J. Fluid Mech. (1988), *vol.* 197, *pp.* 349–388
Printed in Great Britain

On the instability of geostrophic vortices

By GLENN R. FLIERL

Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

(Received 26 January 1988 and in revised form 23 May 1988)

The instabilities of barotropic and baroclinic, quasi-geostrophic, f-plane, circular vortices are found using a linearized contour dynamics model. We model the vortex using a circular region of horizontally uniform potential vorticity surrounded by an annulus of uniform, but different, potential vorticity. We concentrate mostly upon isolated vortices with no circulation in the basic state outside the outer radius b. In addition to linear analyses, we also consider weakly nonlinear waves. The amplitude equation has a cubic nonlinearity and, depending upon the sign of the coefficient of the cubic term, may give nonlinear stabilization or nonlinear enhancement of the growth. Barotropic isolated eddies are unstable when the outer annulus is narrow enough; on the other hand, if the scale of the whole vortex is sufficiently small compared to the radius of deformation of a baroclinic mode, the break up may be preferentially to a depth-varying disturbance corresponding to a twisting and tilting of the vortex. As the vortex becomes more baroclinic, we find that large-scale vortices show an elliptical mode baroclinic instability as well which is relatively insensitive to the scale of the outer annulus. When the baroclinic currents in the basic state dominate, the twisting mode disappears, and we see only the instabilities associated with either strong enough shear in the annular region or sufficiently large vortices compared with the deformation radius. The finite amplitude results show that the baroclinic instability mode for large enough vortices is nonlinearly stabilized while in most cases, the other two kinds of instability are nonlinearly destabilized.

1. Introduction

Transient motions in geophysical flows often have the character of relatively strong vortices embedded in weaker more turbulent eddy fields. Gulf Stream rings, blocking highs, and the 'spots' on Jupiter all seem to have strong recirculating flows and closed streamlines. In addition, the motions are often strongly baroclinic. In this paper, we consider the stability of circular barotropic and baroclinic vortices for quasi-geostrophic motions. We have investigated both the linear stability and the finite amplitude steady state perturbed configurations existing near the neutral stability boundary; the latter allows us to infer the behaviour as an unstable perturbation grows to finite (though still small) amplitude. We use the 'contour dynamics' model (Zabusky, Hughes & Roberts 1979) and derive the methods for applying this to linear and weakly nonlinear stability problems in a baroclinic fluid.

Our results show that isolated vortices (having $V = 0$ for $r > b$) have three instabilities:

(i) For vortices with sharp shear at the outer edge (a relatively narrow region in which the velocity decreases to zero – comparable in size to the distance from the

centre to the velocity maximum), an elliptical mode ($m = 2$) instability occurs. This tends to be subcritical and grows until the vortex breaks up into dipolar structures.

(ii) For vortices with enough baroclinic flow and which are larger than about twice the deformation radius with a fairly broad outer shear region, baroclinic instability is the dominant mechanism. The mode is elliptical and supercritical. At larger amplitudes, dipole formation occurs (Helfrich & Send, 1988).

(iii) When the deep flow is co-rotating with the upper flow, there is a 'twisting' mode instability involving tilting of the vortex axis and generation of more barotropic motions. This occurs for vortices smaller than the deformation radius with fairly broad regions in which the flow decreases to zero. The instability is usually subcritical.

Our results suggest that stable baroclinic vortices will have the velocity maximum at around 1.5 times the deformation radius with the outer shear region being wider than the radius to the maximum velocity region. This is quite consistent with the results of the warm core rings study (Joyce 1984; Olson *et al.* 1985). Our model lacks many elements, of course. The structure of the vortex is very simple, although the method can be easily extended to much more complex cases. Most of the geophysical situations mentioned above, of course, do involve background gradients of potential vorticity – the β-effect and perhaps sheared zonal flows. We do not consider this complication here; the steady vortex structures are much more complex (cf. Malanotte-Rizzoli 1982; Flierl 1987 for reviews) and the contour dynamics method does not easily handle large-scale gradients of potential vorticity. But we feel that the stability boundaries derived still give insight into the β-plane problem in the following sense. We expect the distortions induced by Rossby waves will trigger the instability if it exists. Otherwise, the vortex will propagate and evolve, shedding and entraining water with different potential vorticities. In essence, the β-effect, or, for that matter, external shears or eddies, provide a forcing upon the vortex. Its response depends upon the natural modes of oscillation of the vortex. For strong vortices ($U \gg \beta R^2$), then, we anticipate that the stability for the f-plane structure will still be an important property to estimate. Our study provides a thorough exploration of the parameter space for baroclinic vortices.

2. Linearized perturbations

We shall use the quasi-geostrophic f-plane equations with the assumption that the potential vorticity at each depth is piecewise constant. In the unperturbed state, the boundaries between various potential vorticity values ω_p will be taken to be at non-dimensional radii $r = 1$ and $r = b$. In the perturbed state, we have

$$q = \begin{cases} q_1(z) + q_b(z), & r < 1 + \eta(z), \\ q_b(z), & 1 + \eta(z) < r < b + \tau(z), \\ 0, & b + \tau(z) < r, \end{cases}$$

(see figure 1). Note that q_1 is defined to be the difference between the potential vorticity in the inner and middle regions and q_b is the difference between the middle and outer regions: the q terms represent potential vorticity 'gradients', not the potential vorticity values themselves. Adopting the contour dynamics approach, we express the streamfunction $\Psi(r, \theta, z, t)$ as a functional of the interface positions,

$$(\nabla^2 + L_z)\,\Psi = \left(\nabla^2 + \frac{\partial}{\partial z}\frac{f^2L^2}{N^2H^2}\frac{\partial}{\partial z}\right)\Psi = q_1(z)\,\mathscr{H}(1+\eta-r) + q_b(z)\,\mathscr{H}(b+\tau-r), \quad (2.1)$$

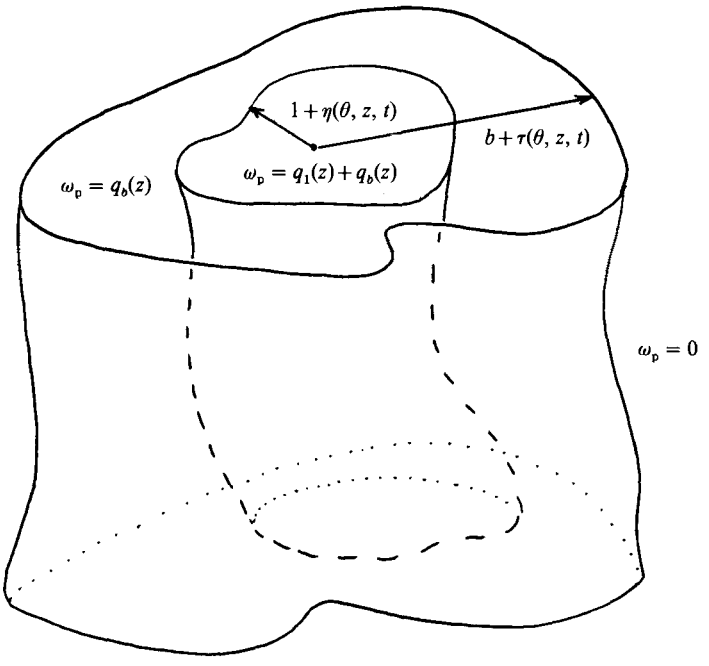

FIGURE 1. Sketch of the perturbed baroclinic vortex, showing the surfaces bounding the regions of horizontally uniform potential vorticity.

where $\mathscr{H}$ is the Heaviside stepfunction. L and H here are the dimensional radii of the unperturbed inner region and the fluid depth, respectively. The interfaces evolve according to the kinematic condition

$$\left.\begin{aligned}\frac{\partial}{\partial t}\eta &= -\frac{1}{1+\eta}\frac{\mathrm{d}}{\mathrm{d}\theta}[\Psi(1+\eta(\theta, z, t), \theta, z, t)],\\ \frac{\partial}{\partial t}\tau &= -\frac{1}{b+\tau}\frac{\mathrm{d}}{\mathrm{d}\theta}[\Psi(b+\tau(\theta, z, t), \theta, z, t)],\end{aligned}\right\} \tag{2.2}$$

where the total derivative with respect to θ of Ψ on the interface is used.

We can now consider a linearized version of the contour dynamics model: we split Ψ into a basic state $\bar{\psi}(r, z)$ satisfying

$$(\nabla^2+L_z)\bar{\psi} = q_1\mathscr{H}(1-r)+q_b\mathscr{H}(b-r), \tag{2.3}$$

and the perturbation

$$(\nabla^2+L_z)\psi = q_1[\mathscr{H}(1+\eta-r)-\mathscr{H}(b-r)]+q_b[\mathscr{H}(b+\tau-r)-\mathscr{H}(b-r)].$$

Linearizing by expanding the right-hand side gives

$$(\nabla^2+L_z)\psi = q_1\eta\delta(1-r)+q_b\tau\delta(b-r), \tag{2.4}$$

while the linearized kinematic equations are just

$$\left.\begin{aligned}\frac{\partial}{\partial t}\eta+\bar{V}(1)\frac{\partial}{\partial\theta}\eta &= -\psi_\theta(1, \theta, t),\\ \frac{\partial}{\partial t}\tau+\frac{\bar{V}(b)}{b}\frac{\partial}{\partial\theta}\tau &= -\frac{1}{b}\psi_\theta(b, \theta, t).\end{aligned}\right\} \tag{2.5}$$

Thus ψ is related to the Green function for the Laplacian operator in polar coordinates. Notice that the equation for $\bar{V}$, determined by differentiating (2.3),

$$\left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}-\frac{1}{r^2}+L_z\right)\bar{V}=-q_1\delta(1-r)-q_b\delta(b-r), \tag{2.6}$$

is very clearly related to the perturbation structure equation (2.4). We now assume ψ, η and τ are all proportional to $\exp(im(\theta-\Omega t))$ to find

$$\left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}-\frac{m^2}{r^2}+L_z\right)\psi=q_1\eta\delta(1-r)+q_b\tau\delta(b-r), \tag{2.7}$$

$$(\bar{V}(1)-\Omega)\eta=-\psi(1,z), \tag{2.8}$$

$$\left(\frac{\bar{V}(b)}{b}-\Omega\right)\tau=-\frac{1}{b}\psi(b,z). \tag{2.9}$$

Equations (2.6), (2.7)–(2.9) or their modally truncated (in the vertical) forms will be the basis for our analysis. The physics in these is quite straightforward: perturbations in the boundaries produce anomalies in potential vorticity near $r=1$ and $r=b$. The anomalies in turn produce flows which move the interfaces.

3. Barotropic vortex

For a barotropic vortex, subjected to barotropic perturbations, the linear results quoted below have been obtained previously (Michalke & Timme 1967; Childress 1984; Flierl 1984) although we will focus on somewhat different aspects. In particular, we wish to emphasize the instability properties of the 'isolated' vortex, having no net integrated vorticity. Gent & McWilliams (1986) have solved numerically for the instabilities of smoother barotropic circular eddies; the results here provide analytic examples confirming their findings.

When the basic flow is barotropic, q_1 and q_b independent of z, we can separate variables in (2.7)–(2.9) to find

$$\left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r}-\frac{m^2}{r^2}-\lambda^2\right)\psi=q_1\eta\delta(1-r)+q_b\tau\delta(b-r), \tag{3.1}$$

replacing (2.7). Here ψ, η, τ are proportional to one of the vertical eigenmodes $F(z)$ and λ^2 is the non-dimensional eigenvalue L^2/R_{d}^2 where R_{d} is the dimensional deformation radius for the mode in question. Solving the Green function equation (3.1) allows us to write $\psi(1)$, $\psi(b)$ as linear combinations of η, τ:

$$\begin{pmatrix}\psi(1)\\ \psi(b)/b\end{pmatrix}=-\begin{pmatrix}I_m(\lambda)K_m(\lambda) & bI_m(\lambda)K_m(\lambda b)\\ I_m(\lambda)K_m(\lambda b)/b & I_m(\lambda b)K_m(\lambda b)\end{pmatrix}\begin{pmatrix}q_1 & 0\\ 0 & q_b\end{pmatrix}\begin{pmatrix}\eta\\ \tau\end{pmatrix}, \tag{3.2}$$

or

$$\boldsymbol{\psi}=-\boldsymbol{M}_m\boldsymbol{Q}\boldsymbol{\eta}.$$

We shall consider both the baroclinic form given above and the barotropic case, which is the limit as λ goes to 0 of (3.2). Note that

$$\lim_{\lambda\to 0} I_m(\lambda\alpha)K_m(\lambda\beta)=(\alpha/\beta)^m/2m,$$

so that the matrix $\boldsymbol{M}_m$ becomes

$$\boldsymbol{M}_m \to \begin{pmatrix} \dfrac{1}{2m} & \dfrac{1}{2mb^{m-1}} \\ \dfrac{1}{2mb^{m+1}} & \dfrac{1}{2m} \end{pmatrix}.$$

The kinematic equations (2.8) yield

$$\begin{pmatrix} \bar{V}(1) & 0 \\ 0 & \bar{V}(b)/b \end{pmatrix} \begin{pmatrix} \eta \\ \tau \end{pmatrix} + \begin{pmatrix} \psi(1) \\ \psi(b)/b \end{pmatrix} = \Omega \begin{pmatrix} \eta \\ \tau \end{pmatrix},$$

$$\boldsymbol{V}\boldsymbol{\eta} + \boldsymbol{\psi} = \Omega \boldsymbol{\eta}.$$

Note that the matrix $\boldsymbol{V}$ satisfies the equation

$$\boldsymbol{V}\mathbf{1} = \boldsymbol{M}_1|_{\lambda \to 0}\, \boldsymbol{Q}\mathbf{1},$$

which can be used to find either the velocity matrix from the potential vorticity differences matrix or vice versa (by inverting the $\boldsymbol{M}_1$ matrix), recalling that both are diagonal matrices.

Altogether then we have the matrix eigenvalue problem

$$[\boldsymbol{V} - \boldsymbol{M}_m \boldsymbol{Q}]\boldsymbol{\eta} = \Omega \boldsymbol{\eta}. \tag{3.3}$$

The dispersion relationship

$$\begin{aligned} \Omega = {} & \tfrac{1}{2}[\bar{V}(1) - q_1 I_m(\lambda) K_m(\lambda) + \frac{\bar{V}(b)}{b} - q_b I_m(\lambda b) K_m(\lambda b)] \\ & \pm \frac{1}{2}\left\{\left[\bar{V}(1) - q_1 I_m(\lambda) K_m(\lambda) - \frac{\bar{V}(b)}{b} + q_b I_m(\lambda b) K_m(\lambda b)\right]^2 + 4 q_1 q_b I_m^2(\lambda) K_m^2(\lambda b)\right\}^{\frac{1}{2}}, \end{aligned} \tag{3.4}$$

gives the necessary condition for stability

$$q_1 q_b < 0,$$

that the potential vorticity gradient must change sign (cf. Howard & Gupta 1962). However the geometrical constraints (integral m) on the problem imply that the necessary condition is not sufficient; therefore we must examine the stability criterion,

$$\left[\bar{V}(1) - q_1 I_m(\lambda) K_m(\lambda) - \frac{\bar{V}(b)}{b} + q_b I_m(\lambda b) K_m(\lambda b)\right]^2 < -4 q_1 q_b I_m^2(\lambda) K_m^2(\lambda b), \tag{3.5}$$

in more detail. We can reduce the parameter space dimension to 4 by choosing $\bar{V}(1) = 1$ ($q_1 + q_b = 2$); we then must explore the stability criterion (3.5) and growth rates versus m, $\Delta = q_b/(q_1 + q_b)$ (the ratio of the outer potential vorticity to the inner potential vorticity), b and λ. From the circulation theorem or the equation relating $\boldsymbol{V}$ to $\boldsymbol{Q}$ above, we can express $\bar{V}(b)$ in terms of these parameters:

$$\bar{V}(b) = \frac{1-\Delta}{b} + \Delta b$$

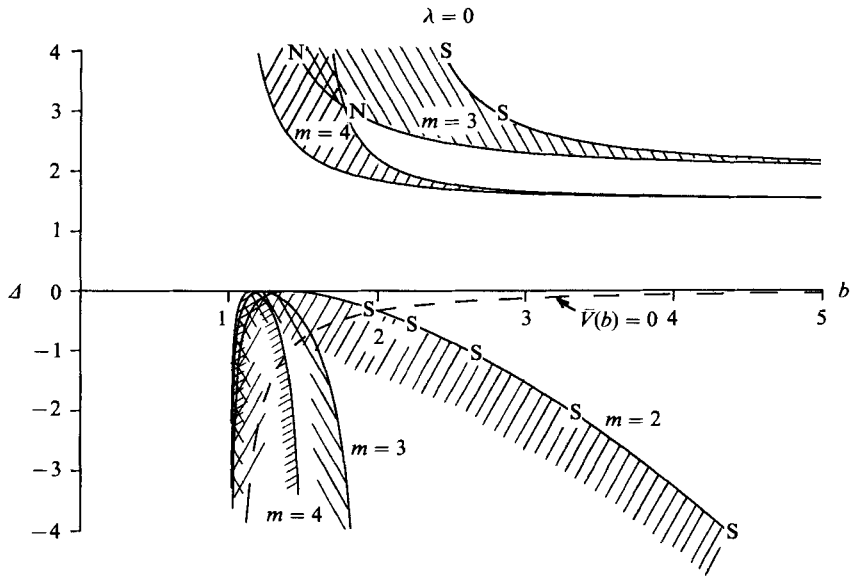

FIGURE 2. Boundaries in parameter space of the unstable regions for barotropic perturbations to a barotropic vortex. Shading indicates unstable regions. Instability boundaries are shown for various azimuthal mode numbers, m. The dashed line shows the relationship between the parameters Δ and b such that $\bar{V}(b) = 0$. The N and S labels show points where the bifurcation is normal or subcritical, respectively (see §6).

and (3.5) becomes

$$\left[1-2(1-\Delta)I_m(\lambda)K_m(\lambda)-\frac{1-\Delta}{b^2}-\Delta+2\Delta I_m(\lambda b)K_m(\lambda b)\right]^2 < -16(1-\Delta)\Delta I_m^2(\lambda)K_m^2(\lambda b). \quad (3.6)$$

3.1. *Barotropic perturbations*

When the perturbations are barotropic ($\lambda = 0$), the stability criterion, found by the same substitution, $I_m(\lambda\alpha)K_m(\lambda\beta)$ becomes $(\alpha/\beta)^m/2m$, is

$$\left[\Delta\left(\frac{2}{m}-1+\frac{1}{b^2}\right)+1-\frac{1}{m}-\frac{1}{b^2}\right]^2 < \frac{4\Delta(\Delta-1)}{m^2b^{2m}}. \quad (3.7)$$

This is plotted for various values of m in figure 2.

