J. Fluid Mech. (1996), *vol.* 313, *pp.* 147–180

The influence of surface topography on rotating convection

By PETER I. BELL AND ANDREW M. SOWARD†

Department of Mathematics and Statistics, University of Newcastle upon Tyne, Newcastle upon Tyne, NE1 7RU, UK

(Received 19 January 1995 and in revised form 2 October 1995)

Busse's annulus is considered as a model of thermal convection inside the Earth's liquid core. The conventional tilted base and top are modified by azimuthal sinusoidal corrugations so that the effects of surface topography can be investigated. The annulus rotates rapidly about its axis of symmetry with gravity directed radially inwards towards the rotation axis. An unstable radial temperature gradient is maintained and the resulting Boussinesq convection is considered at small Ekman number. Since the corrugations on the boundaries cause the geostrophic contours to be no longer circular, strong geostrophic flows may be driven by buoyancy forces and damped by Ekman suction. When the bumps are sufficiently large, instability of the static state is dominated by steady geostrophic flow with the convection pattern locked to the bumps. As the bump size is decreased, oscillatory geostrophic flow is possible but the preferred mode is modulated on a long azimuthal length scale and propagates as a wave eastwards. This mode only exists in the presence of bumps and is not to be confused with the thermal Rossby waves which are eventually preferred as the bump height tends to zero. Like thermal Rossby waves, the new modes prefer to occupy the longest available radial length scale. In this long-length-scale limit, two finite-amplitude states characterized by uniform geostrophic flows can be determined. The small-amplitude state resembles Or & Busse's (1987) mean flow instability. On losing stability the solution jumps to the more robust large-amplitude state. Eventually, for sufficiently large Rayleigh number and bump height, it becomes unstable to a long-azimuthal-length-scale travelling wave. The ensuing finite-amplitude wave and the mean flow, upon which it rides, are characterized by a geostrophic flow, which is everywhere westward.

1. Introduction

The problem of buoyantly driven rotating convection is of fundamental interest in connection with motion inside the Earth's fluid core. There, relative to a frame rotating at angular velocity Ω with the Earth, motion is slow and at lowest order geostrophic:

$$2\rho\boldsymbol{\Omega} \times \boldsymbol{u} = -\nabla P. \tag{1.1}$$

Here $\boldsymbol{u}$ and P are the fluid velocity and pressure, while the density ρ is assumed constant in the Boussinesq approximation which we adopt. Such flow is limited by the constraints of the Proudman–Taylor theorem which states that motion is independent of the coordinate parallel to the rotation axis. As Greenspan (1968) explains, this

† Present address: Department of Mathematics, University of Exeter, Laver Building, North Park Road, Exeter, EX4 4QE, UK.

means that columns of fluid move following geostrophic contours on the upper and lower boundaries separated by constant height. In the case of spherical boundaries inside the Earth's fluid core, the geostrophic contours are circles centred on the rotation axis and motion lies on the geostrophic circular cylinders generated by them.

When small viscous effects are considered as measured by the Ekman number

$$E := \nu/\Omega D^2 (\ll 1), \tag{1.2}$$

where ν is the kinematic viscosity and D is an appropriate radial core length scale (see the remarks above (1.7)), some small driving force is necessary to maintain motion. In the case of thermal or compositional buoyancy, convective motion in a sphere or spherical shell necessarily has a radial component and so is not purely geostrophic. Instead, as Roberts (1968) and Busse (1970) explained, convection at the onset of instability takes the form of quasi-geostrophic axial rolls with short azimuthal length scale of order $E^{1/3}$. On this short length the constraints of the Proudman–Taylor theorem are relaxed by a balance of viscous and buoyancy forces. Exactly like Rossby waves, the rolls propagate azimuthally due to inertia and the influence of the slope of the top and bottom boundaries. The situation changes when the role of the magnetic field permeating the electrically conducting fluid core is considered. Then for sufficiently strong magnetic fields, the constraints of the Proudman–Taylor theorem can be overcome by the Lorentz force and motion occurs on much longer length scales. In this paper, however, only non-magnetic convection, which attracts continuing interest (see, for example, Zhang 1992), will be considered.

The early convective models mentioned above concerned instabilities to spherically symmetric configurations. Nevertheless, Hide (1967) suggested that weak horizontal temperature variations and horizontally extensive topographic features at the core–mantle boundary (CMB) might produce hydrodynamical effects. Only recently has the possible importance of these inhomogeneities on convection been appreciated (see, for example, Gubbins & Richards 1986). The effect of thermal boundary anomalies in a rapidly rotating spherical shell was subsequently investigated by Zhang & Gubbins (1992, 1993) and more recently by Sun, Schubert & Glatzmaier (1994). The configuration is such that weak convection is always forced even when there is no adverse density gradient. As the Rayleigh number Ra (see (1.7) below) measuring that gradient is increased, an 'imperfect bifurcation' sets in at some critical value of Ra. They found that the preferred mode is convective rolls, which cease to propagate azimuthally but remain stationary, locked instead to the thermal inhomogeneities. This is an interesting feature which might have important implications for the geodynamo.

