J. Fluid Mech. (1983), *vol.* 130, *pp.* 279–298
Printed in Great Britain

Baroclinic instability over wavy topography

By R. A. DE SZOEKE
School of Oceanography, Oregon State University, Corvallis, OR 97331

(Received 27 July 1982 and in revised form 9 November 1982)

The occurrence of new unstable modes of quasigeostrophic baroclinic oscillation of rotating stratified shear flow over a wavy bottom is examined. To obtain a tractable mathematical problem, the bottom topography is considered as a perturbation modifying the oscillations. It is found that combinations of a top-intensified and a bottom-intensified Eady mode, each stable without topography, can be destabilized by topography if certain resonant conditions are met. These are that (i) the two modes possess the same wavefrequency, and (ii) topography possesses a wavenumber $\boldsymbol{c}$ bridging the gap between the wavenumbers of the modes $\boldsymbol{a}$, $\boldsymbol{b}$, i.e. $\boldsymbol{c} = \boldsymbol{a} - \boldsymbol{b}$. Growth rate of this instability (called type A) is proportional to the amplitude of the topographic component. There are two special cases: (i) when one of the basic modes is a marginally neutral mode – according to the classical analysis without topography – the instability is stronger (type M) with growth rate proportional to the $\frac{2}{3}$-power of topographic amplitude; (ii) when both modes are marginally neutral the instability is even stronger (type M^2) with growth rate proportional to the square root of topographic amplitude.

These topographic instabilities, like classical baroclinic instability, draw their energy from the available potential by transporting buoyancy down the mean gradient associated with the geostrophic shear flow.

1. Introduction

Baroclinic instability – discovered first by Charney (1947) and independently by Eady (1949) – is a potent mechanism for spontaneously generating velocity and density oscillations in a rotating stratified fluid by drawing on mean potential energy available from the mass distribution associated with a vertically sheared mean geostrophic flow. It is invoked in meteorology and physical oceanography alike to explain, for example, the occurrence of synoptic disturbances in the midlatitude Westerlies and the meanders or eddies of ocean currents like the Gulf Stream or the Antarctic Circumpolar Current. Because of the obvious practical importance of the phenomenon, a considerable literature (for reviews see Hide & Mason 1975; Pedlosky 1979) has been built up by authors using both analytic and numerical methods to study the linearized stability of a multitude of specific basic flows, to establish general conditions for stability, to extend the treatment to encompass nonlinear effects, and so on.

In this paper, I wish to focus attention on the effects of quasiperiodic ('wavy') topography on the simple baroclinic flow studied by Eady (1949): i.e. linear density stratification, and linear velocity profile balanced geostrophically by the slope of isopycnals, as in figure 1. Aspects of the topography question have been addressed before. Blumsack & Gierasch (1972) considered the effect of uniform cross-stream slope on baroclinic instability. De Szoeke (1975) considered the effects of general

weak-amplitude cross-stream topography. These studies were linearized analyses (as this paper is). The question of topography acquired added piquancy recently from the work of Charney & Straus (1980), who, in a layered nonlinear model of planetary flow over topography found multiple wavelike stationary equilibrium states, some strongly resembling the atmospheric situation of blocked flow. The 'energy of these states comes from the potential energy of the mean flow, not from kinetic energy transfer via the mountain torque' (Charney & Straus 1980), suggesting strongly a baroclinic-instability-like mechanism of origin. The major result to be established in this paper is that topography gives rise to new modes of instability which draw their energy from the available potential energy of the mean flow in the classical manner. These new unstable modes are constructed from the superposition, in a fixed ratio, of top-intensified and bottom-intensified modes of the same frequency, but differing wavenumbers that would be stable without topography. For instability, bottom topography must bridge the wavenumber gap between the modes. The heavily truncated spectral model of Charney & Straus (1980) cannot support enough wavenumber interactions to exhibit the kind of instability discussed in this paper, which therefore serves warning that a more elaborate spectral model than Charney & Straus's may sustain a qualitatively richer set of instabilities and equilibria.

This paper will be limited to consideration of quasigeostrophic baroclinic instability. The quasigeostrophic approximation, originally worked out by Charney (1947), assumes that the primary momentum balance of the fluctuations is geostrophic, and thereby greatly simplifies the vorticity balance, though not trivially, because it is derived by consistently retaining certain ageostrophic terms. For an excellent modern derivation see Pedlosky (1979).

There are certain restrictions on the kinds of topography that we can consider. First, as is well-known (Pedlosky 1979), topographic slope cannot exceed the Rossby number of the flow. Were this not so, the quasigeostrophic approximation would be violated because horizontal velocities impinging on such strong slopes would cause vortex stretching too intense to be balanced by increasing quasigeostrophic relative vorticity. Stated another way, treatment of strong topographic slopes requires abandonment of the quasigeostrophic approximation.

Beyond this, to obtain an analytically tractable mathematical problem we expand solutions in power series of ϵ†, a parameter typifying topographic slope *after* scaling by Rossby number. In this paper, we typically obtain and discuss only the first $O(\epsilon)$† corrections to the solutions for no topography, though the method of calculation for higher-order corrections will be obvious. In any case, uniform convergence for a finite range of ϵ can be proved in this type of problem (McIntyre 1970). Hence the first-order solutions we derive here are tentatively offered as initial thrusts at the problem of the effects of finite topography (in the sense of order-Rossby-number slope) on quasigeostrophic baroclinic instability.

Within these limits on amplitude, topography is represented by a sum over two-dimensional spatially periodic components. Though there does not seem to be any essential difficulty in the extension of these methods to cover nonperiodic topographies representable by continuous spatial Fourier integrals, I prefer, in order to fix ideas, to present the analysis for discrete sums representing periodic or quasiperiodic ('wavy') topography.

† Occasionally, expansions in fractional powers of ϵ are necessary.

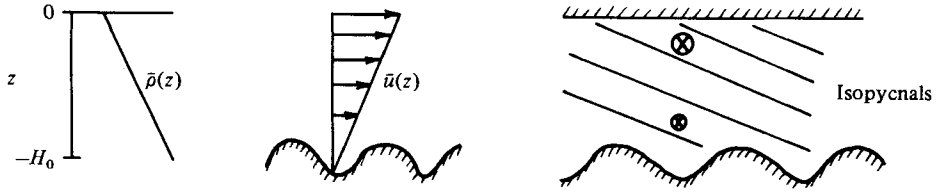

FIGURE 1. Schematic of basic mean flow.

2. Mathematical analysis

The specific mean flow and density field whose stability in the presence of topography is considered in this paper is the same as studied by Eady (1949) (figure 1), namely

$$\bar{u}(z) = \frac{u_0(z+H_0)}{H_0}, \quad \bar{\rho}(y,z) = \rho_0(\alpha y - \gamma z), \tag{1a,b}$$

where

$$\frac{fu_0}{H_0} = g\alpha \quad \text{(geostrophic balance)}, \tag{1c}$$

$$g\gamma = N^2 = \text{constant}; \tag{1d}$$

that is, linear density stratification, and linear shear balanced geostrophically by a linear slope of isopycnals. The mean flow at the average bottom depth $z = -H_0$ is taken to be zero, so that the mean flow by itself is unaffected by the wavy bottom. The horizontal extent of the region occupied by the flow is unbounded in both directions. The dynamical equations and boundary conditions governing perturbations are linearized about this basic mean flow. Because their derivation is familiar, they are merely stated here.

2.1. *Equations of motion*

Disturbances of the steady motion of a rotating stratified inviscid fluid, such as atmosphere or ocean, are governed by the potential-vorticity balance, which, subject to the quasigeostrophic and Boussinesq approximations, can be written (Pedlosky 1979)

$$(\partial_t + \bar{u}(z)\,\partial_x)\,[\partial_x^2 + \partial_y^2 + \partial_z f_0^2 N^{-2}\,\partial_z]\,\psi + Q_y \psi_x = 0, \tag{2}$$

where the symbols have the following meanings:

x, y, z eastward, northward, vertical coordinates,
t time,
$f = f_0 + \beta y$ Coriolis parameter,
$\beta = \dfrac{\mathrm{d}f}{\mathrm{d}y}$ planetary vorticity gradient,
$N(z) = \left(\dfrac{-g}{\rho_0}\dfrac{\partial\bar{\rho}}{\partial z}\right)^{\frac{1}{2}}$ Brunt–Väisälä frequency,
$\bar{u}(z)$ basic horizontal shear flow,
$Q_y = \beta - (f_0^2 N^{-2}\bar{u}_z)_z$ basic potential-vorticity gradient,
ψ disturbance 'stream function',
u, v, w components of disturbance velocity,
$p = f_0\psi$ disturbance pressure,
ρ_0 reference density,
$\bar{\rho}(y,z)$ mean density,
ρ disturbance density,
$b = \dfrac{-g\rho}{\rho_0}$ disturbance buoyancy.