From this diagram and the formulae above, we find:

(i) The elliptical mode ($m = 2$) grows only for $\Delta < 0$ (opposite vorticities in the core and outer regions); in fact, the equation above shows that we need

$$\Delta < -\frac{(\frac{1}{2}b^2-1)^2}{b^2-1}.$$

Note that this result was not discovered by Michalke & Timme, since they considered only the $\Delta > 1$ case.

(ii) The higher modes are unstable for $\Delta < 0$ and b values in the proper range. For ordinary values of Δ, the outer annulus must be pretty narrow: $b < \sim 2$ for instability, with even sharper transition zones required for the higher modes.

(iii) The modes for $m > 2$ are unstable also when $\Delta > 1$ and b is in a band centred at higher values than for the $\Delta < 0$ case. For each mode, the unstable region extends down and toward the right (as b increases), narrowing into a thin region around the curve

$$\Delta = \frac{1 - \frac{1}{m} - \frac{1}{b^2}}{1 - \frac{2}{m} - \frac{1}{b^2}}.$$

Note that the unstable regions for the higher m values extend in under the regions for the lower values (but may be very narrow in extent, i.e. only a limited set of values will be unstable). Thus we expect that a small increase in Δ above the critical value $\Delta = 1$ will lead to the appearance of a high azimuthal mode.

(iv) The $m = 1$ mode is neutral. But when the eddy is isolated in the sense that $\bar{V}(b) = 0$, so that there is no net vorticity, the two roots are degenerate. Both modes have $\Omega = 0$; Stern (private communication) and Biebuyck (1986) have shown that this degeneracy leads to steady propagation of the vortex when the inner and outer circles are displaced in opposite directions. This could be viewed as a linear, rather than exponential, instability. When $\bar{V}(b) \neq 0$, the structure moves in a circle with radius proportional to the (perturbation amplitude)/$\bar{V}(b)$.

The growth rate $\sigma = m \operatorname{Im}(\Omega)$ shows rather different behaviours in the $\Delta < 0$ and $\Delta > 1$ regions of parameter spaces. Michalke & Timme showed the behaviour of σ as Δ increases above 1: for positive Δ, the higher modes enter first as the shear across the outer region is increased. The highest modes are unstable only for a narrow Δ band (when b is fixed) and for large Δ, where several modes may be unstable, the growth rate peaks at an intermediate wavenumber. In contrast, in the negative Δ region, the lower modes are the first to become unstable, the elliptical mode entering for the smallest $|\Delta|$ when $b > 1.3$. In both cases, when b is small enough (e.g. $b = 1.2$) and Δ is large enough, the intermediate wavenumbers are the most unstable.

As a final example we consider the growth rates of perturbations on the isolated vortex $\bar{V}(b) = 0$, $\Delta = -1/(b^2 - 1)$. This restriction makes our barotropic parameter space one-dimensional, shown by the dashed line in figure 2. The instabilities occur when

$$m = 2, \quad b < 2,$$

$$m = 3, \quad b < (1 + \sqrt{2})^{\frac{1}{2}} = 1.554,$$

$$m = 4, \quad b < \tfrac{1}{3}(2 + 10^{\frac{1}{3}}) = 1.385.$$

The growth rates in this case, figure 3, show both that the elliptical mode appears first as we consider narrower and narrower (and also higher vorticity) outer regions and that the higher modes appear with increasing growth rates as b decreases. For vortices with narrow outer regions, the growth rate is highest for intermediate wavenumbers.

Modal structures. The structure of the unstable modes can be drawn from knowledge of the eigenvectors for a growing mode. As an example, we show in figure 4 the positions of the contours $r = 1 + \eta$, $b + \tau$ with an arbitrary amplitude being assigned to the perturbation. Note the phase shifts in the orientation of the vorticity contours. Melander & Zabusky (private communication) have argued that a perturbed contour will elongate when there is a phase shift between the streamfunction and the perturbations in the vorticity contour such that the flow has an outward component. If we consider the motions induced by an elliptical

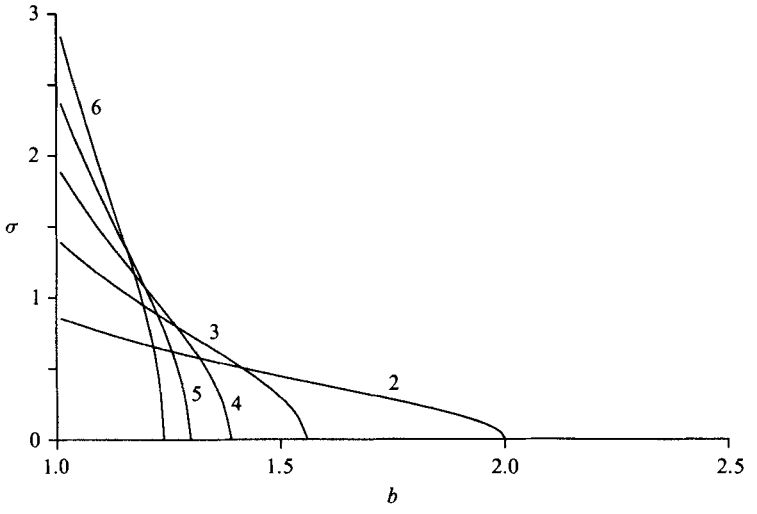

FIGURE 3. Growth rates $\sigma = m\,\mathrm{Im}\,(\Omega)$ for barotropic perturbations of an isolated barotropic vortex as a function of the outer radius b.

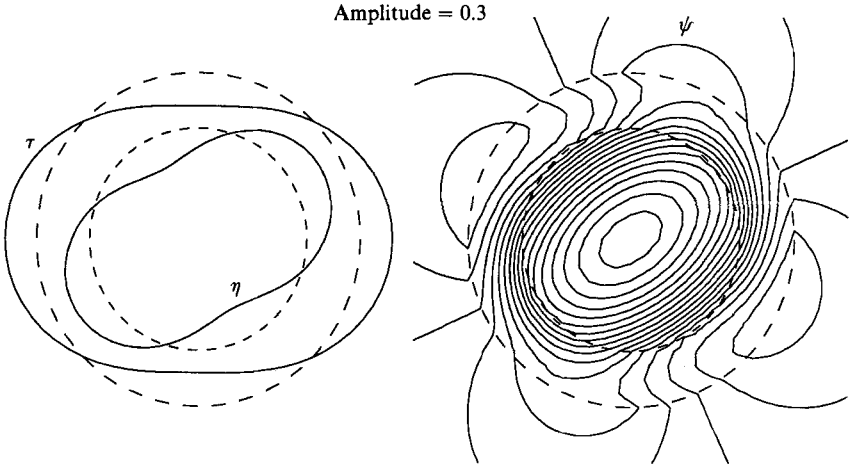

FIGURE 4. Sketch of the contours of vorticity boundaries and streamfunction calculated from the linearized solution for a growing mode. The basic state is an isolated vortex with $b = 1.5$.

deformation of the outer contour (figure 5*a*), they impart a procession to the outer ellipse in the direction indicated but will not, in and of themselves, generate stretching or squashing of the ellipse. But if we orient the elliptically distorted contours in the sense shown (figure 5*b*), the motions induced by the outer vorticity anomalies will tend to increase the distortion of the inner vorticity contour. We next mark the vorticity anomalies induced by the distortion of the inner circle and the circulations induced. At this point, the importance of the Rayleigh criterion becomes clear; in the case $0 < \Delta < 1$, the outer distortion will be reduced if the inner is increased and vice versa. For $\Delta < 0$ or $\Delta > 1$, the two disturbances can both be increasing at the same time so that instability is possible. This method (cf. Hoskins, McIntyre & Robertson 1985) of demonstrating Rayleigh's theorem brings out the

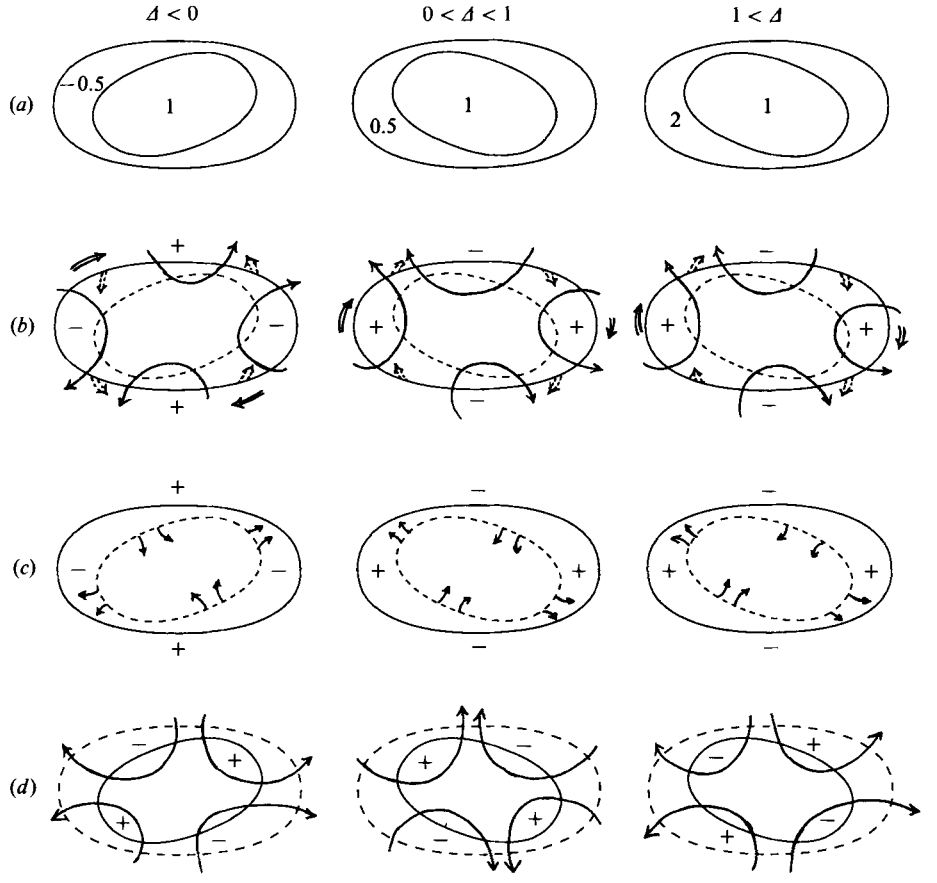

FIGURE 5. (*a*) Sketches of deformed contours for three different types of vortex (different Δ values). The orientations are chosen so that the perturbation flow associated with the outer vortex anomalies will reinforce the deformation of the inner contour (*b*, *c*). (*d*) Flows induced by the deformation of the inner contour, showing that it can reinforce the deformation of the outer contour when $\Delta < 0$ or $\Delta > 1$.

physical meaning clearly. In addition to being able to establish the proper phase relationships at $t = 0$, however, it is necessary to maintain them. Thus the azimuthal propagation rates of the two disturbances must also match. The propagation rates depend on the basic state velocity, the wavenumber of the disturbances, the vorticity jumps, and the separation between the contours. The latter is important because the waves in one contour generally have their phase speeds altered by the waves in the other contour. We can now see that the parameter regimes where Rayleigh's criterion for instability is met, yet the vortex is not unstable, correspond to cases where the waves cannot phase lock. Or one could view the problem as being that the wavelengths required for phase locking in a case where $\bar{V}(b) = 0$ and $b > 2$, for example, are just too long to fit around the circumference of the eddy. This example of the insufficiency of the Rayleigh criterion is interesting since it stems from an inability to meet a quantization condition (cf. Lindzen & Tung 1978) because of the structure of the basic state rather than anything to do with the geometry of the container for the fluid.

We can also use the perturbed eddy structure to guess at the finite amplitude

evolution. If we replace each lobe of each wave with a point vortex having $1/m$ of the vorticity in the area, we can fairly easily see the further evolution. In the case $\Delta < 0$, the opposite-signed vortices will pair and move out away from the original centre. In the $\Delta > 0$ case, however, the vortices all have the same signs so that they will rotate around each other and remain fairly confined. The former case will be discussed in more detail below; Dritschel (1986) has solved the latter case numerically using contour dynamics and demonstrates that the annular ring tends to break up into a set of vortices which interact and cycle around each other.

3.2. *Baroclinic perturbations*

When the perturbations are baroclinic ($\lambda > 0$), the stability properties are somewhat different. In general, the stability boundaries for baroclinic perturbations lie at smaller b values than the critical curves for the barotropic mode. However there are two notable exceptions: first, there is now an unstable band for the elliptical $m = 2$ mode when $\Delta > 1$ and, secondly, the barotropic eddy becomes unstable in previously stable regimes to the $m = 1$ perturbations (figure 6) for $\Delta < 0$. Gent & McWilliams discovered this property in their numerical eigenvalue calculations; for the model considered here, we can prove it analytically the small Δ version of (3.6):

$$\left[1-2I_m(\lambda)K_m(\lambda)-\frac{1}{b^2}\right]^2+2\Delta\left[1-2I_m(\lambda)K_m(\lambda)-\frac{1}{b^2}\right]$$
$$\times\left[2I_m(\lambda)K_m(\lambda)+\frac{1}{b^2}-1+2I_m(\lambda b)K_m(\lambda b)\right]<-16\Delta I_m^2(\lambda)K_m^2(\lambda b).$$

The critical curve (equality rather than inequality in the above equation) grazes the $\Delta = 0$ axis at

$$b_0 = [1-2I_m(\lambda)K_m(\lambda)]^{-\frac{1}{2}},$$

which is finite for $m = 1$ and $\lambda > 0$. If we expand around this value of b, the instability criterion becomes

$$\Delta < -\frac{(b-b_0)^2}{4b_0^6 I_m^2(\lambda)K_m^2(\lambda b_0)}.$$

For small, but finite λ,

$$b_0 \approx \begin{cases} \left(1-\dfrac{1}{m}\right)^{-\frac{1}{2}} & (m \neq 1), \\ (-\frac{1}{2}\lambda^2 \ln \frac{1}{2}\lambda)^{-\frac{1}{2}} & (m = 1), \end{cases} \tag{3.8}$$

so that the instability to baroclinic modes can occur even when the outer boundary is quite far out and has only weak negative vorticity. Figure 6*a* shows how the degenerate $m = 1$ roots around the $\bar{V}(b) = 0$ line open into an unstable region when $\lambda > 0$. Note, however that the whole structure must be small compared to R_d for (3.8) to hold, so that it is perhaps best to think of this case as having very compact cores. This analytic proof of the instability to $m = 1$ modes is confirmed by the instability diagrams (figure 6).

