2a}\right]l^2\pi^2 E$$

$$+ \left(b + \frac{a}{P}\right)\left[\left(b + \frac{a}{P}\right)(b+a) + 2(b+a)^2 + \frac{l^2\pi^2 E}{2a}N^2\right] = 0. \quad \text{(A 4)}$$

If $b > 0$, we see that $(b+a)^2 - |fQ|$ has to be negative. Therefore, if there are complex-conjugate roots with real part $\text{Re}s = b > 0$, then $b < |fQ|^{1/2}$.

Appendix B. Weakly diffusive scale selection for the inertial instability of an arbitrary shear flow

By Stephen D. Griffiths

Department of Physics, University of Toronto, ON, M5R 2S4, Canada

We start from equations (2.1)–(2.5), where we now allow U to vary arbitrarily with y. Taking N^2 to be constant, the homogeneity in z allows us to consider disturbances

of the form

$$(u', \psi, \sigma) = \mathrm{Re}[(\tilde{u}(y), \tilde{\psi}(y), \tilde{\sigma}(y)) \exp(st) \exp(\mathrm{i}mz)],$$

where ψ is the streamfunction (2.9), and $\sigma = -g\rho'/\rho_0$ is a fluctuating buoyancy acceleration. Differentiating (2.7) with respect to t, substituting from (2.1) and (2.4), and dividing by m^2 gives

$$\frac{(N^2 + s^2)}{m^2} \frac{\mathrm{d}^2\tilde{\psi}}{\mathrm{d}y^2} - (s^2 + fQ)\tilde{\psi} = \frac{\nu}{m^2} \left(\frac{\mathrm{d}^2}{\mathrm{d}y^2} - m^2 \right) \left[s \left(\frac{\mathrm{d}^2}{\mathrm{d}y^2} - m^2 \right) \tilde{\psi} + \mathrm{i}fm\tilde{u} + \frac{\kappa}{\nu} \frac{\mathrm{d}\tilde{\sigma}}{\mathrm{d}y} \right],$$

$$(B\,1)$$

where $Q(y) = f - \mathrm{d}U/\mathrm{d}y$ is the absolute vorticity.

B.1. *The non-diffusive system*

When $\nu = \kappa = 0$, it is well-known that instability is only possible when $fQ < 0$ somewhere in the flow. Denoting the cross-stream location of the global minimum of fQ by $\overline{y}$, and its value by $\overline{fQ} < 0$, the maximum growth rate of disturbances is given by $|\overline{fQ}|^{1/2}$. This can be established by multiplying (B 1) by $\tilde{\psi}^*$ and integrating across the domain, after setting the right-hand side of (B 1) to zero. Furthermore, this growth rate is approached as $|m| \to \infty$, by disturbances which become highly localized in the cross-stream direction around $y = \overline{y}$. The cross-stream localization occurs on a length scale $\sim |m|^{-1/2}$, for both unstratified flows (e.g. Bayly 1988) and stratified flows (e.g. Griffiths 2007). Thus, the lateral boundary conditions – whether they be for periodic, bounded or unbounded flow – become irrelevant in this limit.

To study these most unstable modes, it is necessary to introduce a rescaled cross-stream coordinate Y, the optimal choice being

$$Y = (a|m|)^{1/2}(y - \overline{y}), \qquad \text{where} \qquad a = \left(\frac{\overline{(fQ)''}}{2(N^2 + |\overline{fQ}|)} \right)^{1/2}, \qquad (B\,2)$$

and where $\overline{(fQ)''} = \mathrm{d}^2(fQ)/\mathrm{d}y^2$ evaluated at $y = \overline{y}$ is assumed to be non-zero. The limiting dynamics are obtained by making a Taylor expansion for fQ around $y = \overline{y}$ (justifiable because of the cross-stream localization of the modes), and by looking for small deviations of s^2 from its limiting value:

$$fQ \sim \overline{fQ} + \frac{\overline{(fQ)''}}{2a|m|}Y^2 + \frac{\overline{(fQ)'''}}{6(a|m|)^{3/2}}Y^3 + \cdots, \quad s^2 \sim |\overline{fQ}| - \left(\frac{a}{|m|} \right) s_1^2 + \cdots, \quad |m| \to \infty.$$

