On Non-Geostrophic Baroclinic Stability

PETER H. STONE
Harvard University, Cambridge, Mass.

(Manuscript received 18 November 1965, in revised form 28 March 1966)

ABSTRACT

Eady's (1949) model is used to study the non-geostrophic baroclinic stability problem. Growth rates for various types of perturbations are found as a function of the Richardson number, Ri. The results indicate that the conventional baroclinic instabilities dominate if Ri>0.95; symmetric instabilities dominate if $\frac{1}{4}$<Ri<0.95; and Kelvin-Helmholtz instabilities dominate if Ri<$\frac{1}{4}$. It is suggested that symmetric instabilities may play an important role in the dynamics of the atmospheres of the major planets of the solar system.

1. Introduction

Previous studies of the stability of a baroclinic zonal current have revealed two basically different types of instabilities. One is the conventional baroclinic instability which draws its energy from the potential energy of the unperturbed zonal flow [e.g., see Eady (1949)]. The other is a symmetric instability (no longitudinal variations), which occurs when the Richardson number, Ri, is less than unity, and draws its energy from the kinetic energy of the unperturbed zonal flow [e.g., see Eliassen and Kleinschmidt (1957, p. 64)]. The first type of instability has received the most attention in the meteorological literature because of its importance to the dynamics of the earth's atmosphere. However, observations of the atmospheres of the major planets of the solar system reveal a very marked preference for the type of symmetry displayed by the second type of instability, and it is possible that these symmetric instabilities play an equally important role in the dynamics of these atmospheres.

Such a role is possible only if there is some range of conditions under which the symmetric instabilities have the largest growth rates, so that they will dominate the other types of instabilities which may occur. No published studies answer the question of whether there is in fact such a range of conditions. Since the symmetric instabilities can only occur if Ri<1, an analysis of the non-geostrophic baroclinic stability problem is required to answer this question. The published studies of the conventional baroclinic instabilities under non-geostrophic conditions have neglected the latitudinal variation of the perturbations [see Phillips (1964) for a discussion of this neglect] and therefore exclude *a priori* any symmetric instabilities. Similarly the published studies of symmetric instabilities have neglected the longitudinal variations of the perturbations and therefore also exclude *a priori* the conventional baroclinic instabilities. Both latitudinal and longitudinal variations must be taken into account if one is to determine which type of instability has the largest growth rates. In addition one cannot exclude the possibility that instabilities different from the two kinds mentioned above may occur under non-geostrophic conditions.

In an attempt to determine if there is some range of conditions under which symmetric instabilities will dominate, we present in this paper a study of the non-geostrophic baroclinic stability problem. Because of the mathematical complexity of the problem when both latitudinal and longitudinal variations of the perturbations are included, we will use the simplest possible model, i.e., that due to Eady (1949). This model neglects dissipative and curvature effects and uses the Boussinesq approximation for compressibility effects. Consequently this model is too simplified a one to apply directly to the atmospheres of the major planets (which are probably very deep), and we will leave a detailed discussion of the dynamics of these atmospheres to another paper.

If x, y and z are the dimensionless longitudinal, latitudinal, and vertical coordinates, and u, v and w are the corresponding dimensionless velocity components, then in Eady's model the unperturbed wind and dimensionless potential temperature, θ, have the forms

$$\left.\begin{aligned} u&=z\\ v&=w=0\\ \theta&=z-\frac{y}{\mathrm{Ri}} \end{aligned}\right\}, \qquad (1.1)$$

and the thermal wind relation is satisfied. The perturbations of this flow are assumed to be small and to have time and space dependences of the form $\exp[i(\sigma t+kx+\lambda y)]$. Thus σ is the complex frequency (instability will occur if the imaginary part of σ is negative), k the longitudinal wave number and λ

the latitudinal wave number. The resulting eigenvalue problem is [see Eady (1949, Eq. II, 11)]

$$[1-(\sigma+kz)^2]\frac{d^2w}{dz^2}-\left[\frac{2k}{\sigma+kz}-2i\lambda\right]\frac{dw}{dz}-\left[\mathrm{Ri}(k^2+\lambda^2)+\frac{2ik\lambda}{\sigma+kz}\right]w=0, \quad (1.2)$$

and

$$w=0 \quad \text{when} \quad z=0, 1, \quad (1.3)$$

where w is the perturbation vertical velocity.

The dimensionless σ, k and λ are related to the same quantities with dimensions, $\bar{\sigma}$, $\bar{k}$ and $\bar{\lambda}$, by the relations

$$\left.\begin{aligned}\bar{\sigma}&=f\sigma\\ \bar{k}&=\frac{fk}{u_0}\\ \bar{\lambda}&=\frac{f\lambda}{u_0}\end{aligned}\right\}. \quad (1.4)$$

Here f is the coriolis parameter and u_0 is the magnitude of the unperturbed zonal wind. The Richardson number is defined as

$$\mathrm{Ri}=\frac{gz_0^2}{u_0^2}\frac{1}{\bar{\theta}}\frac{\partial\bar{\theta}}{\partial\bar{z}}, \quad (1.5)$$

where g is the gravitational acceleration, z_0 is the distance between the two horizontal boundaries of the model, and $\bar{z}$ and $\bar{\theta}$ are the dimensional analogues of z and θ.

Eq. (1.2) shows that the stability properties of the baroclinic zonal flow in Eady's model are completely determined by the magnitude of Ri. Eq. (1.2) has been solved when Ri$\gg$1 by Eady (1949); when Ri$=\lambda=0$ by Fjørtoft (1950); and when $\lambda=0$ by Årnason (1963). For each different value of Ri, σ will be a different function of the horizontal wave numbers of the perturbation, k and λ, and it is convenient to think of σ for a fixed Ri as varying over a k, λ plane with k as the abscissa and λ as the ordinate. All the instabilities contained in the solution to (1.2) are, strictly speaking, baroclinic instabilities, but in the literature those perturbations which grow when $k=0$ have generally been referred to as symmetric instabilities, while those which grow when $\lambda=0$ have been referred to as baroclinic instabilities. Consequently we will refer to the λ-axis as the symmetric axis and to the k-axis as the baroclinic axis. The reader may find it useful to refer to Fig. 5 in following the mathematical development.

