J. Fluid Mech. (1985), *vol.* 150, *pp.* 41–87
Printed in Great Britain

The influence of the horizontal component of Earth's angular velocity on the instability of the Ekman layer

By S. LEIBOVICH AND S. K. LELE
Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca, New York

(Received 29 March 1984 and in revised form 12 July 1984)

A systematic study is made of the effect of latitude on the linear, normal mode stability characteristics of the laminar barotropic Ekman layer. The outcome depends upon the *direction* of the geostrophic flow (in the case of flows modelling the atmospheric Ekman layer) or, alternatively, upon the direction of the applied stress (in the case of flows modelling the oceanic Ekman layer). The minimum critical Reynolds number R_c is a function of latitude. For the atmospheric Ekman layer $R_c = 30.8$ for all latitudes less than 26.2° and increases monotonically with latitude to 54.2. At a latitude of 45° N, R_c is 33.9 and arises for a geostrophic wind directed towards a compass heading 252° (clockwise from north), corresponding to rolls with axes pointing due west and having wavenumber k (with unit of length taken to be the Ekman layer depth) of 0.594. The minimum R_c for the oceanic boundary layer is 11.6 for latitudes less than 81.1°, and increases with latitude to 11.8. At 45° N latitude, the critical condition arises for a surface-current compass heading of 345.2°, roll axis of 351° and a wavenumber $k = 0.33$. The results for R_c are all symmetric about the equator, with roll axes and associated basic flow directions rotated by 180°. As the Reynolds number R increases, the effects of the perturbation Coriolis acceleration on the instability diminish, as has been previously shown, and the error caused by neglect of the horizontal component of angular velocity therefore decreases. The high Reynolds number limit is systematically explored. It is shown that the lower branch of the neutral curve is not inviscid as $R \to \infty$; rather $kR \to$ constant. The upper branch is inviscid in the limit $R \to \infty$, and corresponds to a regular or singular neutral mode depending on whether the angle ϵ between the outer geostrophic flow and the roll axis is greater or less than 15.93°. 'Inflectional' modes, thought to be relevant by some investigators, do not exist for $\epsilon < 15.93°$. Lastly, the most unstable inviscid mode corresponding to zero phase speed, a condition to which certain well-known experiments are sensitive, occurs at $\epsilon = 11.8°$ with wavenumber $k = 0.6$. This is in good agreement with published experimental data.

1. Introduction

The classical steady Ekman layer (originating with Ekman 1905) is a motion everywhere parallel to a plane; in it, deviations from a state of rigid rotation exist with Coriolis accelerations and pressure gradients balanced by viscous forces. The Coriolis accelerations that enter are those associated with Ω_N, the component of the angular velocity of the reference frame that is normal to the planes of motion. In the geophysical context, the motion is interpreted as a local one, the defining plane being that tangent to the planet's surface at the latitude λ of interest, the curvature of the planet surface being neglected. If the angular speed of rotation about the planetary axis is Ω, then $\Omega_N = \Omega \sin\lambda$; there is in addition a component of the angular velocity, $\Omega_H = \Omega \cos\lambda$, parallel to the tangent plane. The neglect of the parallel component is accepted as an approximation when the vertical thickness of the fluid layer is much smaller than the horizontal lengthscales of the motion (thereby causing vertical speeds in such flows to be much smaller than horizontal speeds). The Ekman layer solution then arises in the limit of zero Rossby number. The solution alternatively may be viewed as the exact solution of the Navier–Stokes equations in a rotating half-space, constrained to be a function only of vertical distance from a surface plane, even allowing for a horizontal component of $\boldsymbol{\Omega}$. Furthermore, in laboratory 'dishpan' experiments (Faller 1963; Faller & Kaylor 1966*a*; Tatro & Mollo-Christensen 1967; Caldwell & Van Atta 1970; Weidman 1976) designed to investigate the Ekman layer, and in particular its stability, there is no horizontal component of angular velocity.

When considering the instability of the Ekman layer, the horizontal component of $\boldsymbol{\Omega}$, Ω_H, assuming it is non-zero, cannot be neglected, although this has been done in nearly all of the previous theoretical work on the problem (Stern 1960; Faller & Kaylor 1966*a*, *b*, 1967; Lilly 1966; Barcilon 1965; Brown 1970, 1972; Kaylor & Faller 1972; Asai & Nakasuji 1971; Iooss, Nielsen & True 1978; Nielsen & True 1979; Weber 1980; Melander 1983; Spooner 1983; Spooner & Criminale 1982). We will call such treatments the 'traditional problem' of Ekman layer stability.† Vertical and horizontal perturbation speeds are comparable, and consequently the horizontal component of the angular velocity of the reference frame is activated in the stability problem, although it does not enter into the basic Ekman layer velocity field. Thus the previous analyses of the instability of the Ekman layer over a no-slip rigid surface (modelling the atmospheric Ekman layer – we refer to this model as the 'atmospheric' case) and the Ekman layer with a free surface having a prescribed surface stress (modelling the oceanic Ekman layer, and which we call the 'oceanic' case; see Spooner 1983; Faller & Kaylor 1967; Iooss *et al.* 1978) are generally incorrect (more or less depending on latitude) when applied to geophysical problems. Wippermann (1969) first pointed out the possible importance of the horizontal component of angular velocity to the stability of the atmospheric Ekman layer; further work by Etling (1971) confirmed it for both stratified and homogeneous atmospheres at a latitude of 45° N, and, for this latitude, Etling showed that the most-unstable perturbations take the form of rolls oriented due west. Additional investigations by Etling & Wippermann (1975) and Wipperman, Etling & Kirstein (1978) incorporate the findings of Etling (1971) – by including the horizontal angular velocity, restricting attention to the particular case of 45° N and rolls due west – in studies devoted to the effects of stratification and turbulence on stability of the atmospheric Ekman layer. This demonstration by Wippermann and Etling of the importance of the horizontal component of angular velocity is not properly recognized in the literature. Their work was brought to our attention by Professor A. J. Faller, to whom we are grateful, after a draft of this paper was submitted for publication.

The geophysical problem depends mainly on three ingredients: turbulent transport of momentum and buoyancy, buoyant force due to density stratification, and Coriolis force. In addition to the work of Wippermann & Etling, stratification effects have been considered, for the atmospheric case under the traditional approximation, by Faller & Kaylor (1972*a*, *b*), Brown (1972) and Asai & Nakasuji (1973), assuming

---

† Professor J. W. Miles has pointed out to us that Eckart (1960, p. 96) referred to the neglect of Ω_H as the 'traditional approximation'.

quasi-laminar (constant eddy viscosity) flow. Spooner (1983) includes stratification through a two-layer model of the oceanic problem, also under the traditional approximation. The complete problem, including proper treatment of the three dominant effects, is not tractable. We confine attention here to barotropic laminar or quasi-laminar flows, and aim at a complete and systematic exploration of the effects of latitude on stability of both the 'atmospheric' and 'oceanic' cases engendered by both components of angular velocity. Other effects, although acknowledged to be important, are ignored henceforth.

The Coriolis acceleration enters into the Ekman layer stability problem in two ways, which we may call indirect and direct. First, the Coriolis acceleration associated with Ω_N controls the basic (unperturbed and horizontal) Ekman-layer flow. Thus perturbations about this state are influenced by the Coriolis acceleration indirectly, through the form of the basic Ekman flow profiles. Secondly, disturbances to the basic state are dynamically influenced directly through the perturbation Coriolis acceleration, containing contributions from both the vertical and horizontal components of the angular velocity. If one completely ignores the perturbation Coriolis accelerations, the problem for normal modes reduces to an Orr–Sommerfeld equation. Lilly (1966) examined the stability of the atmospheric Ekman layer model based upon this approximation as well as the traditional problem. He found a critical value of R, R_c, of about 93 when perturbation Coriolis accelerations were neglected, and about 55 for the traditional problem. The latter value has been found more accurately by Iooss *et al.* (1978) and Melander (1983) to be 54.2, and this is the correct value of R_c at $\lambda = 90°$. In the present work we find R_c to be reduced further for the atmospheric Ekman layer, to 30.8 for latitude of 26.2° and less, when the full perturbation Coriolis acceleration is included (we refer to this as the 'full' problem). As $\lambda \rightarrow 0$ the Ekman layer is degenerate, since its limiting depth is infinite, furthermore, the (singular) limiting stability problem has no solution.

Dudis & Davis (1971) predict an energy-stability limit $R_E \approx 18.3$ for the Ekman layer with rigid surface. Since the energy-stability analysis does not depend on the direct effects of Coriolis acceleration, it still stands. The present results therefore reduce the gap between the global and linear stability limits for the atmospheric Ekman layer.

The traditional Ekman layer stability problem depends upon ϵ, the angle between the geostrophic wind and the roll axis (alternatively, the constant phase surfaces of the normal-mode perturbation); R, the Reynolds number based upon the geostrophic speed and the Ekman layer depth $D = (\nu/\Omega_N)^{\frac{1}{2}}$; and the wavenumber k of the normal mode. To this list one must add $\mu = -\cot\lambda\,\cos\beta$, where β is the counterclockwise angle made by the roll axis measured from due east, when dealing with the full problem. The full problem and the traditional problem both require the solution of similar sixth-order differential systems in which Coriolis terms are $O(R^{-1})$ compared with inertial terms. As $R \rightarrow \infty$ therefore, both full and traditional problems are identical. For fixed values of ϵ (which enters through the basic state), the neutral-stability curve has two branches, as is familiar in parallel-flow stability problems.

In the atmospheric case, slightly supercritical Reynolds numbers correspond to travelling waves, sometimes called the 'type II' mode of instability in experimental work on the problem (for the traditional problem see Faller & Kaylor 1966*a*, *b*; Lilly 1966; Greenspan 1968). The phase fronts are oriented at angles of about 20° to the right of the geostrophic flow direction. A second mode of instability, 'type I', appears at a Reynolds number on the order of 100, with very much smaller phase speeds and larger wavenumbers. Type I disturbances have phase fronts (roll axes) aligned more closely to the geostrophic-flow direction and to the left of it (in the northern hemisphere). The two types of disturbances coexist for a range of Reynolds number in the sense that there are two local maxima of growth rate in the (k, ϵ)-plane, with the type II ultimately disappearing at Reynolds numbers in the range 150–200. This process is nicely illustrated by Melander (1983, figure 4). The type II mode has been called the 'parallel instability' and the type I mode 'inviscid' or 'inflectional' in some theoretical work. The qualitative behaviour found here for the full problem is similar, except that the critical Reynolds number is smaller, the band of unstable modes at slightly supercritical conditions is much wider, and, for sufficiently large μ, only one maximum of growth rate exists in the (k, ϵ)-plane for moderate Reynolds numbers.

Numerical treatments of the traditional stability problem for the oceanic case (Iooss *et al.* 1979; Spooner 1983) reveal the development of two instability modes for a range of supercritical Reynolds numbers. In contrast with the traditional atmospheric problem, these two unstable modes can be traced in wavenumber, as there are bands of wavenumbers for which these modes are distinct eigenvalues for the *same* values of k, R and ϵ. The qualitative behaviour of the full problem is again similar. In this case, the critical Reynolds number is only slightly lower for the full problem, but growth rates of marginally unstable modes are much larger, and the band of excited wavelengths is much wider, than for the traditional problem.

Barcilon (1967) has considered the asymptotic limit $kR \rightarrow \infty$, considering only neutral modes. His analysis is based upon two crucial approximations that he realized were questionable (retention of only one term of a Heisenberg expansion in k^2, and the application of the boundary condition for $z \rightarrow \infty$ by $z = \pi$); unfortunately both approximations are seriously inadequate. The limit $kR \rightarrow \infty$ is governed by the Rayleigh equation, where the velocity component that enters is that parallel to the perturbation wavenumber vector. This component is not monotonic, and has an infinite number of inflection points. The numerical evidence from our solutions of the Rayleigh equation, confirmed by solutions of the viscous problem that we have obtained for R up to 32000, suggests that there are a number of unstable modes in the limit $kR \rightarrow \infty$ and we conjecture that there are, in fact, an infinite number of such modes for each angle ϵ exhibiting instabilities. For $\epsilon \geqslant 15.93°$ a regular neutral mode is associated with the inflection point closest to the surface, but there are additional unstable modes contiguous to singular neutral modes. For $\epsilon < 15.93°$ no regular neutral modes exist; all neutral modes are singular. Since there are many (perhaps infinitely many) unstable modes, each having a (singular) neutral mode, the location of the asymptote of the upper branch of the neutral curve as $R \rightarrow \infty$ is an open question. In addition, we compute the unstable modes and the linearly most unstable mode for various ϵ. We find that the most unstable mode having zero phase speed corresponds to roll angle $\epsilon = 11.8°$, and wavenumber $k = 0.6$ corresponding to a wavelength of $L = 10.5D$, where D is the Ekman-layer depth. This is in better agreement with experimental data than the regular neutral modes to which this, Faller's type I disturbance, has been compared (Faller 1963). Thus Faller finds experimentally that $L = 10.9D$, and $\epsilon = 14.5° \pm 2°$, while we show that the regular neutral mode having zero phase speed has wavelength $L = 4.55D$.

The asymptotic behaviour, as $R \rightarrow \infty$, of the lower branch of the neutral curve displays an uncommon feature for a wall-bounded shear flow. We find that as $R \rightarrow \infty$, $kR \rightarrow$ constant (depending on ϵ); this limiting case as $R \rightarrow \infty$ is an infinitely long disturbance dominated by viscosity. This asymptotic behaviour on the lower branch has not been recognized before in the context of Ekman layer stability, but has been found in the (mathematically) related problem of the ‘buoyancy boundary layer’ (BBL) by Gill & Davey (1969).

It is of interest and importance to us that the stability problem for the BBL for unit Prandtl number coincides with the traditional Ekman layer problem for $\epsilon = 0$, as was noted by Gill & Davey in their treatment of the BBL. Their analysis, particularly in the asymptotic limit $R \to \infty$, provides us with accurate solutions which we have used to check the numerical method we employ in the inviscid limit. Gill & Davey use a shooting method, while we use a combination of shooting and matrix (Galerkin) methods. This allows us to identify additional modes of instability missed by Gill and Davey.

Our problem and methods of solution are described in §2. Our results for finite Reynolds numbers are described in §3, for both the atmospheric and oceanic models. Solutions for the limit $R \to \infty$ are reported in §4. Section 5 covers an analysis of the energetics of the instability. Concluding remarks are placed in §6, and a discussion of numerical issues is given in the Appendix, including the identification of failures of matrix methods when applied to (nearly) singular problems.

2. Problem formulation and methods of solution

We adopt a right-handed Cartesian coordinate system (X, Y, z) with plane $z = 0$ tangent to the earth at latitude λ and rotating with it. Lengths are assumed to be dimensionless, with unit of length taken to be the Ekman layer depth $D = (\nu/\Omega\,|\sin\lambda|\,)^{\frac{1}{2}}$, where Ω is the rotational speed of the earth about its axis. For simplicity, we will assume motion in the northern hemisphere only (so $\lambda > 0$). Let X increase to the east and Y to the north, then z is directed vertically and we let $\boldsymbol{e}_X$, $\boldsymbol{e}_Y$, $\boldsymbol{e}_z$ be the corresponding unit vectors. The angular-velocity vector of the rotating frame is

$$\boldsymbol{\Omega} = \Omega(\cos\lambda\,\boldsymbol{e}_Y + \sin\lambda\,\boldsymbol{e}_z).$$

Choosing a velocity U_0 (to be specified subsequently) as unit of velocity, D/U_0 as unit of time, and neglecting the curvature of the earth, the dimensionless equations of motion may be written (see e.g. Holton 1979) as

$$\left.\begin{aligned} \boldsymbol{u}_t + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + R^{-1}[2(\hat{w}\cot\lambda - \hat{v})\,\boldsymbol{e}_X + 2\hat{u}\boldsymbol{e}_Y - 2\hat{u}\cot\lambda\,\boldsymbol{e}_Z - \nabla^2\boldsymbol{u}] + \nabla p = 0, \\ \nabla\cdot\boldsymbol{u} = 0, \end{aligned}\right\} \tag{1}$$

where p is a modified pressure, including the gravitational potential and the centrifugal acceleration of the reference frame, and made dimensionless by reference to ρU_0^2. In (1) the Reynolds number R is based on U_0 and D, and the velocity vector in the (X, Y, z) frame is $(\hat{u}, \hat{v}, \hat{w})$.

2.1. *Basic Ekman flow: atmospheric model*

Here the motion takes place in the half-space $z \geqslant 0$, and we assume that $\boldsymbol{u} = 0$ on a rigid plane $z = 0$. If the velocity field is independent of t, X and Y then continuity requires $\hat{w} = 0$. We assume that the flow is geostrophic as $z \to \infty$, with velocity vector oriented at an angle α (measured in the usual counterclockwise sense) to the X-axis, or

$$\boldsymbol{u} \to (\cos\alpha\,\boldsymbol{e}_X + \sin\alpha\,\boldsymbol{e}_Y)$$

as $z \to \infty$; thus the velocity scale U_0 chosen to render the problem dimensionless is the geostrophic speed. The corresponding (modified, dimensionless) pressure field balancing the Coriolis acceleration is

$$p = \frac{2}{R}(\sin\alpha\ X - \cos\alpha\ Y + \cos\alpha\ \cot\lambda\, z) + \text{constant}.$$

The Ekman layer then has the velocity vector

$$\boldsymbol{u} = \hat{U}(z)\,\boldsymbol{e}_X + \hat{V}(z)\,\boldsymbol{e}_Y, \tag{2}$$

where

$$\left.\begin{aligned} \hat{U}(z) &= \cos\alpha - \mathrm{e}^{-z}\cos(z-\alpha),\\ \hat{V}(z) &= \sin\alpha + \mathrm{e}^{-z}\sin(z-\alpha). \end{aligned}\right\} \tag{3}$$

Since, by hypothesis, there is no vertical motion, the horizontal angular velocity appears only in the vertical force balance, thereby modifying the pressure, which is given by

$$p = \frac{2}{R}[\sin\alpha\ X - \cos\alpha\ Y + \cot\lambda\,\{\cos\alpha\ z + \tfrac{1}{2}\mathrm{e}^{-z}(\cos(z-\alpha) - \sin(z-\alpha)\}] + \text{constant}.$$

2.2. *Basic Ekman flow: oceanic case*

The fluid in this case is confined to the half-space $z \leqslant 0$, and a stress is applied to the plane $z = 0$ at an angle $\alpha + \frac{1}{4}\pi$ to the X-axis. As $z \to -\infty$ it is assumed that $\boldsymbol{u} \to 0$.

For this case the Ekman layer has velocity vector (2) with components

$$\left.\begin{aligned} \hat{U}(z) &= \mathrm{e}^{z}\cos(z+\alpha),\\ \hat{V}(z) &= \mathrm{e}^{z}\sin(z+\alpha). \end{aligned}\right\} \tag{4}$$

If the applied stress is written as ρu_*^2, where u_* is the friction velocity, then the velocity scale U_0 implied by (4) and used to make the problem dimensionless is

$$U_0 = u_*^2/(2\nu\Omega|\sin\lambda|)^{\frac{1}{2}}.$$

Associated with (4) is the (modified) pressure field

$$p = \frac{\cot\lambda}{R}\mathrm{e}^{z}[\cos(z+\alpha) + \sin(z+\alpha)] + \text{constant}.$$

2.3. *Perturbation equations: the eigenvalue problem*

We intend to attack the stability problem by a normal mode analysis, with perturbation quantities having X-, Y-dependence in the form $\exp \mathrm{i}(\kappa_1 X + \kappa_2 Y)$. As observed by Lilly (1966), the resulting eigenvalue problem may be simplified by rotation of coordinates. We therefore first rotate about the z-axis through an angle β to a new (x, y, z)-coordinate system, as shown in figure 1, and assume that all perturbations are independent of x. Normal modes will then be in the form $\exp iky$, where $\kappa_1 = -k\sin\beta$, $\kappa_2 = k\cos\beta$ and, since we allow β to vary, we may take $k > 0$ without loss of generality. Velocity components with respect to the rotated coordinates are written without the circumflex. In this reference frame, (1), linearized about the relevant Ekman layer flow, are

$$\left.\begin{aligned} \boldsymbol{v}_t + V\boldsymbol{v}_y + wU' + \boldsymbol{\nabla}p = R^{-1}\{\nabla^2\boldsymbol{v} - 2(w\cot\lambda\,\cos\beta - v)\,\boldsymbol{e}_x + 2(w\cot\lambda\,\sin\beta - u)\,\boldsymbol{e}_y&\\ + 2(u\cot\lambda\,\cos\beta - v\cot\lambda\,\sin\beta)\,\boldsymbol{e}_z\},&\\ \boldsymbol{\nabla}\cdot\boldsymbol{v} = 0,& \end{aligned}\right\} \tag{5}$$

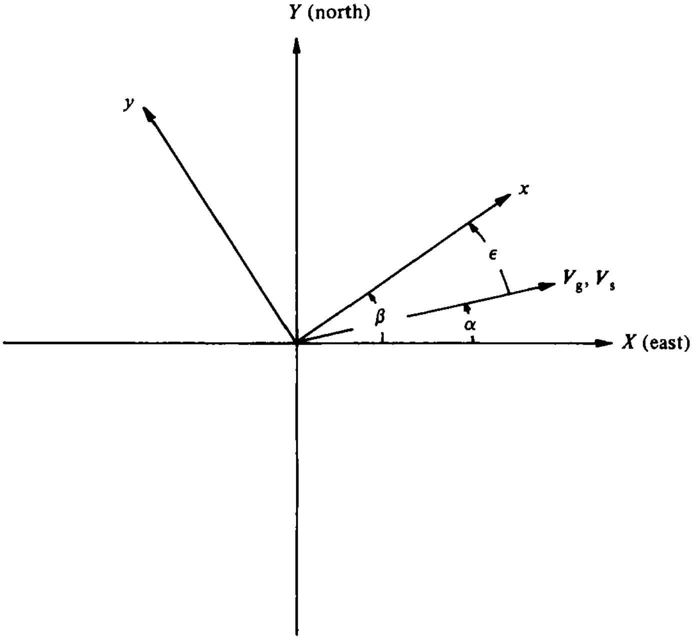

FIGURE 1. The coordinate system used in the present study. X-axis points due east and x-axis points in the direction of the rolls (wave crests). V_g is the direction of the geostrophic flow for the atmospheric problem. V_s is the direction of the surface current for the oceanic problem.

where $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ are the unit vectors in the (x, y, z)-system, $\boldsymbol{v}$ is the perturbation velocity vector $\boldsymbol{v} = u\boldsymbol{e}_x + v\boldsymbol{e}_y + w\boldsymbol{e}_z$, and $\boldsymbol{U} = U\boldsymbol{e}_x + V\boldsymbol{e}_y$ is the basic Ekman-layer velocity vector. We have assumed in (5) that $\boldsymbol{v}$ and p are independent of x, and a prime represents differentiation with respect to z.