We can understand the mechanism for the 'twisting' mode $m = 1$ instability by again considering the vorticity anomalies and their influence upon each other (figure 7). If we displace a single vortex in a baroclinic fashion – giving it a tilt – we produce anomalous vorticity as shown in figure 7(*a*); these patches tend to cause the tilted vortex to precess. The larger the displacement, the larger the areal extent of the anomalies (or the stronger the circulation associated with the anomalies) and the

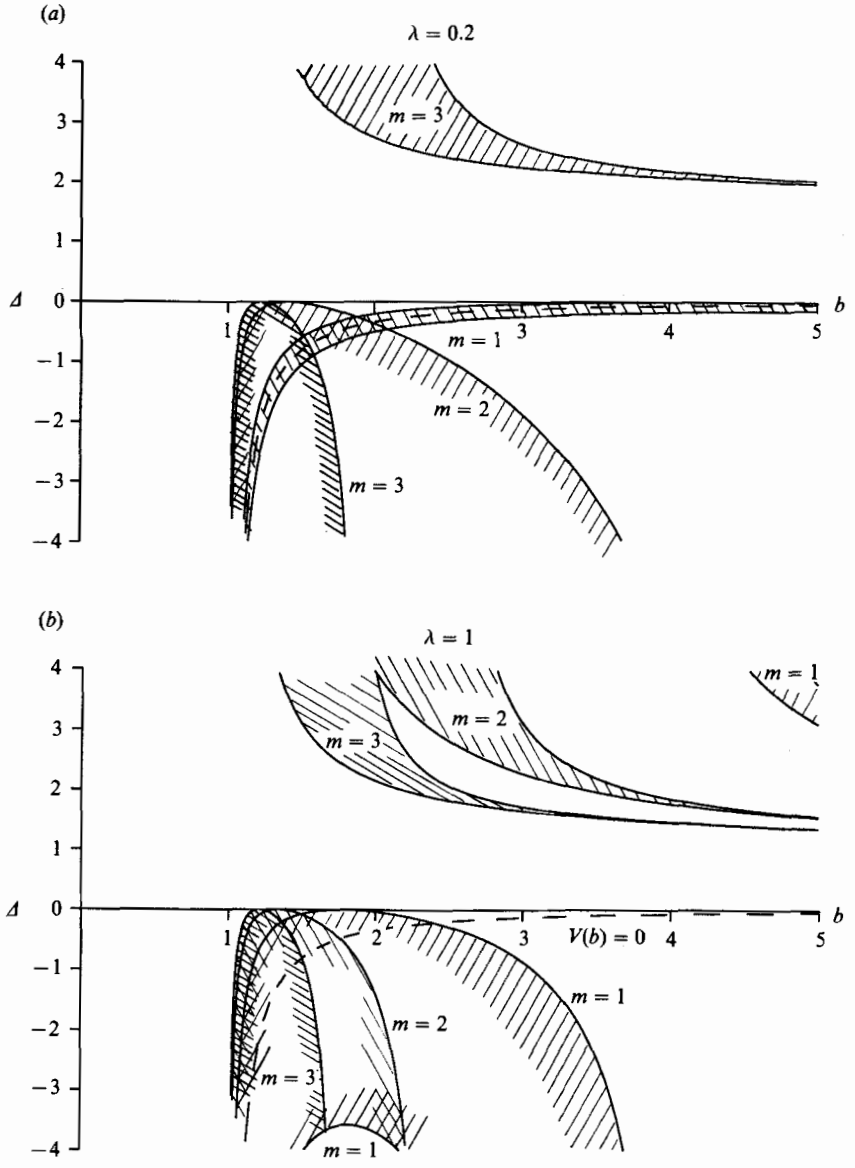

FIGURE 6. Instability boundaries for baroclinic perturbations on a barotropic vortex for various azimuthal modes m and two different ratios of vortex radius to deformation radius $\lambda = 0.2$ and 1.0. Note the destabilization of the 'twisting' mode $m = 1$.

faster the rotation of the tilted axis. (We are, of course, still in the linear regime so that we can interpret the last sentence as saying that the $m = 1$ wave progresses with a frequency Ω implying that a point on the axis on the upper surface moves at a speed $\frac{1}{2}\Omega d$ where d is the horizontal distance between the upper and lower centres. Thus the speed is proportional to d.) Next consider shifting the inner and outer regions relative to each other at one depth (figure 7*b*). Because the regions have opposite signs, the vortex anomalies now tend to produce translation rather than rotation (Stern, private communication; Biebuyck 1986). Again the translational velocity is proportional to the distance between the displaced centres. If we now

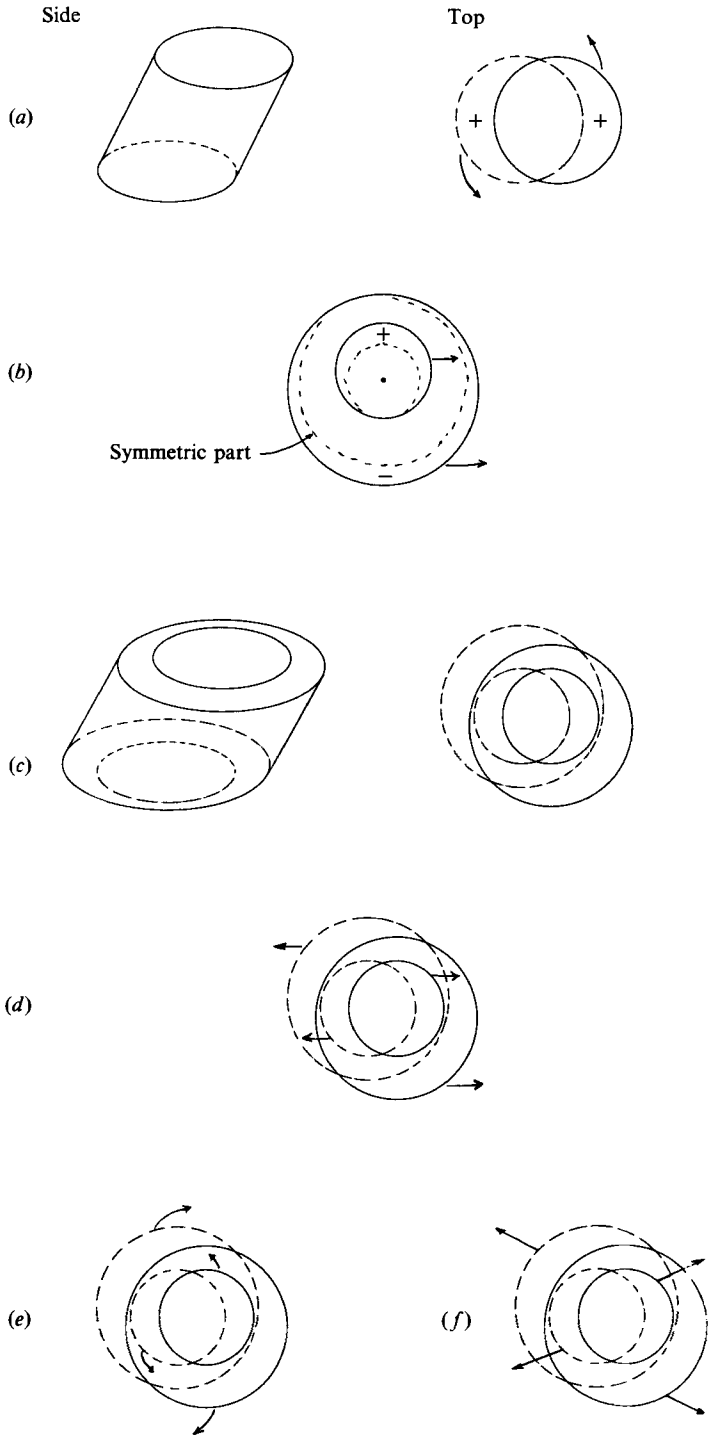

FIGURE 7. Mechanism for twisting mode instability: (*a*) rotation of a tilted vortex, combined with (*b*) propagation when inner and outer boundaries displaced relative to each other. When the inner and outer boundaries are tilted in different directions (*c*), the surface and deep propagation tendencies (*d*) and rotation tendencies (*e*) sum up to a motion enhancing the original perturbation (*f*).

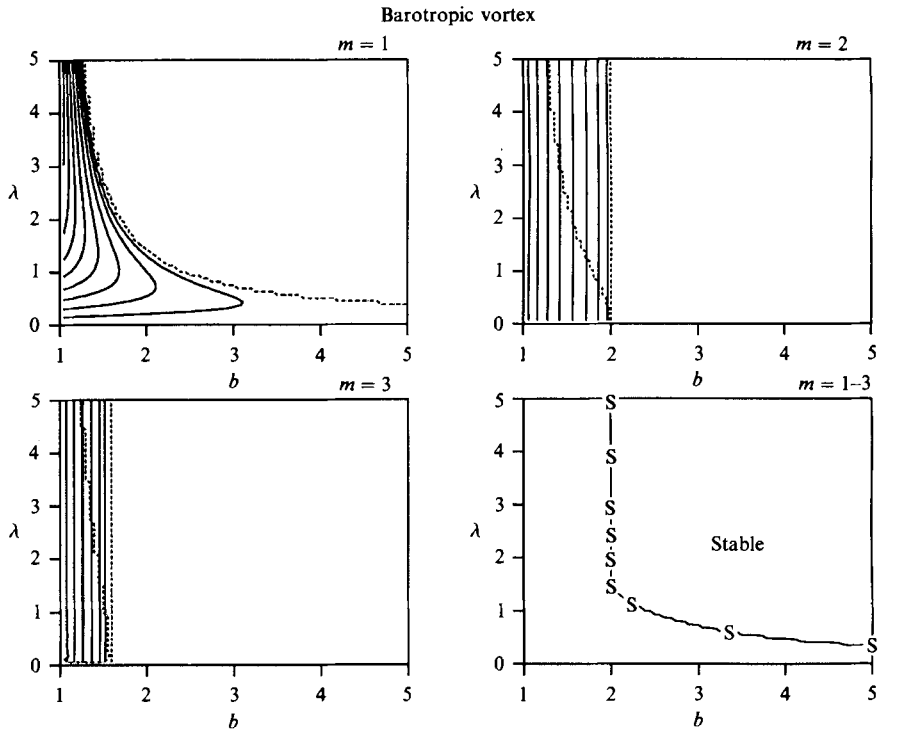

FIGURE 8. Contours of growth rate for the isolated barotropic vortex as a function of outer radius b and the ratio λ of the inner radius to the deformation radius of the perturbation. The first three figures show the growth rates for the individual modes. Contour intervals are 0.05, 0.2, 0.2 for $m = 1, 2, 3$ respectively. Dotted lines delineate regions with different numbers of unstable modes; e.g. for $m = 2$, there are two unstable modes (one barotropic and one baroclinic) for $b = 1.5$ and $\lambda = 1$ (although the most rapidly growing mode for $m = 2$, 3 is the barotropic one). The fourth figure shows the boundaries of the unstable region. The S indicates a nonlinearly subcritical transition.

apply vertical tilt to both the inner and outer boundaries, but in different directions (figure 7*c*), we can arrange that the horizontal interactions of the vorticity anomalies cause propagation of the upper and lower patterns in different directions (figure 7*d*). Meanwhile the interactions in the vertical tend to rotate the inner and outer centres in opposite directions (figure 7*e*); for proper choice of the initial arrangements, then, the net result is a tendency to increase the separation of the four centres (figure 7*f*) (and possibly a rotation of the structure as well; we have not been overly specific here about the relative strengths of various terms, including mean advection, and have considered only some of the interactions among the four anomalies.) But for vortex patches, unlike point vortices, the interactions get stronger as the separations of the centres increase (as long as they are still small enough for the linear approximation to hold). Thus the rate of expansion of the pattern is proportional to the separations and the separations will increase exponentially. So the instability comes about because of the propagation tendency arising from horizontal offsets in the centres of the core and the outer region and the rotation tendencies from the vertical offsets. We anticipate that this twisting mode instability will grow until the interior regions are sufficiently separated that the vertical interactions become weak; simultaneously, the propagation tendency will decrease as the outer boundaries become distorted and zero vorticity fluid is entrained (Stern, private communication; Biebuyck 1986).

3.3. *Discussion*

We summarize the results for the isolated barotropic vortex by contouring the growth rates versus b and λ in figure 8. These plots show the growth rate of the more rapidly growing mode and indicate by dotted lines the regions in which there are two unstable modes. The maps show that the baroclinic mode $m = 1$ will be the only unstable mode for $b > 2$ and small enough $\lambda (< 1.4)$. These results in general agree with Gent & McWilliams: for vortices which are small enough compared to the deformation radius and which have fairly weak shear in the outer region, the instability tends to generate baroclinic modes. This 'twisting' mode instability does not occur for rectilinear flows (or Rossby waves, Fu & Flierl 1980) and, as Gent & McWilliams point out, probably represents an important limitation on the types of eddies which one could expect to persist in a more complicated flow – small-scale barotropic motions will break apart vertically. We shall see below that baroclinic eddies, in contrast, will not be stable if too large.

4. Baroclinic vortex

When the basic state vortex is baroclinic with q_1, q_b (and therefore $\bar{V}(1)$, $\bar{V}(b)$) functions of z, the vertical structures of η and τ are no longer simple functions of z. Instead we expect to find the normal modes to be complicated, with phase shifts in the vertical and critical layers; we shall avoid most of these problems by using a truncated model expansion in the vertical. In the horizontal, of course, the fact that the potential vorticity is piecewise constant likewise eliminates critical layers. Using a modal formulation (cf. Flierl 1978), rather than a layer model, has the advantage that the eigenvalue, Ω, will appear only on the diagonal of the final matrices, so that standard eigenvalue solvers can be employed. We can illustrate this procedure by considering the case with a single interface

$$\left.\begin{aligned}(\nabla^2 + L_z)\,\psi &= q\eta\delta(r-1),\\ (\bar{V}(1)-\Omega)\,\eta &= -\psi(1, z).\end{aligned}\right\} \tag{4.1}$$

This problem was treated in the two-layer model by Pedlosky (1985); we also include a barotropic component to the flow. We shall choose, for simplicity, to make the mean vortex have only barotropic and first baroclinic modal amplitude so that

$$\bar{V}(1, z) = V_0 + F(z).$$

Henceforth $F(z)$ and λ will be used to denote the first baroclinic eigenfunction and eigenvalue of the vertical structure equation

$$\frac{\partial}{\partial z}\frac{f^2L^2}{N^2H^2}\frac{\partial}{\partial z}F_n = -\lambda_n^2 F_n,$$

$$F_{nz}(0) = F_{nz}(1) = 0,$$

and the subscripts will be used when necessary to indicate the other eigenmodes. The F terms will be normalized so that they have vertically-averaged squared value equal to one. The ratio of barotropic to baroclinic kinetic energy in the basic state vortex at $r = 1$ is just $\bar{V}_0^2$.

For this velocity field, the potential vorticity is given by

$$q = 2V_0 + [I_1(\lambda)\,K_1(\lambda)]^{-1}F(z).$$

The form of the vorticity equation suggests that we can expand ψ and η in the eigenfunctions of L_z

$$\left.\begin{aligned}\psi(r, z) &= \Sigma\, p_n(r)\, F_n(z),\\ \eta(z) &= \Sigma\, h_n F_n(z).\end{aligned}\right\} \tag{4.2}$$

Substituting these into (4.1) and projecting the equations on $F_j(z)$, we find

$$p_j(1) = -2V_0 I_m(\lambda_j) K_m(\lambda_j) h_j - \frac{I_m(\lambda_j) K_m(\lambda_j)}{I_1(\lambda) K_1(\lambda)} \sum_{k=0}^{\infty} \xi_{1jk} h_k, \tag{4.2}$$

and

$$\sum_{k=0}^{\infty} (\xi_{1jk} + V_0 \delta_{jk} - \Omega \delta_{jk}) h_k + p_j(1) = 0, \tag{4.3}$$

where

$$\xi_{jkl} = \int_0^1 \mathrm{d}z\, F_j(z) F_k(z) F_l(z).$$

This infinite matrix eigenvalue problem can be solved by truncation to various degrees. In Flierl (1978), it was demonstrated that the two-mode truncation, equivalent to a two-layer model, reproduced the results of a ten-mode truncation fairly closely.

The two-mode truncation matrix eigenvalue problem,

$$\begin{pmatrix} V_0 - \dfrac{V_0}{m} & 1 - \dfrac{1}{2mI_1K_1} \\ 1 - \dfrac{I_m K_m}{I_1 K_1} & V_0 + \xi - 2V_0 I_m K_m - \xi \dfrac{I_m K_m}{I_1 K_1} \end{pmatrix} \begin{pmatrix} h_0 \\ h_1 \end{pmatrix} = \Omega \begin{pmatrix} h_0 \\ h_1 \end{pmatrix},$$

(where the argument of all the I and K terms is λ) leads to a quadratic dispersion relationship with 4 parameters V_0, λ, m and $\xi = \xi_{111}$. We select two ξ values: 0 (atmospheric constant N or equal layer depth model) and 1.8 (oceanic case). The former case corresponds to a baroclinic mode which has equal and opposite positive and negative lobes – $\cos(\pi z/H)$, while the positive values of ξ will be obtained when the near surface positive lobe dominates over the deep negative lobe – i.e. the surface layer is relatively shallow. The value of 1.8 was obtained from numerical solutions of the vertical structure equation for a typical oceanic N^2 profile (Flierl 1978).

From the equation above, the condition for stability is

$$\left[V_0\left(2I_m K_m - \frac{1}{m}\right) + \xi\left(\frac{I_m K_m}{I_1 K_1} - 1\right)\right]^2 < -4\left(1 - \frac{I_m K_m}{I_1 K_1}\right)\left(1 - \frac{1}{2mI_1 K_1}\right), \tag{4.4}$$

which is shown in V_0, λ space for various m in figure 9. We also contrast this stability criterion to the necessary criterion for stability

$$\left(2V_0 + \frac{1}{\sqrt{\delta} I_1 K_1}\right)\left(2V_0 - \frac{\sqrt{\delta}}{I_1 K_1}\right) < 0,$$

for the layer version of the two-mode truncation where

$$\sqrt{\delta} = \tfrac{1}{2}((\xi^2 + 4)^{\frac{1}{2}} - \xi),$$

and δ is the ratio of layer depths. For the ξ values used above, we have $\delta = 1$ and $\delta = 0.19$, respectively. The criterion simply is the condition that the potential vorticity be opposite in sign in the two layers. The stability criterion for the purely baroclinic vortex is given in Pedlosky (1985); the addition of barotropic circulation

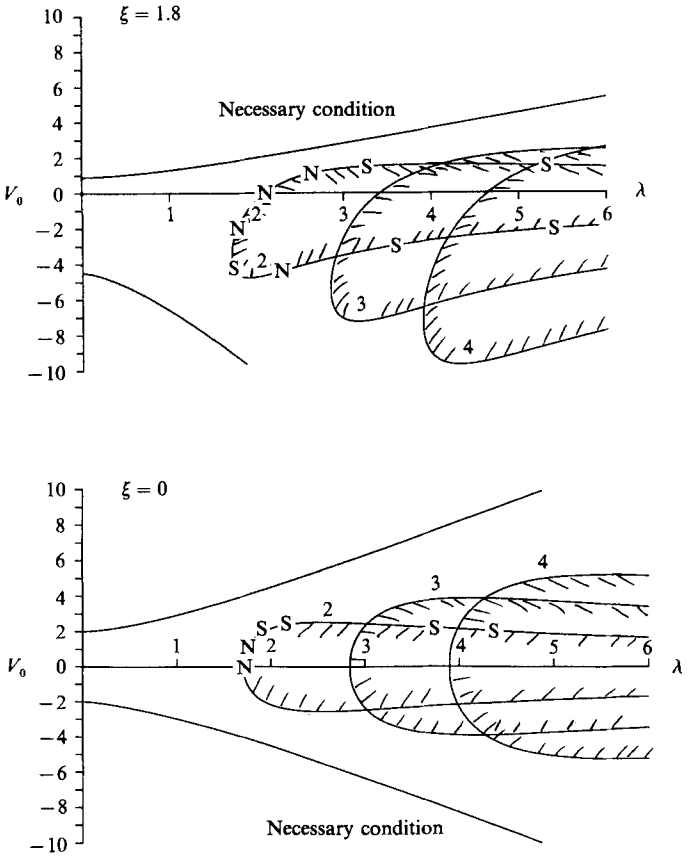

FIGURE 9. Stability boundaries for a baroclinic vortex for various modes. The vortex has only a single boundary (no outer annulus). The parameters are the amount of barotropic flow and the radius compared to the deformation radius. The 'necessary criterion' curves mark the locations where the potential vorticity step has the same sign in the upper and lower layers. Two different ξ values are used, corresponding to equal layer depths ($\xi = 0$) or a ratio of 0.19 ($\xi = 1.8$).

does make some interesting modifications. (Helfrich & Send 1988 have considered this problem also.) The figure shows that for the baroclinic vortex:

(i) The vortex becomes unstable if it is sufficiently large compared to the deformation radius and there is not too much barotropic circulation. For the pure baroclinic vortex, the critical values of λ are

$$\lambda = \begin{cases} 1.705, & m = 2, \quad \xi = 0, \\ 2.840, & m = 3, \quad \xi = 0, \\ 2.096, & m = 2, \quad \xi = 1.8, \\ 3.417, & m = 3, \quad \xi = 1.8. \end{cases}$$

(ii) The necessary condition for instability is far from sufficient.