Substituting in (B 1), and writing $\mathrm{d}^2/\mathrm{d}y^2 = a|m|\mathrm{d}^2/\mathrm{d}Y^2$, the $O(1)$ terms cancel, and at $O(|m|^{-1})$ we have

$$(N^2 + |\overline{fQ}|)\frac{a}{|m|} \left[\frac{\mathrm{d}^2\tilde{\psi}}{\mathrm{d}Y^2} + \left(\frac{s_1^2}{N^2 + |\overline{fQ}|} - Y^2 \right)\tilde{\psi} \right] = 0. \qquad (B\,3)$$

This equation does indeed describe localized solutions, provided

$$s_1^2 = (2n + 1)(N^2 + |\overline{fQ}|), \qquad n = 0, 1, 2, \ldots,$$

with corresponding eigenfunctions $\tilde{\psi} = H_n(Y) \exp(-Y^2/2)$, where H_n is the nth order Hermite polynomial. Thus, the limiting growth rate is given by

$$s^2 \sim |\overline{fQ}| - \frac{(2n + 1)}{|m|} \left(\frac{\overline{(fQ)''}(N^2 + |\overline{fQ}|)}{2} \right)^{1/2} + \cdots, \quad |m| \to \infty. \qquad (B\,4)$$

Although higher-order corrections to s^2 and $\tilde{\psi}$ may be derived, (B 4) is sufficient for our purposes since it captures the inviscid monotonic growth rate increase as $|m| \to \infty$. From (2.1) and (2.4), the remainder of the leading-order solution is

$$\tilde{u} \sim -\frac{im\overline{Q}}{|\overline{fQ}|^{1/2}}\tilde{\psi}, \qquad \tilde{\sigma} \sim -\frac{N^2(a|m|)^{1/2}}{|\overline{fQ}|^{1/2}}\frac{\mathrm{d}\tilde{\psi}}{\mathrm{d}Y}. \qquad (\text{B } 5a, b)$$

B.2. *The diffusive system*

We now evaluate the effect of diffusion on the inertial instability when the Prandtl number $P = \nu/\kappa$ is of order unity. We study the regime where ν and κ are small, so that the inertial instability remains strong, with the large vertical wavenumber limit remaining appropriate. We anticipate that (B 4) will be appropriate for moderately large $|m|$, describing an almost inviscid growth rate increase, whilst as $|m| \to \infty$ more appropriate would be a relationship describing an almost completely diffusive decay. We need to assess for what scaling of $|m|$, in terms of ν and κ, the inviscid growth and diffusive decay terms balance, which will yield the most unstable modes.

Since the inviscid growth term of (B 4) appears at $O(|m|^{-1})$ in (B 1), we simply need to calculate when the terms on the right-hand side of (B 1) scale like $|m|^{-1}\tilde{\psi}$, as $|m| \to \infty$. Using (B 5a, b), the largest terms on the right-hand side of (B 1) as $|m| \to \infty$ are $s\nu m^2\tilde{\psi}$ and $-i\nu f m\tilde{u}$, and are those originating from vertical diffusion in the two horizontal momentum equations. Using (B 4) and (B 5a), to leading order these two terms sum to

$$2\nu m^2|\overline{fQ}|^{1/2}\tilde{\psi}. \qquad (\text{B } 6)$$

Using (B 5b), the term $-\kappa\,\mathrm{d}\tilde{\sigma}/\mathrm{d}y$ (originating from the vertical diffusion of heat) scales like $\kappa m\tilde{\psi}$, and is much smaller than (B 6), as are all terms originating from horizontal diffusion $(\partial^2/\partial z^2 \gg \partial^2/\partial y^2$ under the scaling (B 2)). Thus the regime of interest occurs when (B 6) scales like $|m|^{-1}\tilde{\psi}$, i.e. when $|m| \sim \nu^{-1/3}$. With this understanding, (B 6) may be added to the right-hand side of (B 3), yielding a modified solvability condition