In the rotated coordinate system the atmospheric. Ekman layer velocity vector $\boldsymbol{U}$ has components

$$\left.\begin{aligned} U &= \cos\epsilon - \mathrm{e}^{-z}\cos(z+\epsilon), \\ V &= -\sin\epsilon + \mathrm{e}^{-z}\sin(z+\epsilon), \end{aligned}\right\} \tag{6}$$

where
$$\epsilon \equiv \beta - \alpha \tag{7}$$

is the angle of the rolls measured from the direction of the geostrophic flow. Similarly, the model oceanic Ekman layer has velocity components

$$\left.\begin{aligned} U &= \mathrm{e}^{z}\cos(z-\epsilon), \\ V &= \mathrm{e}^{z}\sin(z-\epsilon), \end{aligned}\right\} \tag{8}$$

where ϵ is the angle of the rolls measured from the direction of the surface velocity.

Under the assumptions we have made, we may introduce a perturbation stream function ψ such that

$$w = \frac{\partial\psi}{\partial y}, \quad v = -\frac{\partial\psi}{\partial z}, \tag{9}$$

and assume normal modes, with

$$\left.\begin{aligned} \psi &= \phi(z)\exp[\mathrm{i}k(y-ct)], \\ u &= \chi(z)\exp[\mathrm{i}k(y-ct)]. \end{aligned}\right\} \tag{10}$$

Introducing (9) and (10) into (5) and eliminating pressure leads to the equations

$$(V-c)(\phi''-k^2\phi)-V''\phi = (\mathrm{i}kR)^{-1}[\phi^{\mathrm{iv}}-2k^2\phi''+k^4\phi+2(-\mathrm{i}k\mu\chi+\chi')], \tag{11a}$$

$$(V-c)\chi+U'\phi = (\mathrm{i}kR)^{-1}[\chi''-k^2\chi-2(-\mathrm{i}k\mu\phi+\phi')], \tag{11b}$$

where
$$\mu \equiv -\cot\lambda\,\cos\beta. \tag{12}$$

Boundary conditions on perturbations for the atmospheric Ekman layer are no slip,

$$\phi=\phi'=\chi=0 \quad \text{at } z=0, \tag{13a}$$

and vanishing normal velocity and shear stress as $z\rightarrow\infty$,

$$(\phi,\phi'',\chi')\rightarrow 0 \quad \text{as } z\rightarrow\infty. \tag{13b}$$

For the oceanic case Spooner (1983) has shown that deflection of the free surface may be ignored. Thus conditions of zero normal velocity and vanishing shear stress are assumed at $z=0$:

$$\phi(0)=\phi''(0)=\chi'(0)=0 \tag{14a}$$

and
$$(\phi,\phi',\chi')\rightarrow 0 \quad \text{as } z\rightarrow-\infty. \tag{14b}$$

If χ did not appear, (11a) would be the Orr–Sommerfeld equation, as if the flow were rectilinear with profile $V(z)$; this connection has been explored by Barcilon (1965) and Lilly (1966).

The problems for (11) subject to either (13) or (14) are regarded as eigenvalue problems with eigenvalue

$$c(k;\epsilon;R;\mu). \tag{15}$$

On writing
$$c=c_{\mathrm{r}}+\mathrm{i}c_{\mathrm{i}},$$

instability arises if $c_{\mathrm{i}}>0$ for any physically accessible point in the (k,ϵ,R,μ)-space. For λ fixed, variation of μ corresponds to variation of β, and the joint variation of k and β is equivalent to the variation of the wavenumber vector (magnitude and direction in the horizontal plane) of the perturbation.

Calculations at finite values of kR are carried out with a matrix method employing Galerkin expansions in Chebyshev polynomials. Numerical details, including assessments of accuracy of computed eigenvalues, are given in the Appendix. Since we use a matrix method, we are able to compute an approximation to a sizeable number of eigenvalues, and to determine in this way if more than one unstable mode exists.

2.4. *The inviscid limit*

As $R\rightarrow\infty$ for fixed k and μ, the equations (11) approach the Rayleigh equation

$$(V-c)(\phi''-k^2\phi)-V''\phi=0. \tag{16}$$

In this case, the viscous boundary conditions (either no-slip or no-stress) at $z=0$ must be dropped. Further, since $V''\rightarrow 0$ exponentially fast for both atmospheric and oceanic cases, $\phi\sim\exp(-k|z|)$ as $|z|\rightarrow\infty$, and so both atmospheric and oceanic cases satisfy the boundary conditions

$$\phi(0)=0, \quad \phi\rightarrow 0 \quad \text{as } |z|\rightarrow\infty, \tag{17}$$

and differ only in the form of V. By a simple transformation, both problems can be reduced to that of the atmospheric case. Thus temporarily writing V_{oc} and V_{atm} for (7) and (6) respectively, we have

$$V_{\mathrm{oc}} = -\mathrm{e}^{-\zeta}\sin(\zeta+\epsilon), \tag{18}$$

where $\zeta = -z$. Setting

$$\epsilon = \hat{\epsilon}+\pi, \quad c = \sin\hat{\epsilon}+\hat{c}, \tag{19\textit{a}, \textit{b}}$$

the oceanic problem for c given ϵ is reduced to the atmospheric problem for $\hat{c}$ given $\hat{\epsilon}$. Now, however, the phase speed in the oceanic case is

$$c_{\mathrm{r}} = -\sin\epsilon+\hat{c}_{\mathrm{r}}. \tag{20}$$

Thus we need solve only the atmospheric case for $\hat{c}$, the oceanic case having

$$c_{\mathrm{oc}}(k,\epsilon) = -\sin\epsilon+\hat{c}_{\mathrm{r}}(k,\epsilon-\pi)+\mathrm{i}\hat{c}_{\mathrm{i}}(k,\epsilon-\pi). \tag{21}$$

Another transformation between the oceanic and atmospheric problems is possible, and it leads to a fundamental symmetry in the eigenvalues of the Rayleigh equation for these problems. This second transformation also starts from (18), but replaces (19) simply by

$$c = -(\sin\epsilon+\hat{c}). \tag{22}$$

This also reduces the oceanic problem for c given ϵ and k to the atmospheric problem for $\hat{c}$ given ϵ and k. But now (21) and (22) must agree, which shows that

$$\hat{c}(k,\epsilon) = -\hat{c}(k,\epsilon-\pi). \tag{23}$$

Since eigenvalues of the Rayleigh equation arise as complex-conjugate pairs, the rule (23) shows that, to an unstable mode with given ϵ and k, there corresponds an unstable mode with the same growth rate and wavenumber, but with the negative of the phase speed, at $\epsilon-\pi$. Thus

$$\left.\begin{aligned}\hat{c}_{\mathrm{r}}(k,\epsilon) &= -\hat{c}_{\mathrm{r}}(k,\epsilon-\pi),\\ \hat{c}_{\mathrm{i}}(k,\epsilon) &= \hat{c}_{\mathrm{i}}(k,\epsilon-\pi),\end{aligned}\right\} \tag{24}$$

and the exploration of a π-interval of ϵ establishes results for all possible ϵ. We have confirmed (24) by computing c for $\epsilon = 20°$ and for $\epsilon = 200°$ at a given k.

The Rayleigh equation, as is well known, is singular for neutral modes which are the limit of unstable modes. Matrix methods are satisfactory when dealing with problems such as these, provided the growth rate is large enough to assure that $|V-c|$ is not small; a brief discussion can be found in Leibovich & Stewartson (1983). When $|V-c|$ is small in $0 \leqslant z < \infty$ matrix methods fail badly. The errors not only are uncontrolled but there need be no obvious problem with them; that is, an apparent convergence of solutions as the size of the approximation (number of Galerkin basis functions) is increased can be illusory. Growth of the error is detectable, however, by the difference between the eigenvalues computed from the direct and the adjoint problem. We discuss the numerical difficulty encountered with our Galerkin solutions for the Rayleigh equation in the Appendix.

To avoid this problem, we switch to a shooting method for most of our calculations of solutions of the Rayleigh equation, numerically deforming the contour of integration in the (complex) z-plane to avoid the singularity existing at $V = c$. Our procedure is similar to that of Gill & Davey (1969) and of Leibovich & Stewartson (1983); the path of integration is deformed to pass below the real axis if $V'(z_{\mathrm{c}}) > 0$, where z_{c} is a real value of z at which $V(z_{\mathrm{c}}) = c_{\mathrm{r}}$, and above the real axis if $V'(z_{\mathrm{c}}) < 0$. Details of this procedure and of our other numerical schemes are given in the Appendix.

3. Onset of instability

This section is concerned with the determination of conditions for the onset of instability and the characteristics of instabilities affected by viscosity. It is divided into two main parts: the first part concerns the atmospheric Ekman layer, and the second part the oceanic Ekman layer. Critical values for linear instability are given for each case as a function of latitude.

Detailed instability results are given for $\lambda = 45°$, this being thought sufficient to establish the effects of latitude. Contour maps of growth rate as a function of wavenumber k and the angle ϵ of the rolls from the geostrophic or surface velocity are presented at certain sequences of R and β. Cross-plots are also presented, giving the maximum growth rate versus Reynolds number, thus identifying the linearly most-unstable modes. These are also given as a function of the geostrophic or surface-velocity direction.

All results given in this section are found by a Galerkin numerical procedure in which each dependent variable ϕ and χ is expanded in a series of N basis functions, each being a combination of Chebyshev polynomials chosen to satisfy boundary conditions. Most of the results are found for $N = 25$, giving rise to an algebraic eigenvalue problem with 50×50 matrices, and thereby yielding 50 eigenvalues. On some calculations, as many as 50 terms were used, yielding a 100×100 system. The adjoint eigenvalue problem was also calculated in a few cases. The exact eigenvalues of the direct and adjoint problems are the same, and differences in their numerical approximations are one measure of the error in the results; this comparison also serves to identify spurious eigenvalues.

3.1. *The atmospheric model*

3.1.1. *The atmospheric model: onset*

Critical conditions for onset of instability are computed in the following way. With existing data for $\mu = 0$ in hand, we compute c for a sequence of other values of μ. Holding μ and R fixed, c is computed for the range of ϵ and k encompassing the unstable modes. We decrease R at fixed β until the maximum value of c_i (over k and ϵ) vanishes. This establishes R_c, the critical Reynolds number for each μ, and the corresponding critical values of k_c and ϵ_c. Repeating the procedures for other values of μ establishes $R_c(\mu)$, $k_c(\mu)$ and $\epsilon_c(\mu)$.

The results for R_c, k_c, ϵ_c and the corresponding phase speeds c_r are given as functions of μ in table 1; the function $R_c(\mu)$ is also plotted in figure 2. The minimum value of $R_c = R_c^*$ over all μ occurs for $\mu = \mu^* = 2.03$. Since $\mu = -\cot\lambda \cos\beta$, and $|\cot\lambda| > \mu^*$ for $|\lambda| < \lambda^* = 26.2°$, we can infer that $R_c = R_c^* = 30.8$ for all latitudes $|\lambda| < \lambda^*$, and is attained at roll angles $\beta_c = \cos^{-1}(-\mu^* \tan\lambda) = \cos^{-1}(-2.03 \tan\lambda)$. Furthermore, for latitudes $|\lambda| > \lambda^*$, critical conditions occur at the largest accessible value of μ; this implies that the most dangerous disturbances correspond to rolls with $|\cos\beta| = 1$. Thus, in the Northern Hemisphere for $\lambda = \lambda^*$, $\beta_c = 180°$, and in the Southern Hemisphere, for $\lambda < -\lambda^*$, $\beta_c = 0$. In either hemisphere the geostrophic wind producing critical conditions for onset has direction $\alpha_c = \beta_c - \epsilon_c$, where the angle α_c is measured counterclockwise from the East.

3.1.2. *The atmospheric model: supercritical instability characteristics*

The results in this subsection are for values of $\mu \leqslant 1$. They are conveniently interpreted as obtaining for $\lambda = 45°$ N ($\cot\lambda = 1$), allowing β, the roll angle, to vary over all possible angles. We adopt this point of view in the exposition, but remind

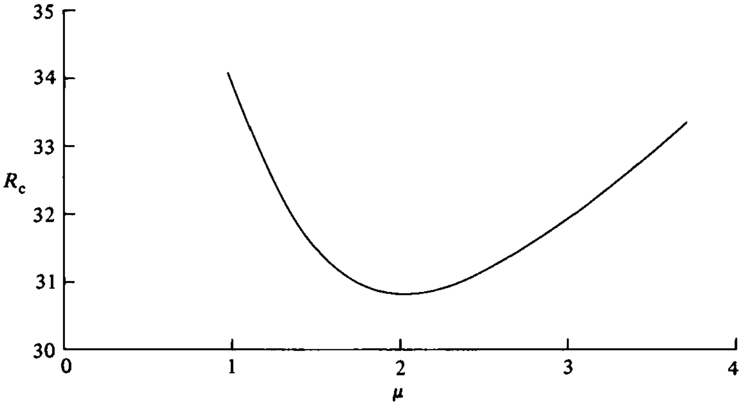

FIGURE 2. Critical Reynolds number R_c as a function of μ for the atmospheric problem.

$\mu = \cot\lambda \cos\beta$	λ	R_c	k_c	ϵ_c	c_r
0	90°	54.16	0.316	−23.3°	0.616
1	45°	33.9	0.59	−18°	0.506
2.033	15°	30.8	0.89	−9°	0.375
3.732	15°	33.3	1.20	−2°	0.290
2.033	26.2°	30.8	0.89	−9°	0.375

TABLE 1. Onset conditions: atmospheric case

the reader that the results are applicable either in part, if $|\lambda| > 45°$, or in their entirety, if $|\lambda| < 45°$, provided the range of angles β is restricted such that $|\cos\beta| < |\tan\lambda|$.

In our computations for moderate values of R (up to about 15 times critical), only one unstable mode was found for each parameter set (ϵ, k, R, μ), for the atmospheric model. This is in accord with Lilly's (1966) findings, and those of subsequent workers on the traditional problem. As we shall see in §4, however, multiple modes of instability arise for $kR \gg 1$, even for the traditional problem.

Figure 3 is a sequence of contour plots of growth rate as a function of (ϵ, k) for a sequence of values of R, holding $\mu = 0$ fixed. This is the traditional problem, and plots of this kind have been presented by Lilly (1966) and, following him, by others. We present them to facilitate comparison with the full problem. Recall that onset of instability occurs at $R = 54.2$.

In these figures, solid contours correspond to positive growth rates, broken lines to negative growth rates. At $R = 100$, the first plot in the panel, the peak amplification rate kc_i (which we will label as σ henceforth) is 0.0051 at $k = 0.27$, $\epsilon = -13°$, with phase speed $c_r = 0.414$. This is the continuation of the instability mode responsible for onset of instability at $R = 54.2$, sometimes called the 'parallel-instability' mode, or the type II mode seen in dishpan experiments.

The second panel in figure 3 is at $R = 130$, a value slightly above the onset of the 'type I' instability, sometimes called the inflectional mode. The entrance of this mode is seen by the second peak at $k \approx 0.55$, $\epsilon \approx 7°$, $c_r \approx 0.09$. The onset of this mode occurs at $R = 112.8$, with small phase speed, and it can be traced to infinite Reynolds number. The type II mode coexists with it for a limited range of R, as the subsequent panels of figure 3 show. Note that, while the modes may be simultaneously present,

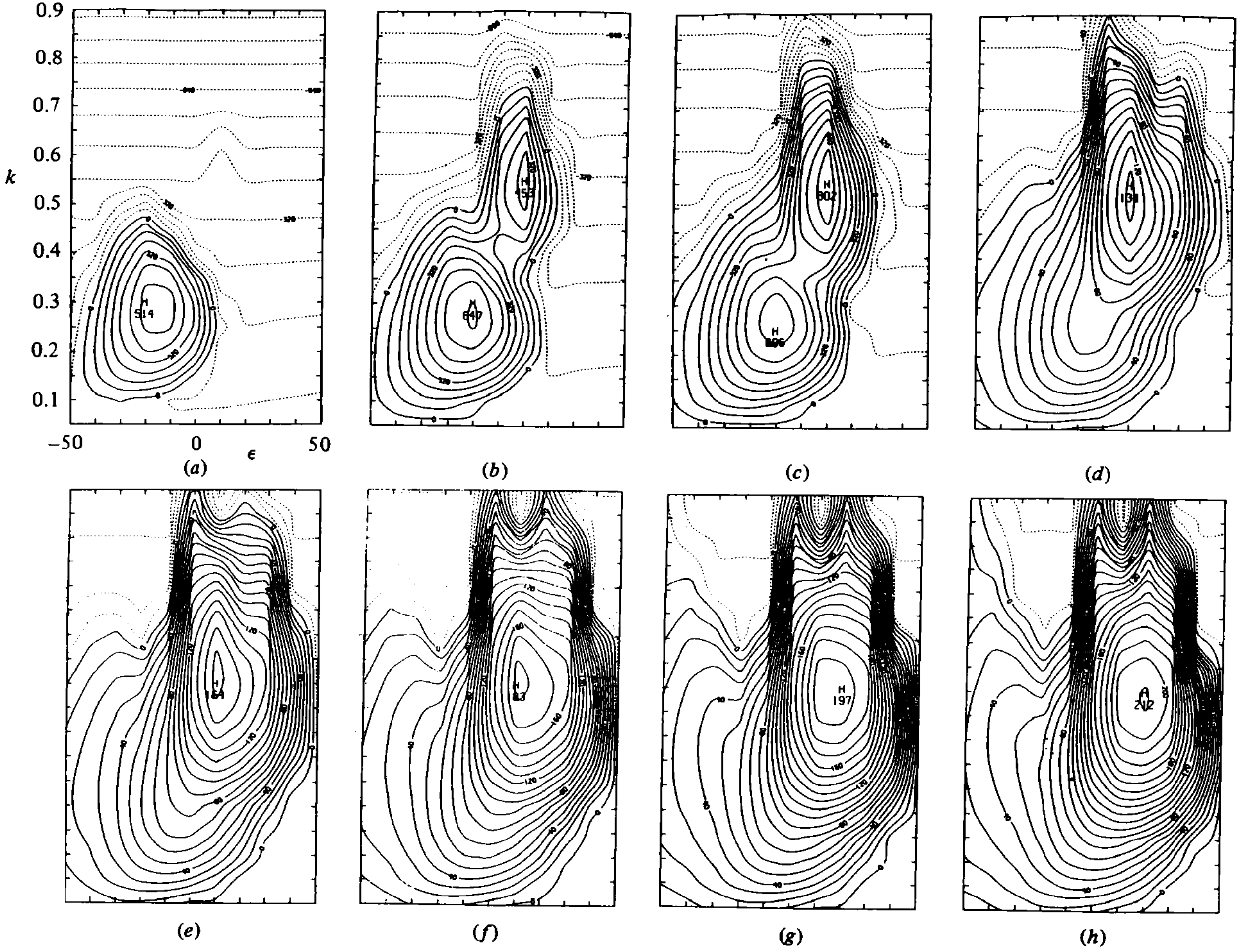

FIGURE 3. Constant-growth-rate contours as a function of roll direction ϵ and wavenumber k for various Reynolds numbers for the traditional problem for the atmospheric model ($\mu = 0$). ϵ varies from $-50°$ to $50°$ and k varies from 0.05 to 0.9 in all the panels. The Reynolds numbers for panels (*a*)–(*h*) are 100, 130, 150, 200, 250, 300, 350, 400 respectively. Note the multiple maxima in panels (*a*)–(*c*).

they exist at different locations in the (ϵ, k)-plane. They are not multiple modes in the sense used later; that is, they do not correspond to the eigenvalues of the same eigenvalue problem. Here we perpetuate the disconcerting confusion in the hydrodynamic-stability literature, in which the use of the word 'mode' of instability seems to have several meanings. At $R = 200$ the (type II) peak at lower wavenumbers has disappeared, leaving only one local maximum corresponding to the type I instability. The region of instability in the (ϵ, k)-plane continues to enlarge, as seen in the remaining panels; in particular, the portion identified at lower R with the type II mode continues to support instability.