It is a consequence of geostrophy and hydrostasy that

$$u = -\psi_y, \quad v = \psi_x, \quad b = f_0 \psi_z, \tag{3a,b,c}$$

while vertical velocity can be determined from the motion of isopycnals, i.e.

$$w = -f_0 N^{-2} [(\partial_t + \bar{u}\,\partial_x)\,\psi_z - \bar{u}_z \psi_x]. \tag{4}$$

The condition of vanishing vertical velocity at the top boundary can be written

$$(\partial_t + \bar{u}(z)\,\partial_x)\,\psi_z - \bar{u}_z \psi_x = 0 \quad (z = 0). \tag{5}$$

The condition of no normal flow at the bottom boundary $z = -H_0 + h(x, y)$ is

$$f_0 N^{-2} [(\partial_t + \bar{u}(z)\,\partial_x)\,\psi_z - \bar{u}_z \psi_x] = h_x \psi_y - h_y \psi_x \equiv J(h, \psi) \quad (z = -H_0 + h). \tag{6}$$

For the Eady model (figure 1), $\bar{u}$ and $\bar{\rho}$ are given by (1), N is constant, and β is neglected; hence $Q_y = 0$. Scaling the vertical coordinate (and depth variation) by H_0, horizontal coordinates by the baroclinic Rossby radius of deformation $\lambda = NH_0/f_0$, time by $NH_0/f_0 u_0$, the depth variation h by δH_0, the non-dimensional differential equation becomes

$$\nabla^2 \psi \equiv (\partial_x^2 + \partial_y^2 + \partial_z^2)\,\psi = 0, \tag{7}$$

while the boundary conditions become

$$(\partial_t + \partial_x)\,\psi_z - \psi_x = 0 \quad (z = 0), \tag{8}$$

$$(\partial_t + (1+z)\,\partial_x)\,\psi_z - \psi_x = J(h, \psi) \quad (z = -1 + \delta h). \tag{9}$$

To make the coefficient on the right of (9) unity, we had to choose $\delta = u_0/NH_0$. For typical oceanic parameters, $N = 10^{-3}$ s^{-1}, $H_0 = 5$ km, $u_0 = 50$ cm s^{-1}, we obtain $\delta = 0.1$. Accordingly, consideration should be formally limited to topographic variations $\lesssim O(10\,\%)$ of H_0, for consistency with this scaling. Note that $\delta = u_0/f_0\lambda$, which is a Rossby number based on λ. Its smallness is consistent with the quasigeostrophic assumptions that underpin (2) (Pedlosky 1979).

The method of solution will be to characterize variations in scaled topography by a small parameter ϵ,†

$$h = \epsilon h_1, \tag{10}$$

so that (9) becomes

$$(\partial_t + O(\epsilon\delta)\,\partial_x)\,\psi_z - \psi_x = \epsilon J(h_1, \psi) \quad (z = -1 + O(\epsilon\delta)). \tag{9'}$$

We may consistently neglect $O(\epsilon\delta)$ terms in preference to the $O(\epsilon)$ term – this entails imposing the bottom condition (9′) at the mean level $z = -1$. The crucial disappearance of the mean advective term in (9′) depends on the vanishing of the mean flow at the bottom, $\bar{u}(-1) = 0$.

2.2. *Power-series expansion in topography parameter*

Exploiting the smallness of topography ($\epsilon \lesssim 1$) amounts to perturbing the Eady problem. This parallels quite closely McIntyre's (1970) treatment of perturbations of the Eady problem by vertical and horizontal variations in the velocity profile of figure 1, and de Szoeke's (1975) treatment of weak cross-stream topography. We seek a solution as a power expansion in ϵ:

$$\psi = \psi_0 + \epsilon\psi_1 + \epsilon^2\psi_2 + \ldots. \tag{11}$$

† Hence the dimensional scale of topography variation is $\epsilon\delta H_0$. We shall anticipate that the perturbation formalism to be developed below is valid within a finite radius of convergence, $\epsilon < \epsilon_0$, where $\epsilon_0 = O(1)$. (See McIntyre (1970) for methods of estimating ϵ_0.)

It is apparent that the equations are separable in time, so ψ and its successive approximations $\psi_0, \psi_1, \ldots$ can be supposed to behave like $e^{-i\omega t}$ in time. Hence ∂_t in (8), (9′) can be replaced by $-i\omega$. These equations then pose an eigenvalue problem for ω. The allowable eigenvalues will depend in general on ϵ, i.e.

$$\omega = \omega_0 + \epsilon\omega_1 + \ldots. \tag{12}$$

2.3. *Lowest-order solution*

Substituting (10)–(12) into (7)–(9′), we obtain for the lowest-order approximation

$$\nabla^2\psi_0 \equiv (\partial_x^2 + \partial_y^2 + \partial_z^2)\,\psi_0 = 0, \tag{13}$$

$$B_0\psi_0 \equiv \psi_{0z} - (-i\omega_0 + \partial_x)^{-1}\psi_{0x} = 0 \quad (z = 0), \tag{14}$$

$$B_{-1}\psi_0 \equiv \psi_{0z} - (-i\omega_0)^{-1}\psi_{0x} = 0 \quad (z = -1). \tag{15}$$

The operators B_0, B_{-1} are written in this way so that the problem is self-adjoint, a useful property which will be exploited later.

Equations (13)–(15) state the familiar quasigeostrophic Eady (1949) problem, whose solution is well-known (see e.g. McIntyre 1970):

$$\psi_0 = e^{-i\omega t + i\boldsymbol{k}\cdot\boldsymbol{x}}\chi_0^k(z), \tag{16}$$

$$\boldsymbol{k} = (k_x, k_y), \quad k = (k_x^2 + k_y^2)^{\frac{1}{2}}, \tag{17}$$

$$\chi_0^k(z) = kc_0 \cosh k(z+1) - \sinh k(z+1), \tag{18}$$

$$\frac{\omega_0}{k_x} = c_0(k) = \tfrac{1}{2} \pm \tfrac{1}{2}[1 + 4k^{-2} - 4k^{-1}\coth k]^{\frac{1}{2}}. \tag{19}$$

The dispersion relation represented by (19) is most familiarly displayed in figure 2. Clearly seen are an unstable region for $k < k_M = 2.40$ – in which complex-conjugate pairs of values of $\omega_0 = k_x c_0$ are possible – and a stable region for $k > k_M$. In the latter, two neutral stable modes are possible, one of which – the top mode – is evanescent from the upper boundary in the limit of large wavenumber k, while the other – the bottom mode – is evanescent from the lower boundary. This evanescence is obscured for wavenumbers nearer k_M, though we shall retain the nomenclature, top mode or bottom mode, to distinguish them.

The dispersion relation (19) can also be thought of as specifying families of different vector wavenumbers $\boldsymbol{k}$ that give the same stable wave frequency ω_0 – a representation which will become useful shortly. These constant-frequency families are shown as curves on the $\boldsymbol{k}$-plane of figure 3. Each curve has two branches, one representing top modes (indicated by solid lines), the other representing bottom modes (dashed lines). The two branches join on the circle $k = k_M$. Because $c_0 = \frac{1}{2}$ at $k = k_M$, whatever the frequency, the branches join at $k_x = 2\omega_0$, for given ω_0. Because $\coth k \sim 1 + O(e^{-2k})$ for large k, it can be shown from (19) that

$$\frac{\omega_0}{k_x} \sim 1 - k^{-1}, k^{-1} \quad \text{as} \quad k \to \infty. \tag{20}$$

This means that $k_x \sim \omega_0$ for the branch corresponding to the top mode, while $k_x/k \sim \omega_0$ for the bottom mode. The latter relation shows that the bottom-mode branch, for given ω_0, heads to infinity along a line inclined at arcos ω_0 to the abscissa in figure 2. Clearly, this makes sense only for $\omega_0 \leqslant 1$. For $1 < \omega_0 < \frac{1}{2}k_M = 1.20$, the bottom-mode branch cannot approach infinity; instead, it intersects the k_x axis

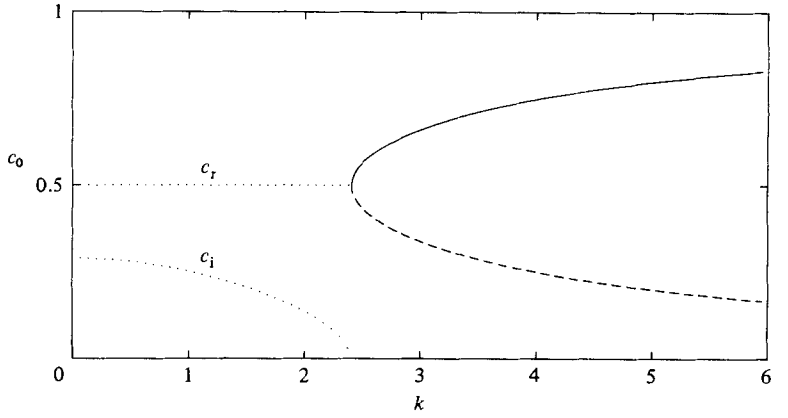

FIGURE 2. Dispersion relation $\omega_0/k_x = c_0(k) = c_r + \mathrm{i}c_i$ for Eady instability. Note real and imaginary parts for $k < k_M = 2.40$; stable top-intensified mode (solid) and table bottom-intensified mode (dashed) for $k > k_M$.

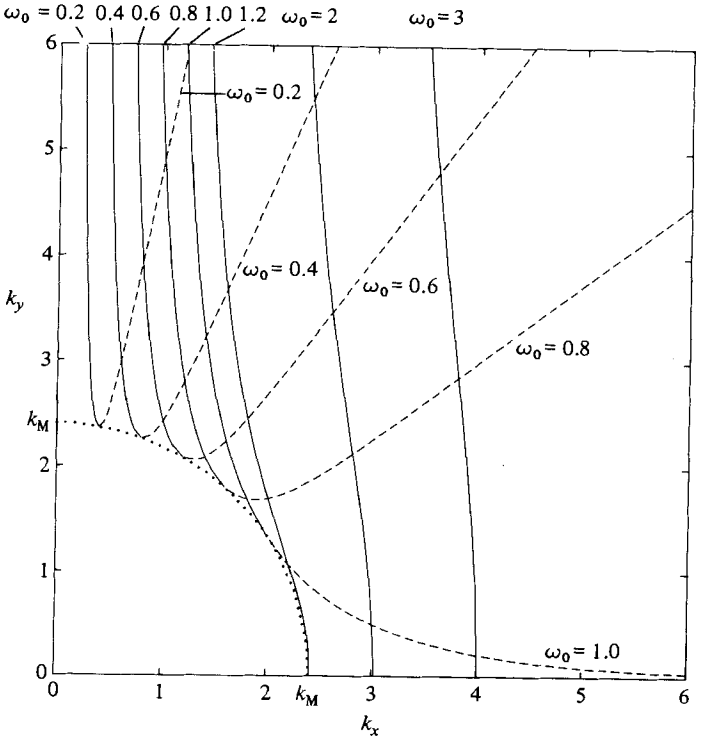

FIGURE 3. Stable modes of figure 2 replotted. Each curve represents the family of wavenumbers that give the same frequency ω_0. Solid (dashed) lines are top- (bottom-) intensified modes.

perpendicularly at some finite value of k_x. For $\omega_0 > \frac{1}{2}k_M$ there can be no bottom mode. Figure 3 displays only one quadrant of the $\boldsymbol{k}$-plane, in which all of k_x, k_y, ω_0 are positive. The $(k_x > 0, k_y < 0)$ quadrant is obtained by simply reflecting figure 3 in the k_x axis; still $\omega_0 > 0$. The remaining quadrants, in which $k_x <$ (are obtained by reflecting the first two in the k_y axis, though frequency should be taken negative, $\omega_0 < 0$, in these quadrants. The latter symmetry, that $\omega_0(\boldsymbol{k}) = -\omega_0(-\boldsymbol{k})$, permits the construction of real solutions from complex ones like (16) by addition of complex

conjugates. The former symmetry, that $\omega_0(k_x, k_y) = \omega_0(k_x, -k_y)$, permits the construction, from simple progressive plane waves, of solutions standing in the y-direction.