2. Growth rates near the symmetric axis

In order to find how σ depends on λ and Ri when $k\ll1$, we must know the power of the leading term of the series expansion of σ in powers of k. An examination of the various possibilities shows that (1.2) will have nontrivial solutions either if $\sigma\sim1$ or if $\sigma\sim k$ as $k\rightarrow0$. Årnason's (1963) solutions for $\lambda=0$ give the result that $\sigma\sim k$ as $k\rightarrow0$. Therefore to begin with we shall replace σ by the phase speed,

$$c=\frac{\sigma}{k}, \quad (2.1)$$

and assume that $c\sim1$ as $k\rightarrow0$. Eq. (1.2) then reduces to the form

$$\frac{d^2w}{dz^2}-\left(\frac{2}{c+z}-2i\lambda\right)\frac{dw}{dz}-\left(\mathrm{Ri}\lambda^2+\frac{2i\lambda}{c+z}\right)w=0. \quad (2.2)$$

This is a circular Bessel equation and has as its general solution

$$w=e^{-i\lambda(c+z)}[a(\sin x-x\cos x)+b(\cos x+x\sin x)], \quad (2.3)$$

where

$$x=\lambda(1-\mathrm{Ri})^{\frac{1}{2}}(c+z). \quad (2.4)$$

Applying boundary conditions (1.3) and setting the determinant of the coefficients a and b equal to zero, we obtain the equation.

$$\frac{x_1-\tan x_1}{1+x_1\tan x_1}=\frac{x_2-\tan x_2}{1+x_2\tan x_2}, \quad (2.5)$$

where

$$x_1=\lambda(1-\mathrm{Ri})^{\frac{1}{2}}c, \quad (2.6)$$

and

$$x_2=\lambda(1-\mathrm{Ri})^{\frac{1}{2}}(c+1). \quad (2.7)$$

Making use of the trigonometric identity

$$\left.\begin{aligned}\frac{x-\tan x}{1+x\tan x}&=\tan(\eta-x)\\ \eta&=\arctan x\end{aligned}\right\}, \quad (2.8)$$

we can reduce (2.5) to a quadratic for c, whose solution is

$$c=-\tfrac{1}{2}\pm\tfrac{1}{2}\left[1-4\left(\frac{1-\alpha\,\mathrm{ctn}\,\alpha}{\alpha^2}\right)\right]^{\frac{1}{2}}, \quad (2.9)$$

where

$$\alpha=\lambda(1-\mathrm{Ri})^{\frac{1}{2}}\equiv i\beta. \quad (2.10)$$

The behavior of $(c+\frac{1}{2})^2$ as a function of α is shown in Fig. 1 for Ri$=2$, 1, and $\frac{1}{2}$. For Ri>1 the imaginary part of c is a maximum at $\lambda=0$ and falls to zero at $\beta=2.40$; for Ri$=1$ c is a constant; and for Ri<1 the imaginary part of c has a minimum at $\lambda=0$ and singu-

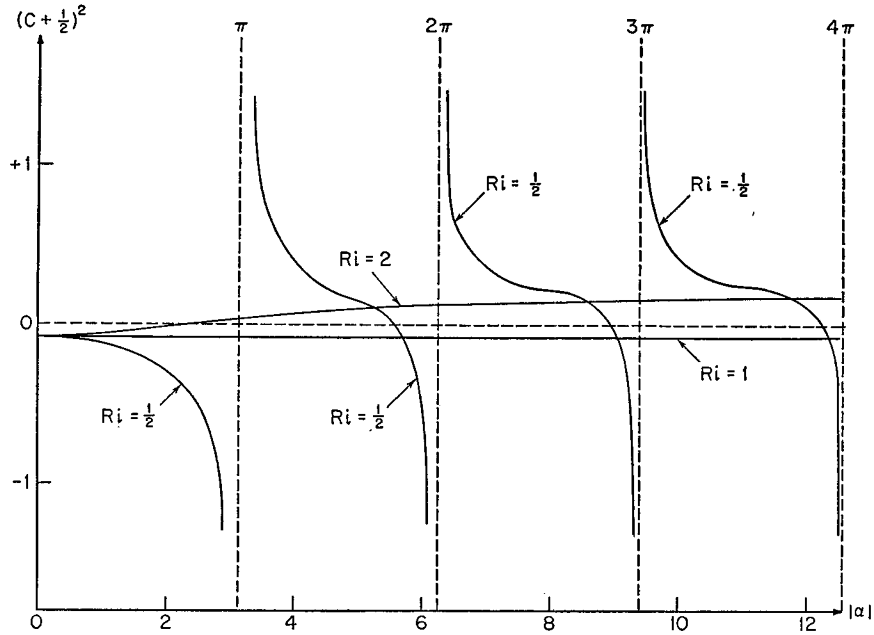

FIG. 1. $(c+\frac{1}{2})^2$ for small k vs. α when Ri=2, 1 and $\frac{1}{2}$, as calculated from Eq. (2.9).

larities at $\alpha=m\pi$, $m=1, 2, 3, \cdots$. If we let

$$\alpha=m\pi-\epsilon, \tag{2.11}$$

and assume that $\epsilon\ll 1$, then (2.9) simplifies to

$$c^2\cong-\frac{1}{m\pi\epsilon}. \tag{2.12}$$

Since we assumed $c\sim 1$ in deriving (2.9), it will not be valid for small ϵ. In particular, the neglected terms in (1.2) are small compared to all the retained terms only if $kc^2\ll 1$. Therefore (2.9) ceases to be valid for $k\ll 1$ if $\epsilon\sim k$, since then $c\sim k^{-\frac{1}{2}}$.

To find the solution as $k\to 0$ if $\epsilon\sim k$ we adopt a new scaling, $c\sim k^{-\frac{1}{2}}$. Retaining only terms with contributions of order unity, $k^{\frac{1}{2}}$, or k in (1.2), we obtain

$$(1-k^2c^2)\frac{d^2w}{dz^2}-\left(\frac{2}{c+z}-2i\lambda\right)\frac{dw}{dz}-\left(\mathrm{Ri}\lambda^2+\frac{2i\lambda}{c+z}\right)w=0. \tag{2.13}$$

If we now let $w=uv$, choose v so that the equation for u assumes normal form, and again only retain terms which have contributions of order unity, $k^{\frac{1}{2}}$, or k, we obtain

$$\frac{d^2u}{dz^2}+\left(\alpha^2+\mathrm{Ri}k^2\lambda^2c^2+2\alpha^2k^2c^2-\frac{2}{c^2}\right)u=0. \tag{2.14}$$

The general solution of this equation is

$$u=ae^{r_+z}+be^{r_-z}, \tag{2.15}$$

where

$$r_\pm=\pm\left(\frac{2}{c^2}-\alpha^2-\mathrm{Ri}k^2\lambda^2c^2-2\alpha^2k^2c^2\right)^{\frac{1}{2}}. \tag{2.16}$$

Applying boundary conditions (1.3) to solution (2.15), we obtain the following requirement for a non-trivial solution,

$$r_+-r_-=2m\pi i,\ m=1, 2, 3, \cdots \tag{2.17}$$

This equation is a quadratic in c^2 and its solution is

$$c^2=\left\{m^2\pi^2-\alpha^2\pm\left[(m^2\pi^2-\alpha^2)^2+8\alpha^2k^2\left(\frac{2-\mathrm{Ri}}{1-\mathrm{Ri}}\right)\right]^{\frac{1}{2}}\right\}\frac{1-\mathrm{Ri}}{2\alpha^2k^2(2-\mathrm{Ri})}. \tag{2.18}$$

This solution is consistent with our assumption that $c\sim k^{-\frac{1}{2}}$ only if $\epsilon\ll 1$, i.e., only near the singular points of solution (2.9). Replacing α by $m\pi-\epsilon$ in (2.18) and assuming that $\epsilon\ll 1$, we obtain

$$c^2=\left\{\epsilon\pm\left[\epsilon^2+2k^2\left(\frac{2-\mathrm{Ri}}{1-\mathrm{Ri}}\right)\right]^{\frac{1}{2}}\right\}\frac{1-\mathrm{Ri}}{m\pi(2-\mathrm{Ri})k^2}. \tag{2.19}$$