(iii) When $\xi = 0$ (e.g. uniform stratification) a barotropic circulation will stabilize the vortex. When $\xi \neq 0$, barotropic flow which reinforces the deeper layer circulation can destabilize the vortex. The least stable case (smallest unstable vortex) has strong deep flow and weaker shallow flow in the same sense as the deep circulation – not the kind of eddy found in oceanic data.

(iv) When $V_0 = 0$ (pure baroclinic flow), mode 2 always becomes unstable first; when $V_0 \neq 0$, however, it is possible that a higher azimuthal mode may be excited first.

(v) The $m = 1$ mode is stable but has degenerate eigenvalues if either $V_0 = 0$ or $I_1(\lambda) K_1(\lambda) = 0.5$. The latter cannot occur so that only pure baroclinic vortices give rise to steadily propagating solutions (see Appendix A). These translating states are less singular versions of the Hogg & Stommel (1985) 'hetons', although we here only obtain the cases where the centres are separated by a small amount. Polvani (private communication) has computed the states with large amplitudes and finds they are similar.

Analysis of the growth rates show that

$$\operatorname{Im} \Omega \sim m^{\frac{1}{2}},$$

for $V_0 = 0$ and large λ. In the barotropic case, if $b \approx 1 + \epsilon$, the systems looks like a shear layer and the most rapidly growing modes have a scale order ϵ. Here, for large λ, the vortex has flow only near $r = 1$ and the shear occurs over a characteristic scale $1/\lambda$ (dimensionally, the deformation radius); the most rapidly growing modes have a scale on this order, corresponding to large m – many waves around the perimeter.

4.1. *Discussion*

The characteristics of the laboratory instabilities observed by Saunders (1973) and Griffiths & Linden (1981) are reproduced fairly well by this linear calculation. As in the experiments, we find that increasing the scale of the vortex relative to the deformation radius leads to higher mode instabilities. Indeed, figure 9(*a*) suggests that the radius at which wavenumber m is neutral increases roughly with m, as in Saunders' experiments. From the stability criterion (4.4), we find that for $\xi = 0$ and $V_0 = 0$ the waves with $m < [2I_1(\lambda) K_1(\lambda)]^{-1}$ are unstable – a nearly linear relationship. Oceanic vortices ($\xi > 0$, corresponding to a shallow upper layer) are destabilized by counter-rotating barotropic flow, although the least stable configuration (requiring the smallest eddy radius) needs to have barotropic flow so strong that the velocity in the upper layer is reversed from the sense of circulation of the vertical shear. Thus the deep flow in this configuration has the same direction of circulation but is stronger than the shallow flow; this kind of flow is certainly not among the common paradigms for oceanic eddies, so that this extreme case is of less importance than the stabilization or destabilization occurring for small co- or counter-rotating barotropic circulations.

The mechanism for the baroclinic instability is essentially the same as that sketched in figure 5, except that the vorticity anomalies associated with the two deformed contours now are in different layers. Again, it is possible for perturbation flows associated with the distortion of one of the contours to enhance the distortion of the other (and vice versa) only if the vorticity jumps from the eddy to the outside are opposite in sign. The outward tendency must be strong enough to overcome the tendency for the upper and lower waves to lose their phase relationships as the waves propagate and are advected around the vortex.

5. Isolated baroclinic vortex

When we have two interfaces at $r = 1$ and $r = b$, we can examine mixed barotropic/baroclinic instabilities. The dispersion relationship becomes fourth order in Ω and therefore can still be solved exactly. The matrix equations are of the same form as (3.2)–(3.3) with the definitions:

$$\boldsymbol{\psi} = \begin{pmatrix} p_0(1) \\ p_0(b)/b \\ p_1(1) \\ p_1(b)/b \end{pmatrix}, \quad \boldsymbol{\eta} = \begin{pmatrix} h_0(1) \\ h_0(b) \\ h_1(1) \\ h_1(b) \end{pmatrix},$$

$$\boldsymbol{M}_m = \begin{pmatrix} 1/2m & 1/2mb^{m-1} & 0 & 0 \\ 1/2mb^{m+1} & 1/2m & 0 & 0 \\ 0 & 0 & I_m(\lambda)K_m(\lambda) & bI_m(\lambda)K_m(\lambda b) \\ 0 & 0 & I_m(\lambda)K_m(\lambda b)/b & I_m(\lambda b)K_m(\lambda b) \end{pmatrix},$$

$$\boldsymbol{Q} = \begin{pmatrix} q_1^{\mathrm{bt}} & 0 & q_1^{\mathrm{bc}} & 0 \\ 0 & q_b^{\mathrm{bt}} & 0 & q_b^{\mathrm{bc}} \\ q_1^{\mathrm{bc}} & 0 & q_1^{\mathrm{bt}} + \xi q_1^{\mathrm{bc}} & 0 \\ 0 & q_b^{\mathrm{bc}} & 0 & q_b^{\mathrm{bt}} + \xi q_b^{\mathrm{bc}} \end{pmatrix}$$

and

$$\boldsymbol{V} = \begin{pmatrix} V^{\mathrm{bt}}(1) & 0 & V^{\mathrm{bc}}(1) & 0 \\ 0 & V^{\mathrm{bt}}(b)/b & 0 & V^{\mathrm{bc}}(b)/b \\ V^{\mathrm{bc}}(1) & 0 & V^{\mathrm{bt}}(1) + \xi V^{\mathrm{bc}}(1) & 0 \\ 0 & V^{\mathrm{bc}}(b)/b & 0 & V^{\mathrm{bt}}(b)/b + \xi V^{\mathrm{bc}}(b)/b \end{pmatrix}.$$

The only real difficulty is that the parameter space is very large: growth rates in the truncated model depend on 7 parameters ($V^{\mathrm{bt}}(1)$, $V^{\mathrm{bc}}(b)$, $V^{\mathrm{bt}}(b)$, λ, ξ, b and m, for $V^{\mathrm{bc}}(1) = 1$). To limit the scope of the discussion we shall consider only isolated vortices $\bar{V}(b) = 0$ and two ξ values $\xi = 0$ and $\xi = 1.8$ as above, leaving a manageable number of parameters: $V_0 = V^{\mathrm{bt}}(1)/V^{\mathrm{bc}}(1)$, λ, b and m. As the discussion below makes clear, the instabilities appear to be modified forms of those found previously – the $m = 1$ and $m = 2$ instabilities from §3 and the $m = 2$ baroclinic instability from §4. In figure 10, we illustrate the growth rates found for different mode numbers m on the (λ, b)-plane. Two values of V_0 have been selected to show the differences when the deep fluid is counter-rotating versus co-rotating with the upper fluid. In figure 11, we show the boundary of the stable region for modes $m = 1$ to 3 and a whole series of V_0 values. Two ξ values 1.8 and 0, have been used. From these figures and study of the amplitudes and phases of the perturbations of the upper and lower layer inner and outer interface displacements, we conclude:

(*a*) The purely baroclinic isolated vortex ($V_0 = 0$) is unstable to modes $m \geqslant 2$ and shows two distinct types of instability. There is a slowly growing mode (with large lower layer amplitudes) which exists for small enough λ and b, and there is a rapidly growing mode which exists for small enough b or large enough λ. At large b, this latter mode looks exactly like the baroclinic instability found in §4. For both modes,

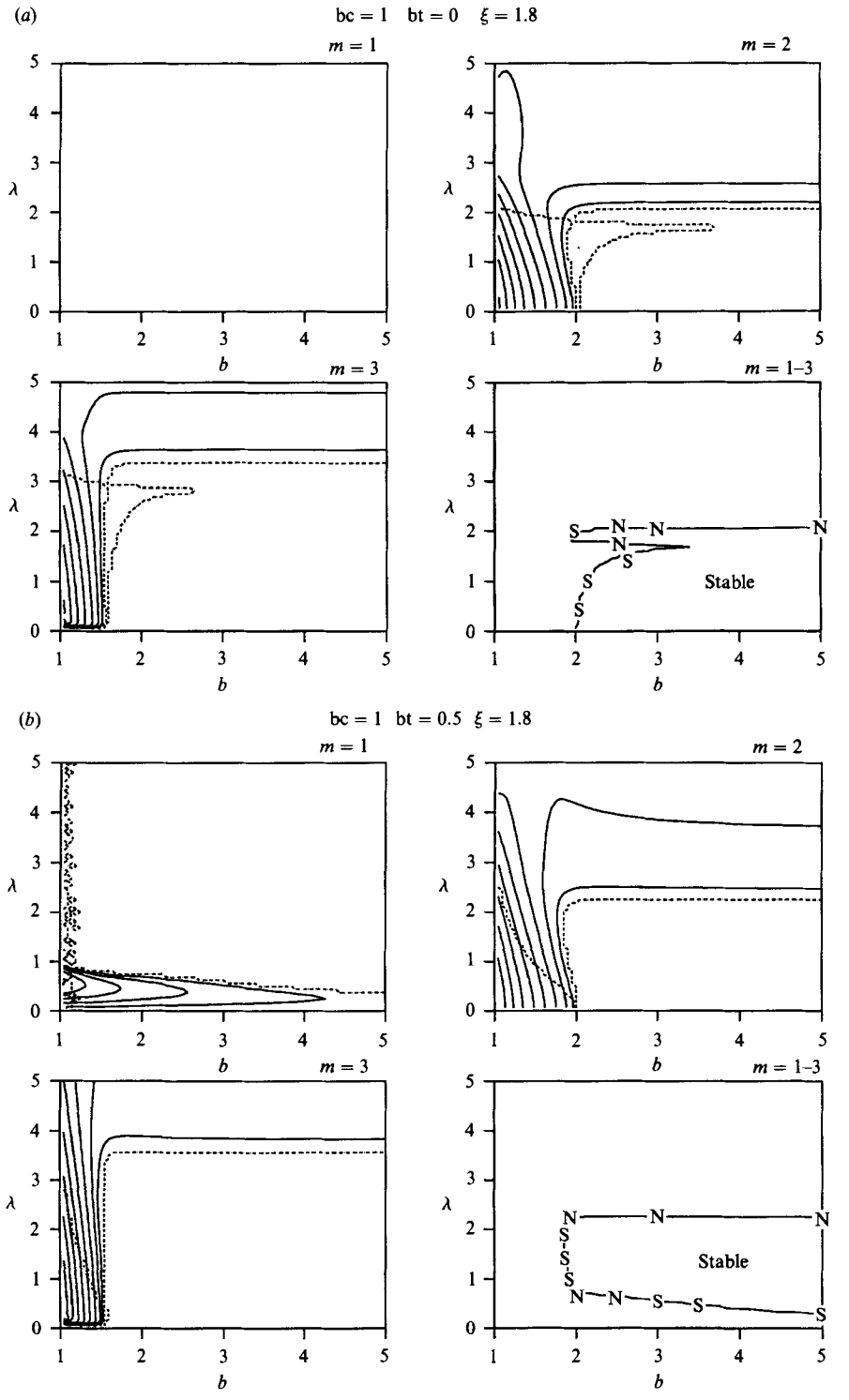

FIGURE 10. Growth rates versus outer radius and the ratio of the inner radius to the deformation radius for various azimuthal mode numbers and increasing ratios of barotropic to baroclinic flow at $r = 1$. The figures are organized as in figure 8. The N indicates a normal, supercritical transition. The bc and bt values indicate the amplitudes of the baroclinic and barotropic velocities, respectively. Contour intervals are (*a*) $V_0 = 0$, $m = (1, 2, 3)$, contoured by (not unstable, 0.2, 0.4), respectively, (*b*) $V_0 = 0.5$, $m = (1, 2, 3)$, contoured by (0.01, 0.25, 0.4).

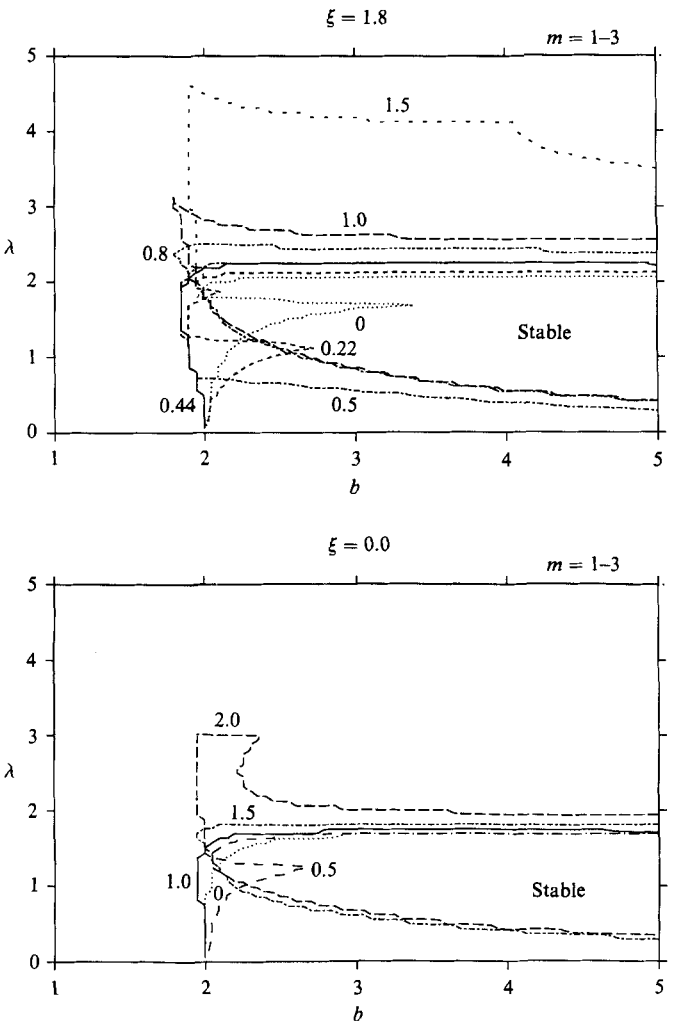

FIGURE 11. Boundaries separating unstable from stable regions for the combination of modes m = 1 to 3. (*a*) For $\xi = 1.8$ and $V_0 = 0, \frac{1}{2}\sqrt{\delta}$ (= 0.22), $\sqrt{\delta}$ (= 0.44), 0.5, 0.8, 1.0, 1.5. (*b*) For $\xi = 0$ (equal layer depths) and $V_0 = 0$, 0.5, 1.0, 1.5, and 2.0.

however, the phase relationships suggest that the mode is growing by extracting kinetic and potential energy from the basic state circular vortex. At small b, the modes appear to be essentially like the barotropic instability of §3, with one mode being dominated by upper layer displacements and the other by lower layer deviations; however, the contribution from the baroclinicity apparently still contributes to the growth of the perturbation. The small out-jutting region of instability near $\lambda = 1.6$, $b = 2.5$ for $\xi = 1.8$ is a region of relatively weakly growing instability; it cuts off above $\lambda = 2$. This mode has fairly large amplitude in the lower layer.

(*b*) Similar results obtain whenever the lower layer velocity at $r = 1$ is in the opposite direction to the upper layer flow; that is, when $V_0 < \delta^{\frac{1}{2}}$, the vortex is stable only for large enough b and small enough λ.

(*c*) When the basic state velocities are zero on both lower layer interfaces, ($V_0 = \delta^{\frac{1}{2}}$ – though the lower layer velocity is not everywhere zero), only one mode is

unstable. This result should be comparable to those obtained by Hart (1974), who used a model with zero flow in the lower layer of a two-layer fluid. The basic state profiles are rather different – Hart has no relative vorticity near the centre; furthermore, he uses a linear rather than a polar coordinate system. However, our results resemble Hart's along a line $b \approx 1.9$. Vortices small compared to the deformation radius are unstable to an essentially barotropic mechanism. Intermediate scale vortices are neutral, while the larger ones become unstable to a baroclinic instability. He does not see the region of parameter space ($b > 2$) for which the small vortices are stable.

(*d*) When the vortex has a large enough barotropic part, ($V_0 > \delta^{\frac{1}{2}}$), we see both the baroclinic $m = 1$ mode and the barotropic higher m modes discussed in §3. Initially, when the lower layer flow is weak but co-rotating with the upper layer circulation, the $m = 1$ mode grows only for small vortices. In any case, this mode is dominant only for $\lambda < 1.5$, $b > 2$ – vortices smaller or order of the deformation radius in size, with a fairly broad region of decay of the basic state velocity field.

5.1. *Discussion*

Our results can be most directly compared to Ikeda's (1981) numerical calculations of the linearized stability properties for Gaussian vortices in a two-layer model. He sees a similar dependence upon λ, but because the barotropic Gaussian is only very weakly unstable (Carton, private communication and our own calculations using multiple contours to approach a smooth Gaussian profile), he does not see the instability for narrow outer shear regions revealed in this study. In effect his b value is nearly two. Also, Ikeda did not consider the $m = 1$ mode. He does, however, a much more complete exploration of the dependence upon the lower layer velocity (compared to the upper layer flow), which demonstrates that the growth rate associated with the baroclinic instability of §4 decreases with stronger lower layer flow. He also finds two modes of instability for larger eddies, whereas we have only one unstable mode in this case. The difference presumably comes from the different potential vorticity distributions.