The following identity can be derived from (18) and (19) and will be found useful (McIntyre 1970):

$$\tfrac{1}{2}[(z+1-c_0)^{-2}\{\chi_0^k(z)\}^2]_{-1}^0 = k^4(c_0-\tfrac{1}{2})[k^2(1-c_0)^2-1]^{-1} \equiv \lambda. \tag{21}$$

Since $\coth k > 1$ $$[1+4k^{-2}-4k^{-1}\coth k]^{\frac{1}{2}} < 1-2k^{-1}, \tag{22}$$

as long as the left-hand side of the inequality is real, i.e. $k \geqslant k_{\mathrm{M}} = 2.40$. Using this in (19), we see that

$$k^{-1} < c_0^-(k) \leqslant c_0^+(k) < 1-k^{-1} \quad (k \geqslant k_{\mathrm{M}}). \tag{23}$$

Hence $$(1-c_0)^2k^2-1 > 0, \quad c_0^2k^2-1 > 0, \tag{24}$$

where either c_0^+ or c_0^- can be taken for c_0. From this it follows that the denominator in (21), the definition of λ, is always positive. Hence λ is positive or negative as $c_0 > \frac{1}{2}$ or $c_0 < \frac{1}{2}$, that is, for top modes or bottom modes respectively:

$$\operatorname{sgn}(\lambda) = \operatorname{sgn}(c_0-\tfrac{1}{2}). \tag{25}$$

This important fact will be used later.

2.4. *The first-order problem*

Next we show how to calculate ω_1, ψ_1, which contain the first effects of topography. The equations for ψ_1 are

$$\nabla^2\psi_1 = 0, \tag{26}$$

$$\mathrm{B}_0\psi_1 = \mathrm{i}\omega_1(-\mathrm{i}\omega_0+\partial_x)^{-1}\psi_{0z} \quad (z=0), \tag{27}$$

$$\mathrm{B}_{-1}\psi_1 = \mathrm{i}\omega_1(-\mathrm{i}\omega_0)^{-1}\psi_{0z}+(-\mathrm{i}\omega_0)^{-1}J(h_1,\psi_0) \quad (z=-1). \tag{28}$$

The operators $\mathrm{B}_0, \mathrm{B}_{-1}, \nabla^2$ are defined in (13)–(15).

Multiplying (26) by ψ_0^* and integrating by parts with respect to x, y and z over a cylinder of cross-section A extending from top to bottom boundaries, we obtain

$$\iint[\psi_0^*\psi_{1z}-\psi_{0z}^*\psi_1]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y = 0. \tag{29}$$

Substitution of (14) and (15) gives

$$\iint[\psi_0^*\{(-\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{1x}\}-\{(\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{0x}^*\}\psi_1]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y + \iint[\psi_0^*(\mathrm{B}\psi_1)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y = 0, \tag{30}$$

where $\mathrm{B} = \mathrm{B}_0, \mathrm{B}_{-1}$ at $z = 0, -1$ respectively. Writing

$$\psi_0^* = (\mathrm{i}\omega_0+(z+1)\partial_x)\,(\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_0^*,$$

we see from integrating by parts that

$$\iint[\psi_0^*\{(-\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{1x}\}]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y = \iint[\{(\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{0x}^*\}\psi_1]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y. \tag{31}$$

Hence (30) becomes

$$\iint[\psi_0^*(\mathrm{B}\psi_1)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y = 0. \tag{32}$$

The integral in (32) as $A \to \infty$ is formally divergent, but the limit of $A^{-1}\iint_A(\ldots)\mathrm{d}x\,\mathrm{d}y$ has a strict meaning. Using (27) and (28) to express $\mathrm{B}\psi_1$ at $z = 0, -1$ in terms of ψ_0,

given by (16)–(19), we can rewrite (32) as

$$\frac{2\lambda\omega_1}{k_x}-\sigma_0=0, \tag{33}$$

where λ is given by (21), while

$$\sigma_0=(-\mathrm{i}\omega_0)^{-1}A^{-1}\iint_A\{\psi_0^* J(h_1,\psi_0)\}_{-1}\,\mathrm{d}x\,\mathrm{d}y. \tag{34}$$

For $\psi_0\sim\exp[\mathrm{i}(\boldsymbol{k}\cdot\boldsymbol{x}-\omega t)]$ and sinusoidal h_1, it is readily seen than $\sigma_0=0$, so that

$$\omega_1=0. \tag{35}$$

In this case a solution of (26)–(28) for ψ_1 is easily found. If $h_1=h_\Delta\cos(\boldsymbol{c}\cdot\boldsymbol{x})$, it is comprised of wavenumbers $\boldsymbol{k}\pm\boldsymbol{c}$. As it has little interest, we shall not trouble to write it down. Clearly this method can be extended to any topography that can be represented by a Fourier sum or a Fourier integral over wavenumber components. Successive higher corrections ψ_2, ψ_3 etc. can be found in the obvious manner (McIntyre 1970).

2.5. *Wavenumber degeneracy; type-A instability*

The above procedure needs to be modified when the ground state ψ_0 is *degenerate*. By degeneracy we mean that there exist two downstream vector wavenumbers $\boldsymbol{a},\boldsymbol{b}$ with the same zeroth-order frequency:

$$\omega_0(\boldsymbol{a})=\omega_0(\boldsymbol{b}),\quad\text{i.e. } a_x c_0(a)=b_x c_0(b). \tag{36}$$

Figure 3 illustrates that there exist whole families of wavenumbers for which this may occur. The reason for only considering pairwise, or binary, degeneracies will become apparent. In this case the wavefunction can be written as a sum

$$\psi_0=\psi_0^a+\gamma\psi_0^b, \tag{37}$$

where γ is an arbitrary constant, and ψ_0^a,ψ_0^b are given by (16)–(19) with $\boldsymbol{k}=\boldsymbol{a},\boldsymbol{b}$. Formally then, the equations for ψ_1 are identical with (26)–(28), but since there are two independent ground states there are two orthogonality conditions like (29), each leading in turn to a condition like (32):

$$\iint[\psi_0^{a*}(\mathrm{B}\psi_1)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y=0,\quad\iint[\psi_0^{b*}(\mathrm{B}\psi_1)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y=0, \tag{38a,b}$$

where $\mathrm{B}_0\psi_1,\mathrm{B}_{-1}\psi_1$ are given by (27), (28). We notice that

$$\iint[\psi_0^{a*}(-\mathrm{i}\omega_0+(z+1)\,\partial_x)^{-2}\psi_{0x}^b]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y=\iint[\psi_0^{b*}(-\mathrm{i}\omega_0+(z+1)\,\partial_x)^{-2}\psi_{0x}^a)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y$$
$$=0, \tag{39}$$

because $\boldsymbol{a}\neq\boldsymbol{b}$, and

$$\iint\psi_0^{a*}J(h_1,\psi_0^a)\,\mathrm{d}x\,\mathrm{d}y=\iint\psi_0^{b*}J(h_1,\psi_0^b)\,\mathrm{d}x\,\mathrm{d}y=0, \tag{40}$$

similar to (34) and its sequel. Hence (38) can be written

$$\frac{2\lambda_a\omega_1}{a_x}-\gamma\sigma_{ab}=0,\quad\frac{2\gamma\lambda_b\omega_1}{b_x}-\sigma_{ba}=0, \tag{41a,b}$$

where λ_a,λ_b are given by expressions analogous to (21) for λ with $\boldsymbol{a},\boldsymbol{b}$ replacing $\boldsymbol{k}$, and

$$\sigma_{ab}=(-\mathrm{i}\omega_0)^{-1}A^{-1}\iint_A\psi_0^{a*}J(h_1,\psi_0^b)|_{-1}\,\mathrm{d}x\,\mathrm{d}y. \tag{42}$$

From a similar expression for σ_{ba}, it can be shown that $\sigma_{ba} = \sigma_{ab}^*$. If h_1 contains a component $h_\Delta \cos(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{x} = \frac{1}{2}h_\Delta \exp[\mathrm{i}(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{x}] + \text{c.c.}$, then

$$\sigma_{ab} = -\mathrm{i}\omega_0^{-1}\tfrac{1}{2}h_\Delta a c_a b c_b (a_x b_y - a_y b_x), \tag{43}$$

using (16) and (18). Otherwise $\sigma_{ab} = 0$. Hence (41a,b) can be solved for ω_1, γ:

$$\omega_1 = \pm\Omega_1, \quad \Omega_1 = |\sigma_{ab}| a_x b_x (4\lambda_a\lambda_b)^{-\frac{1}{2}}, \tag{44}$$

$$\gamma = \pm\Gamma, \quad \Gamma = \frac{|\sigma_{ab}|}{\sigma_{ab}}\left(\frac{b_x}{a_x}\frac{\lambda_a}{\lambda_b}\right)^{\frac{1}{2}}. \tag{45}$$

So topography can split the degeneracy in the ground state, determining shifted frequencies $\omega_0 \pm \epsilon\Omega_1$ and corresponding linear combinations $\psi_0^a \pm \Gamma\psi_0^b$ of the unperturbed degenerate states.