The most unstable configuration in the interval $0 \leqslant \mu \leqslant 1$ of μ occurs at $\mu = 1$ (or $\beta = 180°$, for $\lambda = 45°$). Therefore the effect of the horizontal component of rotation at $\lambda = 45°$ is most clearly seen for $\beta = 180°$; and figure 4, containing the same information as figure 3, has been prepared for this case. The first panel ($R = 50$) shows an instability clearly established below the onset for $\mu = 0$, and at a somewhat higher wavenumber. The peak growth rate $\sigma_{\max} = 0.0087$ occurs at $\epsilon = -18°$, $k = 0.55$ and $c_r = 0.508$. This most unstable disturbance resembles the type II instability of the traditional problem in being the continuation in R of the onset mode (with critical conditions $R_c = 33.9$, $\epsilon_c = -18°$, $k_c = 0.594$, $c_r = 0.506$) and having a phase speed significantly different from zero. On the other hand, it has a wavenumber close to that of the type I disturbance of the traditional problem. As R is increased, the phase speed of the most-unstable mode decreases, while the wavenumber decreases slightly and then increases again, to $k = 0.5$ at $R = 400$. The angle ϵ at which the most-unstable disturbance occurs increases smoothly to $\epsilon = 17°$ at $R = 400$. In contrast with the traditional problem, only one local maximum of growth rate occurs over this range of R.

The variations of the most-unstable modes with $R \leqslant 400$ for both $\mu = 0$ and $\mu = 1$ are traced in figures 5 and 6. Figure 5 shows the growth rate σ as a function of R. As a guide, some points are marked and labelled by a pair of numbers; the first of these gives the angle ϵ and the second the wavenumber k at which the most-unstable disturbance is found. For the traditional problem, $\mu = 0$, we track both local maxima, so long as both exist, even though the high-wavenumber peak is not the most unstable over the entire range of R. In this figure computed points are shown by dots, and the continuous curve is drawn between them. Smaller increments of R were taken for the case $\mu = 1$ in the interval $100 \leqslant R \leqslant 170$ to rule out the possibility of existence of two local maxima of σ in a relatively small interval of R.

The phase speeds of the most-unstable distances described in figure 5 are displayed in figure 6. The two branches for the $\mu = 0$ case correspond to the two local growth-rate maxima. Both figures 5 and 6 show the $\mu = 0$ and $\mu = 1$ cases approaching each other at the larger values of R. This reflects the vanishing effect of the perturbation Coriolis force as $R \rightarrow \infty$, evident also in figures 3 and 4.

Figure 7 shows two neutral curves k versus R. One curve is for $\mu = 0$, and is drawn for $\epsilon = -23.3°$, the critical value of ϵ for this case; the other is for $\mu = 1$ and $\epsilon = -18°$, the critical angle for that case. Phase speeds are given in figure 8. The upper branches of the neutral curves of figure 7 correspond to the lower branches of figure 8, and *vice versa*. Two important features are evident in comparing these neutral curves: R_c is lower for the full problem, and the band of unstable wavenumbers is wider at moderate supercritical values of R.

Both upper and lower branches of the two neutral curves in figure 7 appear to be approaching each other for large R, and the lower branches in fact appear to coincide.

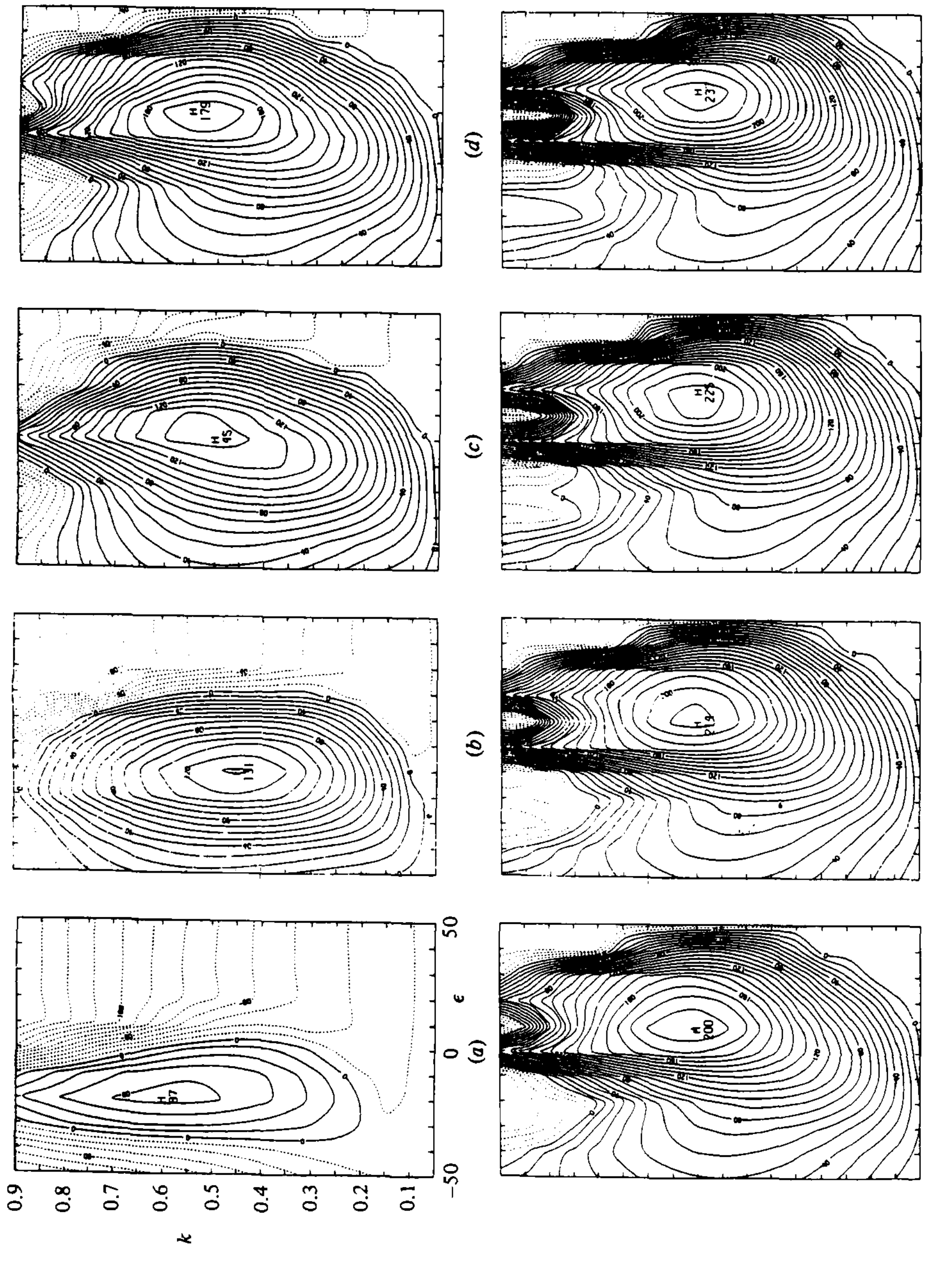

FIGURE 4. Constant growth-rate contours as a function of ϵ and k for $\mu = 1$ for the atmospheric model. The Reynolds numbers in panels (*a*)–(*h*) are 50, 100, 150, 200, 250, 300, 350, 400. Note the absence of multiple maxima in all cases.

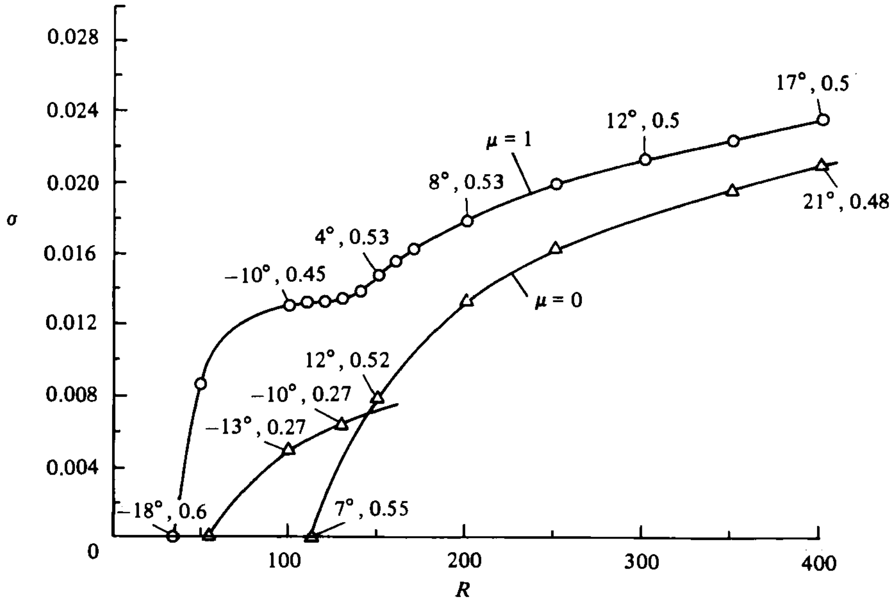

FIGURE 5. Growth rate of the most-unstable mode as a function of Reynolds number for the atmospheric model. Values that tag the points on the curves indicate the roll angle and wavenumber respectively for the most-unstable mode. Circles mark the computed points for $\mu = 1$ and triangles mark the computed points for $\mu = 0$ (traditional problem). Note that with $\mu = 1$ there is only one mode of instability, as opposed to two modes present when $\mu = 0$.

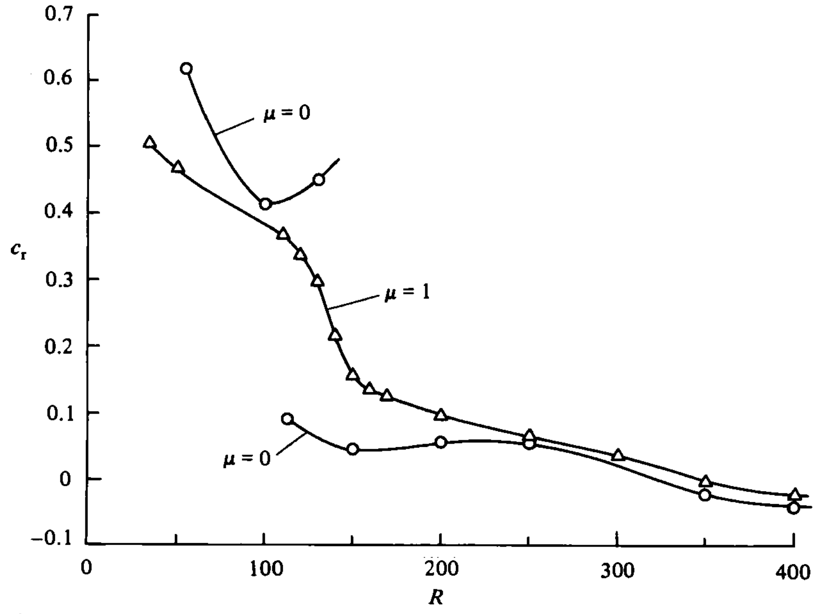

FIGURE 6. Phase speed of the most-unstable mode as a function of Reynolds number for the atmospheric model. Circles mark the computed points for $\mu = 1$ and triangles mark the computed points for $\mu = 0$ (traditional problem). Note that with $\mu = 1$ there is only one mode of instability, as opposed to two modes present with $\mu = 0$.

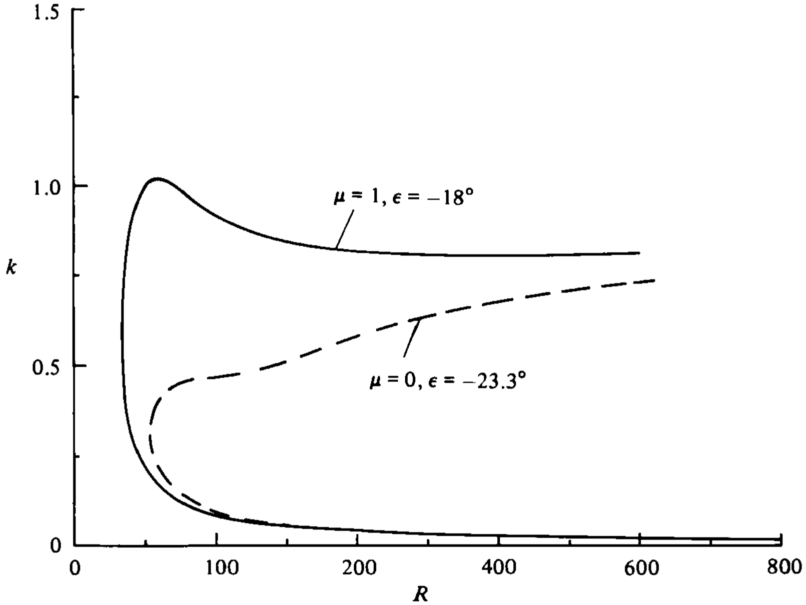

FIGURE 7. Neutral stability curves for the atmospheric model: ——, $\mu = 1$, $\epsilon = -18°$; ---, $\mu = 0$, $\epsilon = -23.3°$ (traditional problem). Note that with $\mu = 1$ the critical Reynolds number is considerably reduced and the range of unstable wavenumbers for supercritical conditions increased.

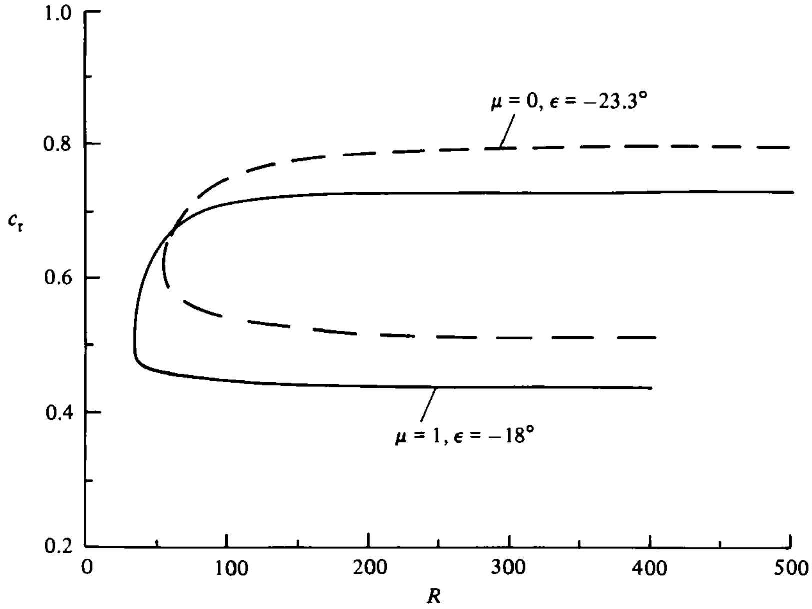

FIGURE 8. Phase speed for the neutral modes displayed in figure 7: ——, $\mu = 1$, $\epsilon = -18°$; ---, $\mu = 0$, $\epsilon = -23.3°$ (traditional problem). Note that the branches with larger values of phase speed correspond to the lower branches of the neutral-stability curves shown in figure 7.

The behaviour of the lower branches in the limit $R \to \infty$ is better displayed in plots of k versus kR; the data in figure 7 are replotted in figure 9. Although one cannot confirm it from these figures, the neutral curves do not in fact coalesce as $R \to \infty$; they turn out to be very close because the two values of ϵ_c for $\mu = 0$ and $\mu = 1$ differ only slightly. In the limit $R \to \infty$, both upper and lower branches of the neutral curves depend only on ϵ. On the lower branch, $R \to \infty$, $k \to 0$, and figures 7–9 suggest that $kR \to$ constant in the limit. This is confirmed by computing the eigenvalue problem for the limit equation set

$$(V-c)\phi'' - V''\phi = (\mathrm{i}kR)^{-1}[\phi^{\mathrm{iv}} + 2\chi'], \tag{25a}$$

$$(V-c)\chi + U'\phi = (\mathrm{i}kR)^{-1}[\chi'' - 2\phi']. \tag{25b}$$

This problem is independent of μ. For each value of ϵ, variation of the parameter kR determines the regions of instability and a neutral value of kR. The results for both the atmospheric and oceanic models are shown in table 2. We note that the atmospheric case for $\epsilon = 0$ has been previously computed by Gill & Davey (1969) for the buoyancy boundary layer with the same result. It is worth noting that the phase speeds in table 2 are either greater than the maximum of $V(z;\,\epsilon)$ over z or less than its minimum; there are no critical layers.

The behaviour of the system for fixed R and variable μ is illustrated in figure 10. Contour plots of growth rate as functions of k and ϵ are given for μ varying from $\mu = -1$ to $\mu = 1$ with R held fixed at 130. Thinking of $\lambda = 45°$, this range of values of μ corresponds to β varying from 0° (rolls due east) to 180° (rolls due west). (Rolls 'east' or 'west' do not sound distinct; however, the system of coordinates employed here implies a disturbance wavenumber vector directed 90° counterclockwise to the roll 'direction', so 'east' rolls have a wavenumber vector directed towards the north, while 'west' rolls have a wavenumber vector directed towards the south.) Only one weak unstable peak is seen in the first plot ($\beta = 0°$) of the panel, centred near the location of the type II disturbance, but a second stable peak exists at higher wavenumber rather near the location of type I disturbances. The second peak gathers strength as β increases, and has achieved positive values at $\beta = 60°$. At $\beta = 90°$, coinciding with the traditional problem, the higher-wavenumber peak is well established, but is still lower than the original unstable peak, and the islands of instability have merged. Only one peak remains at $\beta = 120°$, higher than the two it replaces and closer to the onset peak than to the type I peak. The peak continues to grow and to drift towards the larger ϵ and k regions associated with type I disturbances as β increases to 180°. The problem is symmetric in β about 180°, and the range $0 \leqslant \beta \leqslant 180°$ therefore covers all possibilities; thus the roll direction $\beta = 180°$ is most unstable.

The basic flow is fixed by R and α, the direction of the geostrophic flow, not by R and β. It is therefore of interest to track the most-unstable (highest peak only) disturbance as functions of these physical 'input' parameters and λ. This is done in figure 11 for $\lambda = 45°$ N and $R = 130$. The figure shows the system response parameters as a function of α, giving the growth rate and phase speed of the most-unstable mode, and a pair of numbers giving the associated values of ϵ and k. The response is periodic in α, with a half-period shown.

Figure 12 is the same as 11, but for $R = 300$. Growth rates are larger, and phase speeds are smaller, and dependence on α decreases, in accord with the diminishing influence of Coriolis accelerations.

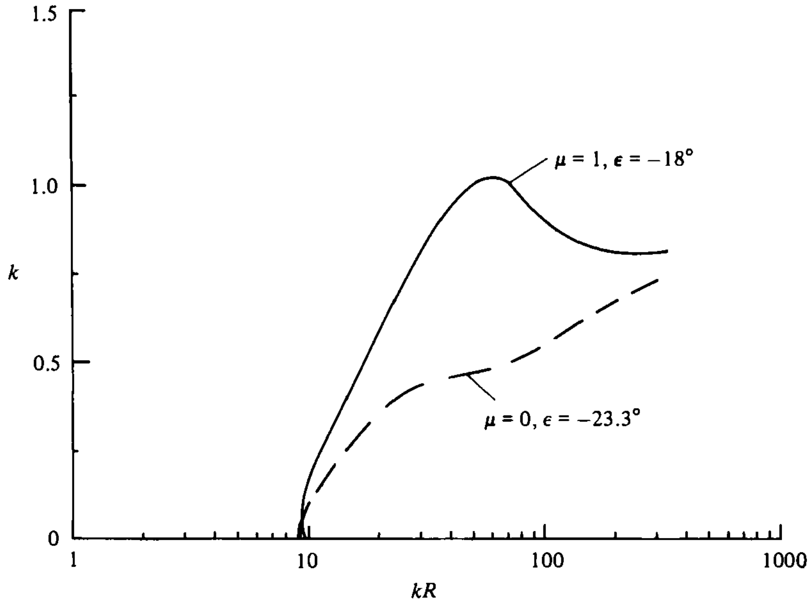

FIGURE 9. Neutral-stability curves for the atmospheric model showing the special feature of the lower branch: ——, $\mu = 1$, $\epsilon = -18°$; – – –, $\mu = 0$, $\epsilon = -23.3°$ (traditional problem). Note that the asymptotic value of kR as $k \rightarrow 0$ is different for the two cases.