6. Nonlinear calculations

In this section, we apply an amplitude expansion to determine steadily rotating finite-amplitude structures bifurcating from the dispersion curve of the linear theory. The calculations are involved and there are subtle differences between the approach here and that of Pedlosky (1985) as discussed below; in fact, slightly different results are obtained for the problem he treated (see Appendix B). The computation leads to an estimate for Ω as a function of A^2, the perturbation amplitude squared, and other parameters such as b and λ which determine the stability. In fact we obtain $\mathrm{d}\Omega/\mathrm{d}A^2$; for parameter values and Ω near coalescence points where unstable roots appear, $\mathrm{d}\Omega/\mathrm{d}A^2$ will reveal whether the transition is subcritical (and may not equilibrate) or supercritical (with equilibration likely). The argument runs as follows (see figure 12): the solution for the displacements of the interfaces behaves like $\eta \sim A(T) \exp(-\mathrm{i}\Omega_0 t)$ and the amplitude satisfies a second-order cubic equation

$$A_{TT} = f(b)\, A + n(b)\, A\, |A|^2,$$

where b, for example, is the bifurcation parameter and T represents a slow timescale. For steady rotating solutions, $A \sim \exp(-\mathrm{i}\delta\Omega)$ and the rotation rate is approximately $\Omega_0 \pm \delta\Omega = \Omega_0 \pm (-f - n\,|A|^2)^{\frac{1}{2}}$. In the upper diagram of figure 12, corresponding to

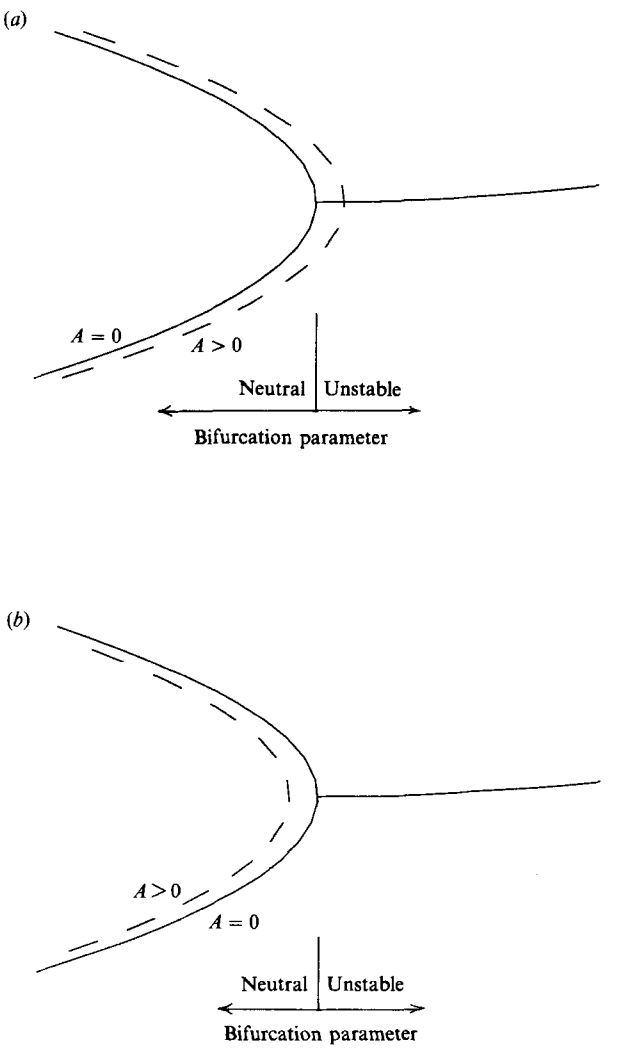

FIGURE 12. Relationship between rotation frequency of neutral mode as a function of amplitude and bifurcation parameter in (*a*) a normal or supercritical bifurcation and (*b*) a subcritical bifurcation.

$\delta\Omega(\mathrm{d}\Omega/\mathrm{d}A^2) > 0$, we have n negative; in the lower diagram, n is positive. The sign of n is calculated in the stable region, where $f(b)$ is negative; however, n varies only slowly as b passes through the stability boundary into the region where f becomes positive, so we expect the sign will not be changed. Examination of the evolution equation above shows that the nonlinearity will retard the growth when $n < 0$ and $f > 0$, leading to a nonlinear vacillation around the solution $|A|^2 = -f/n$. When $n > 0$, the nonlinearity enhances the instability as the amplitude increases and no equilibrium will be found.

We shall first describe the conceptual framework for the calculation, without making the small-amplitude approximation. Details of the amplitude expansion are left to Appendices B (rectilinear, baroclinic, single interface problem analogous to Pedlosky's) and C (circular problem). The important conceptual points can be made in the full finite amplitude context.

The procedure used is based again on contour dynamics and takes advantage of the linear relationship between potential vorticity and streamfunctions. The two-layer, rather than two-mode, formalism is more natural here, since the boundaries will reside in one or the other of the layers and since potential vorticity in each layer is conserved, whereas the potential vorticity in a mode is not.

There are two steps in the procedure:

(i) Given the locations of the boundaries between various regions of constant potential vorticity, calculate the streamfunction on these boundaries. To aid in discussion of this calculation, let us define a vortex patch as a region $\mathscr{D}$ of constant potential vorticity in one layer, with zero potential vorticity outside $\partial\mathscr{D}$ and in the other layer. (Zero potential vorticity is being used in the quasi-geostrophic sense here, as distinguished from the form this quantity takes in the shallow-water model.) In our problem, the potential vorticity is the sum of the potential vorticities of four vortex patches: in the upper layer (indicated by superscript 1), $\mathscr{D}_2$ with area πb^2 and potential vorticity $q_b^{(1)}$ and $\mathscr{D}_1$ with area π and potential vorticity $q_1^{(1)}$. Note that q_1 still indicates the difference between the full potential vorticity in the inner region and that in the outer annulus. Thus we consider the inner regions as being additional anomalies of strength q_1 stacked on top of anomalies of strength q_b. The full potential vorticity in the upper layer inner region is found by summing the contributions from the two patches, $q_1^{(1)}+q_b^{(1)}$. In the lower layer, we have two more patches $\mathscr{D}_4$ and $\mathscr{D}_3$. Step (i) then has three parts: (ia) for a given vortex patch, calculate the streamfunction $\Psi^{(1)}$ and $\Psi^{(2)}$ associated with the vorticity in that patch as a function of r and θ. Then, (ib), the upper layer streamfunction is evaluated along the curves $\partial\mathscr{D}_1$ and $\partial\mathscr{D}_2$ and the lower layer one along $\partial\mathscr{D}_3$ and $\partial\mathscr{D}_4$. Finally, (ic), the contributions to the value of Ψ on the curve $\partial\mathscr{D}$ from each of the four vortex patches are summed. (Note that the step (ia) of obtaining Ψ from a vortex patch should be done in such a way that the flow is obtained as a functional of the boundary shape; the same formulae can then be used for each of the four patches.)

Let us discuss this approach somewhat further for the baroclinic problem, since step (i) is where the distinction lies between our results and Pedlosky's. Consider three procedures for calculating the flow field as sketched in figure 13. The first method (figure 5.3a) follows the procedure outlined above: we use the vorticity distribution in the upper layer [$\omega_p^{(1)} = 1.0$ for $r < 1+\rho$ and 0 for $r > 1+\rho$; $\omega_p^{(2)} = 0$] to calculate Ψ along both the upper layer interface $r = 1+\rho$ and the lower layer interface $r = 1+\mu$. (We use ρ and μ as shorthand notations for $\eta^{(1)}$ and $\eta^{(2)}$, the displacements of the upper and lower layer interfaces, respectively. The potential vorticity values of 1 and -1 are chosen for example only.) In doing this, we compute the flows in the inner and outer regions and match across $r = 1+\rho$. We then calculate the streamfunctions induced by the lower layer vorticity distribution [$\omega_p^{(1)} = 0$; $\omega_p^{(2)} = -1.0$ for $r < 1+\mu$ and 0 for $r > 1+\mu$]. These two components are summed to yield the streamfunction along each interface as a functional of the shape of the two interfaces. In practice, the calculations are carried out to order amplitude cubed.

The second procedure (figure 13b, figure 2 of Pedlosky 1985), seems to be more intuitive: we define the solutions inside and outside of the interfaces from the full vorticity distribution ($\omega_p^{(1)} = 1.0, 0$, $\omega_p^{(2)} = -1.0, 0$, for the inner, outer regions). Since we are Taylor-expanding around $r = 1$ for each order in amplitude, we will match the inner and outer solutions at $r = 1+\rho$ for the upper layer and $r = 1+\mu$ for the lower layer by estimating them using Taylor expansions.

The third procedure (figure 13c) accounts explicitly for the region which is neglected in the above analysis, that is the region between $r = 1+\rho$ and $r = 1+\mu$. In

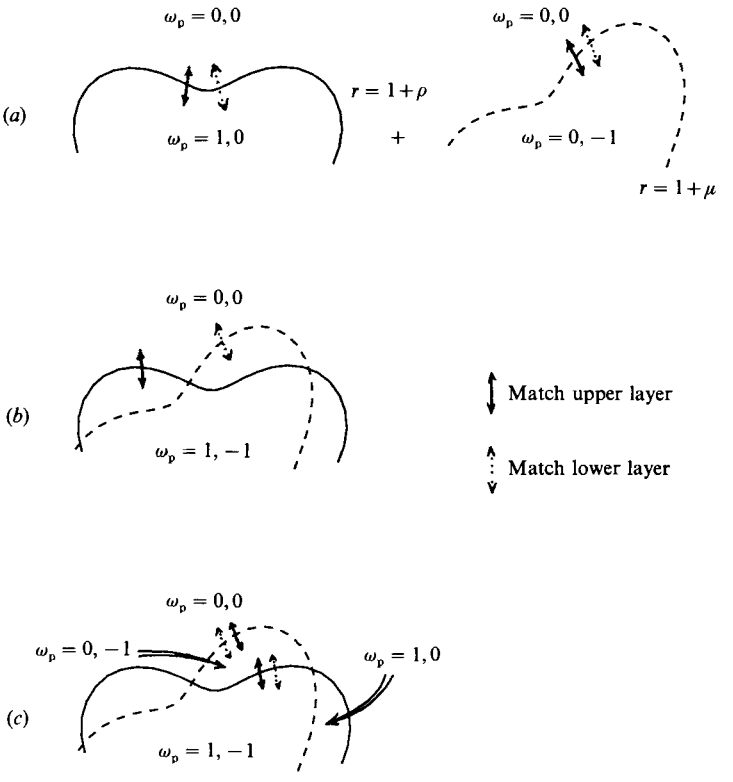

FIGURE 13. Three procedures for calculating flow of perturbed vortex. Arrows indicate matching of inner and outer solutions occurs across indicated curve. In (*a*), the flow is first calculated with the upper layer potential vorticity distribution and uniform PV in the lower layer. Then the calculation of flow is carried out with potential vorticity anomalies only in the lower layer. These two results are added. In (*b*), solutions in the inner and outer regions are matched directly. The intermediate region, in contrast to the emphasis placed upon it in method (*c*), is not considered.

the half of the domain where $\rho > \mu$, we write the streamfunction solution in the three regions shown using the potential vorticity distributions $(\omega_p^{(1)}, \omega_p^{(2)}) = (1, -1)$, $(0, -1)$, $(0, 0)$ for regions 1, 2 and 3, respectively). We then match the solutions in region 1 to those in region 2 for both layers at $r = 1+\rho$ and we match region 2 to region 3 (both layers again) at $r = 1+\mu$. By this procedure we again find the streamfunctions on the interfaces. The middle region is only order amplitude in width, but it turns out that the contribution from this region does enter in the cubic-order term. While this procedure in principle will give the same results as the first, it seems difficult to handle the time-dependent changes in the geometry of region 2 in an unsteady case. Numerical approaches to multilayer contour dynamics (Helfrich & Send 1988; Polvani, Zabusky & Flierl 1988) use the first procedure.

The distinction with Pedlosky's calculation is evident; his procedure essentially neglects the intermediate region (region 2). The explicit calculation in Appendix B demonstrates that different results are obtained by this method in comparison with approaches (*a*) or (*c*). Thus the error in neglecting the intermediate region does affect the amplitude cubed estimate of Ψ and thereby $\mathrm{d}\Omega/\mathrm{d}A^2$; we use approach (*a*) in our circular vortex calculations.

(ii) Now that we have Ψ on each boundary $\partial\mathscr{D}_k$ as a functional of the four

boundary shapes, we can apply the condition that each boundary $\partial\mathscr{D}_k$ be a streamline in the co-rotating frame,

$$\Psi = \Omega\tfrac{1}{2}r^2 + \text{const}_k \quad \text{on} \quad \partial\mathscr{D}_k.$$

This requirement (or really these four requirements, one for each boundary) gives a nonlinear eigenvalue (Ω)–eigenvector ($\partial\mathscr{D}_k$) problem from which we can compute the rotation rate as a function of the amplitude of the boundary disturbance. Also we find the shapes of the various boundaries.

The major goal is simply the calculation of the sign of $\mathrm{d}\Omega/\mathrm{d}A^2$, so we will leave any further detail to the Appendices and summarize the conclusions here. We remark, however, that this quantity is very sensitive to slight procedural or numerical errors. We have checked the code against the linear problem, against self-consistency between upper and lower layer perturbations, against the Kirchhoff vortex, and against the steadily rotating two-layer solutions derived by contour dynamics when vorticity appears in only one patch in the upper layer (Polvani *et al.* 1988).

6.1. *Results*

The most convenient way to present the results of the nonlinear analyses is to indicate on figures 2, 8, 9 and 10 the points where such calculations were carried out and the type of bifurcation found. (Similar calculations were carried out for most of the parameter values shown in figure 11; the comments below are to some degree based also on these other experiments.) For each point we have fixed either λ or b and found the value of the other parameter which is just on the stable side of a transition (at least two significant figures beyond the decimal point). The values of $\mathrm{d}\Omega/\mathrm{d}A^2$ for the higher and lower Ω branches were calculated. If this quantity is positive for the higher-frequency branch and negative for the lower branch, then the transition is a supercritical bifurcation and is expected to equilibrate. This is indicated by an N [normal] on the figure. If the signs are opposite, we have a subcritical, non-equilibrating transition, indicated by S. (See again figure 12.)

For the purely barotropic vortex, we find:

(i) The finite amplitude states for the $m = 2$ instability for $\Delta < 0$ lie outside the linearly unstable region; thus this instability would not be expected to equilibrate at low amplitude. Numerical calculations (figure 14 or Gent & McWilliams 1986) show that indeed the elliptical mode tends to break up into dipolar structures as the form of the linear mode and the simple point vortex calculation suggested. The simulations shown use the fully nonlinear contour dynamics (Zabusky *et al.* 1979), with points inserted as needed. The code was provided by S. Meacham. These show that the cases with initial b values near the critical case (2.0), break into dipoles, while vortices with smaller b values tend to form structures with a fair amount of vorticity from the inner circle left in a band rather than forming two patches. Simulations with a pseudo-spectral code suggest that tripolar structures may be formed by these compact initial conditions when the transitions in vorticity are fairly smooth.

(ii) The baroclinic $m = 1$ transition when $\Delta < 0$, however, is generally also a subcritical bifurcation, although we did find small regions (near $\Delta = -0.4$ and $\lambda = 1$, for example) which appeared to be a standard fork bifurcation.

(iii) For the $\Delta > 0$ modes, we find that the finite amplitude curve is shifted to the right, so that the transition caused by decreasing b is subcritical; this agrees with the results of Benny & Maslowe (1975) that the shear-layer instability does not equilibrate at small amplitudes. The other transition, when the vortex becomes stable again at smaller b is a normal supercritical one; however, at smaller outer

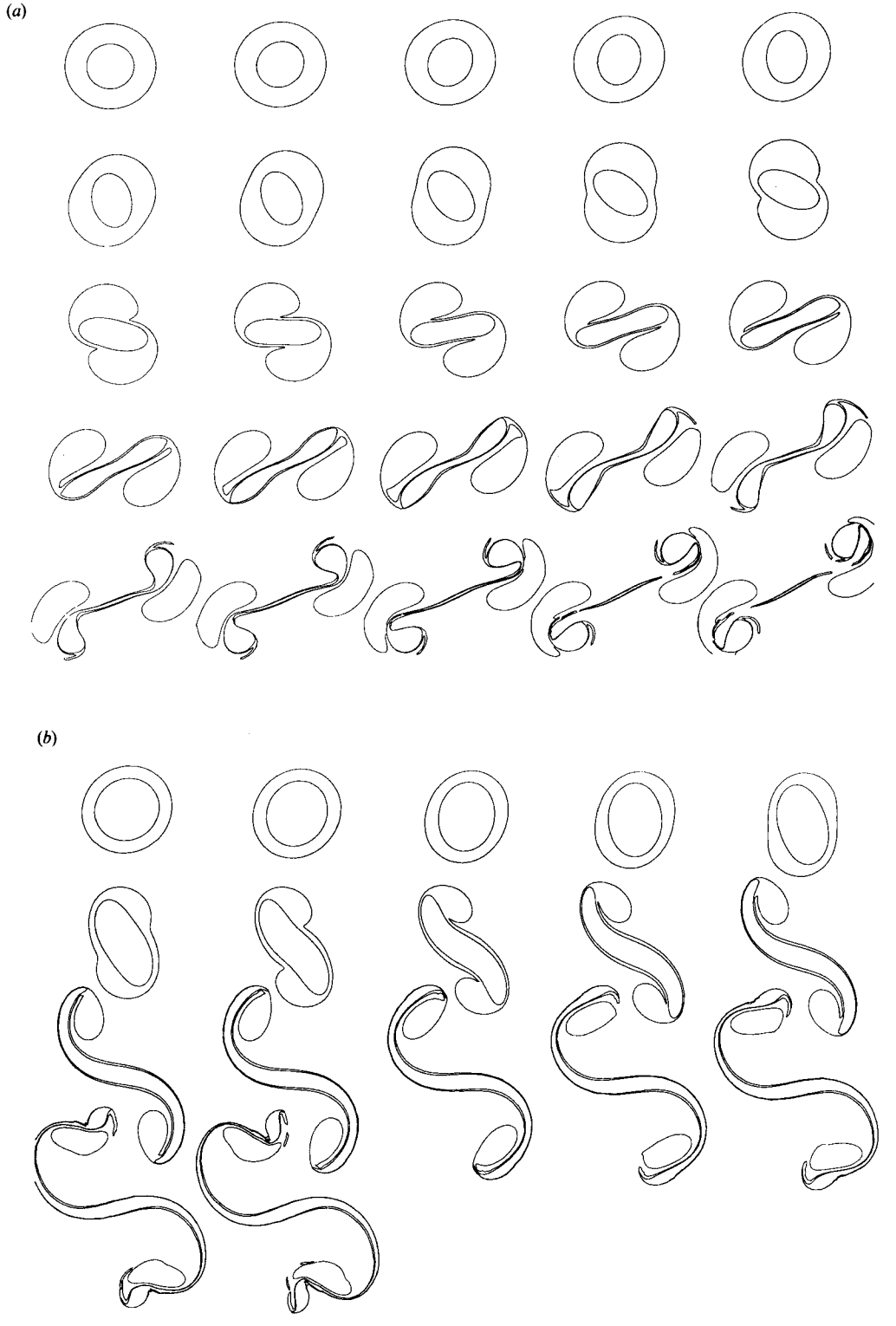

FIGURE 14. Break-up barotropic vortex to an $m = 2$ barotropic instability. (*a*) Corresponds to $b = 1.9$ and (*b*) to $b = 1.4$.

annulus sizes, we expect the higher modes to be more important and these, again, are likely to grow until the shear layer breaks into 'beads'. Dritschel (1986) demonstrates numerically that a finite amplitude perturbation of the neutral vortex at $\Delta = \infty$ and $b = 2$ (the right-hand side boundary of the instability region) is unstable and grows; this is consistent with the findings of the weakly nonlinear theory.