If we keep ϵ fixed and let $k\to 0$, the plus root of (2.19) simplifies to

$$c^2\cong\left\{\begin{array}{ll}\dfrac{2\epsilon}{m\pi k^2}\left(\dfrac{1-\mathrm{Ri}}{2-\mathrm{Ri}}\right), & \text{if } \epsilon>0\\[2ex] -\dfrac{1}{m\pi\epsilon}, & \text{if } \epsilon<0\end{array}\right\}, \tag{2.20}$$

while the minus root simplifies to

$$c^2 \cong \begin{cases} = \dfrac{1}{m\pi\epsilon}, & \text{if} \quad \epsilon>0 \\ \dfrac{2\epsilon}{m\pi k^2}\left(\dfrac{1-\mathrm{Ri}}{2-\mathrm{Ri}}\right), & \text{if} \quad \epsilon<0 \end{cases}. \tag{2.21}$$

These expressions show that the plus root of solution (2.19) is the analytic continuation of solution (2.9) on the left ($\epsilon<0$) side of its singularities, while the minus root of (2.19) is the analytic continuation of (2.9) on the right ($\epsilon>0$) side of its singularities. Thus the different branches of the curve shown in Fig. 1 for $\mathrm{Ri}=\frac{1}{2}$ do not join each other. If we move along one of the branches through one of the singularities, (2.20) and (2.21) show that we come out on the other side of the singularities with values of c^2 which behave like k^{-2} instead of k^{-1}. Therefore, if ϵ is not $\ll 1$, on the side of the singularities where (2.9) is not a valid description of the various branches of c, a third scaling is indicated, i.e., $c\sim k^{-1}$.

Again returning to Eq. (1.2), and this time assuming $c\sim k^{-1}$ ($\sigma\sim 1$) and letting $k\to 0$, we now obtain the equation

$$(1-\sigma^2)\frac{d^2w}{dz^2}+2i\lambda\frac{dw}{dz}-\mathrm{Ri}\lambda^2 w=0. \tag{2.22}$$

This equation has a solution of the form (2.15), but in this case $r_\pm$ are given by

$$r_\pm=\frac{-2i\lambda\pm[4\,\mathrm{Ri}\lambda^2(1-\sigma^2)-4\lambda^2]^{\frac{1}{2}}}{2(1-\sigma^2)}. \tag{2.23}$$

Applying condition (2.17) for a non-trivial solution, we obtain a quadratic for σ^2 which has the roots

$$\sigma^2=k^2c^2=1+\left(\frac{\mathrm{Ri}\lambda^2}{2m^2\pi^2}\right)\left[1\pm\left(1+\frac{4m^2\pi^2}{\mathrm{Ri}^2\lambda^2}\right)^{\frac{1}{2}}\right], \tag{2.24}$$

where $m=1, 2, 3, \ldots$.

The plus root of (2.24) is always of order unity and is not connected with the previous solutions. However, the minus root has zeros at $\alpha^2=m^2\pi^2$, so our assumption that $c\sim k^{-1}$, breaks down near these points, as expected. If we replace α by $m\pi-\epsilon$ in the minus root of (2.24) and assume $\epsilon\ll 1$, we find

$$c^2\cong\frac{2\epsilon}{m\pi k^2}\left(\frac{1-\mathrm{Ri}}{2-\mathrm{Ri}}\right). \tag{2.25}$$

Comparing (2.25) with (2.20) and (2.21) we see that the minus root of (2.24) is indeed the analytic continuation of the plus root of (2.19) for $\epsilon>0$, and the analytic continuation of the minus root of (2.19) for $\epsilon<0$.

Our three different solutions for c as $k\to 0$, (2.9), (2.19) and (2.24), all valid in different but overlapping ranges of λ, together give us the solution for all λ when $k=0$. There are two infinite sets of solutions. One is the plus root of (2.24), i.e.,

$$\sigma^2=1+\frac{\mathrm{Ri}\lambda^2}{2m^2\pi^2}\left[1+\left(1+\frac{4m^2\Pi^2}{\mathrm{Ri}^2\lambda^2}\right)^{\frac{1}{2}}\right], \qquad m=1, 2, 3, \cdots. \tag{2.26}$$

This set of solutions for σ^2 is always positive and therefore does not give rise to any instabilities. Collecting our results for how the other roots join, we find that the other set of solutions behaves as follows:

if $\mathrm{Ri}\geq 1$,

$$\sigma^2=\begin{cases} 0, & \text{(the } n=0 \text{ mode)} \\ 1+\alpha(n)\left[1-\left(1+\dfrac{2}{\mathrm{Ri}\alpha(n)}\right)^{\frac{1}{2}}\right], & n=1, 2, 3, \cdots \end{cases}; \tag{2.27a}$$

if $\mathrm{Ri}<1$,

$$\sigma^2=\begin{cases} 1+\alpha(n)\left[1-\left(1+\dfrac{2}{\mathrm{Ri}\alpha(n)}\right)^{\frac{1}{2}}\right], & \text{for} \quad 0\leq\lambda\leq\left[\dfrac{n\pi}{(1-\mathrm{Ri})^{\frac{1}{2}}}\right] \\ 0, \quad \text{for} \quad \left[\dfrac{n\pi}{(1-\mathrm{Ri})^{\frac{1}{2}}}\right]\leq\lambda\leq\left[\dfrac{(n+1)\pi}{(1-\mathrm{Ri})^{\frac{1}{2}}}\right] & \\ 1+\alpha(n+1)\left[1-\left(1+\dfrac{2}{\mathrm{Ri}\alpha(n+1)}\right)^{\frac{1}{2}}\right], & \text{for} \quad \left[\dfrac{(n+1)\pi}{(1-\mathrm{Ri})^{\frac{1}{2}}}\right]\leq\lambda, \quad n=0, 1, 2, \cdots \end{cases}; \tag{2.27b}$$

where

$$\alpha(n)=\frac{\mathrm{Ri}\lambda^2}{2n^2\pi^2}. \tag{2.27c}$$

When $\mathrm{Ri}<1$ these solutions for σ^2 are zero for a finite range of λ, positive for smaller values of λ, and negative for larger values of λ. Consequently these solutions correspond to stable perturbations for small enough λ and unstable perturbations for large enough λ. The solutions for σ^2 when $m=1$ and 2 in (2.26) and when

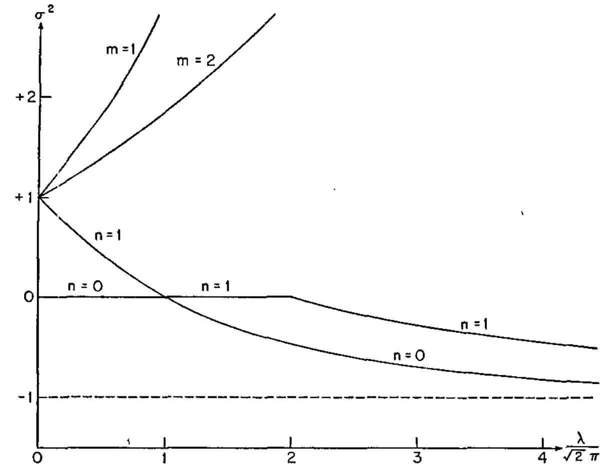

FIG. 2. σ^2 on the symmetric axis vs. λ when $\mathrm{Ri}=\frac{1}{2}$ for the four roots $m=1$ and 2 [Eq. (2.26)] and $n=0$ and 1 [Eq. (2.27)].