	Atmospheric case		Oceanic case	
ϵ	kR	c_r	kR	c_r
−40°	8.7	0.99	0.71	−2.8
−20°	9.3	0.76	0.63	−3.3
0°	13.3	0.47	1.0	−2.2
20°	26.6	0.18	1.7	−1.6
40°	67.5	−0.69	3.4	−1.2

TABLE 2. Lower-branch asymptotes

3.2. *The oceanic model*

The free surface in the oceanic model exerts a weaker restraint on possible disturbance motions, therefore leading to a system more prone to instability – and thus to the low value of critical Reynolds number, $R_c = 11.816$, found by Iooss *et al.* (1978) for the traditional problem. With a critical Reynolds number this low, there is little scope for further destabilization, and the additional effects of the horizontal component of angular velocity are, as expected, quite weak. Table 3 contrasts data for $\mu = \mu^* = 0.156$, the value of μ yielding the smallest value of R_c $(= R_c^*)$, with those for the traditional problem $\mu = 0$.

μ	R_c	k_c	ϵ_c	c_r
0	11.8	0.32	−6.1	−0.561
0.156	11.6	0.33	−5.8	−0.554

TABLE 3

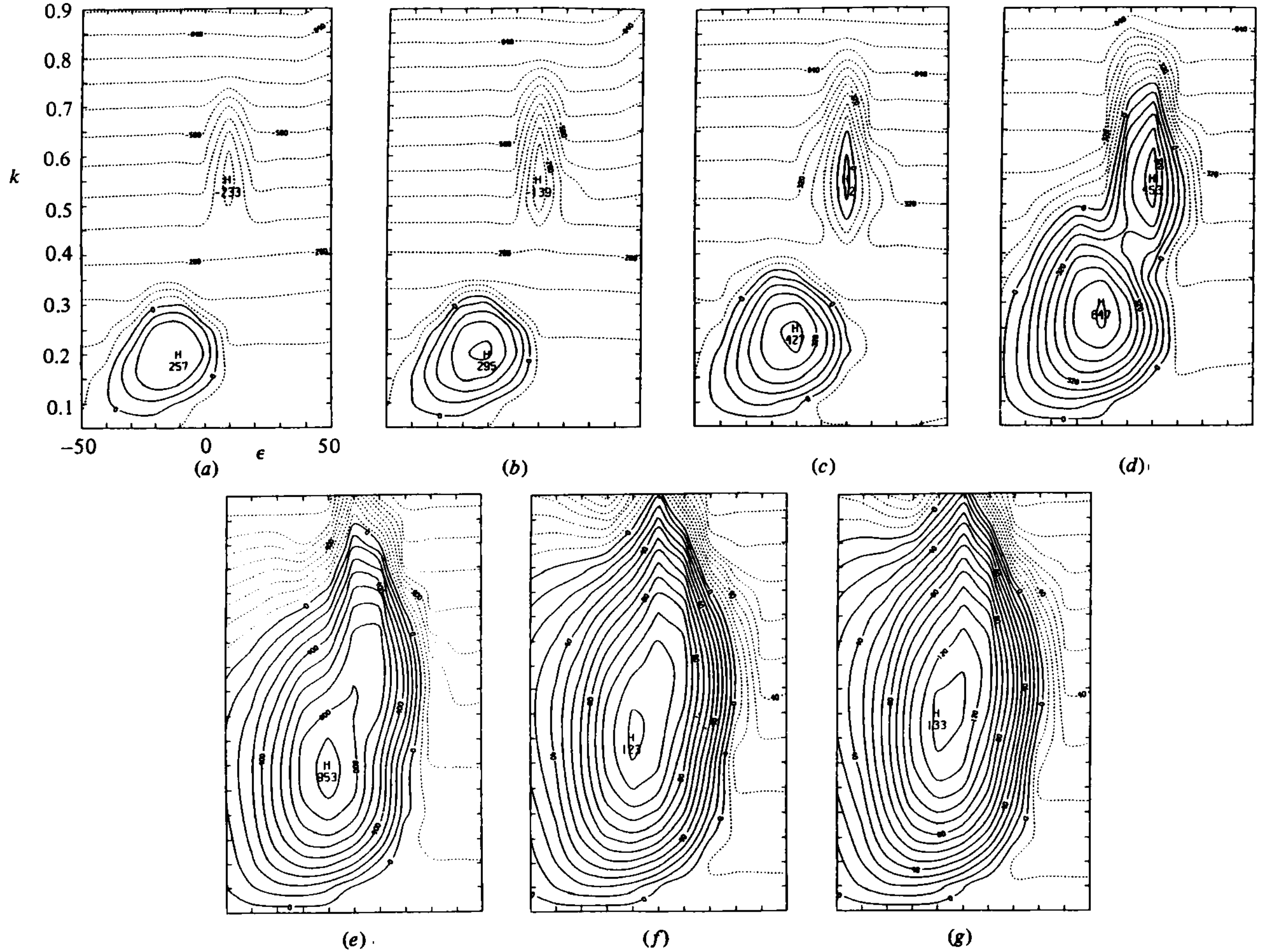

FIGURE 10. Constant-growth-rate contours as a function of ϵ and k for the atmospheric model at a fixed Reynolds number as β is varied. ϵ varies from $-50°$ to $50°$ and k varies from 0.05 to 0.9 in all the panels. $R = 130$ and $\beta = 0$, 30°, 60°, 90°, 120°, 150°, 180° in panels (*a*)–(*g*). $\lambda = 45°$ in all cases. Note the presence of local maxima in (*a*) and (*b*) corresponding to stable modes which lead to multiple positive maxima in (*c*) and (*d*). Note that (*d*) also corresponds to the traditional problem.

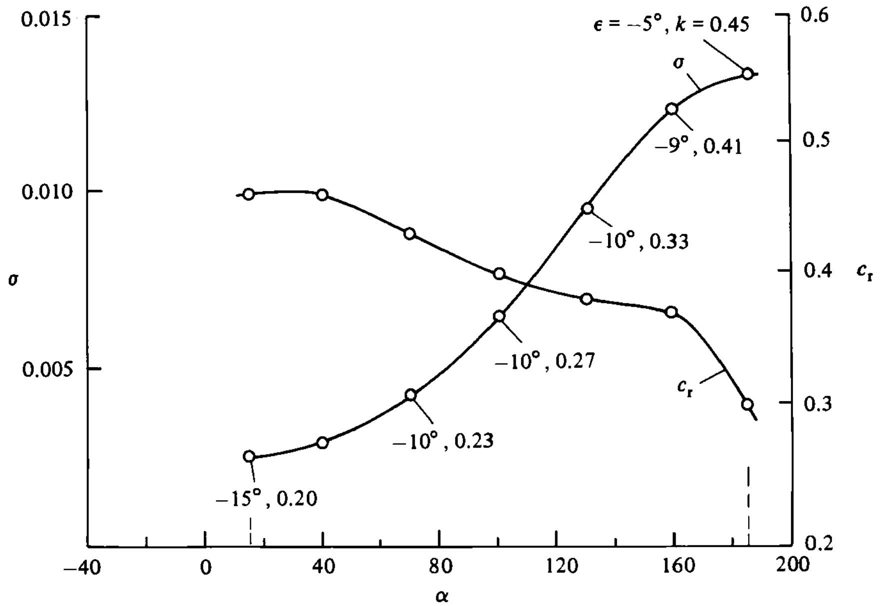

FIGURE 11. Growth rate and phase speed for the most-unstable mode for the atmospheric model as a function of the angle α of the geostrophic flow from due east. In all cases $R = 130$ and $\lambda = 45°$. The values tagging the points indicate the direction ϵ of the rolls and wavenumber k respectively. The curve shown is symmetric about the marks indicated.

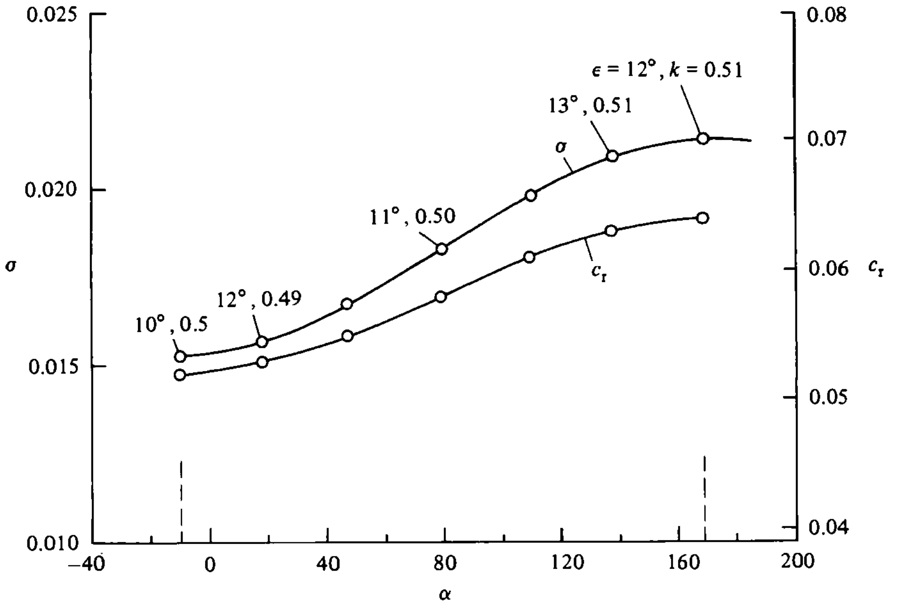

FIGURE 12. Growth rate and phase speed for the most-unstable mode for the atmospheric model as in figure 11, but now $R = 300$ and $\lambda = 45°$. The values tagging the points indicate the ϵ and k respectively. The curve is symmetric about the marks indicated.

The roll direction at onset for this problem is $\beta = 180°$ for latitudes between $|\lambda| = 90°$ and $|\lambda| = \lambda^* = \cot^{-1}\mu^* = 81.1°$. For this narrow range of latitudes, R_c is between 11.6 and 11.8. For $|\lambda| < \lambda^*$, $R_c = 11.6$, and the roll direction at onset is $\beta_c = \cos^{-1}(-\mu^* \tan\lambda)$.

The critical Reynolds number depends so weakly on μ one would think that, in contrast with the atmospheric case, there is no significant dependence of unstable modes on μ. This is not the case. Figure 13, the analogue of figure 5, shows that growth rates of the most-unstable disturbance for $\mu = 1$ are 50 % or more larger for slightly supercritical R than growth rates for $\mu = 0$. Note that $\mu = 1$ is not the most-unstable configuration of the system.

Figure 13 shows the existence of two unstable modes for R exceeding a certain value depending on μ. The most-unstable disturbance switches from one mode to the other at $R \approx 132$ for $\mu = 1$, and at $R \approx 140$ for the traditional problem $\mu = 0$. The existence of two local maxima of σ for $\mu = 0$ is in accord with the atmospheric case, but only one appeared in the atmospheric case for $\mu = 1$. The angle ϵ between the surface velocity vector and the roll direction varies significantly for values of R less than 80, but is virtually constant for each mode for the higher values of R shown in figure 13. Thus the most-unstable mode labelled A for either $\mu = 0$ or 1 settles down to a value of ϵ of 10°, and mode B rapidly settles to a value of about 20°.

Figure 14 shows that the phase velocities of the most-unstable modes are negative and the magnitudes are not small, even for the larger R, in contrast with the atmospheric case. On the other hand, the phase speeds of mode B are small relative to the surface component $V(0)$, the relative propagation speed being about 0.07. The approach to nearly constant values of phase speeds occurs in the range of R at which $\epsilon \rightarrow$ constant, suggesting that the phase speed at higher R depends only on ϵ, a result known to be true for $R \rightarrow \infty$.

Neutral curves, k versus R, are shown for $\mu = 0$ and $\mu = 1$ in figure 15 at $\epsilon = -6°$, close to the critical value of ϵ for the traditional problem $\mu = 0$. This value ϵ does not produce the critical conditions for either $\mu = 0$ or $\mu = 1$, but the shapes of the neutral curves are typical. Although the horizontal component of angular velocity does not change the critical values for instability in the oceanic case, the band of unstable wavenumbers at marginally supercritical values of R is much wider for $\mu = 1$ than it is for $\mu = 0$. Thus the horizontal component produces significantly higher growth rates (figure 13) for a wider range of disturbances than that found under slightly supercritical conditions for the traditional problem.

The data in figure 15 are replotted, k versus kR, in figure 16, to bring out the asymptotic behaviour on the lower branch. This is similar to the atmospheric case, as may be seen by comparing it with figure 9.

The existence of two unstable disturbance forms in the oceanic case (figure 13) appears similar to the traditional atmospheric problem, but there is a fundamental difference. In the atmospheric case only one unstable mode occurs for each (moderate) supercritical value of R at given values of k, ϵ and μ (in §4 we will show that additional disturbance modes enter at very large R). In the oceanic case two unstable modes are found for the same values of k, ϵ and μ for a range of moderate supercritical values of R. This has previously been shown to be true by Spooner (1983). Figure 17 documents the coexistence of two unstable modes for the $\mu = 0$ case; plots for other values of μ are similar.

The most-unstable modes for two supercritical values of R, treated as functions of the Ekman-layer surface-current direction α, are traced in figures 18 ($R = 130$) and 19 ($R = 300$). These are the analogues of figures 11 and 12 for the atmospheric

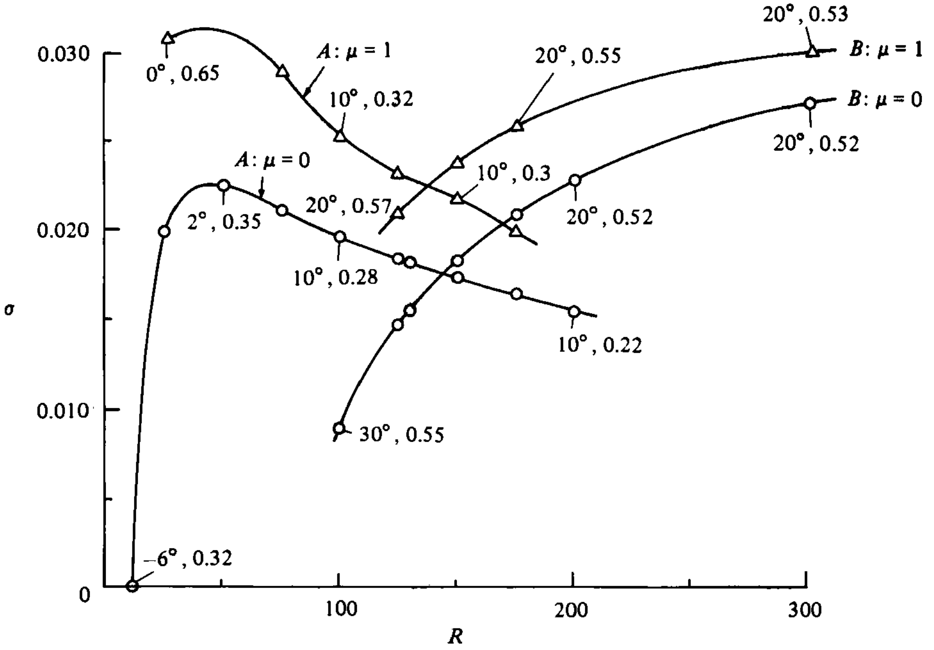

FIGURE 13. Growth rate of the most-unstable modes as a function of Reynolds number for the oceanic model. Values tagging the points indicate the ϵ and k corresponding to the most-unstable modes. Curves marked A are for $\mu = 1$. Curves marked B are for $\mu = 0$ (traditional problem). Note the existence of two modes of instability for both the cases.

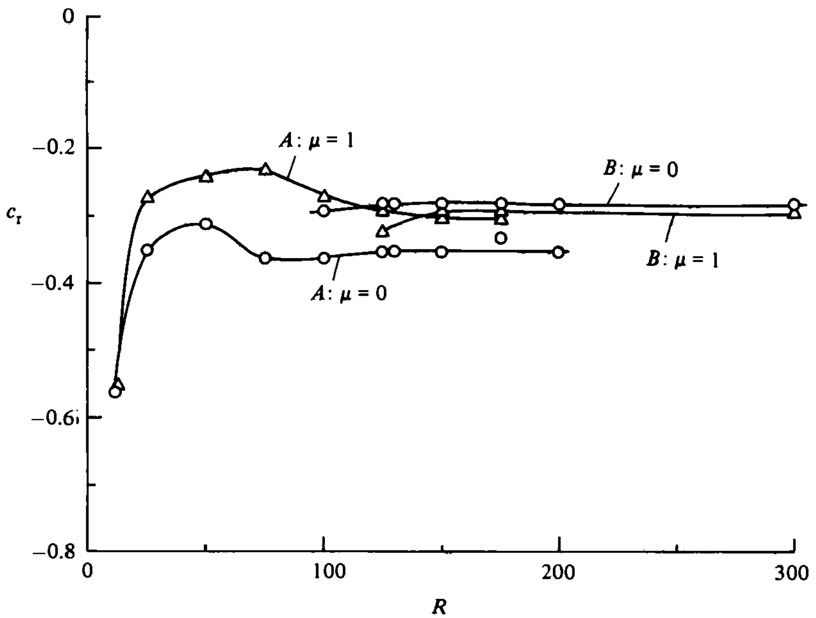

FIGURE 14. Phase speed of the most-unstable modes as a function of Reynolds number for the oceanic model. See figure 13 for explanations.

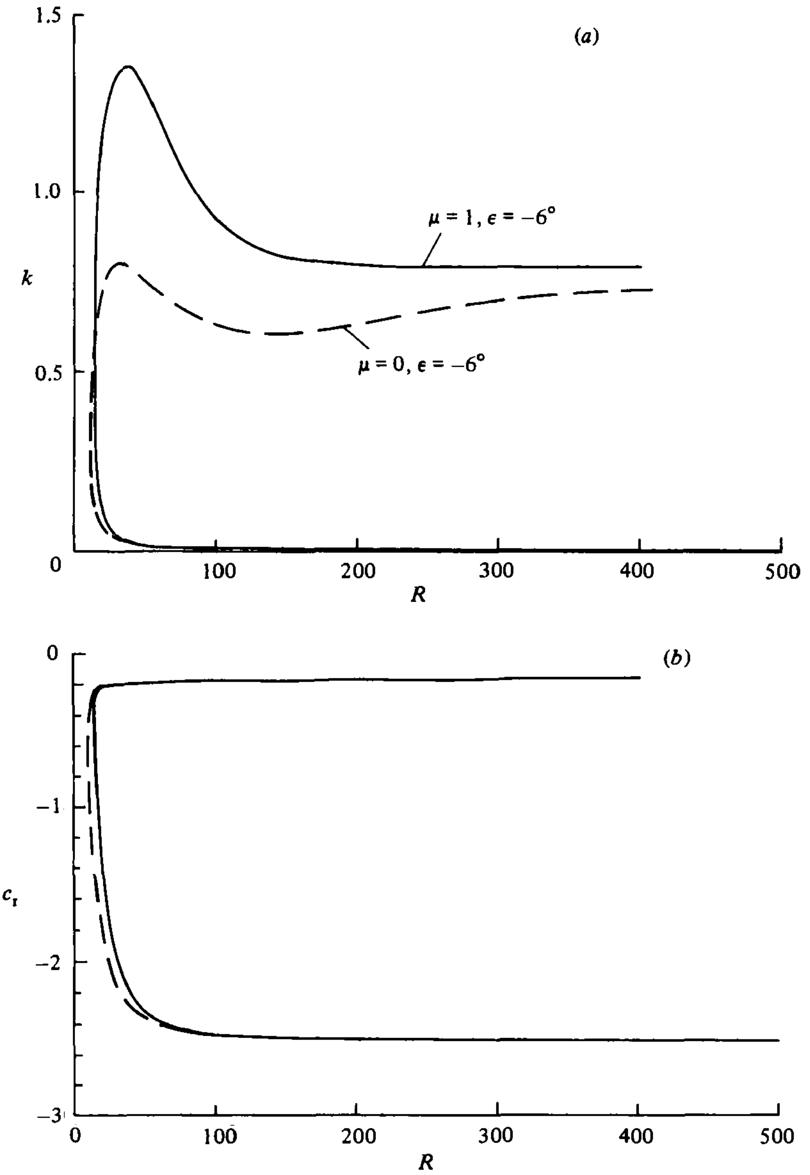

FIGURE 15. (*a*) Neutral-stability curves for the oceanic model: ——, $\mu = 1$, $\epsilon = -6°$; ---, $\mu = 0$, $\epsilon = -6°$ (traditional problem). Note that $\mu = 1$, $\epsilon = -6°$ does not correspond to the critical condition; hence the onset Reynolds number for the traditional problem is smaller than its value for $\mu = 1$, $\epsilon = -6°$. However, the increased range of unstable wavenumbers for supercritical conditions for $\mu = 1$ is clearly seen. (*b*) Phase speed for the neutral modes displayed in (*a*): ——, $\mu = 1$, $\epsilon = -6°$; ---, $\mu = 0$, $\epsilon = -6°$ (traditional problem). Note that the smaller values of phase speed correspond to the upper branch of the neutral-stability curves in (*a*).

case; as in those figures, figures 18 and 19 are drawn for $\lambda = 45°$ N. These curves are periodic in α, with one-half of a period displayed. The growth rates of the most-unstable modes are larger in the oceanic problem than they are in the atmospheric case, reflecting the absence of the stabilizing influence of a rigid wall, but the shapes of the curves are similar at the higher values of R. The wavenumber of the most-unstable mode, the variation between the minimum and maximum values of σ

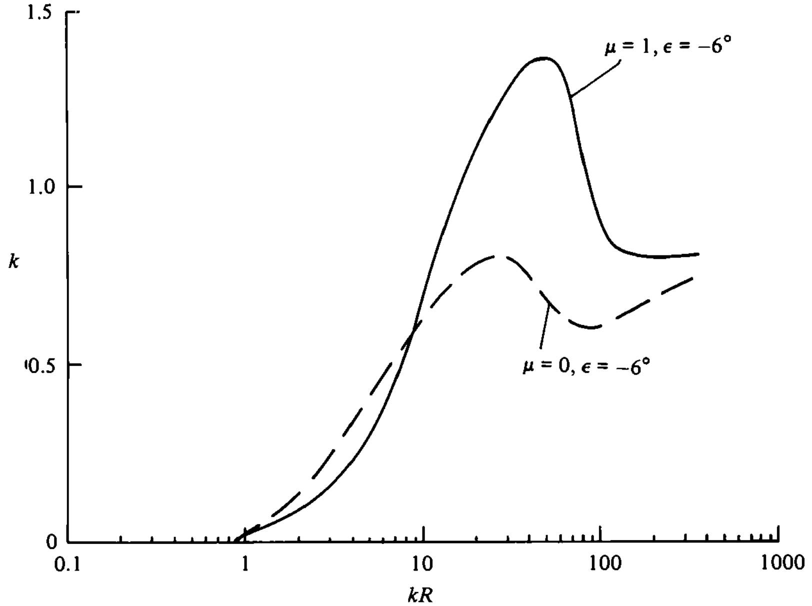

FIGURE 16. Neutral-stability curves for the oceanic model, showing the special feature of the lower branch: ——, $\mu = 1, \epsilon = -6°$; ---, $\mu = 0, \epsilon = -6°$ (traditional problem). Note that the asymptotic values of kR as $k \rightarrow 0$ are equal for the two cases shown.