If we next consider the baroclinic single region 'heton cloud' problem, we find that:

(iv) If there is no barotropic flow, the transition is supercritical; numerical calculations of Helfrich & Send (1988) indicate that equilibration indeed occurs, but only for values of λ very slightly above the critical value. For larger radii, the vortex breaks up into dipolar structures like those observed by Griffiths & Linden (1981).

(v) In the presence of barotropic flow, the results are more complicated: large enough barotropic flow may make the transition become subcritical as λ increases. This occurs fairly near the cutoff associated with large barotropic flow. The transitions to stability for large radii in the presence of barotropic circulation tend to be subcritical; here again though, the higher wavenumber modes must be important.

For the isolated vortex, we find a pattern which is fairly simple:

(vi) For the barotropic, isolated vortex, the instability is subcritical for either the barotropic $m = 2$ or the baroclinic $m = 1, 2$ mode perturbations.

(vii) As the ratio of barotropic to baroclinic velocities at $r = 1$ is reduced, the $m = 1$ mode remains subcritical for small and large values of λ but can be supercritical for intermediate ranges of vortex sizes. The $m = 2$ mode shows supercritical transitions to instability at high λ and subcritical ones for intermediate λ values as the outer radius b decreases.

(viii) When the lower layer flow at $r = 1$ is counter in direction to the upper layer flow ($V_0 < \sqrt{\delta}$), the $m = 1$ twisting mode instability disappears and the remaining $m = 2$ elliptical mode has subcritical transitions in the baroclinic regime (vortices large compared to the deformation radius $\lambda > 2$) when the outer radius b becomes large enough and the vorticity in the outer region is not too large. In the barotropic regime, (vortices small enough compared to the deformation radius $\lambda < 2$ and having a sharp enough shear layer in their outer region $b < 2$ to 2.5), the transitions will generally be subcritical.

In summary, then we generally find three kinds of behaviour:

(*a*) A supercritical (equilibrating) elliptical instability for large baroclinic vortices with wide and weak outer shear zones.

(*b*) A subcritical (non-equilibrating) elliptical instability for eddies intermediate in scale compared to the Rossby radius and having narrow enough (and strongly enough sheared) outer annulli.

(*c*) A subcritical twisting instability for eddies which are small enough compared to the deformation radius, have wide enough outer regions, and have co-rotating flows in the upper and lower layers.

6. Conclusions

We have discussed the instability of three different basic state vortices. For the purely barotropic eddy, we find a non-equilibrating elliptical mode instability when the vorticity in the inner and the outer regions is opposite in sign. Numerical simulations show that such vortices will break up into either dipoles or tripolar

states. In addition, a baroclinic disturbance may grow on the vortex if the size of the eddy is small enough compared to the deformation radius associated with the vertical structure of the baroclinic perturbation and the vortex may be unstable to this mode even when it is stable to barotropic disturbances. When the vorticity is the same sign but larger in the annular region, a non-equilibrating instability can occur when the outer annulus is thin enough.

A baroclinic eddy with only a single patch of anomalous vorticity in each layer in the horizontal – a heton cloud – is unstable when the scale is large enough compared to the deformation radius. The exact stability boundary depends upon the amount of barotropic flow in the vortex, as does the nature of the finite amplitude bifurcation. The cases near the smallest possible unstable radius tend to have a nonlinear equilibration.

Isolated vortices, i.e. those having no flow outside the radius $r = b$, are stable for large enough outer radius b and intermediate values of λ (the ratio of the inner radius to the deformation scale). The instability at small outer radii leads to a non-equilibrating elliptical mode growth and, presumably, to dipole or tripole formation. The instability for large enough scale compared to the deformation radius leads to growth and equilibration of an elliptical disturbance (or, more precisely, vacillation in an inviscid system; cf. Helfrich & Send 1988). Finally, the transition at small scales compared to the deformation radius leads to nonlinearly enhanced growth of a twisting mode, so that the deep vortex separates from the surface one. This type of instability, found originally in Gent & McWilliams (1986), can occur when the deep flow has the same sign of circulation as the surface flow. The growth rates are relatively small, but, like the barotropic case, this mode may grow in regions stable to elliptical mode disturbances.

Ikeda's (1981) work also contains numerical simulations of the evolution of (somewhat sizeable) perturbations. He does see in a low growth rate case, which might be only marginally unstable, a growth and then decay of the perturbation and interprets this as a nonlinear stabilization. It seems that he is working in effectively the $\lambda > 2$ regime, so that his results may be consistent with our analysis. In more strongly growing cases, the vortices do break up into dipoles, similar to the calculations of Helfrich & Send (1988). Direct comparison of our calculations with numerical models is difficult, since it is not really clear to what degree the contour dynamics model successfully mimics the evolution of a smoother vorticity field in the presence of dissipation. In some cases, a model with a few contours can capture most of the main features of a nonlinear interaction, while in others, the smooth fields evolve quite differently, with significant distortion of the eddies and entrainment. We feel that the linear instability calculation probably mimics fairly well the strong instabilities of smooth fields, but lacks the weaker modes associated with critical layer phenomena. In the nonlinear problem, we are encouraged by the dipole formation experiments of Gent & McWilliams (1986) and Flierl (1984) in cases where the theory predicts subcritical bifurcations and the failure to form dipoles in Ikeda's (1981) experiments in the theoretically supercritical regime. However, it is not easy to locate these numerical experiments on the (b, λ)-plane precisely and the theory does have regions in this plane where the sign of the nonlinear coefficient is very sensitive to changes in the parameters; thus, we would not say that the simulations really confirm the analysis.

Our results imply that stable vortices must have relatively weak shear beyond the velocity maximum so that the outer radius is greater than two times the inner radius where the maximum velocities occur, and they must be sufficiently small compared

to the deformation radius (inner radius less than about 2.1 times the deformation radius) if there is any vertical shear. Finally, if the deep flows circulate in the same direction as the shallow flows, it is also necessary that the inner radius (measured by λ) exceed about $1.2/(b-1)$. In terms of the inner radius R, the outer radius R_1, and the deformation radius R_d, these three conditions for stability are $R_1 > 2R$, $R_1 - R > 1.2R_d$ and $R < 2.1R_d$.

Appendix A. The degenerate $m = 1$ mode and vortex propagation

Here we show that the mechanism described by Stern (private communication) and Biebuyck (1986) is applicable to the two-layer baroclinic vortex: the existence of a degenerate $m = 1$ mode implies that vortex propagation will occur. For a vortex moving steadily at speed c in the x-direction, we have

$$(\nabla^2 + L_z)\,\Psi = F(\Psi + cy).$$

For the contour dynamics model F is piecewise constant with the boundaries of each domain being lines of constant $\Psi + cy$. For the single-vortex patch, we have

$$(\nabla^2 + L_z)\,\Psi = q(z)\,\mathscr{H}(1 + \eta(\theta, z) - r), \tag{A 1}$$

and

$$\Psi(1 + \eta(\theta, z), z) + c(1 + \eta(\theta, z))\sin\theta = f(z). \tag{A 2}$$

If we now assume that the vortex is only slightly perturbed from circular and c is also small, (A 1)–(A 2) reduces to

$$\left.\begin{aligned}(\nabla^2 + L_z)\,\psi &= q(z)\,\eta\delta(r-1),\\ \bar{V}(1)\,\eta + \psi(1, \theta, z) + c\sin\theta &= 0.\end{aligned}\right\} \tag{A 3}$$

The homogeneous part of this system is just (4.1) with $\Omega = 0$. However, the $\sin\theta$ term implies that the $m = 1$ mode will be present for a propagating disturbance. Expanding in the vertical and truncating yields

$$\begin{pmatrix} 0 & 1 - \dfrac{1}{2I_1K_1} \\ 0 & V_0(1 - 2I_1K_1) \end{pmatrix} \begin{pmatrix} h_0 \\ h_1 \end{pmatrix} = \begin{pmatrix} -c \\ 0 \end{pmatrix}. \tag{A 4}$$

Clearly we require $V_0 = 0$ (no net circulation) for the propagating solution to exist. Equation (A 4) then gives the amplitude of the baroclinic displacement

$$h_1 = -c\frac{2I_1K_1}{2I_1K_1 - 1},$$

or, alternatively, the speed in terms of the horizontal displacement of the upper and lower centres, that is the tilt of the axis. These approximate solutions are smoother versions of the Hogg & Stommel (1985) 'hetons' in the limit of small displacements of the two centres with respect to the vortex size. (The limit of large displacement is the same as the point vortex case.)

Note that h_0 is arbitrary – barotropic displacement of the interface just

corresponds to a shift in the origin of coordinates. If we look at the associated streamfunction fields, from (4.2), we find

$$p_0 = -\frac{1}{2I_1K_1}\eta_1,$$

$$p_1 = -\eta_0 - \xi\eta_1.$$

Thus there is a barotropic antisymmetric component to the flow fields (the similarity to a barotropic modon (antisymmetric) with a baroclinic rider (symmetric) is clear). The baroclinic component, indeed, is arbitrary in this problem as well.

This argument will be important in the solution presented next for the continuously stratified model. We include it here to demonstrate that the two-layer model does not necessarily smoothly extend to the continuous case. We set $V_0 = 0$ and take ψ and η proportioned to $\sin\theta$. Equation (A 3) becomes

$$\left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r} - \frac{1}{r^2} + L_z\right)\psi = \frac{1}{I_1K_1}\eta F(z)\,\delta(r-1),$$

$$F(z)\,\eta + \psi(1, z) = -c.$$

Eliminating the term $F(z)\,\eta$ from these leads to

$$\left(\frac{1}{r}\frac{\partial}{\partial r}r\frac{\partial}{\partial r} - \frac{1}{r^2} + L_z\right)\psi = -\frac{1}{I_1K_1}[c + \psi(1, z)]\,\delta(r-1). \tag{A 5}$$

Clearly only the barotropic mode is forced by the c terms, though it is possible to add a baroclinic mode as well since it will satisfy (A 5). The barotropic solution is

$$\psi = \begin{cases} \psi(1)\,r & (r < 1), \\ \psi(1)\,\dfrac{1}{r} & (r > 1), \end{cases}$$

where

$$-2\psi(1) = -\frac{1}{I_1K_1}[c + \psi(1)].$$

This equation relates c to the displacements of the interface

$$F(z)\,\eta(z) = -c - \psi(1)$$

$$= -c\frac{2I_1K_1}{2I_1K_1 - 1},$$

so that

$$\eta(z) = -c\left(\frac{2I_1K_1}{2I_1K_1 - 1}\right)\frac{1}{F(z)}.$$

Although this result satisfies the linearized equations (A 3), it is clearly unsatisfactory because of the singularity at the depth where $F(z)$ changes sign. The linearization must break down and a more complicated structure, probably involving patches of anomalous vorticity which close off in the vertical, is required for a fully nonlinear solution. It is also curious that the perturbed columns on the upper and lower water columns lean away from each other as the middle depths are approached – the structure does not look like a single column with an increasing northward shift going toward the surface.

Appendix B. Finite amplitude baroclinic jet

In this Appendix, we consider in detail the instability of a baroclinic jet flow

$$u^{(1)} = \tfrac{1}{2}\,\mathrm{e}^{-|y|},$$

$$u^{(2)} = -\tfrac{1}{2}\,\mathrm{e}^{-|y|},$$

with $\delta = 1$, $\lambda = 1$. We shall discuss three procedures for calculating the nonlinear structures (refer back to figure 13). The vorticity distributions are slightly different: for procedure (*a*), $\omega_{\mathrm{p}}^{(1)} = -0.5$ for $y < \rho$ and 0.5 for $y > \rho$; $\omega_{\mathrm{p}}^{(2)} = 0$ in the first part of the computation and $\omega_{\mathrm{p}}^{(1)} = 0$; $\omega_{\mathrm{p}}^{(2)} = 0.5$ for $y < \mu$ and -0.5 for $y > \mu$ for the second.

In procedure (*b*), we have $\omega_{\mathrm{p}}^{(1)} = \pm 0.5$, $\omega_{\mathrm{p}}^{(2)} = \mp 0.5$ for the north/south regions. Since we are Taylor-expanding around $y = 0$ for each order in amplitude, we will match the north and south solutions in a Taylor-expanded sense at $y = \rho$ for the upper layer and $y = \mu$ for the lower layer. In procedure (*c*), the potential vorticity in the intermediate region, neglected in the analysis (*b*) (i.e. the region between $y = \rho$ and $y = \mu$) is given by $\omega_{\mathrm{p}}^{(1)}$, $\omega_{\mathrm{p}}^{(2)} = -0.5$, -0.5.

We begin the discussion of procedure (*a*) by detailing the calculation of the flow induced by the upper layer potential vorticity structure, a step function at $y = \rho(x, t)$, with zero quasi-geostrophic potential vorticity in the lower layer. This will be summed later with a similar calculation for the flow associated with the lower layer anomalies. These manipulations are tedious; we used the symbolic algebra package MUSIMP to carry them out. We define a vector of upper and lower layer flows north and south of the interface

$$\boldsymbol{\Psi} = \begin{pmatrix} \Psi^{(1)}(y < \rho) \\ \Psi^{(2)}(y < \rho) \\ \Psi^{(1)}(y > \rho) \\ \Psi^{(2)}(y > \rho) \end{pmatrix} = \boldsymbol{R} + \boldsymbol{M}_0 \boldsymbol{A}_0 + \boldsymbol{M}_1(\boldsymbol{A}_1 \phi + \boldsymbol{A}_1^* \phi^{-1}) + \boldsymbol{M}_2(\boldsymbol{A}_2 \phi^2 + \boldsymbol{A}_2^* \phi^{-2}) + \ldots \tag{B 1}$$

where ϕ is used as a shorthand for $\mathrm{e}^{\mathrm{i}kx}$,

$$\boldsymbol{R} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \\ \frac{1}{4} \end{pmatrix} - \begin{pmatrix} \frac{1}{2}y^2 & \mathrm{e}^{y} & 0 & 0 \\ \frac{1}{2}y^2 & -\mathrm{e}^{y} & 0 & 0 \\ 0 & 0 & \frac{1}{2}y^2 & \mathrm{e}^{-y} \\ 0 & 0 & \frac{1}{2}y^2 & -\mathrm{e}^{-y} \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{4} \end{pmatrix},$$

and

$$\boldsymbol{M}_j = \begin{pmatrix} \mathrm{e}^{jky} & \mathrm{e}^{s_j y} & 0 & 0 \\ \mathrm{e}^{jky} & -\mathrm{e}^{s_j y} & 0 & 0 \\ 0 & 0 & \mathrm{e}^{-jky} & \mathrm{e}^{-s_j y} \\ 0 & 0 & \mathrm{e}^{-jky} & -\mathrm{e}^{-s_j y} \end{pmatrix},$$

$$s_j = (1 + j^2 k^2)^{\frac{1}{2}}.$$

The coefficient vectors $\boldsymbol{A}_j$ are found by matching Ψ and Ψ_y for both layers at the interface, $y = \rho$. If we define

$$\boldsymbol{\Delta} = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ \dfrac{\partial}{\partial y} & 0 & -\dfrac{\partial}{\partial y} & 0 \\ 0 & \dfrac{\partial}{\partial y} & 0 & -\dfrac{\partial}{\partial y} \end{pmatrix},$$

we have

$$\Sigma \boldsymbol{\Delta M}_j \boldsymbol{A}_j \phi^j = -\boldsymbol{\Delta R}. \tag{B 2}$$

We use amplitude/Fourier expansion to solve (B 2). The matrices $\boldsymbol{\Delta M}_j$ and the vector $\boldsymbol{\Delta R}$ are Taylor-expanded around $y = 0$. As a shorthand, we will define

$$\boldsymbol{S}^n = \frac{\partial^n}{\partial y^n}(\boldsymbol{\Delta R})|_{y=0},$$

and

$$\boldsymbol{D}_j^n = \frac{\partial^n}{\partial y^n}(\boldsymbol{\Delta M}_j)|_{y=0}$$

This yields

$$\boldsymbol{\Delta R}(\rho) \approx \boldsymbol{S}^0 + \rho \boldsymbol{S}^1 + \tfrac{1}{2}\rho^2 \boldsymbol{S}^2 + \ldots$$

The previous expression, a similar one for the $\boldsymbol{\Delta M}$ matrices and the expansion for ρ,

$$\rho = \rho_1 \phi + \rho_1^* \phi^{-1}, \tag{B 3}$$

are now substituted into (B 2). In more general problems, we would have to include $\rho_2 \phi^2$ terms and (in the circular geometry) even a ρ_0 correction. Here these terms are ruled out because of the north–south symmetry and the conservation of mass.