In contrast, we are interested here in the role of topographic effects. It is well known that this influences core–mantle coupling and there is particular interest in mechanisms which might lead to angular momentum transfer between the fluid core and solid mantle, because of their importance in accounting for changes in the length of the day (see, for example, Hide 1989, and Hide *et al.* 1993). Electromagnetic effects play a crucial role (see, for example, Roberts 1988). Jault & Le Mouël (1989, 1993) have stressed the importance of geostrophic flows in this context and quantify the pressure torques at the CMB (see also Anufriev & Braginsky 1977). Unsteady effects have also been considered. Essentially, the radiation of waves due to large-scale motion over a bumpy boundary can also lead to mean couples. This idea has been explored in the magnetohydrodynamic context by Anufriev & Braginsky (1975), Moffatt & Dillon (1976) and Kuang & Bloxham (1993), all of whom employed two-dimensional topography as we do here. Indeed, our simple sinusoidal variation (1.11) has also been used by Busse & Wicht (1992) to characterize variations in the

electrical conductivity of the mantle at the CMB for the purpose of investigating its dynamo implications. The point which we wish to stress is that other related pioneering investigations have built their studies on comparable simple models.

Like Jault & Le Mouël (1989) our attention focuses on the role of geostrophic flows, but in contrast its influence on convection is our immediate concern. The direct consequence of bumps on the CMB is that the geostrophic cylinders are no longer circular and the buoyancy force $\boldsymbol{F}_B$ has a component tangent to them. When added to the geostrophic balance (1.1), the equation becomes

$$2\rho\boldsymbol{\Omega}\times\boldsymbol{u} = -\nabla P + \boldsymbol{F}_B\,. \tag{1.3}$$

Consider a pair of geostrophic contours on the upper and lower boundaries separated by a constant distance $2H$. Let the unit outward normals to these boundaries be denoted by $\mathbf{1}_{n_+}$ and $\mathbf{1}_{n_-}$ respectively and that of the geostrophic cylinder by $\mathbf{1}_n$. The corresponding unit geostrophic tangent vector is $\mathbf{1}_t := \mathbf{1}_{n_+}\times\mathbf{1}_{n_-}/|\mathbf{1}_{n_+}\times\mathbf{1}_{n_-}|$. If we form line integrals of (1.3) about closed curves $C(H)$ with respect to arc length l parallel to the geostrophic contour and then integrate parallel to the rotation axis, we obtain the result

$$2\rho\Omega\langle\boldsymbol{u}\cdot\mathbf{1}_n\rangle = \langle\boldsymbol{F}_B\cdot\mathbf{1}_t/|\mathbf{1}_\Omega\times\mathbf{1}_t|\rangle, \tag{1.4a}$$

where $\mathbf{1}_\Omega := \boldsymbol{\Omega}/\Omega$ and

$$\langle\ldots\rangle := A^{-1}\int_{A(H)}\ldots\mathrm{d}A \tag{1.4b}$$

denotes the cylinder average over the geostrophic cylinder $A(H)$ of area

$$A = 2H\oint_{C(H)}|\mathbf{1}_\Omega\times\mathbf{1}_t|\mathrm{d}l. \tag{1.4c}$$

The term on the left-hand side of (1.4*a*) is proportional to $A\langle\boldsymbol{u}\cdot\mathbf{1}_n\rangle$ which measures the outward flux of fluid. Since it vanishes, we recover the generalized version of Taylor's condition given by Fearn & Proctor (1992). Taylor (1963) originally stated his condition in the context of the magnetic Lorentz force, but when the geostrophic contours are no longer circular there is a buoyancy contribution to his integral on the right-hand side of (1.4*a*). When, however, this integral fails to vanish, other small effects, so far neglected, must be included. For steady flows, the most important consideration is suction in the Ekman boundary layers. With that addition, the flux $A\langle\boldsymbol{u}\cdot\mathbf{1}_n\rangle$ no longer vanishes and the dominant geostrophic contribution $U\mathbf{1}_t$ to the velocity is determined by (1.4*a*), which yields

$$\rho U = T\langle\boldsymbol{F}_B\cdot\mathbf{1}_t/|\mathbf{1}_\Omega\times\mathbf{1}_t|\rangle, \tag{1.5a}$$

where the parameter

$$T = \frac{A}{(\Omega\nu)^{1/2}}\left\{\oint_{C(H)}\left[|\mathbf{1}_\Omega\cdot\mathbf{1}_{n_+}|^{-1/2} + |\mathbf{1}_\Omega\cdot\mathbf{1}_{n_-}|^{-1/2}\right]\mathrm{d}l\right\}^{-1} \tag{1.5b}$$

has the dimensions of time (see, for example, Roberts & Soward 1992, equation (6.12)).