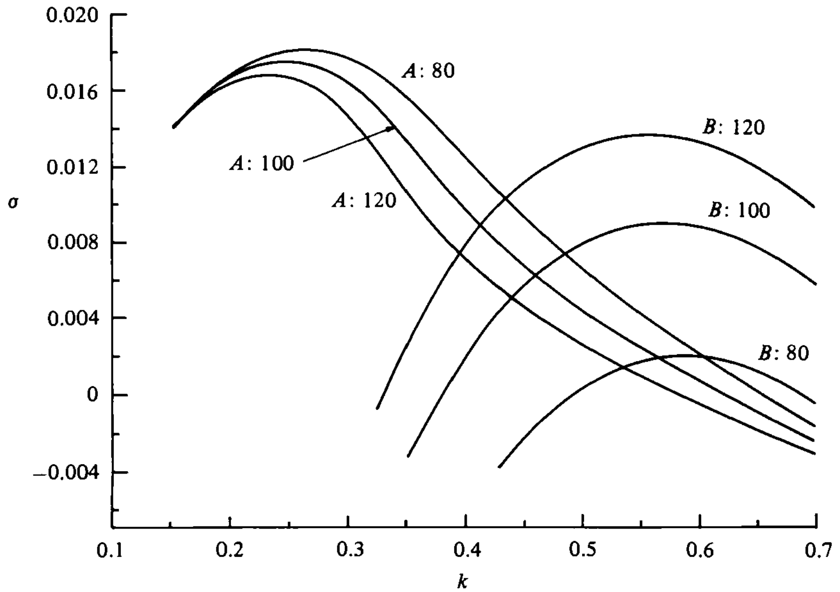

FIGURE 17. Growth rate as a function of wavenumber for the oceanic model for different Reynolds number. $\mu = 0$ in all cases and $R = 80$, 100 and 120 as indicated. Curves marked A correspond to the mode usually called type II or parallel mode. Curves marked B correspond to the mode usually called type I or inflectional mode.

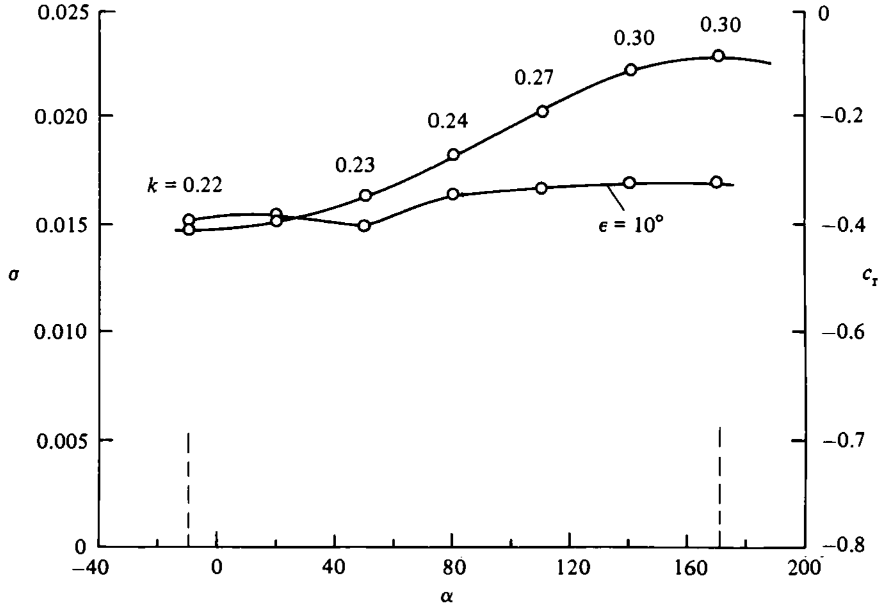

FIGURE 18. Growth rate and phase speed for the most-unstable mode for the oceanic model as a function of the angle (α) of the surface flow from due East. $R = 130$, $\mu = 1$. The value tagging the points is the wavenumber corresponding to the most-unstable mode. In all cases shown the most-unstable has $\epsilon = 10°$.

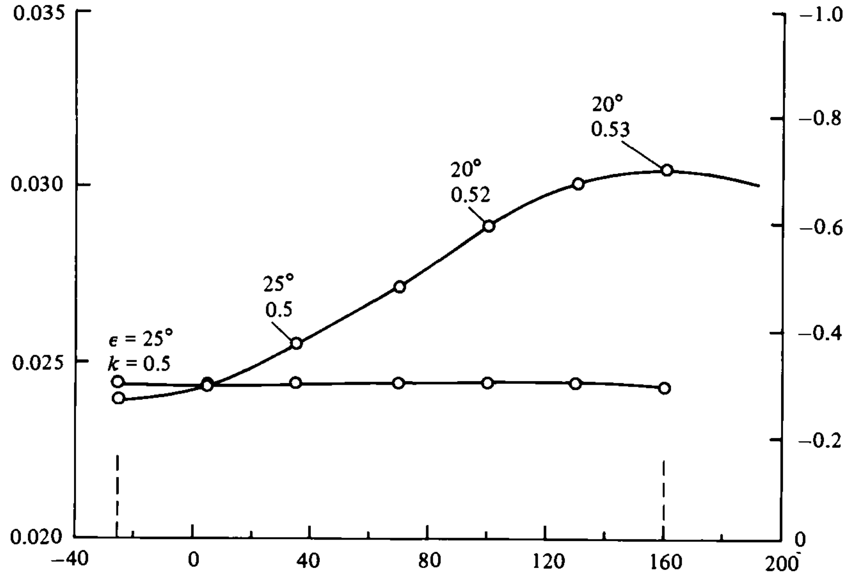

FIGURE 19. Growth rate and phase speed for the most-unstable mode for the oceanic model as a function of α. $R = 300$, $\mu = 1$. The values tagging the points indicate the ε and k corresponding to the most-unstable mode.

and their respective locations in α, are similar at $R = 300$. As we shall see in §4, $\lim_{R\to\infty} \sigma_{\max}(\epsilon, k, R, \alpha)$, known from the results of §2 to be independent of α and to be the same for both the atmospheric and oceanic model, occurs at $\epsilon \approx 23°$ and has value $\sigma_{\max} \approx 0.037$. Comparing figures 12 and 19 suggests that $R = 300$ is 'larger', i.e. closer to the inviscid limit in the oceanic case, than the corresponding atmospheric case.

The phase speeds in figures 18 and 19 vary little with α. The most-unstable modes propagate normal to, and to the right of, the roll axis at a significant speed. The speed relative to the surface current is, however, quite small. Thus, if an unstable roll motion were to be made visible, it would appear to be (nearly) advected with the surface current.

The inviscid limit is considered in §4.

4. The inviscid limit

We turn to the limit $R \to \infty$, k fixed. As discussed in §2, the limit problem requires the solution of Rayleigh's equation (16) subject to boundary conditions (17). Furthermore, as discussed in §2.4, one may restrict attention to the atmospheric case, finding the oceanic results by use of (20). We therefore focus on the atmospheric case $V(z; \epsilon) = -\sin\epsilon + \mathrm{e}^{-z}\sin(z+\epsilon)$.

We recall that the Rayleigh equation captures the behaviour only of unstable modes, $c_i \to 0+$ (Lin 1955), and that non-trivial solutions exist only if V has a point of inflection. Furthermore, it is necessary that $\mathrm{Re}\,(V-c) = 0$ at at least one interior point of the flow. Rayleigh's equation is therefore singular, for *neutral* modes only, at a minimum of one interior point z_c, defined by the condition $V(z_c) = c_r$, unless z_c is also a point of inflection. In the latter case, z_c is a regular point of the differential equation; solutions for which all points are regular are called *regular neutral modes*, following Miles (1961). Neutral modes satisfying Rayleigh's equation with a singular point in the interior are *singular neutral modes*.

The Ekman-layer velocity profile has an infinite number of inflection points at the levels

$$z_n = (n+\tfrac{1}{2})\pi - \epsilon, \quad n = 1, 2, \ldots.$$

If $V - c_r$ vanishes at any inflection point z_n, then it cannot vanish again in $z > z_n$, and no zeros of $V - c_r$ for $z < z_n$ can be at points of inflection. A few profiles are drawn in figure 20, and the inflection points corresponding to $n = 1$ are identified. The $n = 1$ inflection point for $\epsilon = 15.93°$ falls on the z-axis (that is $V(z_1) = 0$). It is clear that the vanishing of $V(z_n) - c_r$, any n, in the range $\epsilon < 15.93°$ requires $V - c_r$ to vanish at least twice, and only one of the intersections can be at an inflection point; thus any neutral modes that may exist for $\epsilon < 15.93°$ must be singular neutral modes. By contrast, if $\epsilon > 15.93°$, there is the possibility of one and only one regular neutral mode, with $c_r = V(z_1)$. Furthermore, by a well-known argument due to Tollmien (see Lin 1955), one expects a regular neutral mode to exist with this value of c_r at an appropriate value of k with contiguous unstable modes existing for neighbouring values of k.

Regular neutral modes are easy to compute when they exist; c_r is known in advance, and the problem devolves to the determination of k. For this calculation, the Rayleigh equation reduces to an ordinary Sturm–Liouville problem. Unstable modes arising at neighbouring values of k are also easy to compute. By contrast, singular neutral modes arise at c_r values that are *a priori* unknown, and the associated eigenfunction is singular at the critical point. To find the singular neutral modes, it

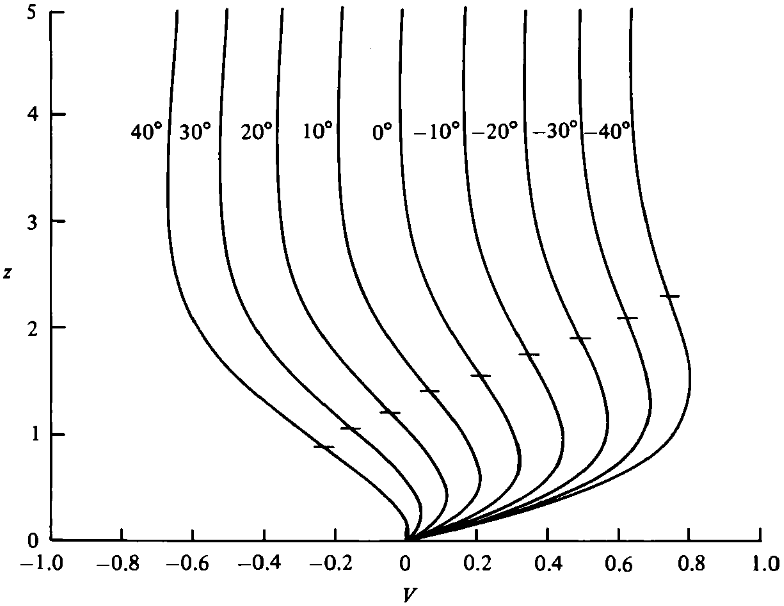

FIGURE 20. Velocity profile of the velocity component V transverse to the roll direction for the atmospheric model for different roll directions. ϵ varies from $-40°$ to $40°$. Location of the inflection point nearest to the ground ($z = 0$) is indicated on the curves.

is preferable first to compute unstable modes, varying k so that $c_i \rightarrow 0+$; at the same time the limit $c_i \rightarrow 0+$ must be computed by analytical continuation to the complex z-plane, deforming the contour of integration to avoid the singular point $z = z_c$. Thus the determination of singular neutral modes is more difficult, and involves the prior determination of unstable modes. We find multiple modes of instability for angles $\epsilon > 15.93°$, for which a regular neutral mode is possible, as well as for angles $\epsilon < 15.93°$, for which no regular neutral modes are possible. Since each unstable mode is presumably associated with a neutral mode as k is varied, it appears that one must anticipate the existence of singular neutral modes for any value of ϵ. With this in mind, we always start with unstable modes, finding neutral modes as special cases: we therefore begin by describing the unstable disturbances.

4.1. *Unstable disturbances*

We may locate unstable modes in one of two ways. For a given value of ϵ we may fix k and compute the complex phase speeds c by an inviscid version of our Galerkin routine. This matrix method produces approximations to a substantial number of eigenvalues, and increments in k may be chosen arbitrarily. Thus a survey in k may be carried out with coarse increments of k. When an approximation to an unstable mode is found the dispersion relation can be found by taking smaller steps in k. (By a 'mode' here, we mean the family of disturbances associated with the eigenvalues c existing as continuous functions of the parameter k. Multiple unstable modes occur if, for fixed ϵ and k, more than one eigenvalue with $c_i > 0$ is found.) Generally, tracing a dispersion relation using a matrix method is inefficient and leads to inaccuracies for small c_i, and we therefore turn to another method to validate Galerkin results, and to trace out the details of the dispersion relations.

The second method that we use is a shooting technique; such methods require a good initial guess for $c(k; \epsilon)$, which we obtain by use of our Galerkin procedure. The

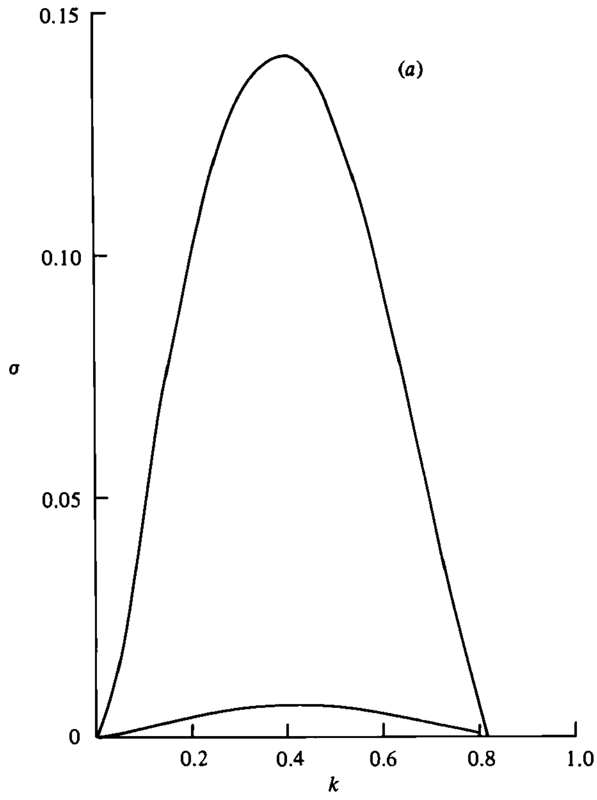

FIGURE 21 (*a*). For caption see facing page.

dispersion relation for each mode can then be traced by taking fine steps in k. An alternative starting procedure that we could have employed is to appeal to the results computed by Gill & Davey (1969) for the BBL, identical with our problem for $\epsilon = 0$. Results can be obtained by continuation in both ϵ and k using shooting methods. Gill & Davey (1969), however, found only one unstable mode; thus this alternative procedure is incomplete and therefore not adopted. Shooting methods are natural for systems that are singular or 'nearly' singular (eigenvalues with small c_i contiguous to singular neutral modes). While matrix methods probably can be modified to treat such problems, they lose those features that give them advantages over shooting methods; thus shooting (or a matrix imitation of shooting) is essential for such problems. Having located a mode, we trace it by continuation in k using a shooting method.

Figure 21 (*a*) shows the growth rates of the first two unstable modes as functions of wavenumber k, for $\epsilon = 0$, and figure 21 (*b*) gives the corresponding phase speeds. The 'primary mode', that giving the largest value of $\sigma = kc_i$, has previously been described by Gill & Davey (1969), and our results agree with theirs; for example, they cite max $\sigma = 0.014$ at $k = 0.39$, which we confirm. The second instability mode is new. For this mode, max σ is 0.00064, considerably smaller, and the band of unstable wavenumbers almost coincides with that of the primary mode, but extends to slightly larger wavenumbers. Figure 21 (*b*) shows that the phase speed of the primary mode is positive, and that of the secondary mode slightly negative. The eigenfunctions for the most unstable of the two modes are shown in figure 23 (page 71); note the higher penetration of the second mode.

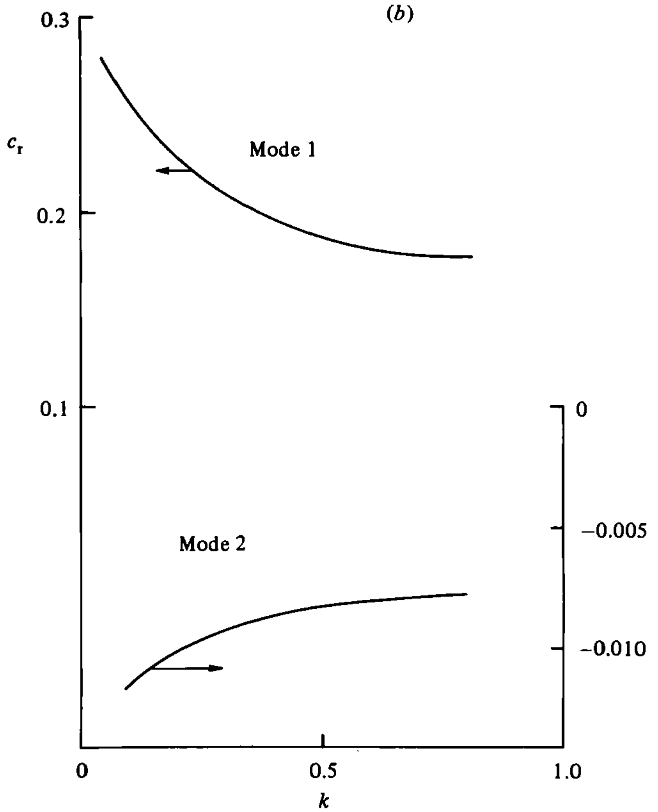

FIGURE 21. (*a*) Growth rate as a function of wavenumber k for the first two modes for the inviscid problem for $\epsilon = 0$. Note that the maximum growth rate for the second mode is far smaller than the first mode. (*b*) Phase speeds for the modes displayed in (*a*). Note that mode 2 has a very small negative phase speed.

The unstable bands of primary and secondary modes do not always show as much overlap as they do for $\epsilon = 0$. The functions $\sigma(k)$, or fragments thereof, are shown in figure 22 (*a*) for $\epsilon = 8°, 9°$ and $10°$; the respective phase speeds are given in figure 22 (*b*). The maximum growth rates of primary and secondary modes shown in this figure do not differ greatly. Furthermore, the high-wavenumber ends of the unstable bands are significantly larger for the secondary modes than for the primary modes.