We then collect like powers of ϕ and like amplitudes, using the amplitude ordering $(\boldsymbol{A}_1 \sim \rho_1)$ and $(\boldsymbol{A}_0 \sim \boldsymbol{A}_2 \sim \rho_2 \sim \rho_1^2)$, to derive a set of matching conditions. We now must solve the linear equations

$$\left.\begin{aligned} \boldsymbol{D}_1^0 \boldsymbol{A}_1 &= -\rho_1 \boldsymbol{S}^1, \\ \boldsymbol{D}_0^0 \boldsymbol{A}_0 &= -\rho_1 \rho_1^* \boldsymbol{S}^2 - \rho_1 \boldsymbol{D}_1^1 \boldsymbol{A}_1^* - \rho_1^* \boldsymbol{D}_1^1 \boldsymbol{A}_1, \\ \boldsymbol{D}_2^0 \boldsymbol{A}_2 &= -\tfrac{1}{2}\rho_1^2 \boldsymbol{S}^2 - \rho_1 \boldsymbol{D}_1^1 \boldsymbol{A}_1, \\ \boldsymbol{D}_1^0 \boldsymbol{A}_1' &= -\tfrac{1}{2}\rho_1^2 \rho_1^* \boldsymbol{S}^3 - \rho_1 \boldsymbol{D}_0^1 \boldsymbol{A}_0 - \rho_1^* \boldsymbol{D}_2^1 \boldsymbol{A}_2 - \rho_1 \rho_1^* \boldsymbol{D}_1^2 \boldsymbol{A}_1 - \tfrac{1}{2}\rho_1^2 \boldsymbol{D}_1^2 \boldsymbol{A}_1^*. \end{aligned}\right\} \tag{B 4}$$

The coefficient $\boldsymbol{A}_1'$ represents the cubic correction to the amplitude vector for the e^{ikx} mode. We could also introduce a ρ_1' term to explicitly represent the order amplitude cubed correction to the interface displacement; however, it is not really needed here and will instead come in when we do a bifurcation analysis. We solve these four equations (noting that although the matrix $\boldsymbol{D}_0^0$ is not invertible, the equation still does have a solution) to find:

$$\boldsymbol{A}_1 = \tfrac{1}{4}\rho 1 \begin{pmatrix} k^{-1} \\ s_1^{-1} \\ k^{-1} \\ s_1^{-1} \end{pmatrix}, \quad \boldsymbol{A}_0 = \tfrac{1}{4}\rho_1 \rho_1^* \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{A}_2 = \tfrac{1}{8}\rho_1^2 \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{A}_1' = \tfrac{1}{8}\rho_1^2 \rho_1^* \begin{pmatrix} k \\ s_1 \\ k \\ s_1 \end{pmatrix}.$$

Now we can evaluate $\Psi^{(1)}(x, \rho(x,t), x, t)$ and $\Psi^{(2)}(x, \mu(x,t), x, t)$ where the lower layer interface displacement is represented by μ. We again use equation (B 1), expanded in amplitude and Fourier modes. For the upper layer, by the previous steps, we have forced $\lim_{y\to\rho} \Psi^{(1)}(y<\rho) = \lim_{y\to\rho} \Psi^{(1)}(y>\rho)$ in the Taylor-expanded sense. For the lower layer, however, we do need to be concerned about switching from the $\Psi^{(2)}(y<\rho)$ form to the $\Psi^{(2)}(y>\rho)$ form when $\mu<\rho$ or $\mu>\rho$, respectively. Although $\Psi^{(2)}$ is continuous everywhere, not all of its derivatives are, so that Taylor expansion could lead to problems. Fortunately, though, the first discontinuity occurs at the level of the fourth derivative (this follows from the equations relating Ψ to ω_p and the fact that $\Psi^{(1)}_{yy}$ is discontinuous). Since in our amplitude expansion we need to use only the first three derivatives of $\Psi^{(2)}$, we can evaluate $\Psi^{(2)}(\mu)$ using either the $y>\rho$ or $y<\rho$ form.

For this version of the expansion, we introduce a matrix

$$\boldsymbol{N} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \\ 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & \mu \end{pmatrix},$$

and represent terms in the following fashion:

$$\boldsymbol{M}_1 \boldsymbol{A}_1 \approx \left(\boldsymbol{M}_1|_{y=0} + \boldsymbol{N}\frac{\partial}{\partial y}\boldsymbol{M}_1|_{y=0} + \tfrac{1}{2}\boldsymbol{N}\boldsymbol{N}\frac{\partial^2}{\partial y^2}\boldsymbol{M}_1|_{y=0} + \ldots \right) \boldsymbol{A}_1. \tag{B 5}$$

We substitute for ρ and μ as in (B 3), and collect like terms to find, finally, that the x-independent and e^{i2kx} modes evaluated on the interfaces are both zero. Thus there is no tendency to excite the second harmonic wave. The results for the primary mode, up to the cubic term, show that indeed the first and third elements of the vector match, as do the second and fourth: this verifies that the conclusion reached in the previous paragraph about the Taylor-expansion is correct. The final expressions are

$$\Psi^{(1)}(\rho) = \tfrac{1}{4}\rho_1\left(\frac{1}{s_1}+\frac{1}{k}-1\right) + \rho_1^2\rho_1^*\tfrac{1}{8}(2k+4s_1-3-s_2),$$

$$\begin{aligned}\Psi^{(2)}(\mu) = {} & \tfrac{1}{4}\rho_1\left(\frac{1}{k}-\frac{1}{s_1}\right) + \tfrac{1}{4}\mu_1 + \tfrac{1}{8}[\rho_1^2\rho_1^*(k-s_1) + 2\rho_1\rho_1^*\mu_1 \\ & + \rho_1^2\mu_1^*(s_2-2k) + 2\rho_1\mu_1\mu_1^*(k-s_1) + \rho_1^*\mu_1^2(k-s_1) + \mu_1^2\mu_1^*].\end{aligned}$$

We now carry out the same operations for the vorticity anomaly in the lower layer, using

$$\boldsymbol{R} = \begin{pmatrix} -\frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{pmatrix} - \begin{pmatrix} \frac{1}{2}y^2 & e^y & 0 & 0 \\ \frac{1}{2}y^2 & -e^y & 0 & 0 \\ 0 & 0 & \frac{1}{2}y^2 & e^{-y} \\ 0 & 0 & \frac{1}{2}y^2 & -e^{-y} \end{pmatrix} \begin{pmatrix} -\frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ -\frac{1}{4} \end{pmatrix},$$

and sum the results. The resultant expression for the streamfunction is then used in the kinematic interface equation to yield the approximate expressions for the time-evolution:

$$\frac{1}{\mathrm{i}k}\frac{\partial}{\partial t}\rho_1 = -\frac{\rho_1}{4}\left(2-\frac{1}{k}-\frac{1}{s_1}\right)-\tfrac{1}{4}\mu_1\left(\frac{1}{k}-\frac{1}{s_1}\right)$$
$$+\tfrac{1}{8}[\rho_1^2\rho_1^*(2k+4s_1-s_2-4)+\rho_1^2\mu_1^*(s_1-k)+2\rho_1\rho_1^*\mu_1(s_1-k)$$
$$-2\rho_1\mu_1\mu_1^*+\rho_1^*\mu_1^2(2k-s_2)+\mu_1^2\mu_1^*(s_1-k)],$$

$$\frac{1}{\mathrm{i}k}\frac{\partial}{\partial t}\mu_1 = \tfrac{1}{4}\rho_1\left(\frac{1}{k}-\frac{1}{s_1}\right)+\tfrac{1}{4}\mu_1\left(2-\frac{1}{s_1}-\frac{1}{k}\right)$$
$$+\tfrac{1}{8}[\rho_1^2\rho_1^*(k-s_1)+\rho_1^2\mu_1^*(s_2-2k)+2\rho_1\rho_1^*\mu_1+2\rho_1\mu_1\mu_1^*(k-s_1)$$
$$+\rho_1^*\mu_1^2(k-s_1)+\mu_1^2\mu_1^*(4+s_2-4s_1-2k)].$$

Standard bifurcation analysis near $k = 1$ (the stability boundary) shows that

(i) At lowest order, $\rho_1 = A(t)$, $\mu_1 = -\rho_1$.

(ii) At next order,

$$\rho_1' + \mu_1' = 4\mathrm{i}\frac{\sqrt{2}-1}{\sqrt{2}}\frac{\partial}{\partial t}A.$$

(iii) At the third order,

$$4\frac{\sqrt{2}-1}{\sqrt{2}}\frac{\partial^2}{\partial t^2}A = -(k-1)A-\tfrac{1}{2}(\sqrt{5}-1)A|A|^2. \tag{B 6}$$

Equation (B 6) shows that the jet is unstable for $k < 1$ and that there is an equilibrium steady solution for

$$|A|^2 = \frac{-2(k-1)}{\sqrt{5}-1}, \tag{B 7}$$

which exists on the unstable side ($k < 1$) of the bifurcation point.

The analysis of this problem with the approach of Pedlosky (1985) (approach (*b*)) differs only slightly from that above. Unfortunately, the problem is very sensitive to small changes. For this method, (B 2) is solved by expansion of the upper layer field at $y = \rho$ and the lower layer flow at $y = \mu$, simultaneously, so that

$$\boldsymbol{\Delta R}(\rho) \approx \boldsymbol{S}^0 + \boldsymbol{N}\boldsymbol{S}^1 + \tfrac{1}{2}\boldsymbol{N}\boldsymbol{N}\boldsymbol{S}^2 + \ldots$$

The linear equations to be solved are basically the same as above with ρ_1 replaced by $\boldsymbol{N}_1$. The only difference in the solution is that

$$\boldsymbol{A}_1' = \begin{pmatrix} -\dfrac{\rho_1\rho_1^*\mu_1}{4k}+\dfrac{\rho_1\mu_1\mu_1^*}{4k}+\dfrac{(k^2+1)\rho_1^2\rho_1^*}{8k}-\dfrac{(k^2+1)\mu_1^2\mu_1^*}{8k}-\dfrac{\rho_1^2\mu_1^*}{8k}+\dfrac{\mu_1^2\rho_1^*}{8k} \\ \dfrac{s_1\rho_1^2\rho_1^*}{8}+\dfrac{s_1\mu_1^2\mu_1^*}{8} \\ -\dfrac{\rho_1\rho_1^*\mu_1}{4k}+\dfrac{\rho_1\mu_1\mu_1^*}{4k}+\dfrac{(k^2+1)\rho_1^2\rho_1^*}{8k}-\dfrac{(k^2+1)\mu_1^2\mu_1^*}{8k}-\dfrac{\rho_1^2\mu_1^*}{8k}+\dfrac{\mu_1^2\rho_1^*}{8k} \\ \dfrac{s_1\rho_1^2\rho_1^*}{8}+\dfrac{s_1\mu_1^2\mu_1^*}{8} \end{pmatrix},$$

whereas the total A_1' from the upper and lower layer vorticity patterns in the previous approach is

$$A_1' = \begin{pmatrix} \frac{1}{8}k(\rho_1^2\rho_1^* - \mu_1^2\mu_1^*) \\ \frac{1}{8}s_1(\rho_1^2\rho_1^* + \mu_1^2\mu_1^*) \\ \frac{1}{8}k(\rho_1^2\rho_1^* - \mu_1^2\mu_1^*) \\ \frac{1}{8}s_1(\rho_1^2\rho_1^* + \mu_1^2\mu_1^*) \end{pmatrix}. \tag{B 8}$$

The different form of A_1' leads, eventually, to an evolution equation very like (B 6) except that the factor $1-\sqrt{5}$ is replaced by $-5\sqrt{5}$. But since this changes the sign of the nonlinear term, the character of the bifurcation is opposite to that predicted by our version.

Solution of the third problem in figure 12 reveals the weakness in the second approach. We follow the procedure outlined for the second approach except that the $\boldsymbol{A}$, $\boldsymbol{S}$ vectors have eight components, corresponding to two matching conditions in each layer at $y = \rho$ and two in each layer at $y = \mu$, the $\boldsymbol{\Delta}$ and $\boldsymbol{M}$ matrices are 8×6 and the $\boldsymbol{D}$ matrices are 8×8. The mean flow vector is given by

$$\boldsymbol{R} = \begin{pmatrix} \frac{1}{2}(1-\mathrm{e}^{y}) \\ -\frac{1}{2}(1-\mathrm{e}^{y}) \\ -\frac{1}{4}y^2 - \frac{1}{2}\sinh(y) \\ -\frac{1}{4}y^2 + \frac{1}{2}\sinh(y) \\ -\frac{1}{2}(1-\mathrm{e}^{-y}) \\ \frac{1}{2}(1-\mathrm{e}^{-y}) \end{pmatrix}.$$

We have calculated the second- and third-order terms only for the case when $\mu = -\rho$ and $k = 1$. The expression for A_1' is the same as for the first procedure: (B 8) with the substitutions just mentioned. We have then looked at the stationary solutions, requiring Ψ = constant on each interface. The results agree with (B 7).

We believe that the second calculation is in error, because neither of the two solutions being used for matching is correct in the region $\rho < y < \mu$ or $\mu < y < \rho$. That is, we cannot directly write an analogous equation to (B 1) if we wish to represent $\Psi^{(1)}(y < \rho \text{ or } y > \rho)$ and $\Psi^{(2)}(y < \mu \text{ or } y > \mu)$, because this representation is not exact unless we explicitly include the regions between the two curves as in the third calculation. Although this region is very narrow, the correction which arises from it enters at the cubic order. The procedure we shall use for the circular problem, the first one in (B 1), does have the flow in this region represented properly to this order as a sum of that arising from the upper layer vorticity distribution and that arising from the lower layer; i.e. (B 1) is an exact representation of that part of the flow from the upper layer vorticity distribution. (It is the case, however, that our solution cannot be pressed to the next order without more explicit concern for this area.)

Appendix C. Circular vortex finite amplitude calculations

The circular vortex problem is very similar to the jet problem considered in Appendix B, except that

(i) There are both θ-independent and second-harmonic corrections to the interface displacement. These must be included explicitly both in the calculation of the

streamfunction as a functional of the perturbation of one of the boundaries and in the calculation of the nonlinear eigenvalue problem.

(ii) We consider four interfaces, inner upper and lower layer boundaries located at $1+\eta^{(1)}$ and $1+\eta^{(2)}$ and outer boundaries at $b+\tau^{(1)}$, $b+\tau^{(2)}$.

(iii) Only steady state solutions are obtained.

First, we outline the calculation of the flow based on a potential vorticity anomaly q within an area $\mathscr{D}$ of radius $a_0+\rho$ in the upper layer. We assume the curve $\partial\mathscr{D}_j$ is nearly circular:

$$\partial\mathscr{D}_j: \quad r = a(\theta) = [a_0]+[\rho_1\phi+\text{c.c.}]+[\rho_0+\rho_2\phi^2+\text{c.c.}]+[\rho_1'\phi+\text{c.c.}+\phi^3 \text{ terms}], \tag{C 1}$$

where $\phi = \exp(im\theta)$. The amplitude parameter A will be taken equal to the absolute value of a_1 for the upper layer, innermost boundary; all of the other first-order terms will be proportional to A. In (C 1), the square brackets show terms collected by amplitude ordering; the θ-independent term at $O(A^2)$ is chosen so that the area with the perturbed patch remains at πa_0^2 including all order A^2 terms. Based on this comment, we have

$$\rho_0 = -\frac{|\rho_1|^2}{a_0}.$$

The equation for $\boldsymbol{\Psi}$ is the same as (B 1) except that we now explicitly order the terms and indicate the A^3 corrections:

$$\boldsymbol{\Psi} = \begin{pmatrix} \Psi^{(1)}(r<a) \\ \Psi^{(2)}(r<a) \\ \Psi^{(1)}(r>a) \\ \Psi^{(2)}(r>a) \end{pmatrix} = [\boldsymbol{R}]+[\boldsymbol{M}_1(\boldsymbol{A}_1\phi+\boldsymbol{A}_1^*\phi^{-1})]+[\boldsymbol{M}_0\boldsymbol{A}_0+\boldsymbol{M}_2(\boldsymbol{A}_2\phi^2+\boldsymbol{A}_2^*\phi^{-2})]$$
$$+[\boldsymbol{M}_1(\boldsymbol{A}_1'\phi+\boldsymbol{A}_1'^*\phi^{-1})+\text{third harmonics}]\ldots). \tag{C 2}$$

The matrices are given by

$$\boldsymbol{M}_j = \begin{pmatrix} r^m & I_m(\lambda r) & 0 & 0 \\ r^m & -\delta I_m(\lambda r) & 0 & 0 \\ 0 & 0 & r^{-m} & K_m(\lambda r) \\ 0 & 0 & r^{-m} & -\delta K_m(\lambda r) \end{pmatrix},$$

with the mode number m being $2j$. We represent the vector $\boldsymbol{R}$ by

$$\boldsymbol{R} = \int^r \boldsymbol{M}_{\frac{1}{2}}\boldsymbol{V}+\text{const. vector}$$

where $\boldsymbol{V}$ is related to the patch vorticity by

$$\boldsymbol{V} = \boldsymbol{D}_{\frac{1}{2}}^{-1}\begin{pmatrix} 0 \\ 0 \\ -q \\ 0 \end{pmatrix}.$$

The matrices $\boldsymbol{D}_j^n$ and vectors $\boldsymbol{S}^n$ will be defined as in Appendix B using the $\boldsymbol{\Delta}$ matrix. (The y derivatives are replaced by r derivatives, of course.) These matrices, when

multiplied by the appropriate amplitude vectors $\boldsymbol{A}$ and Fourier functions ϕ, give the step in nth and $n+$first derivative in the upper and lower halves, respectively.