Our objective is to show how topographic effects can dramatically influence thermal convection and produce the locking previously predicted by Zhang & Gubbins (1992, 1993) for thermal inhomogeneities. To this end we adopt Busse's (1970) simplified annulus model, which we describe relative to cylindrical polar coordinates (s,ϕ,z), where z measures distance parallel to the $\boldsymbol{\Omega}$ rotation axis. His idea was that the almost z-independent quasi-geostrophic motion, which is localized at mid-latitudes close to some geostrophic cylinder $s = s_0$, is largely driven by the radial s-component

of gravity normal to the rotation axis. Accordingly, the model assumptions are that both gravity $\boldsymbol{g}$ and the temperature gradient $\boldsymbol{\beta}_T$ have these characteristics:

$$\boldsymbol{g} := -g\mathbf{1}_s, \qquad \boldsymbol{\beta}_T := -\beta_T \mathbf{1}_s. \tag{1.6a,b}$$

Furthermore, since the radial s-length scale is undefined by low-order theory, it is imposed artificially by the the introduction of boundaries at $s = s_0 \pm D$, where the radial length D is small compared to the core radius. The natural modified Rayleigh number for the system is

$$Ra := \alpha_T \beta_T g D^2 / 2\kappa\Omega, \tag{1.7}$$

where α_T is the coefficient of expansion and κ is the thermal diffusivity. The key additional feature of the model is the slanting top and bottom boundaries of the core, here assumed to be located symmetrically at

$$z_\pm := \pm(H - \eta h), \tag{1.8}$$

where ηh measures small departures from the constant half-height H. Busse (1970) simply adopted a constant slope $\mathrm{d}h/\mathrm{d}s = 1$, from which we can define the small slope parameter

$$S := (D/H)\eta (\ll 1) \tag{1.9}$$

(Ewen & Soward 1994), and in turn the renormalized Rayleigh number

$$\mathscr{R} := Ra/S \equiv \alpha_T \beta_T g D H / 2\eta\kappa\Omega, \tag{1.10}$$

which we find more appropriate in our subsequent analysis. The annulus has proved to be a very useful model, from which a relatively simple understanding of the complex processes in rotating convection can be understood, and it is adopted in many of the theoretical investigations listed in the references. Note also that the radial gravitational field (1.6a) is mimicked in the experimental configuration of Carrigan & Busse (1983).

Curvature of the CMB was incorporated in the studies of Busse & Hood (1982), Busse & Or (1986), Or & Busse (1987). Here, instead, we consider non-axisymmetric bumps on the CMB. For simplicity, we consider the sinusoidal corrugation

$$h := (s - s_0) + (\gamma/k)\sin k s_0 \phi, \tag{1.11a}$$

in which γ is a constant. The crests, of half-height H_{bump}, are radial and separated by a distance $2D_{bump}$:

$$H_{bump} = \eta\gamma/k, \qquad D_{bump} = \pi/k. \tag{1.11b,c}$$

In the absence of bumps, $\gamma = 0$, the geostrophic contours h =constant are circles. On the other hand, when $\gamma \neq 0$ that is no longer the case and the geostrophic cylinders have outward unit normal

$$\mathbf{1}_n := \nabla h/|\nabla h| \qquad \big(\nabla h \doteq (1, \gamma\cos k s_0\phi, 0)\big), \tag{1.11d}$$

showing that γ measures the tilt of the geostrophic contours. Throughout this paper we will only consider small bumps, in the sense that the geostrophic contours are only slightly distorted

$$\gamma \ll 1, \tag{1.12}$$

which undulate over a distance comparable to the radial scale ($kD = O(1)$). At the outer edge of the Ekman boundary layers the relative size of the vertical motion produced by Ekman suction and the slanting boundary is characterized by the small

parameter $E^{1/2}/\eta$. Our asymptotic analysis shows that topography has an important effect when the geostrophic contour tilt γ is of order $(E^{1/2}/\eta)^{1/2}$. Accordingly, the key dimensionless parameter measuring bump size in our theory is

$$\Gamma := E^{-1/4}\eta^{1/2}\gamma, \tag{1.13a}$$

which will be assumed of order unity or more precisely

$$\Gamma^2 = O(kD) \tag{1.13b}$$

(but see (1.14) below). It should be stressed that the quasi-geostrophic approximation made throughout this paper depends on the small size of the slope parameter S (see (1.9)) and not on η.