$n=0$ and 1 in (2.27) are shown in Fig. 2, all for $\mathrm{Ri}=\frac{1}{2}$. The corners in the curves for $n=0, 1, \ldots$ are sharp only when $k=0$. For example, if $k\neq 0$, the curves $n=0$ and $n=1$ shown in Fig. 2 would not touch at $\lambda=2^{\frac{1}{2}}\pi$. All the stable solutions correspond to non-geostrophic inertia-gravity waves. If we let $-\sigma_i$ denote the imaginary part of σ, i.e., σ_i is the growth rate of the perturbations, then σ_i is a maximum for $\lambda=\infty$ and has the value

$$\sigma_i=\left[\frac{1}{\mathrm{Ri}}-1\right]^{\frac{1}{2}} \tag{2.28}$$

for all n. However, for any finite λ the growth rate is largest for $n=0$. The maximum in the growth rate is not sharp, so it is more meaningful to speak of a band of unstable wavelengths rather than one most unstable wavelength (see Fig. 2). The smallest unstable wave number is given by

$$\lambda=\frac{\pi}{(1-\mathrm{Ri})^{\frac{1}{2}}}. \tag{2.29}$$

The dispersion relation for symmetric instabilities assuming $\sigma\sim 1$ has been calculated previously by Eliassen (1949) for a model which differs from Eady's mainly in that it includes some deep atmosphere (non-Boussinesq) effects. Eliassen also found that the most unstable wave number was $\lambda=\infty$, with a growth rate identical to that given by (2.28). However his results yield an expression for the smallest unstable wave number which is the same as (2.29) except that π is replaced by $\frac{1}{2}$. The properties of the symmetric instabilities are the same in both Eliassen's and our model. The circulations consist of a series of rolls parallel to the zonal flow. The sides of the rolls are not vertical but have a slope $dy/dz=\mathrm{Ri}$ when λ is large. The motions draw their energy from the kinetic energy of the initial zonal flow, unlike the conventional baroclinic instabilities, which draw their energy from the potential energy of the initial state.

3. Growth rates far from the baroclinic axis

The results of the preceding section show that on the symmetric axis the growth is a maximum when $\lambda=\infty$. However the maximum on the symmetric axis must also be a local maximum in the k, λ plane if such symmetric instabilities are ever to be observed in a physical situation. In order to determine the growth rates for finite values of k when λ is large, we expand the solution to the eigenvalue problem as a power series,

$$w=w_0+kw_1+k^2w_2+\cdots, \tag{3.1}$$

and

$$\sigma=\sigma_0+k\sigma_1+k^2\sigma_2+\cdots. \tag{3.2}$$

If we denote the operator in the eigenvalue equation (1.2) by L, then the above expansions imply an expansion of L in powers of k and a separation of the eigenvalue problem into an infinite set of eigenvalue problems, i.e.,

$$L=L_0+kL_1+k^2L_2+\cdots, \tag{3.3}$$

where

$$L_0w_0=0, \tag{3.4}$$

$$L_0w_1=-L_1w_0, \tag{3.5}$$

$$L_0w_2=-L_1w_1-L_2w_0,\ \text{etc.}, \tag{3.6}$$

and

$$w_j=0 \quad \text{when} \quad z=0, 1; j=0, 1, 2, \cdots. \tag{3.7}$$

Substituting (3.2) into (1.2) we find

$$L_0=(1-\sigma_0^2)\frac{d^2}{dz^2}+2i\lambda\frac{d}{dz}-\mathrm{Ri}\lambda^2, \tag{3.8}$$

$$L_1=-2\sigma_0(\sigma_1+z)\frac{d^2}{dz^2}-\frac{2}{\sigma_0}\frac{d}{dz}-\frac{2i\lambda}{\sigma_0}, \tag{3.9}$$

and

$$L_2=-(2\sigma_0\sigma_2+\sigma_1^2+2\sigma_1z+z^2)\frac{d^2}{dz^2}+\left[\frac{2(\sigma_1+z)}{\sigma_0^2}\right]\frac{d}{dz}-\mathrm{Ri}+\frac{2i\lambda(\sigma_1+z)}{\sigma_0^2}. \tag{3.10}$$

The solution for the zero-order eigenvalue is given by (2.26), and the corresponding solution for the vertical circulations, obtained by substituting (2.23) into (2.15), is

$$w_0=\exp\left(-\frac{i\lambda z}{1-\sigma_0^2}\right)\sin(n+1)\pi z. \tag{3.11}$$

Applying L_1 to w_0, we find that (3.5) leads to a solution for w_1 of the form

$$w_1=\exp\left(\frac{-i\lambda z}{1-\sigma_0^2}\right)\times\sum_{j=0}^{2}\left[A_j h^j \sin(n+1)\pi z+B_j h^j\cos(n+1)\pi z\right]. \quad (3.12)$$

If we require that (3.12) satisfy (3.5) and apply the boundary conditions (3.7), we find that a non-trivial solution for w_1 exists only if

$$\sigma_1=-\tfrac{1}{2}. \quad (3.13)$$

Since σ_1 is real, it only gives us information about the phase speed of the instabilities, and to determine how the growth rate changes for non-zero k we must go to the next highest order in k. Proceeding in the same manner used to find σ_1, we find that w_2 has the same form as w_1, except j in (3.12) ranges from zero to four. This time the requirement of a non-trivial solution leads to a very complex algebraic equation for σ_2. If we limit ourselves to large values of λ, then we find that the solution for σ_2 is

$$\sigma_2=\left[\frac{(n+1)^2\pi^2-15}{24(n+1)^4\pi^4}\right]\sigma_0\,\mathrm{Ri}^3\lambda^2+\mathrm{O}(\lambda). \quad (3.14)$$

Letting $\lambda\rightarrow\infty$ in (2.26) and collecting our solutions for σ_0, σ_1 and σ_2, we have

$$\sigma=\pm\left[1-\frac{1}{\mathrm{Ri}}\right]^{\frac{1}{2}}+\frac{1}{\lambda^2}\left\{\frac{\pm(n+1)^2\pi^2}{2\,\mathrm{Ri}^{\frac{3}{2}}(\mathrm{Ri}-1)^{\frac{1}{2}}}-\tfrac{1}{2}k\lambda^2 \pm\frac{k^2\lambda^4\,\mathrm{Ri}^{\frac{5}{2}}(\mathrm{Ri}-1)^{\frac{1}{2}}}{24(n+1)^4\pi^4}\left[(n+1)^2\pi^2-15\right] +\mathrm{O}(k^3\lambda^6)\right\}+\mathrm{O}\left(\frac{1}{\lambda^3}\right). \quad (3.15)$$

In particular, at $k=0$,

$$\frac{\partial^2|\sigma_i|}{\partial k^2}=\frac{\mathrm{Ri}^{\frac{5}{2}}(1-\mathrm{Ri})^{\frac{1}{2}}\lambda^2}{12(n+1)^4\pi^4}\left[(n+1)^2\pi^2-15\right]+\mathrm{O}(\lambda). \quad (3.16)$$

This expression is negative for $n=0$, but positive for $n\geq 1$. Therefore for large λ the growth rate at $k=0$ is a maximum for the lowest mode ($n=0$), but a minimum for the higher modes. This result raises the possibility that the largest growth rate belongs to one of the higher modes and does not correspond to a symmetric ($k=0$) perturbation. To examine this possibility an expansion in powers of k is not sufficient.