$$s^2 \sim |\overline{fQ}| - \frac{(2n+1)}{|m|}\left(\frac{\overline{(fQ)''}(N^2 + |\overline{fQ}|)}{2}\right)^{1/2} - 2\nu m^2|\overline{fQ}|^{1/2} + \cdots, \qquad (\text{B } 7)$$

where $|m| \to \infty$, $\nu \to 0$, and $|m| \sim \nu^{-1/3}$. This is a self-consistent procedure, since $\nu m^2 \sim \kappa m^2 \sim |m|^{-1} \ll 1$, so that the approximately non-diffusive balances leading to (B 5a, b) remain valid.

Equation (B 7) describes the diffusive modification of the inertial instability. Although not valid for arbitrary ν and $|m|$, it should describe the most unstable inertial instabilities as $\nu \sim \kappa \to 0$. Maximizing s^2 with respect to $|m|$, the maximum growth rate s_* satisfies

$$s_*^2 \sim |\overline{fQ}| - \frac{3}{2}\{2\nu(2n+1)^2|\overline{fQ}|^{1/2}\overline{(fQ)''}(N^2 + |\overline{fQ}|)\}^{1/3} + \cdots, \quad \nu \to 0, \qquad (\text{B } 8)$$

which is largest for the $n = 0$ mode. The most unstable vertical wavenumber m_* satisfies

$$|m_*| \sim \left\{\frac{2n+1}{4\nu}\left(\frac{\overline{(fQ)''}(N^2 + |\overline{fQ}|)}{2|\overline{fQ}|}\right)^{1/2}\right\}^{1/3} + \cdots, \quad \nu \to 0. \qquad (\text{B } 9)$$

These scalings, with s_* decreasing from $|\overline{fQ}|^{1/2}$ like $\nu^{1/3}$ and $m_* \sim \nu^{-1/3}$, are consistent with the particular cases reported by Griffiths (2003a), and Kloosterziel *et al.* (2007). There is disagreement with the results of §4.1 because it is a case for which $\overline{(fQ)''} = 0$.

The scalings apply when $\nu \neq \kappa$ (unless $\nu \ll \kappa$), but the effects of κ on the inertial instability are rather weak in the large vertical wavenumber regime, as was noted in §4.3 for the uniform shear flow. The terms involving κ appear as an $O(|m|^{-2})$ correction to the growth rate in (B 7), and are not calculated here.

Less cumbersome scalings for m_* can be derived from (B 9) from case to case. We denote the length scale of fQ by L, and take $Q \sim f$, since for instability horizontal shears must be larger than f. Thus:

(*a*) For unstratified flow on the f-plane, (B 9) implies $m_* \sim (f/\nu L)^{1/3}$. Since L must also be the length scale for U, and f is a characteristic shear, a characteristic velocity for the basic flow is just fL, so that we can write the Reynolds number as $Re = fL^2/\nu$, implying $m_* \sim Re^{1/3}L^{-1}$.

(*b*) For hydrostatic flow ($N \gg f$) on the f-plane, (B 9) implies $m_* \sim (N/\nu L)^{1/3} \sim (N/f)^{1/3}Re^{1/3}L^{-1}$. The characteristic vertical wavenumbers are much larger for this hydrostatic case than for the unstratified case, since stratification acts to reduce the vertical extent of the overturning circulations.

(*c*) For hydrostatic flow ($N \gg f$) on the equatorial β-plane, the length scale L of fQ is no longer that of U. Rather, denoting a characteristic shear by Λ, $fQ \approx \beta y(\beta y - \Lambda)$ so that $L \sim \Lambda/\beta$, and $Q \sim f \sim \beta L \sim \Lambda$. Thus $m_* \sim (N\beta/\Lambda\nu)^{1/3} \sim (N/\Lambda)^{1/3}Re^{1/3}L^{-1}$, consistent with the scaling given by Griffiths (2003*a*), as discussed in §4.2.