The paper is organized as follows. In §2 the equations governing the topographic convection are derived. For our new geostrophic modes of convection, motion is separated into two parts. The larger is geostrophic (steady/possibly oscillatory) or quasi-geostrophic (travelling wave, modulated on a long azimuthal length scale); the smaller is ageostrophic (periodic on the bump length scale). The heat convected by these flows leads to the buoyancy forces necessary to drive them; the governing equations are Taylor's condition (2.12), the axial vorticity equations (2.14) and (2.16), and the heat conduction equation (2.17). The stability of the static state is investigated in §3. For sufficiently large bumps, instability first sets in as steady convection. For smaller bumps, oscillatory convection is possible, though a propagating modulated mode is preferred. These low-frequency travelling waves should not be confused with the traditional thermal Rossby mode of convection which has the very short $E^{1/3}$ azimuthal length scale. The existence of the new geostrophic modes relies on the presence of the bumps. Like thermal Rossby waves, however, convection prefers to adopt the longest radial length scale available. In that long-length-scale limit, we determine in §4.1 finite-amplitude states consisting of a uniform geostrophic flow, which is self-maintained by the buoyancy force induced by the temperature distribution that it convects. This has some resemblance to the mean flow instability considered by Or & Busse (1987), Schnaubelt & Busse (1990), which rides on thermal Rossby waves. The validity of the results when the radial length scale is finite is discussed in §4.2. The stability of the uniform geostrophic flow to unmodulated disturbances is tested in §4.3. The case of modulated disturbances is investigated in §5. These uniform finite-amplitude states always come in pairs. The small-amplitude solution has a domain of stability, but when lost the solution jumps to the large-amplitude state, which is generally more robust. It is characterized by a negative geostrophic flow, which corresponds to a Westward motion in the geophysical context. This large-amplitude state is eventually unstable to modulated travelling waves, whose finite-amplitude disturbance is investigated numerically in §6. Despite the wave's complex character, the dominant qualitative feature of systematic westward motion is preserved. Some concluding remarks are added in §7. Many of the results presented here have their roots in Bell's (1993) PhD thesis. Before we proceed, however, some remarks about units and applicability to the geophysical problem are appropriate.

To maintain consistency with earlier studies, we non-dimensionalize on some prescribed radial length scale D. On the other hand, the bump half-wavelength $D_{bump}(=\pi/k)$ is the only relevant length in the long-radial-length-scale limit $D \gg D_{bump}$. That is why the combinations

$$\frac{\Gamma^2}{kD} = \left(\frac{\Omega}{k^2\nu}\right)^{1/2}\eta\gamma^2, \qquad \frac{\mathscr{R}}{kD} = \frac{H\alpha_T\beta_T g}{2\eta k\kappa\Omega}, \tag{1.14a,b}$$

which are independent of D, emerge as the most important dimensionless parameters of our theory. For example, they are the coordinates employed on figures 1 and 2, which illustrate various stability boundaries. In the same spirit, it is often convenient to rescale our variables by amounts which recognize the importance of the bump length scale (see particularly (3.4*c*,*d*) below). Indeed, we could define define D to be D_{bump}/π; then with $kD = 1$, much of the rescaling is superfluous. We avoid this condensation of notation so that we can keep track of the role of the bump length scale in our equations by its explicit appearance as k^{-1}. To help clarify the notation and its usage, a summary of the important parameters and variables is provided in Appendix C. Note, however, that the superscript * introduced briefly at the beginning of §2 to distinguish dimensionless quantities is dropped there as elsewhere in the paper.

In application to the Earth's core, η is indeed of order unity and so (1.9) and (1.13) require

$$D \ll H, \qquad \gamma^2 = O(E^{1/2}kD) \qquad \text{when} \qquad \eta = O(1). \tag{1.15a,b}$$

Use of (1.11) shows that the value of the the key parameter Γ^2/kD (see (1.14*a*)) on the CMB is

$$\left(\frac{\Gamma^2}{kD}\right)_{CMB} = \left(\frac{\Omega R_{core}^2}{\nu}\right)^{1/2}\left(\frac{\pi H_{bump}^2}{\eta R_{core} D_{bump}}\right), \tag{1.16a}$$

where $R_{core} \doteq 3.5 \times 10^3$km is the core radius and the core Ekman number is $\nu/(\Omega R_{core}^2) \doteq 10^{-15}$ for kinematic values of the viscosity. For moderate size undulations on the CMB of order $H_{bump} = 1$ km height extending over distances of order $D_{bump} = 10^3$ km, we obtain upon setting $\eta = 1$ the value

$$\left(\frac{\Gamma^2}{kD}\right)_{CMB} \doteq 30. \tag{1.16b}$$

According to our result (3.8), the steady geostrophic mode sets in when Γ^2/kD exceeds 2, a limitation which is easily met by (1.16*b*). For smaller values, the geostrophic mode oscillates and $\mathscr{R}/kD$ is of order kD/Γ^2 (see (3.7)). This only gives way to the thermal Rossby mode, for which $\mathscr{R}/kD$ is large of order $E^{-1/3}$, at extremely small values of Γ^2/kD. Of course, for larger bump heights and/or shorter bump diameters, the parameter $(\Gamma^2/kD)_{CMB}$ may become very large and asymptotic results for $\Gamma \uparrow \infty$ are then appropriate.