Eq. (3.15) shows that for large λ the natural expansion parameter is not k, but $\theta=k\lambda^2$. This fact suggests that we replace k by θ in our eigenvalue equation before letting $\lambda\rightarrow\infty$. The resulting equation can then be solved to determine the behavior of σ when λ is large, without necessarily restricting the size of θ. Also (3.11) shows that as $\lambda\rightarrow\infty$, w is a rapidly oscillating function of z. Before letting $\lambda\rightarrow\infty$ we must remove this behavior so that the derivatives in the eigenvalue equation are of order unity. This can be accomplished by replacing w by φ, i.e.,

$$w=\exp\left\{\frac{-i\lambda z}{1-\sigma^2}\right\}\varphi. \quad (3.17)$$

In terms of φ and θ Eq. (1.2) is

$$\left[1-\left(\sigma+\frac{\theta z}{\lambda^2}\right)^2\right]\frac{d^2\varphi}{dz^2} +\left\{\frac{2i(\lambda^2\sigma+\theta z)^2}{\lambda^3(1-\sigma^2)}-\frac{2i\lambda\sigma^2}{1-\sigma^2}-\frac{2\theta}{\lambda^2\sigma+\theta z}\right\}\frac{d\varphi}{dz} +\left\{\frac{2\lambda^2}{1-\sigma^2}-\frac{\lambda^2}{(1-\sigma^2)^2}\left[1-\frac{(\lambda^2\sigma+\theta z)^2}{\lambda^4}\right] +\frac{2i\theta\lambda}{(1-\sigma^2)(\lambda^2\sigma+\theta z)}-\mathrm{Ri}\left(\frac{\theta^2+\lambda^6}{\lambda^4}\right) -\frac{2i\theta\lambda}{\lambda^2\sigma+\theta z}\right\}\varphi=0. \quad (3.18)$$

Eq. (3.15) shows that σ can be written in the form

$$\sigma=\pm\left[1-\frac{1}{\mathrm{Ri}}\right]^{\frac{1}{2}}+\frac{\bar{\sigma}(\theta)}{\lambda^2}+\mathrm{O}\left(\frac{1}{\lambda^3}\right). \quad (3.19)$$

Substituting (3.19) into (3.18) and letting $\lambda\rightarrow\infty$, we find that (3.18) reduces to the simple form

$$\frac{d^2\varphi}{dz^2}+(\gamma+\delta z)\varphi=0, \quad (3.20)$$

where

$$\gamma=\pm 2\bar{\sigma}\mathrm{Ri}^{5/2}(\mathrm{Ri}-1)^{1/2}, \quad (3.21)$$

and

$$\delta=\pm 2\theta\mathrm{Ri}^{5/2}(\mathrm{Ri}-1)^{1/2}. \quad (3.22)$$

If we let γ_r, γ_i, φ_r, and φ_i denote respectively the real and imaginary parts of γ and φ, then from (3.19) and (3.21) we see that

$$|\sigma_i|=\left(\frac{1}{\mathrm{Ri}}-1\right)^{\frac{1}{2}}\left[1-\frac{\gamma_r}{2\,\mathrm{Ri}^2(1-\mathrm{Ri})\lambda^2}\right]+\mathrm{O}\left(\frac{1}{\lambda^3}\right). \quad (3.23)$$

Therefore, a minimum γ_r corresponds to a maximum growth rate. A useful expression for γ_r can be derived by separating (3.20) into its real and imaginary parts.

Since δ is imaginary, we find

$$\frac{d^2\varphi_r}{dz^2}+\gamma_r\varphi_r-(\gamma_i+|\delta|z)\varphi_i=0, \tag{3.24}$$

and

$$\frac{d^2\varphi_i}{dz^2}+\gamma_r\varphi_i+(\gamma_i+|\delta|z)\varphi_r=0. \tag{3.25}$$

Multiplying (3.24) by φ_r, (3.25) by φ_i, subtracting the latter from the former, integrating from $z=0$ to $z=1$, and making use of the boundary conditions $\varphi=0$ when $z=0, 1$, we obtain

$$\gamma_r=\frac{\int_0^1\left[\left(\frac{d\varphi_r}{dz}\right)^2+\left(\frac{d\varphi_i}{dz}\right)^2\right]dz}{\int_0^1[\varphi_r^2+\varphi_i^2]dz}. \tag{3.26}$$

Therefore γ_r is always positive, and for a fixed θ the growth rate will be a maximum when $\lambda=\infty$.

From (3.15) we know how γ depends on δ when δ is small. In particular, if $n\geq 1$, γ_r decreases as $|\delta|$ increases. We can find the behavior of γ_r when δ is large by applying Horn's method to find the approximate solution of (3.20) [e.g., see Jeffries (1962, p. 52)]. The approximate solution when δ is large is

$$\varphi=(\gamma+\delta z)^{-\frac{1}{4}}\left[A\exp\left\{\frac{2i}{3\delta}(\gamma+\delta z)^{\frac{3}{2}}+O(\delta^{-\frac{1}{2}})\right\}\right.$$
$$\left.+B\exp\left\{-\frac{2i}{3\delta}(\gamma+\delta z)^{\frac{3}{2}}+O(\delta^{-\frac{1}{2}})\right\}\right]. \tag{3.27}$$

Applying the boundary conditions we find the following requirement for a non-trivial solution,

$$\frac{2}{3\delta}[(\gamma+\delta)^{\frac{3}{2}}-\gamma^{\frac{3}{2}}]=(n+1)\pi+O(\delta^{-\frac{1}{2}}),$$
$$n=0, 1, 2, \cdots. \tag{3.28}$$

The solution of this equation for large δ is

$$\gamma=\frac{|\delta|}{2(3)^{\frac{1}{2}}}-\frac{\delta}{2}+(n+1)\pi(|\delta|)^{\frac{1}{2}}(3)^{\frac{1}{4}}+O(1). \tag{3.29}$$

Therefore when δ is large, γ_r increases as $|\delta|$ increases. Consequently γ_r for the higher modes ($n\geq 1$) must have a minimum (and the growth rate a maximum) for some intermediate value of $|\delta|$.