If ω_1 can be complex, the disturbance ψ can grow like $\exp[\mathrm{Im}(\epsilon\omega_1 t)]$. This can only happen when

$$\lambda_a\lambda_b < 0, \tag{46}$$

that is, λ_a and λ_b must have opposite signs. Recalling the formula (21) for λ, we see that this implies

$$c_a = c_0(a) > \tfrac{1}{2}, \quad c_b = c_0(b) < \tfrac{1}{2}, \tag{47}$$

or, in terms of figure 3, that wavenumbers $\boldsymbol{a}, \boldsymbol{b}$ correspond to top and bottom modes respectively. If $\boldsymbol{a}, \boldsymbol{b}$ are *both* top (or bottom) modes then ω_1 is purely real and instability is impossible. To avert a trivial result σ_{ab} should be non-zero; that is, topography should possess a Fourier component at the bridging wavenumber $\boldsymbol{a}-\boldsymbol{b}$. The growth rate $\mathrm{Im}(\epsilon\omega_1)$ is proportional to the amplitude of this topography component. If (44) is examined to determine the conditions for the largest growth rate, it is seen that ω_1, which can be thought of as depending on ω_0 and the manifold of conceivable $\boldsymbol{a}, \boldsymbol{b}$, is unbounded. It becomes infinite when λ_a (or λ_b) vanishes, which happens according to (21) when c_a (or c_b) $= \frac{1}{2}$, that is, when a (or b) $= k_{\mathrm{M}} = 2.40$. This situation requires special consideration, which we take up in §3.

The necessity for a top-intensified mode to interact with a bottom-intensified mode and topography to produce instability follows from the requirement (46). Physically, the λ-factors in this condition are linked to the meridional diffusion of potential vorticity by

$$-\int \overline{v'q'}\,\mathrm{d}z = \frac{\mathrm{d}}{\mathrm{d}t}\left[\tfrac{1}{2}\bar{u}_z \overline{\frac{\psi}{(\bar{u}-c)^2}}\right]_{-1}^{0} \sim \omega_{\mathrm{i}}\lambda \tag{48}$$

(Bretherton 1966; Pedlosky 1979). Hence the condition (46) requires that the meridional potential-vorticity fluxes of the participating modes be in opposite senses. This can only be accomplished by the combination of a top mode (northward flux) and a bottom mode (southward flux). These fluxes, though offsetting, do not balance precisely. There is a third contribution to potential-vorticity flux $\overline{v'_{-1}h}$ due to the bottom topography itself. All three contributions must sum to zero to satisfy Bretherton's (1966) theorem of no net meridional potential-vorticity flux.

The members of a given constant-frequency family of wavenumbers need only be considered in pairs. This is because at lowest order only a binary degeneracy can be countenanced by a wave–topography–wave interaction.

We call these instabilities type A, of which we immediately distinguish two subtypes, A_1 and A_2. Type A_1 occurs in the situation where $\boldsymbol{a}, \boldsymbol{b}$ belong to the same quadrant of the wavenumber plane; type A_2 occurs when one of $\boldsymbol{a}, \boldsymbol{b}$ belongs to one quadrant and the other to the quadrant reflecting the first in the k_x axis. Figure 4 illustrates the two situations schematically for $\omega_0 > 0$.

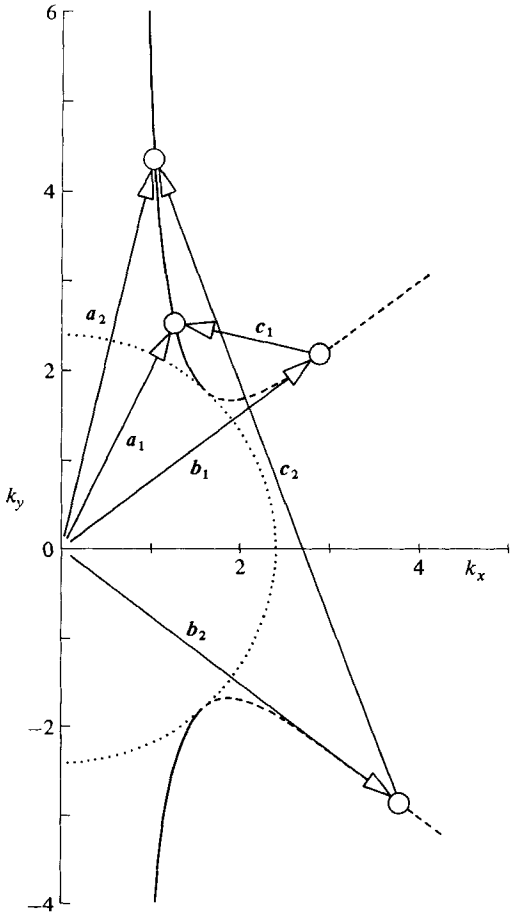

FIGURE 4. Schematic illustrating resonance conditions for type-A topographic instability. $\boldsymbol{a}, \boldsymbol{b}$ are top- and bottom-mode wavenumbers of the same frequency; $\boldsymbol{c}$ is the topographic wavenumber, necessary to trip an instability. Subscripts 1, 2 denote examples of type-A_1 or type-A_2 instabilities.

For a given frequency ω_0, there is a two-dimensional manifold of topographic wavenumbers $\boldsymbol{c}$ that can satisfy the resonance condition (36). Taking the degrees of freedom together, there is a three-dimensional manifold of type-A instabilities.

2.6. *The first-order wavefunction correction*

Having determined ω_1 and γ, we now show how to calculate ψ_1 from (26)–(28) in the degenerate situation. Substitution of (37), (16) and (18) into the right of (27) gives

$$B_0\psi_1 = \frac{\omega_1}{a_x}(1-c_a)^{-2}(ac_a\cosh a - \sinh a)\,\mathrm{e}^{\mathrm{i}(\boldsymbol{a}\cdot\boldsymbol{x}-\omega t)} + \gamma\frac{\omega_1}{b_x}(1-c_b)^{-2}(bc_b\cosh b - \sinh b)\,\mathrm{e}^{\mathrm{i}(\boldsymbol{b}\cdot\boldsymbol{x}-\omega t)} \quad (49)$$

at $z = 0$. The other boundary condition (28) is redundant, having been used in the determination of ω_1 and γ. A little algebra using the identity (21) shows that a particular integral of (26) and (49) is

$$\psi_{1\mathrm{p}} = \frac{\omega_1}{a_x}\chi_1^a(z)\,\mathrm{e}^{\mathrm{i}(\boldsymbol{a}\cdot\boldsymbol{x}-\omega t)} + \gamma\frac{\omega_1}{b_x}\chi_1^b(z)\,\mathrm{e}^{\mathrm{i}(\boldsymbol{b}\cdot\boldsymbol{x}-\omega t)}, \quad (50)$$

where $$\chi_1^a(z) = a(a^2c_a^2-1)\,[a^2(1-c_a)^2-1]^{-1}\cosh a(z+1), \tag{51}$$

with a similar expression for $\chi_1^b(z)$. This contribution automatically balances inhomogeneous terms with wavenumbers $\boldsymbol{a}, \boldsymbol{b}$ on the right of (28), as is guaranteed by the solvability conditions (38a, b). There are further terms contributing to ψ_1, caused by other topographic components interacting with ψ_0. These terms have horizontal wavenumbers other than $\boldsymbol{a}, \boldsymbol{b}$. Their calculation is quite straightforward given a full specification of h_1, but their effect is inconsequential for our purposes. Further, any arbitrary multiple of $\psi_0 = \psi_0^a + \gamma\psi_0^b$ can be added to ψ_1, but this too is inconsequential and can be seen to cause merely a renormalization of ψ_0 by a factor $1+O(\epsilon)$ (McIntyre 1970).

2.7. *Topographic modifications of the classical baroclinic instability*

What happens to the classical baroclinic instabilities when they are modified by topography? That is, what happens to the zeroth-order modes (16)–(19) when ω_0 and $\chi_0^k(z)$ are complex for $k < k_M$? Though the perturbation treatment given above is correct only for real ω_0, the necessary modifications are slight. While the problem posed by (13)–(15) is self-adjoint for ω_0 real, it is not so for ω_0 complex. The consequence of this is that while the adjoint solution of (13)–(15) is ψ_0^* (equation (16)) the adjoint solution when ω_0 is complex is

$$\tilde{\psi}_0 = \chi_0^k(z)\mathrm{e}^{-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}+\mathrm{i}\omega^* t}, \tag{52}$$

which must be used instead of ψ_0^* in (29) and its sequels above. The perturbation analysis of a monochromatic solution like (16) again gives $\omega_1 = 0$ and an undistinguished form for ψ_1.

The case of wavenumber degeneracy (equation (36)) is again special. But now ω_0 is complex, so the real and imaginary parts of both sides of (36) must be equal. Equality of real parts quickly implies, from (19), that $a_x = b_x$. The equality of imaginary parts demands that $a = b$, as can be verified graphically on figure 2. The only non-trivial solution of these two equations is $a_y = -b_y$, that is,

$$\boldsymbol{b} \text{ is the reflection of } \boldsymbol{a} \text{ in the } x\text{-axis.} \tag{53}$$

Hence the only degeneracy possible when ω_0 is complex is a kind of type-A_2 degeneracy.