There are at least two unstable modes for every angle ϵ for which instability obtains, and we conjecture the existence of an infinite number of modes whenever a primary mode exists. The existence of three unstable modes has been verified for $\epsilon = -20°$, $k = 0.5$, with eigenvalues shown in table 4. Also shown in the table are the values of V at the first three inflection points z_n, $n = 1, 2, 3$, the locations of these inflection points, and the locations of the critical levels for each mode. The limit $V(z)$, $z \to \infty$, for this value of ϵ is $V_\infty = 0.342020$: the difference $|V(z_n) - V_\infty| \to 0$ rapidly with n, $V(z_3)$ being 0.34229 and $V(z_4)$ being 0.34201.

The first mode has two critical levels, the higher being slightly greater than z_1; the second has three critical levels, the highest being slightly greater than z_2. The third mode has three critical levels, the third being near z_3. Since c_r is known to a fixed accuracy and the difference $|c_r - V_\infty| \to 0$ as n increases, there is a loss accuracy in the location of the critical levels, and we cannot say if the trend of the first two modes is continued, with the highest critical level just above z_3. In any event, each

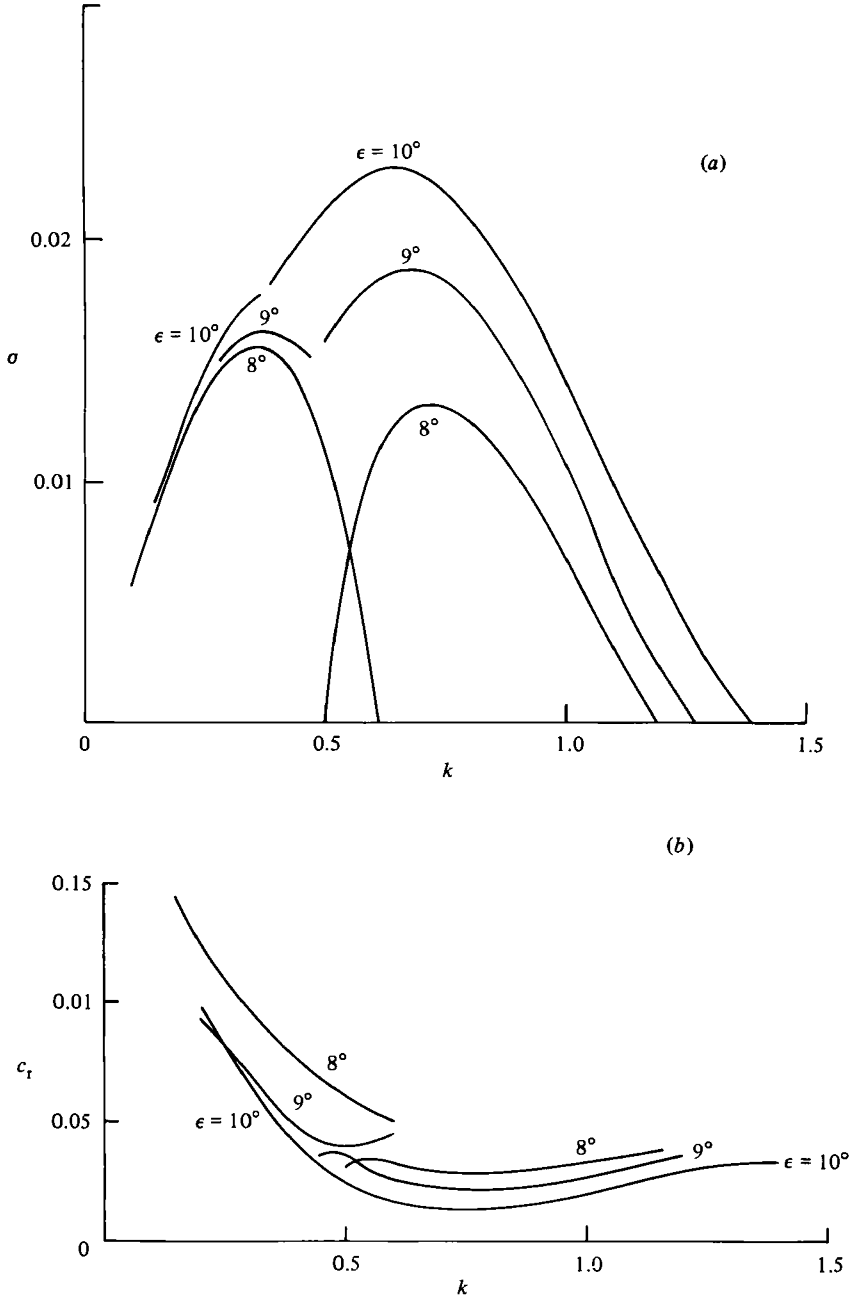

FIGURE 22. (a) Growth rate as a function for wavenumber k for the first two modes for the inviscid problem. The curves are for $\epsilon = 8°$, $9°$ and $10°$ respectively. (b) Phase speeds for the modes displayed in (a).

mode is associated with a critical level 'near' an inflection point, and the next mode to occur is associated with the next (higher) critical level, and has one additional critical level. Our conjecture that an infinite number of unstable modes exists is based upon the observations just made, since the profile $V(z)$ possesses an infinite number of points of inflection.

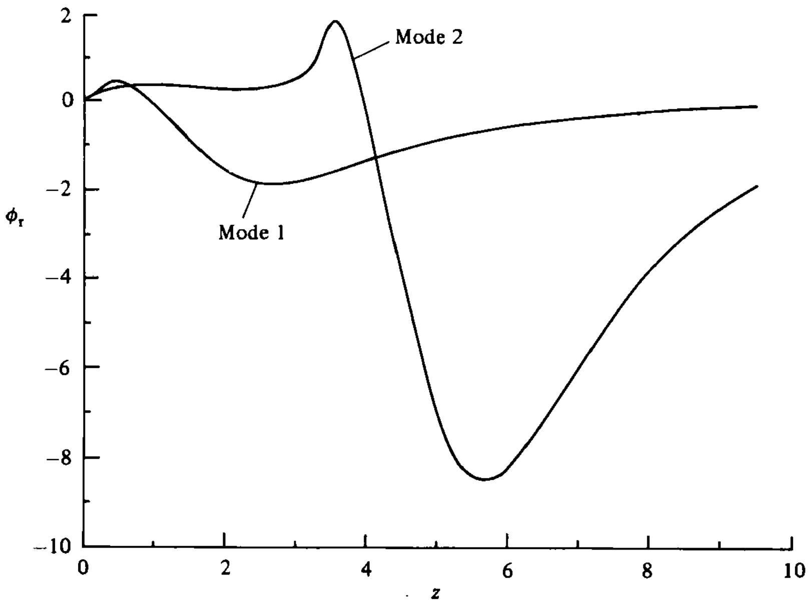

FIGURE 23. Real part of the eigenfunction ϕ for the first two unstable modes for the inviscid problem. $\epsilon = 0$, $k = 0.4$. The eigenfunctions in this figure and in figure 24 are normalized to yield $d\phi/dz = 1$ at $z = 0$. Note the deeper penetration of mode 2.

Mode	c	z_n	$V(z_n)$	z_c
1	0.4751+0.01943i	1.920	0.48865	2.013, 0.592
2	0.3363+0.0008502i	5.061	0.33658	5.16, 3.73, 0.34
3	0.3423+0.00003674i	8.203	0.34229	8.2, 6.9, 3.5, 0.35, 0.3495

TABLE 4. Eigenvalues for $\epsilon = -20°$, $k = 0.5$

The higher the mode, the greater the computing requirements. Since the highest critical level increases by approximately π as one goes from mode n to mode $n+1$, the computational region required to capture each successive mode increases significantly. This is illustrated in figures 23 and 24, which shows the real part of the eigenfunctions for the three modes of table 3; these are normalized by taking $\phi'(0) = 1$. The eigenfunctions have maximum slope near the highest critical level and are maximum in absolute value for values of z above the highest critical level.

To confirm the existence of multiple modes for fixed k, $R \to \infty$, we have computed the viscous $\mu = 0$ equations for $k = 0.3$ for a sequence of Reynolds numbers ranging from 500 to 32000. The results for the atmospheric model for $\epsilon = 0°$ and for the oceanic model for $\epsilon = 20°$ are shown in table 5: for these calculations, each variable was expanded in a 50 term Galerkin series. The upper boundary was taken to be at $z = z_m = 15$ in the atmospheric model and $z = -20$ in the oceanic case. For comparison, the results from the Rayleigh equation are also shown; the Rayleigh-equation phase speed has been adjusted for the oceanic case using the transformation indicated in §2.4. Notice that the Rayleigh equation accurately predicts the primary mode over the entire range of R shown. Although the second mode is present at

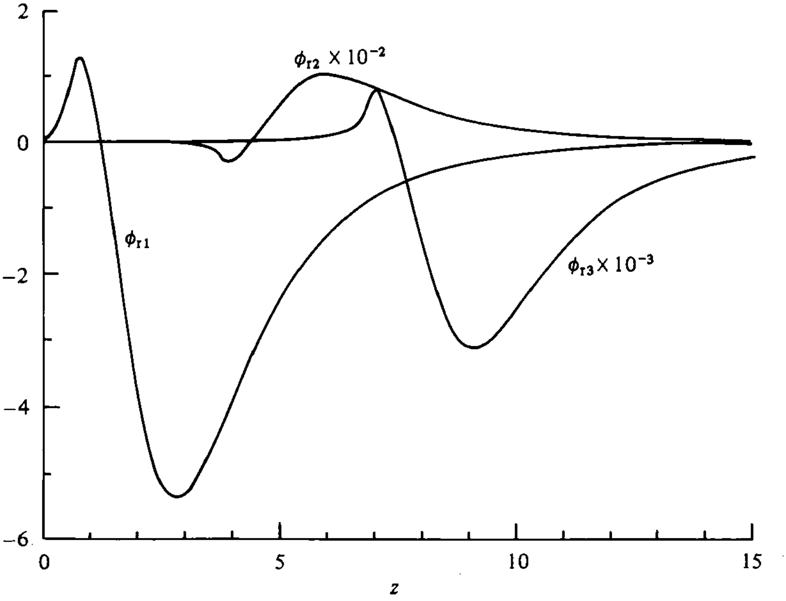

FIGURE 24. Real part of the eigenfunction ϕ for the first three unstable modes for the inviscid problem. $\epsilon = -20°$, $k = 0.5$. Note the successively deeper penetration of higher modes. The subscript on ϕ_r indicates the mode number.

	Atmosphere $\epsilon = 0°$, $k = 0.3$	
R	c (mode 1)	c (mode 2) × 10
500	0.215522+0.0434869i	−0.167226−0.027159i
1000	0.206772+0.0459140i	−0.130228−0.005799i
2000	0.204442+0.0433230i	−0.112328+0.004375i
4000	0.206365+0.0410813i	−0.101881+0.009725i
8000	0.208811+0.0414429i	−0.096005+0.013714i
16000	0.209797+0.0427342i	−0.093137+0.016507i
32000	0.209993+0.0434141i	−0.091452+0.017726i
∞	0.21000 +0.043975i	−0.091645+0.019485i

	Ocean $\epsilon = 20°$, $k = 0.3$	
R	c (mode 1)	c (mode 2) × 10
500	−0.318340+0.077668i	0.205706−0.020657i
1000	−0.311769+0.085859i	0.170093+0.000578i
2000	−0.308277+0.089350i	0.151163+0.010519i
4000	−0.306387+0.090947i	0.139775+0.016862i
8000	−0.305385+0.091705i	0.133906+0.021288i
16000	−0.304865+0.092072i	0.131950+0.022955i
32000	−0.304598+0.092251i	0.134093+0.022782i
∞	−0.30429 +0.092382i	0.1299 +0.027605i

TABLE 5. Eigenvalues for two examples, illustrating the existence of multiple modes of instability at large R

Reynolds numbers as low as 1000 (in the oceanic case), it is not accurately predicted by the Rayleigh equation until R is very large.

In figure 25 the growth rates, wavenumbers and phase speeds of the most-unstable inviscid disturbances are presented as functions of ϵ for the interval $0^\circ \leqslant \epsilon \leqslant 40^\circ$. The data in this figure were generated by identifying the primary mode for each ϵ, then locating its peak over k. The maximum growth rate σ_{m} for any given mode, and the wavenumber k_{m} at which it occurs are continuous functions of ϵ. The break in the curve at $\epsilon \approx 8.5^\circ$ is associated with a crossover of the primary and secondary modes. The continuations of the secondary mode σ_{m} and k_{m} are shown in figure 25 as broken lines.

The most-unstable inviscid mode having zero phase speed is of particular interest. Lilly (1966) has argued that the experimental method (flow visualization using dye) used by Faller (1962) to discover the type I instability is primarily responsive to disturbances with small phase speeds. Faller noted the similarities between his experiment and that of Gregory, Stuart & Walker (1955, GSW), and of Stuart's stability analysis appearing in that paper. The experimental method of GSW is responsive only to disturbances with zero phase speed, and Stuart found reasonable agreement between the band angle (equivalent of ϵ) associated with neutral disturbances with zero phase speed and experimentally observed band angles. Faller (1962), following Stuart, also found reasonable agreement between experimentally observed band angles and those associated with inflection points at small phase speeds. Stuart found poor agreement between the wavelengths predicted by the same method, and those observed experimentally; Faller did not have the corresponding theoretical wavelength information. It would seem that these comparisons are in any event inappropriate; it seems more reasonable to compare experimentally observed patterns with the *most-unstable* disturbances, not with neutrally stable ones. If this is done in the present case, the results are in better agreement with the experimental data. Faller's observed average wavelength was $10.9D$, or $k = 0.58$, and $\epsilon = 14.5^\circ \pm 2^\circ$. We have computed the neutral mode having zero phase speed, and find it to occur at $\epsilon = 15.93^\circ$ and $k = 1.38$, giving a wavelength shorter than experiment by more than a factor of 2. By contrast, we find that the most-unstable mode with zero phase speed (see figure 25) occurs at $\epsilon = 11.8^\circ$ and $k = 0.6$, close to experimental observations.

In GSW, Stuart suggested that the discrepancy in wavenumber between that observed and that predicted for neutral stability for the rotating disk is evidence of the considerable influence of viscosity on the wavenumber. The present quite analogous case suggests that the discrepancy is due to comparing inappropriate situations, and that better agreement is achieved if the experimental data are compared to the most-unstable mode having zero phase speed.

4.2. *Neutral modes*

For each ϵ the growth-rate curve $\sigma(k)$ has two branches for each instability mode, and the neutral wavenumber of the upper branch can be found numerically. For the singular neutral modes which predominate, this must be done by appropriate deformation of the contour of integration. All higher modes have more than one critical level and are necessarily singular.

Since more than one unstable mode appears to exist for each value of ϵ, a collection of neutral wavenumbers k_n of the upper branches of $\sigma_n(k)$ exists, where the index n labels the unstable modes. If the collection of neutral wavenumbers were to be sorted according to wavenumber, with $k_{n+1} > k_n$, then the largest such k defines the asymptote of the neutral curve as $kR \to \infty$ for each ϵ. Unfortunately we cannot track

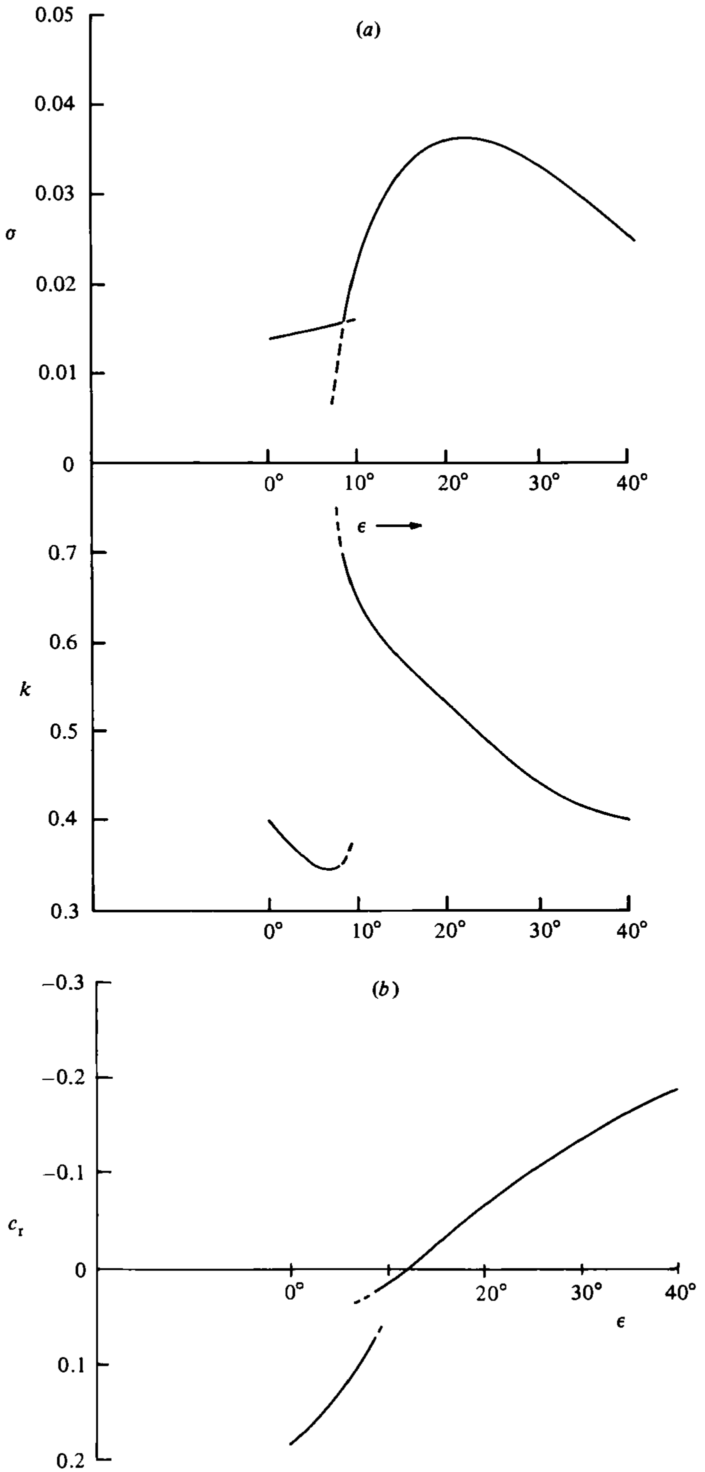

FIGURE 25. (*a*) Growth rate and wavenumber for the most-unstable mode of the inviscid problem as a function of ϵ. Dashed lines indicate the continuation of the most-unstable mode into regions where they no longer correspond to the most-unstable mode. (*b*) Phase speed for the modes displayed in (*a*). See (*a*) for explanations.

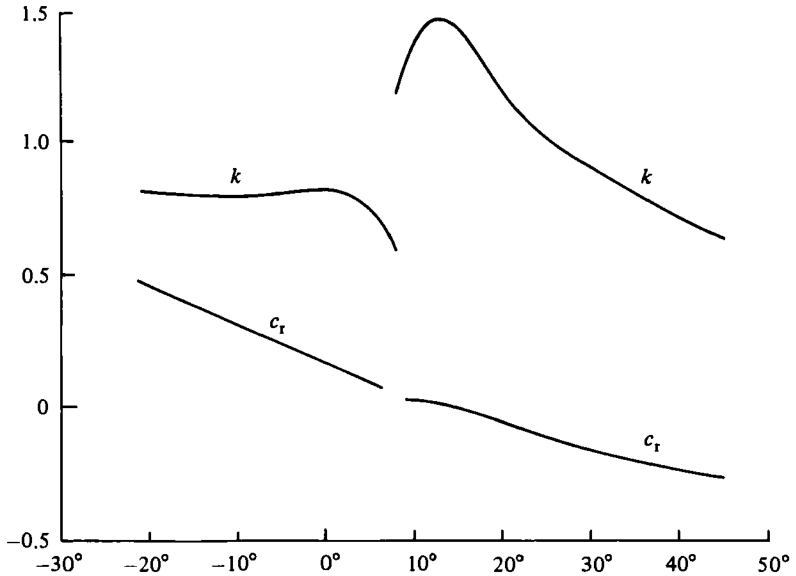

FIGURE 26. Wavenumber and phase speed of neutral modes for the inviscid problem.

more than three unstable modes. We are unable even to determine if the number of unstable modes is finite or infinite, and, if the latter is true, whether the sequence k_n has an accumulation point.

The uncertainty in fixing the asymptote of the neutral curves is, however, largely a question of academic interest. Except in the neighbourhood of certain angles (like $\epsilon \approx 8.5°$, as shown in figure 25) at which primary and secondary modes interchange, the secondary and tertiary instability modes that we have found are extremely weak, and only the primary mode is of significance. In figure 26 we show the upper neutral wavenumber for the primary modes for ϵ between $-20°$ and $50°$, together with the associated phase speeds.