The linear equations giving the amplitudes $\boldsymbol{A}_j$ are quite similar to (B 4) except for the additional corrections to the vortex shape:

$$\left.\begin{aligned}
&\boldsymbol{D}_1^0 \boldsymbol{A}_1 = -\rho_1 \boldsymbol{S}^1, \\
&\boldsymbol{D}_0^0 \boldsymbol{A}_0 = -\rho_0 \boldsymbol{S}^1 - \rho_1^2 \boldsymbol{S}^2 - 2\rho_1 \boldsymbol{D}_1^1 \boldsymbol{A}_1, \\
&\boldsymbol{D}_2^0 \boldsymbol{A}_2 = -\rho_2 \boldsymbol{S}^1 - \tfrac{1}{2}\rho_1^2 \boldsymbol{S}^2 - \rho_1 \boldsymbol{D}_1^1 \boldsymbol{A}_1, \\
&\boldsymbol{D}_1^0 \boldsymbol{A}_1' = -(\rho_0\rho_1 + \rho_2\rho_1)\boldsymbol{S}^2 - \tfrac{1}{2}\rho_1^3 \boldsymbol{S}^3 - \rho_1 \boldsymbol{D}_0^1 \boldsymbol{A}_0 - \rho_1 \boldsymbol{D}_2^1 \boldsymbol{A}_2 - (\rho_0+\rho_2)\boldsymbol{D}_1^1 \boldsymbol{A}_1 - \tfrac{3}{2}\rho_1^2 \boldsymbol{D}_1^2 \boldsymbol{A}_1.
\end{aligned}\right\} \tag{C 3}$$

Here, we have also explicitly used the fact that all fields are real for the equilibrium case when the vortex motion is steady. The $\boldsymbol{D}$ matrices and $\boldsymbol{S}$ vectors are evaluated at $r = a_0$ here. These equations are solved numerically to give the vector proportionality constants relating, for example, $\boldsymbol{A}_2$ to ρ_2 and ρ_1^2. Thus we have

$$\left.\begin{aligned}
\boldsymbol{A}_1 &= \boldsymbol{\alpha}_1 \rho_1, \\
\boldsymbol{A}_0 &= \boldsymbol{\alpha}_0 \rho_0 + \boldsymbol{\alpha}_{01} \rho_1^2, \\
\boldsymbol{A}_2 &= \boldsymbol{\alpha}_2 \rho_2 + \boldsymbol{\alpha}_{21} \rho_1^2, \\
\boldsymbol{A}_1' &= \boldsymbol{\alpha}_{10} \rho_0 \rho_1 + \boldsymbol{\alpha}_{12} \rho_2 \rho_1 + \boldsymbol{\alpha}_{11} \rho_1^3.
\end{aligned}\right\} \tag{C 4}$$

The next step is to relate the streamfunction values on each of the interfaces to the amplitudes $\boldsymbol{A}_j$, analogous to the step (B 5). Because we have four interfaces, we cannot represent the result in quite the same fashion; rather we seek

$$\boldsymbol{\Psi} = \begin{pmatrix} \Psi^{(1)}(r = 1+\eta^{(1)}) \\ \Psi^{(2)}(r = 1+\eta^{(2)}) \\ \Psi^{(1)}(r = b+\tau^{(1)}) \\ \Psi^{(2)}(r = b+\tau^{(2)}) \end{pmatrix}.$$

With the definitions

$$\boldsymbol{N} = \begin{pmatrix} \eta^{(1)} & 0 & 0 & 0 \\ 0 & \eta^{(2)} & 0 & 0 \\ 0 & 0 & \tau^{(1)} & 0 \\ 0 & 0 & 0 & \tau^{(2)} \end{pmatrix},$$

and

$$\boldsymbol{\Phi}^n = \begin{pmatrix} \dfrac{\partial^n}{\partial r^n} \Psi^{(1)}(r = 1) \\ \dfrac{\partial^n}{\partial r^n} \Psi^{(2)}(r = 1) \\ \dfrac{\partial^n}{\partial r^n} \Psi^{(1)}(r = b) \\ \dfrac{\partial^n}{\partial r^n} \Psi^{(2)}(r = b) \end{pmatrix},$$

we can write

$$\boldsymbol{\Psi} = \boldsymbol{\Phi}^0 + \boldsymbol{N}\boldsymbol{\Phi}^1 + \tfrac{1}{2}\boldsymbol{N}\boldsymbol{N}\boldsymbol{\Phi}^2 + \ldots \tag{C 5}$$

The vectors $\boldsymbol{\Phi}^n$ can be evaluated from the coefficients $\boldsymbol{A}_j$, or the $\boldsymbol{\alpha}$ vectors, by matrix multiplication. Thus, we have

$$\left.\begin{aligned}
\boldsymbol{\Phi}^3 &= \boldsymbol{R}_{rrr}|_{r=1,b},\\
\boldsymbol{\Phi}^2 &= \boldsymbol{R}_{rr}|_{r=1,b} + \boldsymbol{P}^2_{1,1}\rho_1(\phi+\phi^{-1}),\\
\boldsymbol{\Phi}^1 &= \boldsymbol{R}_r|_{r=1,b} + \boldsymbol{P}^1_{1,1}\rho_1(\phi+\phi^{-1}) + (\boldsymbol{P}^1_{0,0}\rho_0 + \boldsymbol{P}^1_{0,01}\rho_1^2) + (\boldsymbol{P}^1_{2,2}\rho_2 + \boldsymbol{P}^1_{2,21}\rho_1^2)(\phi^2+\phi^{-2}),\\
\boldsymbol{\Phi}^0 &= \boldsymbol{R}|_{r=1,b} + (\boldsymbol{P}^0_{1,1}[\rho_1+\rho_1'] + \boldsymbol{P}^0_{1,10}\rho_1\rho_0 + \boldsymbol{P}^0_{1,12}\rho_1\rho_2 + \boldsymbol{P}^0_{1,11}\rho_1^3)(\phi+\phi^{-1})\\
&\quad + (\boldsymbol{P}^0_{0,0}\rho_0 + \boldsymbol{P}^0_{0,01}\rho_1^2) + (\boldsymbol{P}^0_{2,2}\rho_2 + \boldsymbol{P}^0_{2,21}\rho_1^2)(\phi^2+\phi^{-2}),
\end{aligned}\right\} \tag{C 6}$$

where the vectors $\boldsymbol{P}^n_{q,s}$ are given by

$$\boldsymbol{P}^n_{q,s} = \left(\frac{\partial^n}{\partial r^n}\boldsymbol{M}_q|_{r=1,b}\right)\boldsymbol{\alpha}_s,$$

and the notation $r = 1, b$ stands for evaluating the first two rows at $r = 1$ and the second two at $r = b$. Thus the superscript indicates the order of the r derivative, and the subscripts are the Fourier mode and a marker of the dependence on the shape of $\partial\mathscr{D}_j$. These $\boldsymbol{P}$ vectors are the fundamental results of the first stage of the computation.

In the final part of the first stage, we next need to combine the results for the four different patches. For each of the patches, we follow the same sequence of steps, except the $\boldsymbol{V}$, $\boldsymbol{R}$, and a values are different. We can now define the sum of the four $\boldsymbol{\Phi}^n$ vectors by, for example,

$$\Sigma\boldsymbol{\Phi}^j = \Sigma\boldsymbol{R}_r|_{r=1,b} + \boldsymbol{P}^j_{1,1}\boldsymbol{\eta}_1(\phi+\phi^{-1}) + (\boldsymbol{P}^j_{0,0}\boldsymbol{\eta}_0 + \boldsymbol{P}^j_{0,01}\boldsymbol{\eta}_1^2) + (\boldsymbol{P}^j_{2,2}\boldsymbol{\eta}_2 + \boldsymbol{P}^j_{2,21}\boldsymbol{\eta}_1^2)(\phi^2+\phi^{-2}).$$

Here the matrices $\boldsymbol{P}$ are constructed by stacking the four corresponding $\boldsymbol{P}$ vectors associated with patches one to four side-by-side to form four by four matrices. The vectors such as $\boldsymbol{\eta}_1^2\boldsymbol{\eta}_2$ will be defined by

$$\boldsymbol{\eta}_1^2\boldsymbol{\eta}_2 = \begin{pmatrix} \eta_1^{(1)}\eta_1^{(1)}\eta_2^{(1)} \\ \eta_1^{(2)}\eta_1^{(2)}\eta_2^{(2)} \\ \tau_1^{(1)}\tau_1^{(1)}\tau_2^{(1)} \\ \tau_1^{(2)}\tau_1^{(2)}\tau_2^{(2)} \end{pmatrix},$$

etc. Altogether, the fourteen $\boldsymbol{P}$ matrices give us all we need for the calculation of the functional dependence of Ψ on each boundary upon the shapes of the various boundaries.

The second stage, satisfying the kinematic boundary conditions, written as

$$\boldsymbol{\Psi} = \Omega(\boldsymbol{B}\boldsymbol{\eta} + \tfrac{1}{2}\boldsymbol{\eta}^2),$$

with

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & b & 0 \\ 0 & 0 & 0 & b \end{pmatrix},$$

proceeds by making the same kind of expansion and gathering of Fourier components, using (C 5) and the matrix form of (C 6). At lowest order, we have

$$\boldsymbol{N}_1 \Sigma \boldsymbol{R}_r + \boldsymbol{P}^0_{1,1}\boldsymbol{\eta}_1 = \Omega \boldsymbol{B}\boldsymbol{\eta}_1.$$

Since $\boldsymbol{N}_1$ is diagonal, we can rewrite the first term by defining a matrix $\boldsymbol{R}_r$ which has the elements of $\Sigma \boldsymbol{R}_r$ on the diagonal and zeros off-diagonal. Then we have an ordinary eigenvalue–eigenvector problem

$$\boldsymbol{U}_1\boldsymbol{\eta}_1 = (\boldsymbol{R}_r + \boldsymbol{P}^0_{1,1})\boldsymbol{\eta}_1 = \Omega \boldsymbol{B}\boldsymbol{\eta}_1$$

and we can find the eigenvalue, eigenvector and the adjoint vector

$$\boldsymbol{\eta}_1^+(\boldsymbol{U}_1 - \Omega \boldsymbol{B}) = 0.$$

This problem can, of course, be converted into (5.1) by a similarity transformation. It appears different only because it is written in terms of layers and streamfunction Fourier mode amplitudes rather than vertical modes and boundary perturbation amplitudes.

At the next order, we have a matrix problem to solve for the second harmonic corrections to the interface shape:

$$(\boldsymbol{P}^0_{2,2} + \boldsymbol{R}_r - \Omega \boldsymbol{B})\boldsymbol{\eta}_2 = \tfrac{1}{2}\Omega \boldsymbol{\eta}_1^2 - \boldsymbol{P}^0_{2,21}\boldsymbol{\eta}_1^2 - \boldsymbol{N}_1 \boldsymbol{P}^1_{1,1}\boldsymbol{\eta}_1 - \tfrac{1}{2}\boldsymbol{R}_{rr}\boldsymbol{\eta}_1^2$$

which is solved numerically for $\boldsymbol{\eta}_2$. The correction to the θ independent interface position is already known by conservation of area to be

$$\boldsymbol{\eta}_0 = -\boldsymbol{B}^{-1}\boldsymbol{\eta}_1^2.$$

At the last order, we pick up the corrections to η_1 (η_1', the cubic nonlinearity terms) and the correction to the frequency Ω. The result is

$$\begin{aligned}(\boldsymbol{M}_1 - \Omega \boldsymbol{B})\boldsymbol{\eta}_1' &+ \boldsymbol{P}^0_{1,10}\boldsymbol{\eta}_1\boldsymbol{\eta}_0 + \boldsymbol{P}^0_{1,12}\boldsymbol{\eta}_1\boldsymbol{\eta}_2 + \boldsymbol{P}^0_{1,11}\boldsymbol{\eta}_1^3 \\ &+ \boldsymbol{N}_1(\boldsymbol{P}^1_{0,0}\boldsymbol{\eta}_0 + \boldsymbol{P}^1_{0,01}\boldsymbol{\eta}_1^2 + \boldsymbol{P}^1_{2,2}\boldsymbol{\eta}_2 + \boldsymbol{P}^1_{0,21}\boldsymbol{\eta}_1^2) + (\boldsymbol{N}_0 + \boldsymbol{N}_2)\boldsymbol{P}^1_{1,1}\boldsymbol{\eta}_1 + \tfrac{3}{2}\boldsymbol{N}_1\boldsymbol{N}_1\boldsymbol{P}^2_{1,1}\boldsymbol{\eta}_1 \\ &+ \boldsymbol{R}_{rr}(\boldsymbol{\eta}_0\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2\boldsymbol{\eta}_1) + \tfrac{1}{2}\boldsymbol{R}_{rrr}\boldsymbol{\eta}_1^3 - \Omega\boldsymbol{\eta}_0\boldsymbol{\eta}_1 - \Omega\boldsymbol{\eta}_2\boldsymbol{\eta}_1 - \frac{\mathrm{d}\Omega}{\mathrm{d}A^2}\boldsymbol{B}\boldsymbol{\eta}_1 = 0.\end{aligned}$$

Multiplying by the adjoint vector $\boldsymbol{\eta}_1^+$ and normalizing by the factor multiplying the last term leads to the final expression for the cubic term $\mathrm{d}\Omega/\mathrm{d}A^2$.

This work was supported by grants from the National Science Foundation (OCE-8504080) and the Office of Naval Research (N00014-86-K-0325).

REFERENCES

Benny, D. J. & Maslowe, S. A. 1975 The evolution in space and time of nonlinear waves in parallel shear flows. *Stud. Appl. Maths* **54**, 181–205.

Biebuyck, G. L. 1986 Self propagation of a barotropic circular eddy. In *Summer Study Program in Geophysical Fluid Dynamics. Woods Hole Oceanographic Institution Tech. Rep.* WHOI-86-45, pp. 193–197.

Childress, S. 1984 A vortex-tube model of eddies in the inertial range. *Geophys. Astrophys. Fluid Dyn.* **29**, 29–64.

Dritschel, D. G. 1986 The nonlinear evolution of rotating configurations of uniform vorticity. *J. Fluid Mech.* **172**. 157–182.

FLIERL, G. R. 1978 Models of vertical structure and the calibration of two-layer models. *Dyn. Atmos. Oceans* **2**, 341–381.

FLIERL, G. R. 1984 The emergence of dipoles from instabilities on the f and beta planes. In *Summer Study Program in Geophysical Fluid Dynamics. Woods Hole Oceanographic Institution Tech. Rep.* WHOI-84-44, pp. 104–110.

FLIERL, G. R. 1987 Isolated eddy models in geophysics. *Ann. Rev. Fluid Mech.* **19**, 493–530.

FU, L. L. & FLIERL, G. R. 1980 Energy and enstrophy transfers in a realistically stratified ocean. *Dyn. Atmos. Oceans* **5**, 1–41.

GENT, P. R. & MCWILLIAMS, J. C. 1986 The instability of circular vortices. *Geophys. Astrophys. Fluid Dyn.* **35**, 209–233.

GRIFFITHS, R. W. & LINDEN, P. F. 1981 The stability of vortices in a rotating stratified fluid. *J. Fluid Mech.* **117**, 343–377.

HART, J. E. 1974 On the mixed stability problem for quasi-geostrophic ocean currents. *J. Phys. Oceanogr.* **4**, 349–356.

HELFRICH, K. R. & SEND, U. 1988 Finite-amplitude evolution of two-layer geostrophic vortices. *J. Fluid Mech.* **197**, 331–348.

HOGG, N. G. & STOMMEL, H. M. 1985 The heton, an elementary interaction between discrete baroclinic geostrophic vortices, and its implications concerning heat flow. *Proc. R. Soc. Lond.* A **397**, 1–20.

HOSKINS, B. J., MCINTYRE, M. E. & ROBERTSON, A. W. 1985 On the use and significance of isentropic potential vorticity maps. *Q. J. R. Met. Soc.* **111**, 877–949.

HOWARD, L. N. & GUPTA, A. S. 1962 On the hydrodynamic and hydromagnetic stability of swirling flows. *J. Fluid Mech.* **14**, 463–476.

IKEDA, M. 1981 Instability and splitting of mesoscale rings using a two-layer quasi-geostrophic model on an f-plane. *J. Phys. Oceanogr.* **11**, 987–998.

JOYCE, T. 1984 Velocity and hydrographic structure of a Gulf Stream warm core ring. *J. Phys. Oceanogr.* **14**, 936–947.

LINDZEN, R. S. & TUNG, K. K. 1978 Wave overreflection and shear instability. *J. Atmos. Sci.* **35**, 1626–1632.

MALANOTTE-RIZZOLI, P. 1982 Planetary solitary waves in geophysical flows. *Adv. Geophys.* **24**, 147–224.

MICHALKE, A. & TIMME, A. 1967 On the inviscid instability of certain two-dimensional vortex-type flows. *J. Fluid Mech.* **29**, 647–666.

OLSON, D. B., SCHMITT, R., KENNELLY, M. & JOYCE, T. 1985 Two-layer diagnostic model of the long-term physical evolution of warm-core ring 82B. *J. Geophys. Res.* **90**, 8813–8822.

PEDLOSKY, J. 1985 The instability of continuous heton clouds. *J. Atmos. Sci.* **42**, 1477–1486.

POLVANI, L. M., ZABUSKY, N. J. & FLIERL, G. R. 1988 Two-layer geostrophic V-states and merger. 1. Constant potential vorticity lower layer. Submitted to *J. Fluid Mech.*

SAUNDERS, P. M. 1973 The instability of a baroclinic vortex. *J. Phys. Oceanogr.* **3**, 61–65.

ZABUSKY, N. J., HUGHES, M. H. & ROBERTS, K. V. 1979 Contour dynamics for the Euler equations in two dimensions. *J. Comp. Phys.* **30**, 96–106.