In the case of the spatially modulated modes, the azimuthal (ϕ) modulation half-wavelength $D_{modulation}$ cannot be too long and is clearly limited by the half-circumference of the core! Furthermore, the bumps need to be on a fairly short length scale (say $D_{bump} = 10$ to 10^2 km) in order to achieve scale separation. For the modes considered here with radial (s) half-wavelength $2D$, corresponding to $\beta = \frac{1}{2}\pi$ (see (3.2*a*)), the parameter $\mathscr{A}$ defined by (3.3*b*) (see also Appendix C) is generally of order unity for the most readily excited modes. For such modes we obtain

$$\frac{D_{modulation}}{\pi R_{core}} = \frac{2\eta}{\mathscr{A}}\left(\frac{\Omega R_{core}^2}{\nu}\right)^{1/2}\left(\frac{2D}{\pi R_{core}}\right)^2 \qquad (\mathscr{A} = O(1)) \tag{1.16c}$$

and so the condition $D_{modulation} \leqslant \pi R_{core}$ imposes a severe restriction on D because of the small size of the Ekman number. The nature of the difficulty can be identified from (2.14) for, as the radial length scale increases, the Ekman suction and buoyancy torques decrease in concert leaving the unbalanced topographic term $2\Gamma^2\partial\overline{\psi}/\partial\zeta$. Since this term is not dissipative, it could, on these longer length scales, be balanced by

the inertia term dropped in (2.14). This is the usual Rossby wave balance; it leaves the crucial balance between viscous dissipation and buoyancy forces mentioned to the next order of approximation. Accordingly, we may expect modulated modes of instability to occur in the Earth's core but, since their azimuthal length scale is necessarily shorter than the optimal scale of our theory, the corresponding critical Rayleigh numbers will accordingly be larger. On the other hand, in a laboratory system with a moderately small Ekman number of order say 10^{-4}, the conditions of our theory are easily met by an annulus with sufficiently small tilt η. One final point: our small Ekman number asymptotics leads to the neglect of viscous effects in the interior of the flow, leaving them only in the Ekman boundary layers. This distinction actually imposes tight restrictions on various length scales. The qualitative features of our results are not sensitive to this distinction (but see further remarks in §2 connected with Bell's (1993) thesis).

The main thrust of the present study is to emphasize the possible importance of the interesting steady geostrophic mode of convection locked solidly to the topography. Our secondary objective is to explore the possible modulated variants with their Rossby wave links. The importance of such modes is highlighted by the results of §3, where this type of instability of the static state is shown to set in when Γ^2/kD is less than $4/\sqrt{3}$. Of course, in the spherical shell or cavity, the modulated modes will not be as potent as the present theory suggests, because of the limitation on the azimuthal length scale.

2. The governing equations

The customary approximations of Busse's annulus are employed. Note, however, that it is usually argued that η is small; as we remarked below (1.13), that is unnecessary provided that the slope parameter S (see (1.9)) is small. The derivation of our governing equations, relies on expansions in terms of the small geostrophic tilt parameter γ (see (1.12)). The actual ordering adopted in (2.2), (2.11), (2.15) is subtle but appropriate to obtain the correct balances in the governing equations for the geostrophic convective modes. In Bell's (1993) original study, he undertook a complete harmonic analysis of the vorticity equation (2.7) without the Ekman suction term $-\frac{1}{2}(E^{1/2}/\eta)\nabla_H^2\psi$ as appropriate to stress-free boundary conditions. His initial numerical study at small but finite Ekman number E revealed the existence of the new modes. Subsequently, he developed an asymptotic theory valid as $E \downarrow 0$, which agreed with his numerical results for small values of E. Of course, other modes exist, for example thermal Rossby waves, which have different orderings; our choice selects those with the smallest critical Rayleigh number for sufficiently large bumps.

Since the horizontal length scale D is small compared to the cylinder radius s_0, we introduce local rectangular Cartesian coordinates (x^*, y^*, z^*) non-dimensionalized on the length D:

$$s - s_0 := Dx^*, \qquad s\phi := Dy^*, \qquad z := Dz^* \qquad (h := Dh^*). \tag{2.1a–c}$$

Our unit of time D^2/κ is based on thermal diffusion:

$$t := (D^2/\kappa)t^*. \tag{2.1d}$$

Accordingly, the fluid velocity $\boldsymbol{u}$ and temperature $T(:= T_0 - \beta_T(s - s_0) + T')$ are rendered dimensionless by the following change of variables:

$$\boldsymbol{u} := (\kappa/D)\boldsymbol{u}^*, \qquad T := T_0 + D\beta_T(-x^* + \gamma\theta^*). \tag{2.2a,b}$$

The corresponding buoyancy force is

$$\boldsymbol{F}_B = \rho\alpha_T\beta_T g D\gamma\theta^*\mathbf{1}_{x^*}. \tag{2.3}$$

Correct to order γ the unit outward normals to the top and bottom boundaries are

$$\mathbf{1}_{n_\pm} = (\eta, \eta\gamma\cos k^* y^*, \pm 1) \qquad (k^* := kD), \tag{2.4a}$$

while the tangent vector to the geostrophic cylinder is

$$\mathbf{1}_t = (-\gamma\cos k^* y^*, 1, 0). \tag{2.4b}$$

Our choice of scaling of $\boldsymbol{u}$ and T is such that the geostrophic contribution $U^*\mathbf{1}_t$ determined by (1.5) satisfies

$$U = -2\mathscr{R}\Gamma^2\langle\theta\cos ky\rangle, \tag{2.5}$$

where here and below the superscript * labelling dimensionless variables is dropped.