In case of such a degeneracy, the lowest-order solution must be thought of as a bimodal like (37). The coefficient γ and wave-frequency correction ω_1 are determined by a procedure analogous to the above. The result is a set of equations just like (41a, b):

$$\frac{2\gamma\lambda\omega_1}{a_x} - \sigma_{ba} = 0, \quad \frac{2\lambda\omega_1}{a_x} - \gamma\sigma_{ab} = 0 \tag{54a, b}$$

((53) implies that $\lambda_a = \lambda_b = \lambda$, by (21)). Given a topography component

$$h_\Delta \cos{(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{x}} = h_\Delta \cos 2a_y y,$$

these further suggest that $$\sigma_{ab} = \sigma_{ba}^* = \mathrm{i}h_\Delta a^2 c_a a_y, \tag{55}$$

upon simplication of (43). Hence the solution of (54a, b) is

$$\gamma = \pm\mathrm{i}, \quad \omega_1 = \frac{\pm h_\Delta a^2 a_x a_y c_a}{2\lambda}. \tag{56}$$

While ω_1 is in general complex, γ is particularly simple. It means that the lowest-order wavefunction is

$$\psi_0 = \psi_0^a + \gamma\psi_0^b = 2\,\mathrm{e}^{\pm\frac{1}{4}\mathrm{i}\pi}\chi_0^a(z)\cos{(a_y y \mp \tfrac{1}{4}\pi)}\,\mathrm{e}^{\mathrm{i}a_x x-\mathrm{i}\omega t}. \tag{57}$$

Given cross-stream topography of wavenumber $2a_y$, any Eady instability with half that wavenumber is split into two modes standing in the y-direction with slightly different $O(\epsilon)$ wave frequencies and growth rates, one shifted $+45°$ in phase from the topography, the other shifted $-45°$.

Topography with downstream variation, so that no topography–wave–wave triplet $\boldsymbol{c}, \boldsymbol{a}, \boldsymbol{b}$ (where $\boldsymbol{b} = \boldsymbol{a} - \boldsymbol{c}$) can fulfil (53), modifies the unstable Eady mode quite unremarkably.

3. Degeneracies involving marginally neutral modes

The analysis of §2 breaks down, or tends to, when either c_a or c_b is $\frac{1}{2}$ or close to it. This happens when a (or b) approaches $k_M = 2.40$, the marginally neutral wavenumber below which classical baroclinic instabilities occur in the absence of topography. In this section an amended analysis is presented to cover these interesting cases, and stronger versions of type-A instability emerge with growth rates of order $\epsilon^{\frac{2}{3}}$ or $\epsilon^{\frac{1}{2}}$ according to whether $a = k_M$ only, or whether both $a = b = k_M$.

If $a = k_M$, then $\lambda_a = 0$ and (41a, b) have no solution. This situation is similar to those encountered by McIntyre (1970) and de Szoeke (1975), where perturbations around marginally neutral states were considered. The resolution of the difficulty lies partially in abandoning the power-series representation in ϵ, (11), (12), of the perturbation solution. The first correction to the zeroth-order solution, it turns out, is larger than $O(\epsilon)$. Instead of (11) and (12) we suppose that the solution of (7)–(9′) can be written

$$\psi = \psi_0 + \psi', \tag{58}$$

$$\omega = \omega_0 + \omega', \tag{59}$$

where ψ', ω' tend to zero as $\epsilon \to 0$, though not necessarily linearly, and

$$\psi_0 = \psi_0^a + \gamma\psi_0^b, \tag{60}$$

$$\omega_0(\boldsymbol{a}) = \omega_0(\boldsymbol{b}), \tag{61}$$

as before, but $a \approx k_M$, so that $c_a = \omega_0(\boldsymbol{a})/a_x \approx \frac{1}{2}$.

Substitution of (58) and (59) into (7)–(9′) gives

$$\nabla^2\psi' = 0, \tag{62}$$

$$\mathrm{B}_0\psi' = \mathrm{i}\omega'(-\mathrm{i}\omega_0 + \partial_x)^{-1}\psi_z \quad (z = 0), \tag{63}$$

$$\mathrm{B}_{-1}\psi' = \mathrm{i}\omega'(-\mathrm{i}\omega_0)^{-1}\psi_z + (-\mathrm{i}\omega_0)^{-1}J(h, \psi) \quad (z = -1). \tag{64}$$

The obvious iterative solution for ψ' fails as $a \to k_M$, for the reasons noted. However, McIntyre (1970) observed that

$$\psi_{\mathrm{I}}^a \equiv \frac{4\omega'}{a_x}(a^2 - 4)^{-1}\psi_{0z}^a \tag{65}$$

is an exact particular integral of (62)–(64), balancing the inhomogeneous terms proportional to $\omega'\psi_{0z}^a$ in (63) and (64) when $a = k_M$ and $c_a = \omega_0/a_x = \frac{1}{2}$ exactly. Accordingly we write

$$\psi' = \psi_{\mathrm{I}}^a + \gamma\psi_{\mathrm{I}}^b + \psi'' \tag{66}$$

when $a \approx k_M$, including the $\gamma\psi_{\mathrm{I}}^b$ term for symmetry. Substituting (66) into (62)–(64), we obtain – after a little manipulation exploiting the properties of ψ_0^a and ψ_0^b, namely that they satisfy (14) and (15)

$$\nabla^2\psi'' = 0, \tag{67}$$

$$\mathrm{B}_0\psi'' = -\mathrm{B}_0(\psi_{\mathrm{I}}^a+\gamma\psi_{\mathrm{I}}^b)+\mathrm{i}\omega'(-\mathrm{i}\omega_0+\partial_x)^{-1}\partial_z(\psi_0^a+\gamma\psi_0^b+\psi_{\mathrm{I}}^a+\gamma\psi_{\mathrm{I}}^b) \quad (z=0), \tag{68}$$

$$\mathrm{B}_{-1}\psi'' = -\mathrm{B}_{-1}(\psi_{\mathrm{I}}^a+\gamma\psi_{\mathrm{I}}^b)+\mathrm{i}\omega'(-\mathrm{i}\omega_0)^{-1}\partial_z(\psi_0^a+\gamma\psi_0^b+\psi_{\mathrm{I}}^a+\gamma\psi_{\mathrm{I}}^b)$$
$$+(-\mathrm{i}\omega_0)^{-1}J(h,\psi_0^a+\gamma\psi_0^b) \quad (z=-1). \tag{69}$$

Note the retention in these equations of contributions to ψ from the approximate particular integrals $\psi_{\mathrm{I}}^a, \psi_{\mathrm{I}}^b$, which are of order $\omega' \ll 1$. In the topographic term, however, only the lowest-order approximation to ψ is retained. Like the higher-order terms in §2, ψ'' must satisfy conditions of orthogonality to ψ_0^a, ψ_0^b that lead to compatibility equations just like (38a, b) with ψ'' replacing ψ_1. Substituting (68) and (69) into these, we encounter terms like

$$\iint[\psi_0^{a*}(\mathrm{B}\psi_{\mathrm{I}}^a)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y; \tag{70}$$

substitution from (65) for ψ_{I}^a, use of the definitions (14) and (15) of the operators B_0 and B_{-1}, and application of (19) and (21) shows that this term vanishes identically, as will a similar term involving $\psi_0^b, \psi_{\mathrm{I}}^b$. The cross terms like $\int\int[\psi_0^{a*}(\mathrm{B}\psi_{\mathrm{I}}^b)]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y$ also vanish, because they are spatial averages of fluctuating quantities like $\exp[\mathrm{i}(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{x}]$. Hence the first group of terms involving ψ_{I}^a and ψ_{I}^b in (68) and (69) make no contribution to the compatibility conditions. From the next group of terms we obtain

$$\mathrm{i}\omega' A^{-1}\iint_A[\psi_0^{a*}(-\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{0z}^a]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y = 2\lambda_a\omega'/a_x, \tag{71}$$

like the term in (41a). As before (equation (39)), the cross-terms vanish identically. The important new contributions to the compatibility equations are terms like

$$\mathrm{i}\omega' A^{-1}\iint[\psi_0^{a*}(-\mathrm{i}\omega_0+(z+1)\partial_x)^{-1}\psi_{\mathrm{I}z}^a]_{-1}^0\,\mathrm{d}x\,\mathrm{d}y$$
$$=\frac{4a^4}{a^2-4}\frac{1-a^2c_a(1-c_a)}{1-a^2(1-c_a)^2}\left(\frac{\omega'}{a_x}\right)^2 \equiv \mu_a\left(\frac{\omega'}{a_x}\right)^2 \tag{72}$$

from (65), (16)–(19) and (21). For $a = k_{\mathrm{M}}$, $c_a = \frac{1}{2}$, $\mu_a = \mu_{\mathrm{M}} = 75.45$.