5. Energetics

The equation governing perturbation kinetic energy may be obtained by forming the scalar product of (5) with the perturbation velocity vector $\boldsymbol{v}$, followed by integration over a volume consisting of one wavelength in the y-direction, all z occupied by the fluid and unit distance in the x-direction. The perturbation Coriolis forces do no net work and therefore do not contribute to the overall energy balance. They do provide coupling, however, allowing energy exchange between the kinetic energy associated with perturbation velocity component along the rolls (x-direction or longitudinal) with those in the cross-plane (y, z) normal to the roll axis. Therefore it is useful to separate out these contributions to the kinetic energy. Let

$$\langle f \rangle \equiv \frac{k}{2\pi} \int_0^{2\pi/k} f(y,z)\,\mathrm{d}y$$

for any function of y and z periodic in y, and let

$$E_1 = \frac{1}{2} \int_a^b \langle u^2 \rangle\,\mathrm{d}z, \tag{26a}$$

$$E_2 = \frac{1}{2}\int_a^b \langle v^2+w^2\rangle \, \mathrm{d}z, \tag{26b}$$

$$I_1 = -\int_a^b \langle uw\rangle \, U' \, \mathrm{d}z, \tag{26c}$$

$$I_2 = -\int_a^b \langle vw\rangle \, V' \, \mathrm{d}z, \tag{26d}$$

$$J = 2\int_a^b \langle uv\rangle \, \mathrm{d}z + 2\mu \int_a^b \langle uw\rangle \, \mathrm{d}z, \tag{26e}$$

$$D_1 = \int_a^b \langle |\nabla u|^2\rangle \, \mathrm{d}z, \tag{26f}$$

$$D_2 = \int_a^b \langle |\nabla v|^2 + |\nabla w|^2\rangle \, \mathrm{d}z = \int_a^b \langle (\boldsymbol{e}_x \cdot \mathrm{curl}\, \boldsymbol{v})^2\rangle \, \mathrm{d}z, \tag{26g}$$

with $a = 0$, $b = \infty$ for the atmospheric model, and $a = -\infty$, $b = 0$ for the oceanic model. The second form of D_2, arising by extracting a divergence with vanishing volume integral, is the more convenient. We then have

$$\frac{\mathrm{d}E_1}{\mathrm{d}t} = I_1 + \frac{1}{R}J - \frac{1}{R}D_1, \tag{27a}$$

$$\frac{\mathrm{d}E_2}{\mathrm{d}t} = I_2 - \frac{1}{R}J - \frac{1}{R}D_2. \tag{27b}$$

The terms on the right-hand side of (27) may be computed for any unstable mode, and we may then try to determine which processes are important to the instability.

In (27) I_1 and I_2 may be identified as energy-production terms, their sum being precisely the production rate of total perturbation kinetic energy; $R^{-1}D_1$ and $R^{-1}D_2$ are rates of dissipation of perturbation kinetic energy, and $R^{-1}J$ represents a transfer from perturbation kinetic energy of longitudinal motion to perturbation kinetic energy in the cross-plane. In the absence of this transfer term, the growth of E_2 would then be independent of the longitudinal perturbation-velocity component, and would derive solely from the production I_2; E_1 would derive solely from work done by the Reynolds stress $-\langle uw\rangle$, which exists by virtue of the growth of cross-plane perturbations. This represents the situation as $R \to \infty$. For finite R, the J-term couples cross-plane and longitudinal motions more intimately, and it is then possible for instabilities to owe their existence to growth of u, with cross-plane motions driven, at least in part, by transfer of energy from E_1 to E_2.

Table 6 lists a number of cases and the corresponding values of the integrals appearing in the energy budgets (27). In the case of the (three) unstable examples, all integrals are proportional to $\exp(2\sigma t)$, and the values shown are the constants of proportionality. The integrands are thought to be accurate to four significant digits, but there is a loss of accuracy in some cases due to truncation errors in the integration procedure (Simpson's rule with 200 points). The first three cases in the atmosphere correspond to onset in the atmospheric model at the respective values of μ, as do the cases 8 and 9 for the oceanic model. Thus cases 1 and 8 correspond to the critical conditions over all latitudes less than λ^*, case 2 corresponds to critical conditions at latitude 45°, cases 3 and 9 correspond to critical conditions for the traditional problem ($\lambda = 90°$), case 4 corresponds to onset of the type II instability in the traditional problem, while cases 5–7 are the most-unstable modes for $\lambda = 45°$.

Case	1	2	3	4	5	6	7	8	9
Type	Neutral	Neutral	Neutral	Neutral	Unstable	Unstable	Unstable	Neutral	Neutral
R	30.8	33.9	54.16	112.76	100.0	200.0	400.0	11.6	11.8
ϵ	$-9°$	$-18°$	$-23.3°$	$7.2°$	$-10°$	$8°$	$17°$	$-6°$	$-6°$
k	0.89	0.59	0.136	0.551	0.45	0.53	0.50	0.33	0.32
μ	2.03	1.0	0.0	0.0	1.0	1.0	1.0	0.156	0.0
E_1	2.329	1.570	7.040	3.505	2.446	3.733	3.960	0.6715	0.6597
I_1	0.3056	0.1175	0.02365	0.07898	0.09097	0.1241	0.1367	0.04905	0.04277
$-R^{-1}D_1$	0.2202	0.08887	0.01610	0.00760	0.04724	0.05157	0.03776	0.2704	0.02278
$-R^{-1}J$	0.08564	0.02873	0.00755	0.00138	0.01161	0.00468	0.00223	0.02206	0.01996
I_2	0.02651	0.01230	0.00174	0.01980	0.01243	0.02641	0.02919	0.00082	0.00032
$R^{-1}D_2$	0.1122	0.04106	0.00929	0.02118	0.01455	0.01479	0.00990	0.02293	0.02027
E_2	1.407	0.8745	0.4717	0.8515	0.7233	0.8965	0.8813	0.5348	0.5306

TABLE 6

In all cases shown $J < 0$, which implies that energy is transferred from the longitudinal motion and deposited in the cross-plane. This transfer is an essential factor in the instability mechanism at lower values of R. A disturbance can only be self-excited if vertical motions develop, but the rate of production of cross-plane perturbation kinetic energy is much smaller than the rate of its dissipation at onset, and it remains smaller for a range of supercritical R. For such cases, illustrated by cases 1–5, 8 and 9 in table 6, self-excitation depends upon production of longitudinal energy E_1, with feedback occurring by transfer of energy to vertical motions in the cross-plane through J, and then back to the production term I_1, which carries the primary role in maintaining the perturbation motion.

At higher values of R, the energy-transfer term is small compared with I_2, and self-excitation of cross-plane motions no longer depends upon production of longitudinal energy I_1. While production I_1 is greater than I_2 for all R, at the larger Reynolds numbers, this serves mainly to increase the longitudinal energy and to overcome longitudinal dissipation. The longitudinal motion is 'slaved', and the instability now depends upon I_2, since it is I_2 that is responsible for the vertical perturbations necessary for instability.

The eigenfunctions and the distributions of Reynolds stress, production, transfer and dissipation with depth are illustrated graphically in figures 27–31 for case 1. The eigenfunction χ for case 1 has a critical layer marked by a minimum in its modulus (figure 27) and a rapid change in its phase (figure 28) near $z = 2.2$. The Reynolds stress $-\langle uw\rangle$ responsible for production is, however, centred near $z = 1.2$, as seen in figure 29. The production itself is maximum (figure 30) at an even smaller value of z ($\simeq 1$), reflecting the weighting by shear U' against which the Reynolds stress works. This relation between the locations of the maxima of Reynolds stress and rate of working is typical. Figure 31 shows that the cross-plane dissipation rate exceeds the production at *all* elevations in this case, but that the transfer term is greater than dissipation rate over a substantial range of elevations. Since this is a neutral case, the sum of the areas under $-R^{-1}J$ and I_2 equals the area under $R^{-1}D_2$.

The eigenfunctions of the oceanic case near onset must satisfy a stress-free boundary condition at the surface $z = 0$, and the magnitude of the vertical motion is larger than that in the atmospheric case, exceeding the maximum modulus of χ, and the depth of penetration is larger, reflecting the smaller value of the Reynolds number. With these exceptions, however, the shapes of the eigenfunctions are similar in the two cases, with χ having a critical layer, and with peaks of the Reynolds stress $-\langle uw\rangle$, and I_1 in the same relationship to the critical layer and to each other

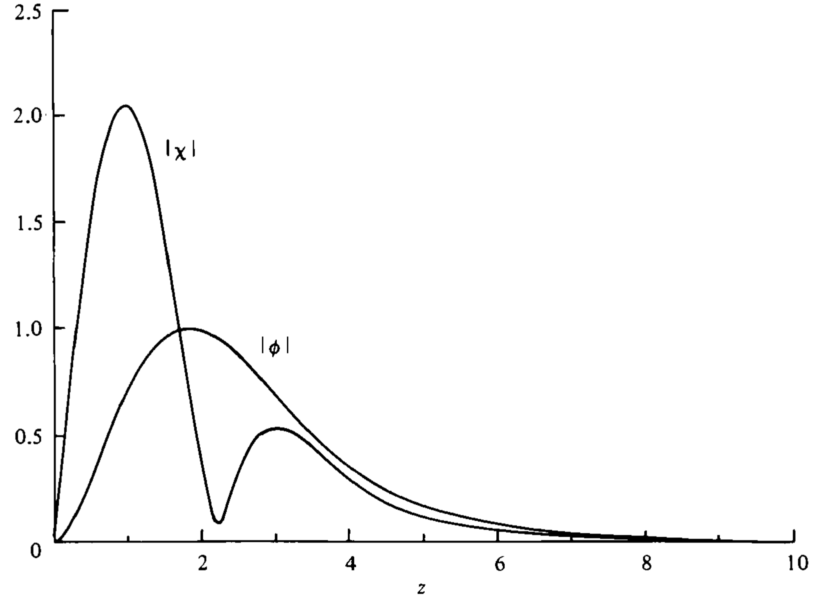

FIGURE 27. Modulus of the eigenfunctions ϕ and χ for the atmospheric model. $R = 30.8$, $\mu = 2.03$, $\epsilon = -9°$, $k = 0.89$.

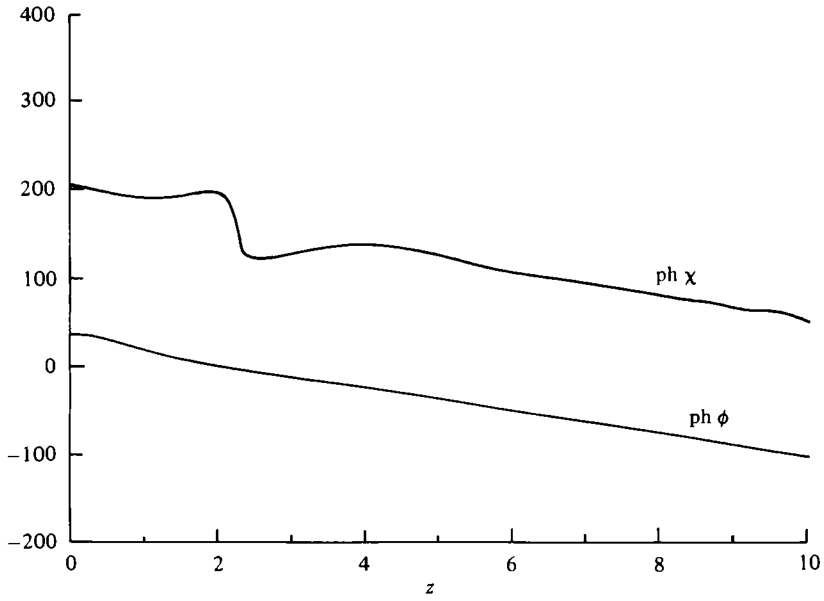

FIGURE 28. Phase of the eigenfunctions ϕ and χ for the atmospheric model. $R = 30.8$, $\mu = 2.03$, $\epsilon = -9°$, $k = 0.89$. The eigenfunctions in this figure and in figures 29–31 are normalized to yield $\max(|\phi|) = 1$.

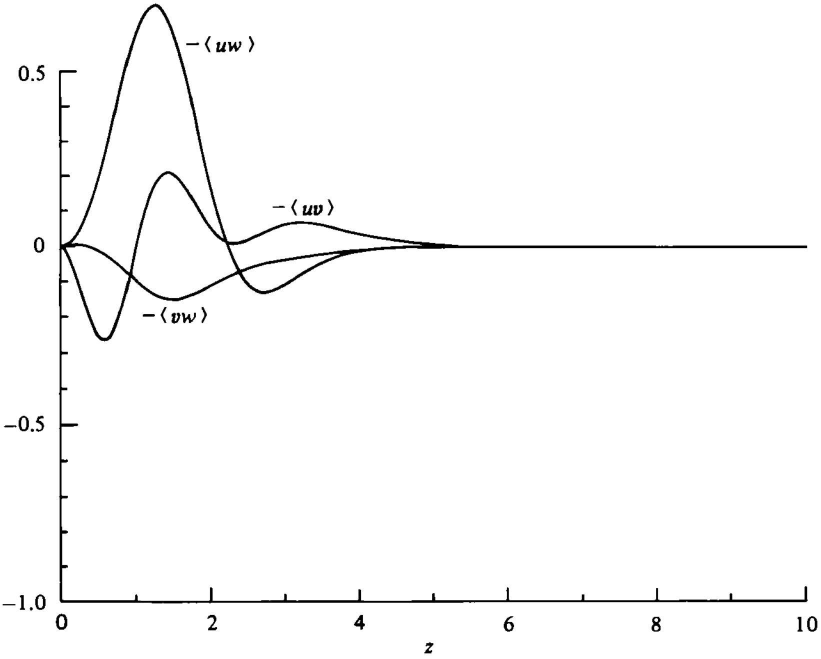

FIGURE 29. Reynolds stresses $-\langle uv \rangle$, $-\langle vw \rangle$, $-\langle uw \rangle$ as functions of z for the atmospheric model. $R = 30.8$, $\mu = 2.03$, $\epsilon = -9°$, $k = 0.89$.

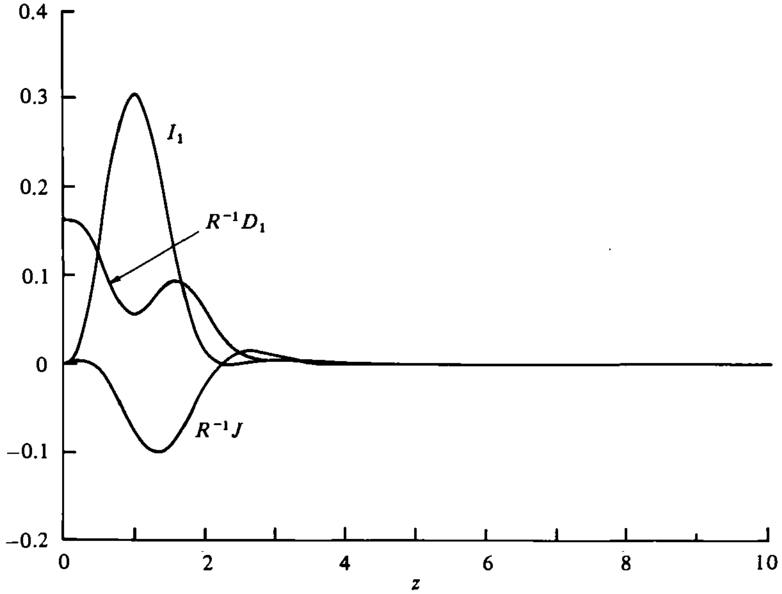

FIGURE 30. Production I_1, transfer $R^{-1}J$ and dissipation $R^{-1}D_1$ terms appearing in the budget of longitudinal kinetic energy, as functions of z for the atmospheric model. $R = 30.8$, $\mu = 2.03$, $\epsilon = -9°$, $k = 0.89$.

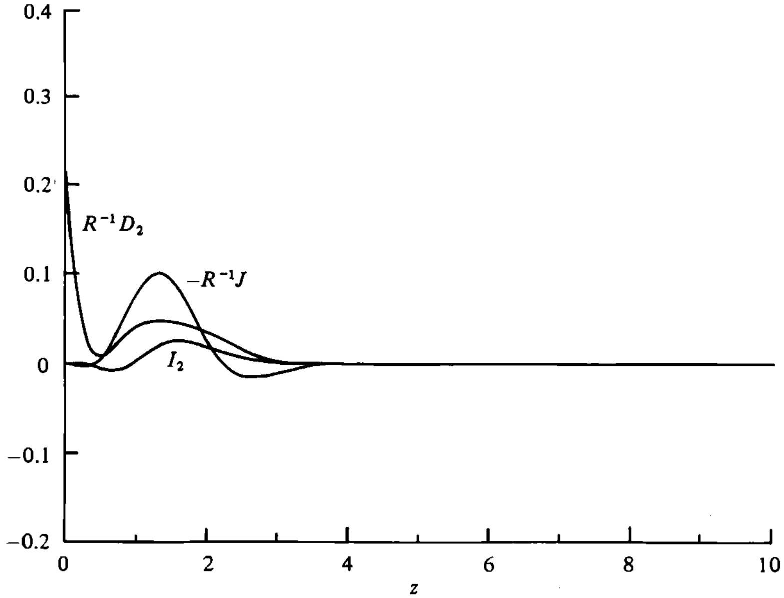

FIGURE 31. Production I_2, transfer $-R^{-1}J$ and dissipation $R^{-1}D_2$ terms appearing in the budget of transverse kinetic energy as functions of z for the atmospheric model. $R = 30.8$, $\mu = 2.03$, $\epsilon = -9°$, $k = 0.89$.

seen for the atmospheric case. A selection of plots of these distributions, as well as others for the atmospheric case, may be found in Lele (1985).

6. Summary

We have shown that the horizontal component of angular velocity is important to the stability of a planetary Ekman layer, acting as a destabilizing agent. It lowers the critical Reynolds number, significantly so for flows with a rigid boundary, less so for flows with a free boundary, where the critical Reynolds number is already very small. The coupling of the energy of longitudinal disturbances to the energy of cross-plane disturbances is enhanced by the horizontal component, thereby strengthening a feedback essential to the onset of instability in Ekman layers.

In addition to reducing the critical Reynolds number, growth rates under slightly supercritical conditions are larger by a significant factor, and the bands of unstable wavenumbers much wider, when the horizontal component of angular velocity is not small. In these cases, the slightly unstable system does not act as an efficient filter, but serves instead to amplify simultaneously a broad range of wavelengths. Thus waves having the greatest growth rates, and therefore preferred on linear grounds, may not be preferred by much under marginally supercritical conditions; mode selection, if the process indeed occurs, may be expected to be dominated by considerations other than linear stability characteristics.

We have also shown that at high Reynolds numbers (where the problem reduces to the traditional one treated at some length in the literature), the instability has features previously undetected. One such feature, that the lower neutral curves are dominated by viscosity for all Reynolds numbers, had previously been discovered by Gill & Davey (1969) for an analogous problem, and we have given an extended account for the Ekman layer.

Another new feature is the discovery of multiple inviscid modes of instability, perhaps an infinite sequence of them, in the limit $kR \rightarrow \infty$, the limit governed by the Rayleigh equation. In addition, we have located the most-unstable inviscid modes. This process is aided by our demonstration that the growth rates are periodic in ϵ with a period of π, and that the phase speeds of unstable modes satisfy $c_r(k, \epsilon) = -c_r(k, \epsilon - \pi)$.

We are concerned here with geophysical phenomena: these are invariably turbulent and usually affected by stable or unstable density stratification. It has been argued (see e.g. Faller 1965; Brown 1970, 1980) that coherent motions in the form qualitatively like the instabilities found here may be interpreted as the large eddies of the complete motion, and that they are the marginally unstable modes that may be extracted from a stability analysis of the turbulent mean flow. Such ideas of marginal instability appear to go back at least to Malkus (1956) and have been invoked by Malkus & Veronis (1958), Lessen (1978) and others. A rational method of identifying the 'big eddies' by a 'proper orthogonal decomposition' of a stochastic flow field has been advanced by Lumley (1967). In its predictive mode, Lumley (1967, 1981) has shown that, to a first approximation, the big eddies in his theory arise as eigenfunctions of a problem equivalent to the linear stability problem, and this equivalence suggests an underlying physical connection between the two conceptionally distinct approaches. In the implementations so far made of both marginal instability and (predictive) proper orthogonal decomposition, the effects of small eddies are subsumed by a (constant) eddy viscosity.

Our assumption of a constant eddy viscosity is made in the same spirit. There are two effects of this choice, with consequences of different import. We believe the use of a constant eddy viscosity for the perturbed motion is not as serious a step as its use in finding the basic Ekman flow. Observed Ekman layers deviate, often significantly, from Ekman's (1905) classical, constant eddy viscosity solution. The mean motion is determined not only by the small eddies, but also by the large ones, and (presumably on both counts) a constant eddy viscosity must be inadequate to capture the details of the mean profiles.