From a more general point of view, we may begin with the quasi-geostrophic flow

$$\boldsymbol{u} := \nabla\times\psi\mathbf{1}_z. \tag{2.6}$$

Under the assumption that ψ is z-independent, the z-average of the axial vorticity equation yields

$$\frac{1}{2}\frac{E}{S}\left\{-\frac{\kappa}{\nu}\left(\frac{\partial}{\partial t}+\boldsymbol{u}\cdot\nabla\right)(\nabla_H^2\psi)+\nabla_H^4\psi\right\}-\frac{1}{2}\left(\frac{\gamma}{\Gamma}\right)^2\nabla_H^2\psi+\left(\frac{\partial\psi}{\partial y}-\gamma\cos ky\frac{\partial\psi}{\partial x}\right)$$
$$=-\gamma\mathscr{R}\frac{\partial\theta}{\partial y} \qquad \left(\nabla_H^2 := \frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial x^2}\right). \tag{2.7}$$

In the absence of bumps, the onset of instability is characterized by thermal Rossby waves which occur on the short width $D(E/S)^{1/3}$ of Munk (1950) shear layers (having some similarity to $E^{1/3}$ Stewartson (1957) layers). In this case, the (topographic) Rossby wave is driven by buoyancy forces and damped by viscous forces in the interior of the flow. The corresponding Rayleigh number $\mathscr{R}$ is of order $E^{-1/3}$. In the presence of bumps a second mode exists, for which convection occurs on the bump length scale D/k or simply D (k is assumed to be of order unity) and oscillates on the thermal diffusion time scale D^2/κ. More precisely, motion is aligned predominantly with the geostrophic contours and so we refer to it as the geostrophic mode. The corresponding critical Rayleigh number is of order Γ^{-2}. Clearly for the order-unity values of Γ considered here, the onset of instability is characterized by the geostrophic mode.

In view of the characteristics of the geostrophic mode just described, we neglect all the terms in (2.7) proportional to E. More precisely, we ignore inertia because the diffusion time D^2/κ is long compared to the Rossby wave time scale $(S\Omega)^{-1}$ so placing a lower limit, albeit very small, on the size of the Prandtl number ν/κ. In addition, we ignore viscous effects in the interior and retain only Ekman damping described by the term proportional to $E^{1/2}/\eta$ $[\equiv(\gamma/\Gamma)^2]$ in (2.7). This approximation is justified if the bump length scale D of the convection is large compared to the $E^{1/4}$ Stewartson (1957) shear layer width $H^{1/2}(\nu/\Omega)^{1/4}$. Indeed the vortex line stretching balance due to Ekman suction and topography, described by the remaining terms on the left-hand side of (2.7), is well known in oceanography and occurs in Stommel (1948) layers in western boundary layer coastal currents. Indeed Pedlosky (1979, p. 286) in his equation (5.5.29) neatly summarizes the balances appropriate to Monk, $E^{1/4}$ Stewartson (frictional sublayer) and Stommel layers. Furthermore, the oceanic

link is even tighter, if we link our averaged buoyancy force on the right-hand side of (2.7) with the surface wind stresses of the oceanic models.

It is worth emphasizing the neglect of interior viscous effects. Though the approximation is common in rotating magnetoconvection, viscous friction in the interior provides a vital force balance for thermal Rossby waves. Indeed Bell's (1993) results were based on that balance and Ekman suction effects were ignored. As we have just explained, that approximation is only valid when the bump scale is smaller than the $E^{1/4}$ Stewartson layer length scale, but that is not the case in our model. Nevertheless, despite the differences, the qualitative features of the two sets of results are very similar.

Only in the heat conduction equation do we retain nonlinear terms and under our non-dimensionalization it reduces to

$$\frac{\partial\theta}{\partial t}+\left(\frac{\partial\psi}{\partial y}\frac{\partial\theta}{\partial x}-\frac{\partial\psi}{\partial x}\frac{\partial\theta}{\partial y}\right)=\frac{1}{\gamma}\frac{\partial\psi}{\partial y}+\nabla^2\theta. \tag{2.8}$$

In order to distinguish carefully between geostrophic and ageostrophic flow, it is helpful to replace the radial coordinate x with the geostrophic coordinate

$$h := x+(\gamma/k)\sin ky \tag{2.9a}$$

(see (1.11*a*) and (2.1)), for which correct to order γ we have

$$\mathbf{1}_n=\nabla h, \qquad \mathbf{1}_t=\mathbf{1}_z\times\mathbf{1}_n. \tag{2.9b,c}$$