A lower bound on γ_r (and therefore an upper bound on the growth rate) can be found from (3.26). It follows from (3.20) and the boundary conditions on φ that $d^2\varphi/dz^2=0$ when $z=0, 1$, and therefore its real and imaginary parts can be expressed as Fourier sine series,

$$\left.\begin{aligned}\frac{d^2\varphi_r}{dz^2}&=\sum_{j=1}^{\infty}A_j\sin j\pi z\\ \frac{d^2\varphi_i}{dz^2}&=\sum_{j=1}^{\infty}B_j\sin j\pi z\end{aligned}\right\}. \tag{3.30}$$

Integrating these expressions and applying the boundary conditions on φ, we obtain the differentiable series

$$\left.\begin{aligned}\varphi_r&=-\sum_j\frac{A_j}{j^2\pi^2}\sin j\pi z\\ \varphi_i&=-\sum_j\frac{B_j}{j^2\pi^2}\sin j\pi z\end{aligned}\right\}. \tag{3.31}$$

Substituting (3.31) into (3.26) we find

$$\gamma_r=\pi^2\frac{\sum_{j=1}^{\infty}(A_j^2+B_j^2)/j^2}{\sum_{j=1}^{\infty}(A_j^2+B_j^2)/j^4}. \tag{3.32}$$

Consequently

$$\gamma_r\geq\pi^2. \tag{3.33}$$

The equality in (3.33) holds only if $A_j=B_j=0$ for $j>1$. Eqs. (3.15), (3.20) and (3.23) show that this is possible only if $\delta=n=0$. Therefore the largest growth rate belongs to the $n=0$ mode, and occurs for $k=0$. Consequently the maximum in the growth rate on the symmetric axis is also a local maximum in the k, λ plane. Substituting (3.29) into (3.21) and neglecting terms of order λ^{-3} and $\delta^{\frac{1}{2}}$ we find that the growth rate for a fixed λ falls to one half its maximum value at

$$k=\left[3\left(\frac{1-\mathrm{Ri}}{\mathrm{Ri}}\right)\right]^{\frac{1}{2}}. \tag{3.34}$$

Fig. 3 shows γ_r as a function of $|\delta|$ for $n=0$, 1 and 2. The values for small $|\delta|$ were calculated from (3.15), and those for large $|\delta|$ from (3.29). The dotted lines are interpolations between the two extremes.

4. Growth rates near the baroclinic axis

The results of the preceding sections show that for small k the maximum growth rates occur at $\lambda=0$ if $\mathrm{Ri}>1$, but at $\lambda=\infty$ if $\mathrm{Ri}<1$. However, the maximum growth rates on the baroclinic axis occur not for small values of k, but for values of k of order unity [see Árnason (1963)]. For these values the effect of the maximum on the symmetric axis will be felt in general for smaller values of Ri than unity. In order to determine at what values of Ri the maximum on the baroclinic axis is also a local maximum in the k, λ plane, and also

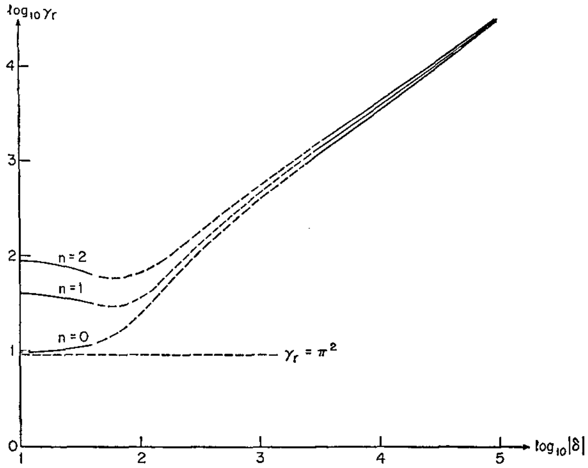

FIG. 3. γ_r for large λ vs. $|\delta|$ for the modes $n=0$, 1, and 2.

to find an analytic expression for the maximum growth rate on the baroclinic axis, we will expand the solution to the eigenvalue problem as a double power series in k and λ.

For this calculation it is convenient to introduce a new dependent variable ψ,

$$w=e^{-i\lambda(c+z)}\psi. \tag{4.1}$$

In terms of ψ Eq. (1.2) becomes

$$L\psi=0,$$

with

$$L=[1-k^2(c+z)^2]\frac{d^2}{dz^2}+2\left[i\lambda k^2(c+z)^2-\frac{1}{c+z}\right]\frac{d}{dz}+\lambda^2(1-\mathrm{Ri})-\mathrm{Ri}k^2+\lambda^2k^2(c+z)^2. \tag{4.2}$$

Since the only unstable root when λ is small is that for which $c\sim 1$, since only k^2, not k, appears in L and since only λ^2 appears when $k^2=0$, we expand as follows,

$$\psi=\psi_{00}+\lambda^2\psi_{02}+\cdots+k^2(\psi_{10}+\lambda\psi_{11}+\lambda^2\psi_{12}+\cdots)+\cdots, \tag{4.3}$$

and

$$c=c_{00}+\lambda^2c_{02}+\cdots+k^2(c_{10}+\lambda c_{11}+\lambda^2c_{12}+\cdots)+\cdots. \tag{4.4}$$

Substituting (4.3) and (4.4) into (4.2) we obtain

$$L_{00}\psi_{00}=0, \tag{4.5}$$

$$L_{00}\psi_{02}=-L_{02}\psi_{00}, \tag{4.6}$$

$$L_{00}\psi_{10}=-L_{10}\psi_{00}, \tag{4.7}$$

$$L_{00}\psi_{11}=-L_{11}\psi_{00}, \tag{4.8}$$

$$L_{00}\psi_{12}=-L_{02}\psi_{10}-L_{10}\psi_{02}-L_{12}\psi_{00},\ \text{etc.}, \tag{4.9}$$

where

$$L_{00}=\frac{d^2}{dz^2}-\left(\frac{2}{c_{00}+z}\right)\frac{d}{dz}, \tag{4.10}$$

$$L_{02}=\left[\frac{2c_{02}}{(c_{00}+z)^2}\right]\frac{d}{dz}+1-\mathrm{Ri}, \tag{4.11}$$

$$L_{10}=-(c_{00}+z)^2\frac{d^2}{dz^2}+\left[\frac{2c_{10}}{(c_{00}+z)^2}\right]\frac{d}{dz}-\mathrm{Ri}, \tag{4.12}$$

$$L_{11}=\left[2i(c_{00}+z)^2+\frac{c_{11}}{(c_{00}+z)^2}\right]\frac{d}{dz}, \tag{4.13}$$

$$L_{12}=-2c_{02}(c_{00}+z)\frac{d^2}{dz^2}+\left[\frac{2c_{12}}{(c_{00}+z)^2}-\frac{4c_{10}c_{02}}{(c_{00}+z)^3}\right]\frac{d}{dz}+(c_{00}+z)^2,\ \text{etc.} \tag{4.14}$$