Despite the use of mean flows expected to differ from Ekman layers in the atmosphere or ocean, we do not expect qualitative differences between the stability characteristics found here and those obtaining for many of the mean flows that are observed to occur, so long as these flows are essentially barotropic. The effects of stratification, ignored here, are often of major importance; this is particularly true at high Reynolds numbers, where Coriolis accelerations may be neglected. At lower Reynolds numbers, traditional treatments of the stratified problem need reconsideration in light of the present findings. In any event, the relative importance of the horizontal and vertical Coriolis accelerations established here is independent of the representation of the turbulence or the stratification conditions.

This work was supported by the Physical Oceanography and Fluid Mechanics Programs of the National Science Foundation under Grants OCE8310624 and MEA8306713. We are indebted to Professor A. J. Faller for directing us to the works of Wippermann and Etling.

Appendix

The significant features of our numerical methods are outlined here; further details of the Galerkin algorithm may be found in Lele (1985). The Galerkin scheme and its convergence characteristics and accuracy are documented first, with particular reference to the difficulties experienced in trying to compute eigenvalues with small imaginary parts. In this connection the differences between eigenvalues computed from the direct and the adjoint problems is found to be a useful measure of the accuracy of the Galerkin approximation. Finally the shooting method we use in most of our inviscid calculations is outlined.

Galerkin scheme

Equation (11) defines the eigenvalue problem we wish to solve. The boundary conditions for the atmospheric model are given by (13) and by (14) for the oceanic model. In our calculations the boundary condition at infinity (or its asymptotic form) is applied at a finite distance z_{m}.

It is convenient to transform the range of the independent variable to a (0, 1)-range by introducing

$$\zeta = \frac{z}{z_{\mathrm{m}}} \quad \text{for the atmospheric problem,}$$

or

$$\zeta = \frac{z + z_{\mathrm{m}}}{z_{\mathrm{m}}} \quad \text{for the oceanic problem.}$$

In terms of the new independent variable ζ, the boundary conditions are in the same form as (13) or (14), but the condition at infinity is applied at $\zeta = 1$ for the atmospheric problem. For the oceanic problem $\zeta = 1$ represents the free surface, and the boundary condition at infinity is transposed to $\zeta = 0$. This latter condition is either applied directly, or, to improve accuracy, is replaced by approximations utilizing the known asymptotic behaviour at large $|z|$. We briefly outline the latter procedure for the atmospheric problem; the corresponding results for the oceanic problem can be obtained in a similar manner. As $z \rightarrow \infty$ the coefficients in (11) approach constant values, and the problem has solutions in the form of exponentials. We take these to be in the forms

$$\phi = \mathrm{e}^{-\lambda z} \phi_{\mathrm{a}}, \quad \chi = \mathrm{e}^{-\lambda z} \chi_{\mathrm{a}},$$

with ϕ_{a} and χ_{a} constants. The rate constant λ satisfies the sixth-order algebraic equation

$$(kR)^{-2} (\lambda^2 - k^2)^3 + 2(\mathrm{i}kR)^{-1} (V_\infty - c)(\lambda^2 - k^2)^2 - (V_\infty - c)^2 (\lambda^2 - k^2) + 4(kR)^{-2} (\lambda + \mathrm{i}k\mu)^2 = 0,$$

where $V_\infty = V(z)$ as $z \rightarrow \infty$. Roots of the above equation having a positive real part correspond to solutions that decay as $z \rightarrow \infty$. Of these roots the one with smallest real part ($\lambda_{\min}$) corresponds to the solution that decays most slowly. We construct our asymptotic boundary conditions using this most-slowly decaying solution in the following manner:

$$\left. \begin{aligned} \frac{\mathrm{d}\phi}{\mathrm{d}z} &= -\lambda_{\min} \phi, \\ \frac{\mathrm{d}\chi}{\mathrm{d}z} &= -\lambda_{\min} \chi \end{aligned} \right\} \quad \text{at } z = z_{\mathrm{m}}.$$

To solve for $\lambda_{\min}$ we need to know the eigenvalue c, and we use a previously generated approximation to c. Such a process requires an iteration with successively more accurate values of c. In the present problem the accuracy gained by using the asymptotic form of the boundary condition is only marginal provided that z_{m} is sufficiently large (10–15). Furthermore, since $\lambda_{\min}$ depends on c, k, μ and R, a new set of Galerkin basis functions (discussed below) is required for each set of the non-dimensional parameters. This makes the numerical implementation inefficient. Owing to these considerations we have used the asymptotic boundary conditions only with a limited set of parameter values.

The dependent variables ϕ and χ are expanded in a series of independent basis functions with unknown (complex) coefficients:

$$\phi(\zeta) = \sum_{i=1}^{N} a_i \hat{\phi}_i(\zeta), \tag{A 1}$$

$$\chi(\zeta) = \sum_{i=1}^{N} a_{N+i} \hat{\chi}_i(\zeta). \tag{A 2}$$

In our calculations the basis functions are combinations of Chebyshev polynomials chosen so that each basis function satisfies all the relevant boundary conditions. For example, the basis function $\hat{\phi}_m$ satisfying

$$\hat{\phi}_m(0) = \mathrm{D}\hat{\phi}_m(0) = 0 \quad \text{and} \quad \hat{\phi}_m(1) = \mathrm{D}^2\hat{\phi}_m(1) = 0$$

is given by

$$\hat{\phi}_m(\zeta) = \frac{m^2+m+1}{2(m+2)} T^*_{m+3}(\zeta) - T^*_{m+2}(\zeta) - \frac{(m+1)(m^2+2m+3)}{m(m+2)} T^*_{m+1}(\zeta) + T^*_m(\zeta) + \left(\frac{m^2+3m+3}{2m}\right) T^*_{m-1}(\zeta),$$

where T^*_m is the shifted Chebyshev polynomial of degree m.

The expansions (A 1) and (A 2) are substituted in the transformed forms of the differential equations (11). Since we always have a finite set of basis functions, the differential equations cannot be satisfied exactly by any choice of the unknown coefficients $a_1, a_2, \ldots, a_{2N}$. In a Galerkin method the coefficients a_i, $i = 1, 2, \ldots, 2N$, are determined by requiring the residuals to be orthogonal to each of the basis functions (Finlayson 1972). This procedure leads to a system of equations given by

$$\boldsymbol{A}\boldsymbol{a} = c\boldsymbol{B}\boldsymbol{a}, \tag{A 3}$$

where $\boldsymbol{a}$ is a vector with $a_1, a_2, \ldots, a_{2N}$ as its $2N$ components and $\boldsymbol{A}$, $\boldsymbol{B}$ are $2N \times 2N$ complex matrices whose elements depend parametrically on the non-dimensional parameters governing the problem, viz R, k, ϵ and μ. Equation (A 3) then presents an algebraic eigenvalue problem for the eigenvalue c. The eigenvalue problem (A 3) was solved by using a standard eigenvalue subroutine EIGZC from the International Mathematical and Statistical Library (IMSL); all calculations were performed in double-precision arithmetic.

For any set of non-dimensional parameters $\{R, k, \epsilon, \mu\}$ the eigenvalue problem (A 3) has $2N$ eigenvalues. These eigenvalues are sorted and numbered in decreasing order of their imaginary part. Let c_n represent the nth eigenvalue in such an ordering. For any specific choice of R, k, ϵ and μ the eigenvalues c_n depend on N (the size of the Galerkin approximation) and z_m (the dimension of the computational domain). Convergence of the eigenvalues c_n to the eigenvalues of the differential equations (11)

z_m	c_r	c_i
5	−0.349913	0.0320019
10	−0.353198	0.0496970
15	−0.353198	0.0498359
20	−0.353199	0.0498414

TABLE 7. Convergence study of Galerkin scheme for $R = 25$, $k = 0.4$, $\epsilon = 0$, $\mu = 0$. Here $N = 30$ and the domain of the computation is increased, eigenvalue c_i is displayed.

N	c_r	c_i
5	−0.287041	0.026867
7	−0.289669	0.0365115
9	−0.322624	0.0528470
10	−0.340251	0.0549264
15	−0.354345	0.0497111
20	−0.353008	0.0502229
30	−0.353198	0.0498413
40	−0.353198	0.0498372
50	−0.353197	0.0498366

TABLE 8. Convergence study of the Galerkin scheme for the case shown in table 7. Here N is increased for fixed z_m

can then be studied by varying both N and z_m. Tables 7 and 8 summarize the results of such a convergence study.

The example presented is for the oceanic problem with $R = 25$, $k = 0.4$, $\epsilon = 0$ and $\mu = 0$. Table 7 displays the variation of the eigenvalue c_i as z_m is increased from 5 to 20 with a fixed number of terms N. It is evident that z_m of 10–15 provides a good approximation to the most unstable eigenvalue. Table 8 displays the variation of the eigenvalue c_i as the size N is increased with z_m fixed at a value of 20. Evidently $N = 30$ provides adequate accuracy. This table also demonstrates that the real part of the eigenvalue converges more rapidly than its imaginary part, a feature found in all of our convergence studies on this scheme (and with our shooting method as well). We note that we found the accurate eigenvalues of the traditional Ekman-layer problem presented by Melander (1983) and of the plane Poiseuille flow by Orszag (1971) of great value in debugging our numerical codes.

Another measure of accuracy of the computed eigenvalues is provided by comparing the eigenvalues with those obtained by computing the eigenvalues of the adjoint system of differential equations. The exact spectra of the two problems are identical, and hence the difference between the two numerical approximants is a measure of the accuracy of the approximation. With N between 25–50 and z_m of 15 the numerical approximations to the largest eigenvalues for the two problems agree very well; they differ in the 5th significant digit for Reynolds number of 4000, and in the 4th significant digit at $R = 32000$. This agreement deteriorates much more rapidly if we attempt to compute those neutral modes that become singular in limit of $R \to \infty$. This feature is best illustrated by a direct computation of the inviscid problem. In figure 32 we present the results of such a calculation, together with results obtained from a shooting technique thought to be accurate to five significant digits. In the range of wavenumbers where the eigenvalue has a significant imaginary part, the results

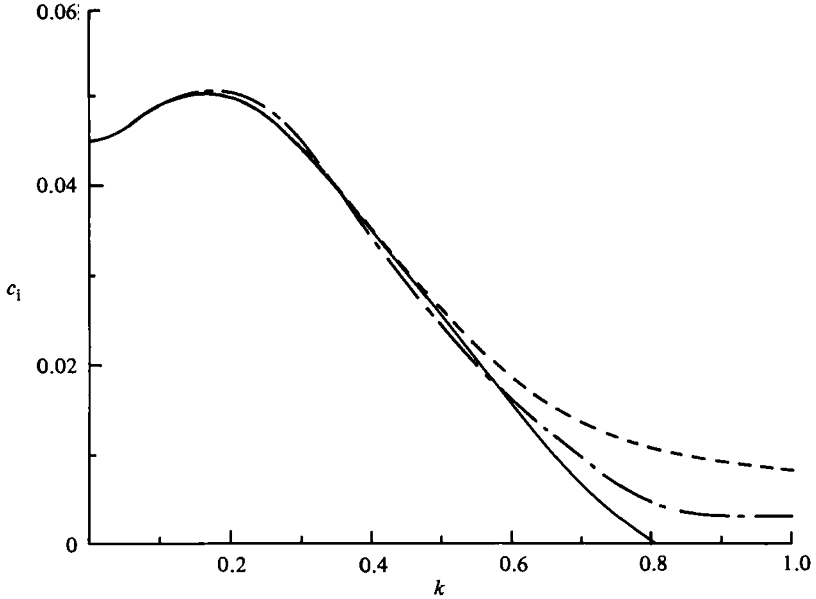

FIGURE 32. Imaginary part of the eigenvalue c_i as a function of k using different methods: ——, shooting method; ---, Galerkin method, direct problem $N = 70$, $z_m = 10$; –·–·, Galerkin method, adjoint problem, $N = 70$, $z_m = 10$.

of all the three methods agree. As the singular neutral mode is approached (by increasing k) the eigenvalues obtained from the direct and adjoint problems begin to deviate from one another and are grossly in error at $(k-k_c)/k_c \approx 0.2$, where k_c is the wavenumber corresponding to the singular neutral mode. We believe that such a loss of accuracy in the Galerkin scheme is due to a progressively slower convergence (in N) as a singular neutral mode is approached.

Such a deterioration of the convergence rate of the Galerkin scheme has unfortunate practical implications. A convergence criterion based on the difference of two Galerkin estimates, one with N and the other with $2N$ terms, will fail to signal the deterioration of the convergence rate (and a build-up of error) as a singular neutral mode is approached; this may well mislead one into believing that accurate eigenvalues have been found. In such cases the role played by the adjoint eigenvalue problem is critical. It not only can signal build-up of errors due to loss of convergence, but is also very useful in eliminating other spurious eigenvalues.

A convincing test of the accuracy of a scheme is to compare it with an independent scheme that is known to be reliable. In our work on the Rayleigh equation, we use a shooting technique in conjunction with the Galerkin method. This not only provides an independent check on the numerical results obtained from the Galerkin scheme, but it also allows the calculation of singular neutral modes. By analytically continuing the coefficients of the Rayleigh equation to the complex z-plane and by deforming the contour of integration around the singular points as indicated in the text, the shooting technique allows computation of singular neutral eigenvalues. Shooting methods have the disadvantage of requiring a good initial guess and converge to only one eigenvalue. In cases where multiple eigenvalues exist (as in the present problem) a shooting method is an inefficient means of conducting the search.

The shooting method employed is similar to that used by Gill & Davey (1969). The path of integration in the complex z-plane is chosen to be either the real z-axis or the contour

$$z = s - \beta \mathrm{i} s \frac{\mathrm{d}V}{\mathrm{d}z}(s),$$

where β is a positive constant. The latter path is needed when the singular neutral modes are approached. It is implemented (when needed) by changing the variable of integration from z to the real variable s, and by analytically continuing $\mathrm{d}V/\mathrm{d}z$ into the complex z-plane, evaluating it on the deformed path. Starting from an initial guess found from our Galerkin program, the differential equations are integrated using a 4th-order Runge–Kutta scheme using the asymptotic boundary condition at z_{m}, viz $\mathrm{d}\phi/\mathrm{d}z = -k\phi$ at $z = z_{\mathrm{m}}$. Typically z_{m} of 15 is adequate, although, as discussed in §4, larger values of z_{m} (up to 25) are required in order to compute the higher modes accurately.

REFERENCES

Asai, T. & Nakasuji, I. 1973 On the stability of Ekman boundary layer flow with thermally unstable stratification. *J. Meteor. Soc. Japan* **51**, 29–42.

Barcilon, V. 1965 Stability of a non-divergent Ekman layer. *Tellus* **17**, 53–68.

Brown, R. A. 1970 A secondary flow model for the planetary boundary layer. *J. Atmos. Sci.* **27**, 742–757.

Brown, R. A. 1972 On the inflection point instability of a stratified Ekman boundary layer. *J. Atmos. Sci.* **29**, 850–859.

Brown, R. A. 1980 Longitudinal instabilities and secondary flows in the planetary boundary layer: a review. *Rev. Geophys. Space Phys.* **18**, 683–697.

Caldwell, D. R. & Van Atta, C. W. 1970 Characteristics of Ekman boundary layer instabilities. *J. Fluid Mech.* **44**, 79–95.

Dudis, J. J. & Davis, S. H. 1971 Energy stability of the Ekman boundary layer. *J. Fluid Mech.* **47**, 405–413.

Eckart, C. 1960 *Hydrodynamics of Oceans and Atmospheres*. Pergamon.

Ekman, V. W. 1905 On the influence of the Earth's rotation on ocean currents. *Ark. Mat. Astr. Fys.* **2**, no. 11.

Etling, D. 1971 The stability of an Ekman boundary layer flow as influenced by the thermal stratification. *Beitr. Phys. Atmos.* **44**, 168–186.

Etling, D. & Wippermann, F. 1975 On the instability of a planetary boundary layer with Rossby-number similarity. *Boundary-Layer Meteor.* **9**, 341–360.

Faller, A. J. 1963 An experimental study of the instability of the laminar Ekman boundary layer. *J. Fluid Mech.* **15**, 560–576.

Faller, A. J. 1965 Large eddies in the atmospheric boundary layer and their possible role in the formation of cloud rows. *J. Atmos Sci.* **22**, 176–184.

Faller, A. J. & Kaylor, R. E. 1966*a* Investigations of stability and transition in rotating boundary layers. In *Dynamics of Fluids and Plasmas* (ed. S. I. Pai), pp. 309–329. Academic Press.

Faller, A. J. & Kaylor, R. E. 1966*b* A numerical study of the instability of the laminar Ekman boundary layer. *J. Atmos. Sci.* **23**, 466–480.

Faller, A. J. & Kaylor, R. E. 1967 Instability of the Ekman spiral with applications to the planetary boundary layer. *Phys. Fluids Suppl.* **10**, 212–219.

Finlayson, B. A. 1972 *The Method of Weighted Residuals and Variational Principles*. Academic.

Gill, A. E. & Davey, A. 1969 Instabilities of a buoyancy-driven system. *J. Fluid Mech.* **35**, 775–798.

Gregory, N., Stuart, J. T. & Walker, W. S. 1956 On the stability of three dimensional boundary layers with application to the flow due to a rotating disk. *Phil. Trans. R. Soc. Lond.* A **248**, 155–199.

Greenspan, H. P. 1968 *The Theory of Rotating Fluids*. Cambridge University Press.

HOLTON, J. R. 1979 *An Introduction to Dynamic Meteorology*, 2nd edn. Academic.

IOOSS, G., NIELSEN, H. B. & TRUE, H. 1978 Bifurcation of the stationary Ekman flow into a stable periodic flow. *Arch. Rat. Mech. Anal.* **68**, 227–256.

KAYLOR, R. & FALLER, A. J. 1972 Instability of the stratified Ekman boundary layer and the generation of internal waves. *J. Atm. Sci.* **29**, 497–509.

LEIBOVICH, S. & STEWARTSON, K. 1983 On the stability of columnar vortices. *J. Fluid Mech.* **126**, 335–366.

LESSEN, M. 1978 On power laws for turbulent jets, wakes and shearing layers and their relationship to the principle of marginal instability. *J. Fluid Mech.* **88**, 535–540.

LELE, S. K. 1985 Some problems of hydrodynamic stability arising in geophysical fluid dynamics. Ph.D. thesis, Cornell University.

LILLY, D. K. 1966 On the instability of Ekman boundary flow. *J. Atmos. Sci.* **23**, 481–494.

LIN, C. C. 1955 *The Theory of Hydrodynamic Stability*. Cambridge University Press.

LUMLEY, J. L. 1967 The structure of inhomogeneous turbulent flows. In *Atmospheric Turbulence and Radio Wave Propagation* (ed. A. M. Yaglom & V. I. Tatarsky), pp. 157–163. Nauka.

LUMLEY, J. L. 1981 Coherent structures in turbulence. In *Transition and Turbulence* (ed. R. E. Meyer). Academic.

MALKUS, W. V. R. 1956 Outline of a theory of turbulent shear flow. *J. Fluid Mech.* **1**, 521–539.

MALKUS, W. V. R. & VERONIS, G. 1958 Finite amplitude cellular convection. *J. Fluid Mech.* **4**, 225–260.

MELANDER, M. V. 1983 An algorithmic approach to the linear stability of the Ekman layer. *J. Fluid Mech.* **132**, 283–293.

MILES, J. W. 1961 On the stability of heterogeneous shear flow. *J. Fluid Mech.* **10**, 496–508.

NIELSON, H. B. & TRUE, H. 1979 Numerical solution of the nonlinear stability of an incompressible Ekman boundary layer. In *Proc. 6th Intl Conf. on Numerical Methods in Fluid Dynamics, 1978* (ed. H. Cabannes, M. Holt & V. V. Rusanor). Lecture Notes in Physics vol. 90. Springer.

ORSZAG, S. A. 1971 Accurate solution of the Orr–Sommerfeld stability equation. *J. Fluid Mech.* **50**, 689–703.

SPOONER, G. F. 1983 Stability of free surface Ekman layers. *J. Phys. Oceanogr.* **13**, 663–677.

SPOONER, G. F. & CRIMINALE, W. O. 1982 The evolution of disturbances in an Ekman boundary layer. *J. Fluid Mech.* **115**, 327–346.

STERN, M. E. 1960 Instability of Ekman flow at large Taylor number. *Tellus* **12**, 399–417.

TATRO, P. R. & MOLLO-CHRISTENSEN, E. L. 1967 Experiments on Ekman layer instability. *J. Fluid Mech.* **28**, 531–543.

WEBER, J. E. 1980 Symmetric instability of stratified geostrophic flow. *Tellus* **32**, 176–185.

WEIDMAN, P. D. 1976 On the spin-up and spin-down of a rotating fluid. Part 2. Measurements and stability. *J. Fluid Mech.* **77**, 709–735.

WIPPERMANN, F. 1969 The orientation of vortices due to instability of the Ekman-boundary layer. *Beitr. Phys. Atmos.* **48**, 30–45.

WIPPERMANN, F., ETLING, D. & KIRSTEIN, H. J. 1978 On the instability of a planetary boundary layer with Rossby number similarity. Part II. The combined effect of inflection point instability and convective instability. *Boundary-Layer Meteor.* **15**, 301–321.