Consequently, when the streamfunction $\psi(=\overline{\psi})$ depends on h alone, the resulting motion is geostrophic and

$$U=-\partial\overline{\psi}/\partial h. \tag{2.10}$$

Since this motion has a radial component, it distorts the isotherms. In the absence of thermal diffusion, the isotherms would coincide with the geostrophic contours, so, like U, the temperature would be a function of h. With the inclusion of diffusion, however, the dominant contribution to the temperature perturbations fluctuates on the period of the bumps. In fact we find that the appropriate expansion of the temperature for the most unstable preferred modes is

$$\theta := \gamma\overline{\theta}+\tfrac{1}{2}\left(\widetilde{\theta}\mathrm{e}^{\mathrm{i}ky}+\widetilde{\theta}^{*}\mathrm{e}^{-\mathrm{i}ky}\right)+O(\gamma^2), \tag{2.11}$$

where $\overline{\theta}$ and $\widetilde{\theta}$ are complex functions of h alone. Here and below the superscript $*$ is used to denote complex conjugate. Substituting (2.11) into Taylor's condition (2.5) yields

$$U=-\tfrac{1}{2}\mathscr{R}\Gamma^2\left(\widetilde{\theta}+\widetilde{\theta}^{*}\right). \tag{2.12}$$

Though (2.11) only contains one harmonic in the azimuthal y-direction, the complete linear solution contains all harmonics proportional to $\exp(\mathrm{i}nky)$ for integer n. So our additional approximations based on the small size of γ are crucial.

In addition to steady modes, we will also consider travelling topographic waves. These are not perturbations of the short-azimuthal-wavelength thermal Rossby waves occurring in the absence of bumps. Rather they are long-wavelength modes whose structure continues to be defined by (2.11) but with coefficients dependent on time t and

$$\zeta := \gamma^2 y \tag{2.13}$$

in addition to h. In other words, the amplitudes are now modulated on the long dimensionless length γ^{-2}. Since $\overline{\psi}$ depends on ζ, the motion it defines is not strictly

geostrophic. Accordingly, we may no longer use Taylor's condition (2.5) but must use (2.7) instead. Correct to lowest order in γ it gives

$$\frac{\partial U}{\partial h} + 2\Gamma^2 \frac{\partial \overline{\psi}}{\partial \zeta} = -\frac{\mathscr{R}\Gamma^2}{2}\left(\frac{\partial \widetilde{\theta}}{\partial h} + \frac{\partial \widetilde{\theta}^*}{\partial h}\right) \tag{2.14}$$

for the mean terms independent of y. Of course, it is the presence at this order of the term $2\Gamma^2 \partial\overline{\psi}/\partial\zeta$ describing motion across the geostrophic contours which motivated the introduction of the stretched y-variable ζ. On the other hand, when there is no dependence on ζ and this term vanishes we can integrate (2.14) with respect to h and recover (2.12) up to an arbitrary constant. The vanishing of that constant cannot be inferred from the axial vorticity equation, but instead depends on averaging the equation of motion following Taylor's recipe.

Of course, the fluctuating temperature perturbation leads to buoyancy forces, which directly drive fluctuating ageostrophic flow. In fact the complete representation for the streamfunction is

$$\psi := \overline{\psi} + \tfrac{1}{2}\gamma\left(\widetilde{\psi} e^{iky} + \widetilde{\psi}^* e^{-iky}\right) + O(\gamma^2). \tag{2.15}$$

Upon substitution into the vorticity equation (2.7), the fluctuating terms proportional to $\exp(\pm iky)$ yield the primary order-one torque balance

$$\widetilde{\psi} = -\mathscr{R}\widetilde{\theta}, \tag{2.16}$$

a result which motivated our choice of scaling for the fluctuating streamfunction. This simple balance has its analogue in the oceanic context mentioned below (2.7) provided that we identify our buoyancy forces with wind stresses. In that context it is known as the Sverdrup (1947) relation (see Pedlosky, 1979, p. 264, particularly equation (5.3.2)). For us, it describes part of the usual balance invoked for thermal Rossby waves, namely that between the topographic effects and buoyancy torques. There, however, inertia and interior viscous effects play a significant role as well (motivating Bell's 1993 original choice), while in our case viscous effects are only manifest through Ekman suction in boundary layers. Perhaps it should be stressed that the balance in (2.14) is delicate in the sense that it arises from the secondary torque balance of order γ.

Upon substitution of (2.11) and (2.15) into the heat conduction equation (2.8), the leading-order-one balance, which ignores all terms proportional to γ and smaller, is

$$\frac{\partial \widetilde{\theta}}{\partial t} + ikU\widetilde{\theta} = ik\widetilde{\psi} - U + \left(\frac{\partial^2}{\partial h^2} - k^2\right)\widetilde{\theta}. \tag{2.17}$$

A significant feature is the convection of the fluctuating temperature by the geostrophic velocity described by $ikU\widetilde{\theta}$, which is the only nonlinear term retained in our theory. The radial convection of the original applied temperature gradient is brought about by both the fluctuating ageostrophic and mean geostrophic velocities as indicated by the term $ik\widetilde{\psi} - U$.