Finally, topography produces non-vanishing cross-terms like $\epsilon\gamma\sigma_{ab}$, (42), provided that it contains spectral components at wavenumber $\boldsymbol{a}-\boldsymbol{b}$. Hence the compatibility equations lead to

$$2\lambda_a\frac{\omega'}{a_x}+\mu_a\left(\frac{\omega'}{a_x}\right)^2-\epsilon\gamma\sigma_{ab}=0, \tag{73a}$$

$$2\gamma\lambda_b\frac{\omega'}{b_x}+\gamma\mu_b\left(\frac{\omega'}{b_x}\right)^2-\epsilon\sigma_{ab}^*=0. \tag{73b}$$

3.1. *Degeneracy between a marginally neutral mode and an ordinary mode; type-M instability*

Suppose that $a = k_{\mathrm{M}}$, so that $\lambda_a = 0$, and neglect the second term in (73b); then γ can be eliminated to give

$$\frac{2\mu_a\lambda_b\omega'^3}{a_x^2 b_x} = \epsilon^2|\sigma_{ab}|^2. \tag{74}$$

The solutions for ω' are

$$\omega' = \left(\frac{a_x^2 b_x}{2\mu_a\lambda_b}\right)^{\frac{1}{3}}|\sigma_{ab}|^{\frac{2}{3}}\epsilon^{\frac{2}{3}}\{1, \tfrac{1}{2}\pm\mathrm{i}\tfrac{1}{2}\sqrt{3}\} \tag{75}$$

The corresponding values of γ are

$$\gamma = \left(\frac{b_x}{2a_x}\right)^{\frac{2}{3}}\mu_a^{\frac{1}{3}}\lambda_b^{-\frac{2}{3}}\left(\frac{|\sigma_{ab}|^{\frac{4}{3}}}{\sigma_{ab}}\right)\epsilon^{\frac{1}{3}}\{1, -\tfrac{1}{2}\pm\mathrm{i}\tfrac{1}{2}\sqrt{3}\}. \tag{76}$$

The second term in (73b) is $O(\epsilon^2)$, the neglect of which in comparison to $\epsilon\sigma^*_{ab}$ is thereby justified. The solutions (75) correspond to a neutral mode and a complex-conjugate pair of amplifying and decaying modes.

If a is not exactly k_{M} it is easy to see that $\lambda_a \sim (a-k_{\mathrm{M}})^{\frac{1}{2}}$, so that the neglect of the first term in (73a) is conditional on

$$a-k_{\mathrm{M}} \lesssim O(\epsilon^{\frac{4}{3}}). \tag{77}$$

In the range

$$O(\epsilon^{\frac{4}{3}}) \lesssim a-k_{\mathrm{M}} \lesssim 1. \tag{78}$$

The first term of (73a) supplants the second in importance as ω' changes from $O(\epsilon^{\frac{2}{3}})$ to $O(\epsilon)$. In the latter case the solutions of (73) reduce smoothly to $\pm\epsilon\Omega_1$, given by (44). The instabilities arising from this analysis, with $O(\epsilon^{\frac{2}{3}})$ growth rates (equation (75)), we shall call type-M disturbances, signifying their association with the marginal neutral point of the non-topographic stability analysis.

Considering the one-dimensional manifold of topographic wavenumber components that can support this instability together with the range of frequency, there is a two-dimensional manifold of such instabilities, disregarding the small freedom of the 'neutral' mode to be within $O(\epsilon^{\frac{4}{3}})$ of strict neutrality.

3.2. *Degeneracy between two marginally neutral modes; type-M^2 instability*

Can the second terms in both of (73a, b) be influential when a and b are both close to k_{M}? If $a \approx b \approx k_{\mathrm{M}}$, then $\omega_0/a_x \approx \omega_0/b_x \approx \frac{1}{2}$, so that $a_x \approx b_x$. Then either (i) $a_y \approx b_y$ or (ii) $a_y \approx -b_y$. We consider these cases in turn. In either case

$$\lambda_a \approx \alpha(a-k_{\mathrm{M}})^{\frac{1}{2}}, \quad \lambda_b \approx -\alpha(b-k_{\mathrm{M}})^{\frac{1}{2}}, \tag{79}$$

where $\alpha = 16.06$.

(i) Suppose $a_y \approx b_y$. Then $\boldsymbol{a} \approx \boldsymbol{b}$. If topography contains a low-wavenumber component $\frac{1}{2}h_\Delta \exp[\mathrm{i}(\boldsymbol{a}-\boldsymbol{b})\cdot\boldsymbol{x}]$ then (42) gives

$$\begin{aligned}\sigma_{ab} &= -\mathrm{i}\omega_0^{-1}\tfrac{1}{2}h_\Delta a c_a b c_b |\boldsymbol{a}\times\boldsymbol{b}| \\ &\approx -\tfrac{1}{8}\mathrm{i}\omega_0^{-1}h_\Delta k_{\mathrm{M}}^3 |\boldsymbol{a}-\boldsymbol{b}| \sin\theta,\end{aligned} \tag{80}$$

where θ is the angle between $\boldsymbol{a}-\boldsymbol{b}$ and $\boldsymbol{b}$ (or $\boldsymbol{a}$). Hence (73a, b) become approximately

$$2\alpha(a-k_{\mathrm{M}})^{\frac{1}{2}}\frac{\omega'}{a_x} + \mu_{\mathrm{M}}\left(\frac{\omega'}{a_x}\right)^2 + \tfrac{1}{4}\mathrm{i}\epsilon\gamma\frac{h_\Delta}{\omega_0}k_{\mathrm{M}}^3|\boldsymbol{a}-\boldsymbol{b}|\sin\theta = 0, \tag{81a}$$

$$-2\gamma\alpha(b-k_{\mathrm{M}})^{\frac{1}{2}}\frac{\omega'}{b_x} + \gamma\mu_{\mathrm{M}}\left(\frac{\omega'}{b_x}\right)^2 - \tfrac{1}{4}\mathrm{i}\epsilon\frac{h_\Delta}{\omega_0}k_{\mathrm{M}}^3|\boldsymbol{a}-\boldsymbol{b}|\sin\theta = 0. \tag{81b}$$

It is clear that there is a complex-conjugate pair of solutions

$$\omega' = O(\epsilon|\boldsymbol{a}-\boldsymbol{b}|^{\frac{1}{2}}), \quad \gamma = O(1) \tag{82}$$

for which the second terms of (81) are negligible. These solutions are merely the continuation of (44) and (45) into the region where $\boldsymbol{a} \approx \boldsymbol{b}$.

(ii) On the other hand, let us consider $a_y = -b_y$ and $a = b = k_{\mathrm{M}}$. (We take the exact equalities, so that $\lambda_a = \lambda_b = 0$ exactly. The approximate equalities are easily taken into account.) Then the necessary topography is oriented across the stream with wavenumber $\boldsymbol{a}-\boldsymbol{b} = 2a_y\hat{\boldsymbol{y}}$, and (73$a, b$) become

$$\mu_{\mathrm{M}}\left(\frac{\omega'}{2\omega_0}\right)^2 - \gamma\epsilon\sigma_{ab} = 0, \tag{83a}$$

$$\gamma\mu_{\mathrm{M}}\left(\frac{\omega'}{2\omega_0}\right)^2 - \epsilon\sigma^*_{ab} = 0. \tag{83b}$$

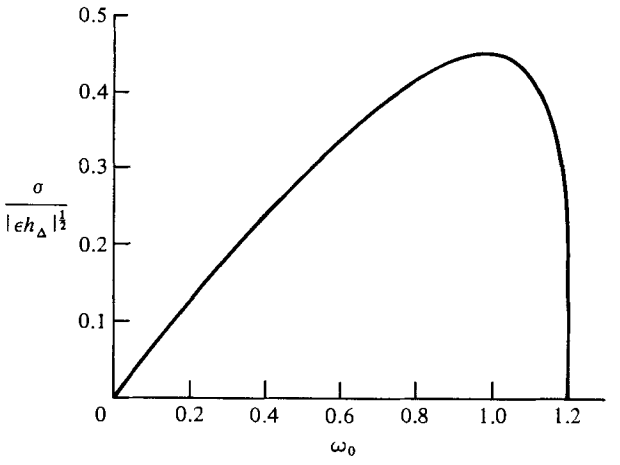

FIGURE 5. Growth rate of type-M² instability.

Eliminating γ, we obtain

$$\mu_{\mathrm{M}}^2\left(\frac{\omega'}{2\omega_0}\right)^4 = \epsilon^2|\sigma_{ab}|^2. \tag{84}$$

the solutions of which are

$$\omega' = 2\omega_0\left(\frac{\epsilon|\sigma_{ab}|}{\mu_{\mathrm{M}}}\right)^{\frac{1}{2}}\{\pm 1, \pm \mathrm{i}\}. \tag{85}$$

The corresponding values of γ are

$$\gamma = \frac{|\sigma_{ab}|}{\sigma_{ab}}\{+1, -1\}, \tag{86}$$

whose magnitude is in any case 1. One of the solutions (85) gives unstable growth. The growth rate is $O(\epsilon^{\frac{1}{2}})$, the largest order of magnitude for any of the topographic instabilities we have discovered. We call this double neutral mode a type-M² instability. It can easily be checked from (73) and (79) that (85) and (86) are good approximations as long as

$$a - k_{\mathrm{M}} \sim b - k_{\mathrm{M}} \lesssim O(\epsilon^{\frac{1}{2}}). \tag{87}$$

Since, for $a = b = k_{\mathrm{M}}$, $c_a = c_b = \frac{1}{2}$, $a_x = b_x = 2\omega_0$, $a_y = -b_y = (k_{\mathrm{M}}^2 - 4\omega_0^2)^{\frac{1}{2}}$, the expression (43) for σ_{ab} can be simplified. We obtain

$$\sigma_{ab} = -\mathrm{i}h_\Delta k_{\mathrm{M}}^2(\tfrac{1}{4}k_{\mathrm{M}}^2 - \omega_0^2)^{\frac{1}{2}}. \tag{88}$$

The unstable member of (85) is then

$$\omega' = 2\mathrm{i}(\epsilon|h_\Delta|)^{\frac{1}{2}}\mu_{\mathrm{M}}^{-\frac{1}{2}}k_{\mathrm{M}}\omega_0(\tfrac{1}{4}k_{\mathrm{M}}^2 - \omega_0^2)^{\frac{1}{4}} \equiv \mathrm{i}\sigma. \tag{89}$$

The growth rate $\sigma(\epsilon|h_\Delta|)^{-\frac{1}{2}}$ is shown in figure 5 as a function of ω_0. It attains its largest value of $0.44k_{\mathrm{M}}^{\frac{5}{2}}\mu_{\mathrm{M}}^{-\frac{1}{2}} = 0.45$ at $\omega_0 = k_{\mathrm{M}}/6^{\frac{1}{2}} = 0.98$, corresponding to $a_x = b_x = 1.96$, $a_y = -b_y = 1.39$, all in the non-dimensional units of the scaling of §2. As only frequency ω_0 can be varied, there is a one-dimensional manifold of type-M² instabilities.