Since L_{00} is an equidimensional operator, the integration of all these equations is straightforward. If we apply the boundary conditions that $\psi=0$ when $z=0$, 1 we obtain the following solutions and eigenvalues (taking the undetermined multiplicative constant to be unity),

$$\psi_{00}=(c_{00}+z)^3-c_{00}^3, \tag{4.15}$$

$$\psi_{02}=\frac{(\mathrm{Ri}-1)}{10}[(c_{00}+z)^5-c_{00}^3(c_{00}+z)^2], \tag{4.16}$$

$$\psi_{10}=\frac{6+\mathrm{Ri}}{10}[(c_{00}+z)^5-c_{00}^3(c_{00}+z)^2], \tag{4.17}$$

$$\psi_{11}=-\frac{i}{3}[(c_{00}+z)^6-c_{00}^6] \tag{4.18}$$

$$\psi_{12}=\left(\frac{\mathrm{Ri}^2+12\,\mathrm{Ri}-18}{140}\right)(c_{00}+z)^7+\left(\frac{\mathrm{Ri}^2-10\,\mathrm{Ri}+14}{20}\right)c_{00}^3(c_{00}+z)^4+\left(\frac{12\,\mathrm{Ri}^2+25\,\mathrm{Ri}-62}{350}\right)c_{00}^3(c_{00}+z)^2+\left(\frac{4\,\mathrm{Ri}^2-5\,\mathrm{Ri}+1}{675}\right)(c_{00}+z), \tag{4.19}$$

$$c_{00}=-\tfrac{1}{2}\pm\frac{i(3)^{\frac{1}{2}}}{6}, \tag{4.20}$$

$$c_{02}=\mp\frac{i(3)^{\frac{1}{2}}}{45}(\mathrm{Ri}-1), \tag{4.21}$$

$$c_{10}=\mp\frac{i(3)^{\frac{1}{2}}}{45}(\mathrm{Ri}+1), \tag{4.22}$$

$$c_{11}=0 \tag{4.23}$$

and

$$c_{12}=\pm i(3)^{\frac{1}{2}}\left(\frac{12\,\mathrm{Ri}^2+25\,\mathrm{Ri}-62}{9450}\right). \tag{4.24}$$

Using these results, we find in particular that at $\lambda=0$,

$$\sigma=-\frac{k}{2}\pm\frac{i}{2(3)^{\frac{1}{2}}}\left[k-\frac{2}{15}k^3(\mathrm{Ri}+1)\right]+O(k^5), \tag{4.25}$$

and

$$\frac{\partial^2\sigma_i}{\partial\lambda^2}=\mp\frac{2i(3)^{\frac{1}{2}}}{45}\left[k(\mathrm{Ri}-1)-k^3\left(\frac{12\,\mathrm{Ri}^2+25\,\mathrm{Ri}-62}{210}\right)\right]+O(k^5). \tag{4.26}$$

Ignoring terms of order k^5 or higher, we find from (4.25) that the growth rate on the baroclinic axis is a maximum when

$$|k|=\left(\frac{5/2}{\mathrm{Ri}+1}\right)^{\frac{1}{2}}, \tag{4.27}$$

and has a value

$$\sigma_i=\left(\frac{5/54}{\mathrm{Ri}+1}\right)^{\frac{1}{2}}. \tag{4.28}$$

Comparison of these last two expressions with the analytical results of Eady (1949) and Fjørtoft (1950) for $\mathrm{Ri}\gg 1$ and $\mathrm{Ri}=0$, and with the numerical results of Årnason (1963) for intermediate values of Ri, shows that the error in (4.27) and (4.28) is only about 1 per cent. If we evaluate (4.26) at the value of k given by (4.27), again ignoring terms of order k^5, we find that $\partial^2\sigma_i/\partial\lambda^2$ changes sign only once, not at $\mathrm{Ri}=1$, but at $\mathrm{Ri}=0.75$. Consequently, to a first approximation the maximum on the baroclinic axis is a local maximum in the k, λ plane if $\mathrm{Ri}>0.75$, but is a saddle point if $\mathrm{Ri}<0.75$. The maximum growth rate for the symmetric instabilities, Eq. (2.28), is compared with that for the conventional baroclinic instabilities, Eq. (4.28), in Fig. 4. The symmetric instabilities have a larger growth rate if $\mathrm{Ri}<0.95$.

5. Growth rates far from the symmetric axis

Eq. (1.2) can have non-trivial solutions as $k\rightarrow\infty$ only if $\sigma\sim k$. Assuming that $c\sim 1$ and that at most $\lambda\sim k$, in the limit $k\rightarrow\infty$, Eq. (1.2) reduces to

$$(c+z)^2\frac{d^2w}{dz^2}+\mathrm{Ri}\left(1+\frac{\lambda^2}{k^2}\right)w=0. \tag{5.1}$$

This equation has the solution

$$w=(c+z)^{\frac{1}{2}}[A(c+z)^p+B(c+z)^{-p}], \tag{5.2}$$

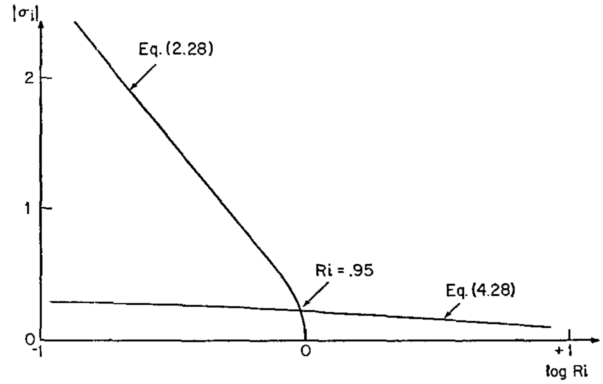

FIG. 4. Maximum growth rates for the symmetric instabilities [Eq. (2.28)] and for the conventional baroclinic instabilities [Eq. (4.28)] as a function of Ri.

where

$$p=\left[\tfrac{1}{4}-\mathrm{Ri}\left(1+\frac{\lambda^2}{k^2}\right)\right]^{\frac{1}{2}}. \tag{5.3}$$

Applying the boundary conditions on w, we find that a non-trivial solution exists if $c=0$, $c=1$, or

$$c=-\frac{1}{2}\pm\frac{i}{2}\,\mathrm{ctn}\frac{m\pi}{2p},\quad m=1,2,3,\cdots. \tag{5.4}$$

Eq. (5.4) shows that unstable solutions exist if p is real. In particular, p will be real inside the region

$$\lambda^2<k^2\left(\frac{1}{4\,\mathrm{Ri}}-1\right),\quad k\gg 1, \tag{5.5}$$

and such a region exists so long as

$$\mathrm{Ri}<\tfrac{1}{4}. \tag{5.6}$$

These instabilities are analogous to the Kelvin-Helmholtz instabilities of a flow with curvature in a non-rotating coordinate system [e.g., see Chandrasekhar, (1961, p. 494 *et seq*)]. Like the symmetric instabilities these instabilities draw their energy from the kinetic energy of the initial zonal flow, but, unlike the symmetric instabilities, they also store up potential energy.

Actually (5.4) implies that

$$c=0,\quad \text{if}\quad \frac{1}{p}=\frac{2n-1}{m},\quad m,n=1,2,3,\cdots, \tag{5.7}$$

and

$$c=\infty,\quad \text{if}\quad \frac{1}{p}=\frac{2n-2}{m},\quad m,n=1,2,3,\cdots. \tag{5.8}$$

However, when c is small the approximate equation (5.1) is not valid, and when c is large it is not possible to neglect the second derivative of the initial zonal flow. For example, if we had not assumed that the undisturbed zonal flow, $u(z)$, was linear, we would have found, when $k \gg 1$ and $\lambda = 0$, the eigenvalue equation

$$(c+u)^2 \frac{d^2w}{dz^2} + \left[\mathrm{Ri} - \frac{d^2u}{dz^2}(c+u) \right] w = 0. \qquad (5.9)$$

The last term, which does not appear in (5.1), should not be neglected when c is very large, since in any physical situation d^2u/dz^2 may be small, but in general will not be zero. Consequently we can only conclude from (5.4) that there exist unstable solutions for which c is of order unity.