REFERENCES

BAYLY, B. J. 1988 Three dimensional centrifugal-type instabilities in inviscid two-dimensional flows. *Phys. Fluids* **31**, 56–64.

BILLANT, P. & GALLAIRE, F. 2005 Generalized Rayleigh criterion for non-axisymmetric centrifugal instabilities. *J. Fluid Mech.* **542**, 365–379.

CATON, F., JANIAUD, B., & HOPFINGER, E. J. 2000 Stability and bifurcations in stratified Taylor–Couette flow. *J. Fluid Mech.* **419**, 93–114.

CHANDRASEKHAR, S. 1961 *Hydrodynamic and Hydromagnetic Stability*. Oxford University Press.

CHARNEY, J. G. 1973 Lecture notes on planetary fluid dynamics. *Dynamic Meteorology* (ed. P. Morel). Reidel.

DRAZIN, P. & REID, W. 1981 *Hydrodynamic Stability*. Cambridge University Press.

DUNKERTON, T. J. 1981 On the inertial instability of the equatorial middle atmosphere. *J. Atmos. Sci.* **38**, 2354–2365.

DUNKERTON, T. J. 1982 The double-diffusive modes of symmetric instability on an equatorial beta-plane. *J. Atmos. Sci.* **39**, 1653–1657.

EMANUEL, K. A. 1979 Inertial instability and mesoscale convective systems. Part I: Linear theory of inertial instability in rotating viscous fluids. *J. Atmos. Sci.* **36**, 2425–2449.

GRIFFITHS, S. D. 2003*a* Nonlinear vertical scale selection in equatorial inertial instability. *J. Atmos. Sci.* **60**, 977–990.

GRIFFITHS, S. D. 2003*b* The nonlinear evolution of of zonally symmetric equatorial inertial instability. *J. Fluid Mech.* **474**, 245–273.

GRIFFITHS, S. D. 2007 The limiting form of inertial instability in geophysical flows. Submitted to *J. Fluid Mech.*

HOSKINS, B. J. 1974 The role of potential vorticity in symmetric stability and instability. *Q. J. R. Met. Soc.* **100**, 480–482.

KLINE, M. 1972 *Mathematical Thought from Ancient to Modern Times*. Oxford University Press.

KLOOSTERZIEL, R. C. & CARNEVALE, G. F. 2003 Closed-form linear stability conditions for rotating Rayleigh-Bénard convection with rigid, stress-free upper and lower boundaries. *J. Fluid Mech.* **480**, 25–42.

KLOOSTERZIEL, R. C. & CARNEVALE, G. F. 2007 Generalized energetics for inertially stable parallel shear flows. *J. Fluid Mech.* **585**, 117–126.

KLOOSTERZIEL, R. C., CARNEVALE, G. F. & ORLANDI, P. 2007 Inertial instability in rotating and stratified fluids: barotropic vortices. *J. Fluid Mech.* **583**, 379–412.

McIntyre, M. E. 1970 Diffusive destabilization of the baroclinic circular vortex. *Geophys. Fluid Dyn.* **1**, 19–58.

Ooyama, K. 1966 On the stability of the baroclinic circular vortex: a sufficient condition for instability. *J. Atmos. Sci.* **23**, 43–53.

Rayleigh, Lord 1916 On the dynamics of revolving fluids. *Proc. R. Soc. Lond.* A **93**, 148–154.

Smyth, W. D. & McWilliams, J. C. 1998 Instability of an axisymmetric vortex in a stably stratified, rotating environment. *Theor. Comput. Fluid Dyn.* **11**, 305–322.

Smyth, W. D. & Peltier, W. R. 1994 Three-dimensionalization of barotropic vortices on the f-plane. *J. Fluid Mech.* **265**, 25–64.

Stone, P. H. 1966 On non-geostrophic stability. *J. Atmos. Sci.* **23**, 390–400.

Thorpe, S. A. 1966 The stability of stratified Couette flow. In *Notes on 1966 Summer Geophys. Fluid Dyn.*, pp. 80–107. Woods Hole Oceanographic Inst.