4. Energetics; buoyancy flux

Like classical baroclinic instabilities, the new topographic instabilities draw their energy from the available potential energy, i.e., from the eddy flux of buoyancy down the geostrophically balanced buoyancy gradient. The way in which the bimodal unstable disturbances are organized in order to accomplish the necessary buoyancy flux is studied in this section.

By integrating the identity

$$\boldsymbol{\nabla}\cdot[\psi(\partial_t+\bar{u}\,\partial_x)\,\boldsymbol{\nabla}\psi] = \psi(\partial_t+\bar{u}\,\partial_x)\,\nabla^2\psi+(\partial_t+\bar{u}\,\partial_x)\tfrac{1}{2}(\boldsymbol{\nabla}\psi)^2+\bar{u}_z\partial_x(\psi\psi_z)-\bar{u}_z\psi_x\psi_z \tag{90}$$

(where $\bar{u}=z+1$, $\bar{u}_z=1$) over a volume V formed by a cylinder of cross-section A extending from top to bottom boundary, and using (7)–(9′), one obtains

$$\iint_A [\psi\psi_x - z\,J(h,\psi)]^0_{-1}\,\mathrm{d}x\,\mathrm{d}y = \partial_t\iiint_V \tfrac{1}{2}(\boldsymbol{\nabla}\psi)^2\,\mathrm{d}V+\iiint_V \partial_x[\tfrac{1}{2}\bar{u}(\boldsymbol{\nabla}\psi)^2+\bar{u}_x\psi\psi_z]\,\mathrm{d}V-\iiint_V \bar{u}_z\psi_x\psi_z\,\mathrm{d}V. \tag{91}$$

Divide both sides by A. Since we can write

$$\psi J(h,\psi) = \partial_x(h\psi\psi_y)-\partial_y(h\psi\psi_x), \tag{92}$$

the integrands on the left, as well as in the middle integral on the right, are composed of terms of the form $\partial_x(\ldots)$, $\partial_y(\ldots)$. If ψ and its derivatives are bounded, these integrals approach 0 as $A\to\infty$. This leaves

$$\partial_t\int_{-1}^{0}\tfrac{1}{2}\overline{(\boldsymbol{\nabla}\psi)^2}\,\mathrm{d}z = \int_{-1}^{0}\bar{u}_z\overline{\psi_x\psi_z}\,\mathrm{d}z, \tag{93}$$

where the bars denote horizontal averages. This equation expresses the well-known integral energy balance principle for baroclinic disturbances: the kinetic plus potential energies of the fluctuations

$$\tfrac{1}{2}\overline{(\boldsymbol{\nabla}\psi)^2} = \tfrac{1}{2}(\overline{u^2}+\overline{v^2}+\overline{b^2}), \tag{94}$$

in order to grow, must be fuelled by a 'meridional' eddy buoyancy flux

$$\overline{\psi_x\psi_z} = \overline{vb} \tag{95}$$

down the mean gradient $\bar{b}_y=-\bar{u}_z=-1$. A more general form of (93) can be obtained by using (2) instead of (7) with quite general $\bar{u}(y,z)$† and $N(z)$ (see Pedlosky 1979); in any case, the two-dimensional bottom topography $h(x,y)$ entering through the bottom condition (9′) does not alter the simplicity of (93).

It is instructive to examine how the topographic instabilities accomplish the necessary buoyancy flux. We shall examine in turn the three types of instability that we have distinguished.

4.1. *Classical baroclinic instability*

For comparison, we first calculate the well-known form of the buoyancy flux for the classical non-topographic baroclinic instability, which occurs for wavenumber magnitudes $k<k_{\mathrm{M}}$ when $\omega=k_x c_0=\omega_{\mathrm{r}}+\mathrm{i}\omega_{\mathrm{i}}$ is complex (equation (19)). The wavefunction is

$$\psi = \chi_0^k(z)\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}-\mathrm{i}\omega t}+\text{c.c.}, \tag{96}$$

† Except that $\bar{u}(y,-1)=0$ is required.

where $\chi_0^k(z)$ is given by (18), and c.c. denotes, here and in the following, the complex conjugate of all terms to its left. Hence

$$\overline{\psi_x \psi_z} = \mathrm{i}k_x(\chi_0^k \chi_{0z}^{k*} - \chi_0^{k*} \chi_{0z}^k)\, \mathrm{e}^{2\omega_i t}. \tag{97}$$

It is a quite general consequence for the baroclinic instability of flows with vanishing internal potential vorticity gradient that the associated buoyancy flux is independent of height (Pedlosky 1979). Hence (97) need only be evaluated at (say) $z = -1$:

$$\overline{\psi_x \psi_z} = 2\omega_i k^2\, \mathrm{e}^{2\omega_i t}. \tag{98}$$

The existence of a non-zero buoyancy flux is tied to the complexity of $\chi_0^k(z)$; for downgradient flux, the phase of the stream function must increase with height.

4.2. *Type-A instability*

How does the type-A instability contrive to transport the buoyancy necessary for its growth? To calculate the buoyancy flux we need the stream function correct to $O(\epsilon)$:

$$\psi = \left[\chi_0^a(z) + \epsilon \frac{\omega_1}{a_x} \chi_1^a(z)\right] \mathrm{e}^{\mathrm{i}(\boldsymbol{a}\cdot\boldsymbol{x} - \omega t)} + \gamma \left[\chi_0^b(z) + \epsilon \frac{\omega_1}{b_x} \chi_1^b(z)\right] \mathrm{e}^{\mathrm{i}(\boldsymbol{b}\cdot\boldsymbol{x} - \omega t)} + \text{c.c.}, \tag{99}$$

from (11), (37), (18), (50 and (51). Note that, to $O(\epsilon)$,

$$\omega = \omega_0 + \epsilon\omega_1, \tag{100}$$

where ω_0 is real (equation (19)), and ω_1 is imaginary (equation (44)). The contributions from each of the wavenumbers $\boldsymbol{a}$, $\boldsymbol{b}$ in (99) to $\overline{\psi_x \psi_z}$ are identical in form; that from wavenumber $\boldsymbol{a}$ is

$$2\epsilon|\omega_1|\, \chi_0^a \chi_{1z}^a \exp(2\epsilon|\omega_1| t). \tag{101}$$

Again, general considerations imply that this quantity is independent of z – a fact which direct calculation confirms. In fact, (101) can be written

$$2\epsilon|\omega_1|\, a^2(a^2 c_a^2 - 1)\,[a^2(1 - c_a)^2 - 1]^{-1} \exp(2\epsilon|\omega_1| t). \tag{102}$$

As stated, there is a similar term, multiplied by $|\gamma|^2$, with b substituted for a. There is no $O(1)$ counterpart to (97), because χ_0^a, χ_0^b are real. The $O(\epsilon)$ terms in (99) add complex structure at wavenumbers $\boldsymbol{a}$ and $\boldsymbol{b}$ to the wavefunction, and so induce an $O(\epsilon)$ phase variation with height at each wavenumber, considered separately.

4.3. *Type-M instability*

We calculate the buoyancy flux by type-M instabilities for the case where $a = k_{\mathrm{M}}$ exactly, though the formula we obtain is accurate as long as condition (77) is fulfilled. The wavefunction can be written

$$\psi = \chi_0^{\mathrm{M}}(z)\, \mathrm{e}^{\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{x} - \mathrm{i}\omega t} + \gamma \chi_0^{\beta}(z)\, \mathrm{e}^{\mathrm{i}\boldsymbol{b}\cdot\boldsymbol{x} - \mathrm{i}\omega t} + 4\frac{\omega'}{a_x}(k_{\mathrm{M}}^2 - 4)^{-1} \chi_{0z}^{\mathrm{M}} \mathrm{e}^{\mathrm{i}\boldsymbol{a}\cdot\boldsymbol{x} - \mathrm{i}\omega t} + \text{c.c.} \tag{103}$$

correct to $O(\epsilon)$, where ω' and γ are given by (75) and (76) and are of order $\epsilon^{\frac{3}{4}}, \epsilon^{\frac{1}{4}}$ respectively; and

$$\chi_0^{\mathrm{M}}(z) = k_{\mathrm{M}} c_{\mathrm{M}} \cosh k(z+1) - \sinh k_{\mathrm{M}}(z+1) = (\tfrac{1}{4}k_{\mathrm{M}}^2 - 1)^{-1} \cosh k_{\mathrm{M}}(z + \tfrac{1}{2}), \tag{104}$$

as can be shown from (18) using (19) and (21). Hence

$$\overline{\psi_x \psi_z} = 4\mathrm{i}(\omega'^* - \omega')\,(k_{\mathrm{M}}^2 - 4)^{-1} \{\chi_0^{\mathrm{M}} \chi_{0zz}^{\mathrm{M}} - (\chi_{0z}^{\mathrm{M}})^2\}\, \mathrm{e}^{2\omega_i t} + O(\epsilon^{\frac{4}{3}}). \tag{105}$$

This simplifies to $$\overline{\psi_x \psi_z} = 2\omega_i k_M^2 e^{2\omega_i t} + O(\epsilon^{\frac{4}{3}}), \tag{106}$$

using (104), and is again independent of z, as it must be. Setting $c_a = c_M = \frac{1}{2}$ and $a = k_M$ in (102) for type-A buoyancy flux, we see that it reduces to (106) – neglecting the $\omega'\gamma^2 \approx \epsilon^{\frac{4}{3}}$ contribution from wavenumber $\boldsymbol{b}$ – though a different approximation for ω_i needs to be used (equation (75)) as we pass from one type to the other. Further, (106) is identical in form with the buoyancy flux for classical instability (98), though for the latter ω_i would be zero without topography! Remarkably, only wavenumber $\boldsymbol{a}$ – the perturbed 'neutral' mode – effects a contribution to lowest-order buoyancy flux, though the other wavenumber $\boldsymbol{b}$ and the bridging topography are necessary for the resonant vortex-stretching mechanism by which the wavenumbers interact and catalyse the instability.