However, $\sigma = kc$, and k is indefinitely large. Therefore, the growth rate of the Kelvin-Helmholtz instabilities is greatest for large k, and is unbounded. Consequently these instabilities will dominate the symmetric and conventional baroclinic instabilities whenever $\mathrm{Ri} < \frac{1}{4}$. Eq. (5.5) shows that the Kelvin-Helmholtz instability is greatest for small λ, and for these perturbations the circulation consists of rolls perpendicular to the zonal flow.

6. Conclusions

The solutions discussed in the preceding sections give the growth rates of those perturbations of a baroclinic zonal flow which have latitudinal or longitudinal scales which are either large or small. The results are summarized schematically in Fig. 5. The growth rate has three distinct local maxima corresponding to three basically different types of instabilities. One cannot rule out the possibility that there are other maxima corresponding to perturbations with both longitudinal and latitudinal scales of order unity. However, the results obtained show that of the three types of instabilities discussed, the baroclinic instabilities first found in geostrophic theory have the largest growth rates if $\mathrm{Ri} > 0.95$, the symmetric instabilities have the largest growth rates if $\frac{1}{4} < \mathrm{Ri} < 0.95$ and the Kelvin-Helmholtz instabilities have the largest growth rates if $\mathrm{Ri} < \frac{1}{4}$.

The results also indicate several weaknesses of the simple model used. The growth rates for the symmetric and Kelvin-Helmholtz instabilities are greatest for perturbations with small scales, and clearly dissipative effects should be included for an accurate determination of these growth rates. Similarly for very large horizontal wave numbers the approximation of hydrostatic equilibrium is no longer valid. Such effects will be more important for the Kelvin-Helmholtz instabilities whose growth rate is peaked at $k = \infty$ than for the symmetric instabilities which have a broad band of equally unstable wave numbers for large λ. Also, as previously pointed out, a proper model of the Kelvin-Helmholtz instabilities should take into account the variability of the shear of the initial zonal flow. In addition Phillips (1964) has pointed out that the latitudinal variations of the unperturbed zonal flow

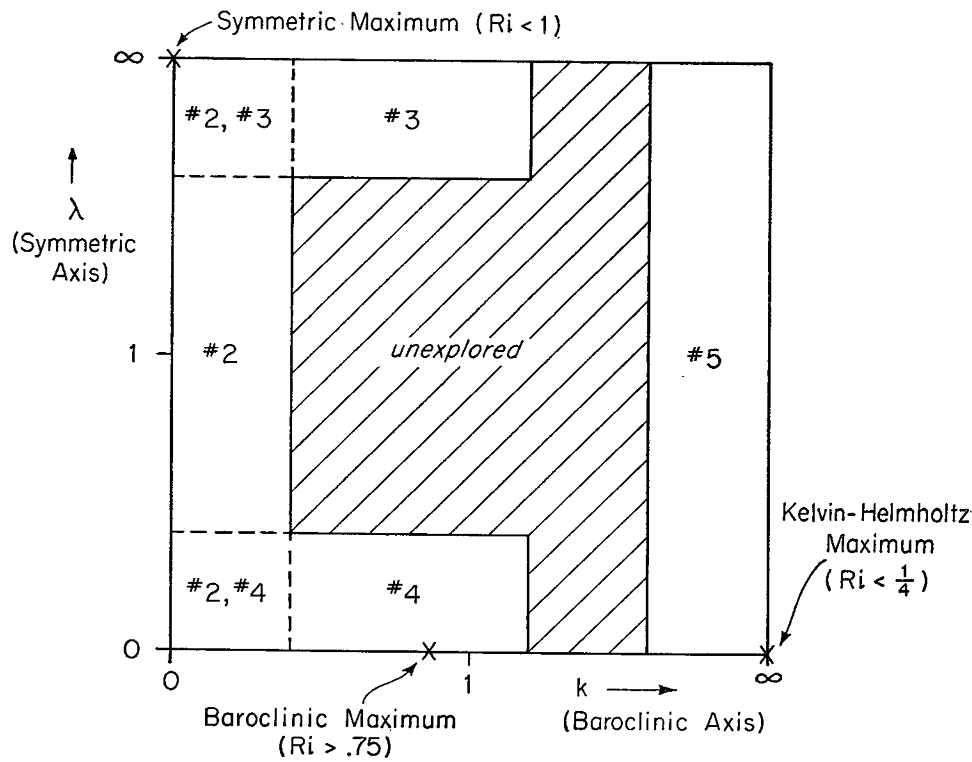

FIG. 5. Schematic diagram of the $k-\lambda$ plane, showing the various regions where the analyses of Sections 2 through 5 apply. The locations of the three local maxima in the growth rate are indicated schematically by x's, and the range of values of Ri for which these maxima exist are also given.

may have an important effect. Our results indicate that such effects could be important for the baroclinic and Kelvin-Helmholtz instabilities, since their greatest growth rates occur for perturbations with large latitudinal scales. However, latitudinal variations are less likely to be important for the symmetric instabilities since their greatest growth rates occur for perturbations with small latitudinal scales. Finally, if one wishes to apply these results to the atmospheres of the major planets, deep atmosphere (non-Boussinesq) effects should be taken into account.

In view of the possible importance of the symmetric instabilities in the dynamics of the atmospheres of the major planets, an experimental verification of our results would be of definite interest. Since the Richardson number is the critical parameter which determines whether or not the symmetric instabilities will dominate, it is essential that not only the horizontal temperature contrast, but also the vertical temperature contrast be controlled in any attempt at experimental verification. Laboratory studies of the conventional baroclinic instabilities have not controlled the vertical temperature contrast, and therefore have only dealt with large values of the Richardson number. Also, in laboratory experiments vertical walls will be present which could affect the results predicted by our unbounded model. However, so long as the separation of any vertical walls is large compared to our predicted cut-off wavelength for symmetric instabilities, such effects should be minimal.

REFERENCES

Árnason, G., 1963: The stability of nongeostrophic perturbations in a barolinic zonal flow. *Tellus*, **15**, 205–229.

Chandrasekhar, S., 1961: *Hydrodynamic and Hydromagnetic Stability*. Oxford, Clarendon Press, 654 pp.

Eady, E. T., 1949: Long waves and cyclone waves. *Tellus*, **1**, 33–52.

Eliassen, A., 1949: The quasi-static equations of motion with pressure as independent variable. *Geofys. Publ.*, **17**, (3), 43 pp.

——, and E. Kleinschmidt, 1957: *Handbuch der Physik*, **48**, Berlin, Springer-Verlag, 1–154.

Fjørtoft, R., 1950: Application of integral theorems in deriving criteria of stability for laminar flows and for the baroclinic circular vortex. *Geofys. Publ.*, **17**, (6).

Jeffries, H., 1962: *Asymptotic Approximations*. Oxford, Clarendon Press, 144 pp.

Phillips, N. A., 1964: An overlooked aspect of the baroclinic stability problem. *Tellus*, **16**, 268–270.