4.4. *Type-M² instability*

We calculate the buoyancy flux for $a = b = k_M$, $a_x = b_x = 2\omega_0$ exactly, though our result will be accurate as long as condition (87) is fulfilled. The wavefunction can be written

$$\psi = \left\{\chi_0^M(z) + \frac{2\omega'}{\omega_0}(k_M^2 - 4)^{-1}\chi_{0z}^M(z)\right\}\{e^{i\boldsymbol{a}\cdot\boldsymbol{x} - i\omega t} + \gamma e^{i\boldsymbol{b}\cdot\boldsymbol{x} - i\omega t}\} + \text{c.c.} \tag{107}$$

correct to $O(\epsilon)$ (cf. (103)), where ω' and γ are given by (85) and (86). Note that $|\gamma| = 1$. Wavenumbers $\boldsymbol{a}$ and $\boldsymbol{b}$ each give a contribution to buoyancy flux analogous to the dominant term of (106), so that the total flux is

$$\overline{\psi_x \psi_z} = 4\omega_i k_M^2 e^{2\omega_i t} + O(\epsilon), \tag{108}$$

where $\omega_i = \operatorname{Im}\omega' \approx \epsilon^{\frac{1}{3}}$.

The comparison of (108) with (106) is interesting; the former is just double the latter. One may think of this as the wavenumber $\boldsymbol{b}$ contribution in (106) – there formally $O(\epsilon^{\frac{4}{3}})$ – emerging from the higher-order gloom, so to speak, as $b \to a$ and reinforcing exactly the wavenumber $\boldsymbol{a}$ contribution.

5. Summary and discussion

This paper has treated the effects of topography on quasigeostrophic instability. For consistency with quasigeostrophy, only topography with slope bounded in order of magnitude by Rossby number times vertical–horizontal aspect ratio could be considered. Beyond this, the treatment proceeded by means of a perturbation in a parameter ϵ typifying topography, *after* scaling by Rossby number and aspect ratio. Hence the first-order effects in ϵ explicitly calculated above are formally valid approximations only for topography *smaller* even than Rossby number times aspect ratio. However, uniform convergence can be established for finite ϵ for the perturbation formalism we have used (McIntyre 1970). Hence the first-order corrections obtained above, although they may need improvement by higher-order terms – which may be straightforwardly if laboriously calculated by the formalism we have developed – for quantitative accuracy, at least give qualitative indication of the effects of order-unity topography (after Rossby-number scaling).

The qualitative effects of topography are quite remarkable. Any periodic component of topography with two-dimensional wavenumber $\boldsymbol{c}$ can destabilize a pair of top- and bottom-intensified waves – stable according to the analysis without topography – with wavenumbers $\boldsymbol{a}, \boldsymbol{b}$ such that

$$\boldsymbol{a} - \boldsymbol{b} = \boldsymbol{c},$$

and the same wave frequency

$$\omega_a = \omega_b.$$

These conditions suggest a situation like that of resonant triads, except that here only two members of the triad are dynamic wave-modes, the third being a feature – having zero frequency – of the spatially inhomogeneous environment. The result of the resonance is not adiabatic exchange of energy among the triad, but the coupled growth of the two dynamic modes at the expense of the available potential energy tied up in the basic flow. For given c, there exists a subrange of frequency ω_0 within $0 < \omega_0 < \frac{1}{2}k_M = 1.20$ at each of which a unique† pair of wavenumbers $\boldsymbol{a}[\omega_0], \boldsymbol{b}[\omega_0]$ exists satisfying the resonance condition. This may be confirmed by a perusal of figures 2 and 3. This general kind of topographic instability we called type-A instability. The growth rate of the instability is proportional, among other things, to the magnitude of the topographic component that catalyses it.

There are certain special instances of this new instability. One, which we called the type-M instability, occurs when one of the dynamic modes is (exactly or nearly) the marginally neutral mode between classical non-topographic baroclinic instability, which occurs for wavenumber magnitudes below $k_M = 2.40$, and the otherwise-stable range above k_M. The other dynamic mode may be either a top-intensified or a bottom-intensified wave. The growth rate of this instability is proportional to the $\frac{2}{3}$ power of the amplitude of the catalysing component of topography.

The second special case, called type-M^2, occurs when both dynamic modes are marginally neutral by the non-topographic analysis. This can only happen for topography oriented cross-stream. When it does, the resulting instability grows at a rate proportional to the square root of the topographic component. If topography is small ($\epsilon \ll 1$), these three types are ranked in order of increasing growth rates. (Classical baroclinic instability is still more potent, possessing growth rates of order unity.) In contrast with their growth rates, these three classes of instability are progressively less populous: the type-A instabilities form a three-dimensional manifold, type-M a two-dimensional manifold, and type-M^2 a one-dimensional manifold.

Like the classical baroclinic instability without topography, the energy of growing disturbances comes from the available potential energy bound up in the mass field balancing the geostrophic mean current. That is, the kinetic plus potential energies of the disturbances are fed by the eddy flux of buoyancy down the mean gradient. Again like the classical instability, the two-mode hybrid topographic instabilities contrive to transport buoyancy by arranging constant-phase surfaces of the stream function of each mode separately to slope upwards and upwind. For type-A instability this slope (from the vertical) is $O(\epsilon)$, and both modes contribute significantly to the buoyancy flux. In type-M instability, the 'neutral' mode is perturbed to have $O(\epsilon^{\frac{2}{3}})$ phase slope, and this mode dominates in contributions to the eddy flux. In type-M^2 instability, both 'neutral' modes are perturbed to have $O(\epsilon^{\frac{1}{2}})$ phase slope, and contribute precisely equally to the buoyancy flux.

Charney & Straus (1980) produced a heavily truncated spectral model of two-layer baroclinic flow over topography. The six spectral wavefunctions permitted in their model of a channel bounded meridionally at $y = 0, \pi$, and zonally periodic in x, were

$$F_A = \cos y, \quad F_K = \cos nx \sin y, \quad F_L = \sin nx \sin y,$$

$$F_C = \cos 2y, \quad F_M = \cos nx \sin 2y, \quad F_N = \sin nx \sin 2y$$

† Occasionally there will be two, one corresponding to type-A_1 instability, the other to type-A_2 (see figure 4).

(n is an integer). Topography of a form similar to F_K was supplied. A baroclinic zonal flow $u_A \approx -\partial F_A/\partial y$ tends to be forced by a prescribed meridional heating distribution $\theta_A^* F_A$. It seems fair to identify this mode, though varying meridionally, with our zonal basic flow $\bar{u}$. The modes F_K, F_L do not represent different wavenumbers but are necessary to represent an arbitrary phase relationship of flow relative to F_K-like topography. A similar thing might be said for the pair F_M, F_N. The mode F_C can only be generated by the binary nonlinear interaction of the pairs F_K, F_N and F_L, F_M, as can be seen from a careful inspection of Charney & Straus's equations. In a linearized analysis this mode would appear to have only secondary importance. This really only leaves the two wavenumber pairs, represented by F_K, F_L on the one hand and F_M, F_N on the other, that can exist as analogues of 'free' modes in the linearized system. And topography has the same structure as one of these. Hence there is not sufficient wavenumber structure possible in the Charney–Straus system to permit the kind of mode–mode–topography triad interaction which has been the subject of this paper. (Though in fairness it must be pointed out that Charney & Straus's calculations are nonlinear with finite-amplitude topography and fluctuations, while our calculation is linearized with weak amplitudes.)

Pedlosky (1981) has abstracted the Charney–Straus mechanism in a linearized model which shows how the Charney–Straus quasistationary equilibria can be seen as an interaction among a stationary (zero-frequency) Rossby wave with $e^{ikx} \sin ly$ structure (cf. F_K), topography of the same structure, and a correction to the zonal flow behaving like $\sin 2ly$ (cf. F_C), generated by self-interaction of the stationary Rossby wave.

The instabilities that have formed the subject of this paper may therefore be expected to permit a much richer class of baroclinic motions interacting with topography than the model of Charney & Straus (1980) can allow.

This work has been supported by grants from the National Science Foundation to the program of International Southern Ocean Studies.

REFERENCES

Blumsack, S. L. & Gierasch, P. J. 1972 Mars: the effect of topography on baroclinic instability. *J. Atmos. Sci.* **29**, 1081–1089.

Bretherton, F. P. 1966 Critical layer instability in baroclinic flows. *Q. J. R. Met. Soc.* **92**, 325–334.

Charney, J. G. 1947 The dynamics of long waves in a baroclinic westerly current. *J. Met.* **4**, 135–163.

Charney, J. G. & Straus, D. M. 1980 Form-drag instability, multiple equilibria and propagating planetary waves in baroclinic, orographically forced, planetary wave systems. *J. Atmos. Sci.* **37**, 1157–1176.

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

Hide, R. & Mason, P. J. 1975 Sloping convection in a rotating fluid. *Adv. Phys.* **24**, 47–100.

McIntyre, M. E. 1970 On the non-separable baroclinic parallel-flow instability problem. *J. Fluid Mech.* **40**, 273–306.

Pedlosky, J. 1979 *Geophysical Fluid Dynamics.* Springer.

Pedlosky, J. 1981 Resonant topographic waves in barotropic and baroclinic flows. *J. Atmos. Sci.* **38**, 2626–2641.

Szoeke, R. A. de 1975 Some effects of bottom topography on baroclinic stability. *J. Mar. Res.* **33**, 93–122.