Though knowledge of the perturbation $\gamma\overline{\theta}$ to the mean temperature is not needed in our theory, we give the equation governing it for completeness. It comes from the mean terms of the heat conduction equation (2.8) which at leading order γ give

$$\frac{\partial \overline{\theta}}{\partial t} + \frac{\partial}{\partial h}\left[\tfrac{1}{4}ik\left(\widetilde{\psi}\,\widetilde{\theta}^* - \widetilde{\psi}^*\widetilde{\theta}\right) + \mathscr{H}\right] = \frac{\partial \overline{\psi}}{\partial \zeta} + \frac{\partial^2 \overline{\theta}}{\partial h^2}, \tag{2.18a}$$

where

$$\mathscr{H} := -\tfrac{1}{4}\left[(\tilde{\psi} + \tilde{\psi}^*) + \mathrm{i}k(\tilde{\theta} - \tilde{\theta}^*)\right]. \tag{2.18b}$$

Note that in view of the result (2.16), the nonlinear heat flux $\frac{1}{4}ik(\tilde{\psi}\,\tilde{\theta}^* - \tilde{\psi}^*\tilde{\theta})$ vanishes leaving only the linear term $\mathscr{H}$. For the spatially periodic solutions of our system, the mean heat flux $\langle\mathscr{H}\rangle$ provides a measure of the amplitude of our solutions.

The boundary conditions on our system need to be considered carefully. In the azimuthal $y(\zeta)$-direction it is natural to assume spatial periodicity due to the absence of boundaries. In the radial direction the issue is more subtle. The boundaries at $x = \pm 1$ are not real and only invoked to define a radial length scale. For simplicity, therefore, we assume spatial periodicity in the radial direction also. Evidently, for the geophysical system, we anticipate that this periodicity is modulated radially in order to localize the solution. Still, in order to understand key processes introduced by the presence of bumps, we believe that our spatially periodic boundary conditions are adequate.

3. Stability of the static state

3.1. *The dispersion relation*

The most general form of perturbation to the static state has the form

$$\tilde{\theta} := \tfrac{1}{2}\left(\hat{\theta}_+ \mathrm{e}^{\mathrm{i}\varphi} + \hat{\theta}_-^* \mathrm{e}^{-\mathrm{i}\varphi}\right)\mathrm{e}^{qt}, \tag{3.1a}$$

$$U := \tfrac{1}{2}\left(\hat{U}\mathrm{e}^{\mathrm{i}\varphi} + \hat{U}^*\mathrm{e}^{-\mathrm{i}\varphi}\right)\mathrm{e}^{qt}, \tag{3.1b}$$

where

$$\varphi := (\beta h + \alpha\zeta) - \omega t \equiv \beta h + \alpha(\zeta - ct) \qquad (c := \omega/\alpha). \tag{3.2a}$$

Here the wavenumber β is fixed by the length scale in the radial direction, the stretched wavenumber α characterizes the slow modulation in the y-direction propagating at the phase velocity c (frequency ω), while

$$p := q - \mathrm{i}\omega \qquad (c_g := \mathrm{i}\,\mathrm{d}p/\mathrm{d}\alpha) \tag{3.2b}$$

is the complex growth rate (group velocity). Upon substitution into (2.14) and making use of (2.10), we obtain the relation

$$(1 + \mathrm{i}\mathscr{A})\hat{U} = -\tfrac{1}{2}\mathscr{R}\Gamma^2\left(\hat{\theta}_+ + \hat{\theta}_-\right), \qquad \mathscr{A} := 2(\Gamma/\beta)^2\alpha. \tag{3.3a,b}$$

In addition, (2.16) and (2.17) yield

$$(k^2\mathscr{P} \pm \mathrm{i}k\mathscr{R})\hat{\theta}_\pm = -\hat{U}, \qquad \mathscr{P} := \mathscr{Q} - \mathrm{i}\mathscr{F}, \tag{3.4a,b}$$

where

$$\mathscr{Q} := (q + \beta^2 + k^2)/k^2, \qquad \mathscr{F} := \omega/k^2. \tag{3.4c,d}$$

Substitution of the result into (3.3*a*) leads to the dispersion relation

$$1 + \mathrm{i}\mathscr{A} = \left(\frac{\mathscr{R}\Gamma^2}{2k^2}\right)\left(\frac{1}{\mathscr{P} + \mathrm{i}(\mathscr{R}/k)} + \frac{1}{\mathscr{P} - \mathrm{i}(\mathscr{R}/k)}\right). \tag{3.5}$$

All points on figures 1 and 2, which illustrate our results, will be identified by their coordinates $(\Gamma^2/k, \mathscr{R}/k)$.

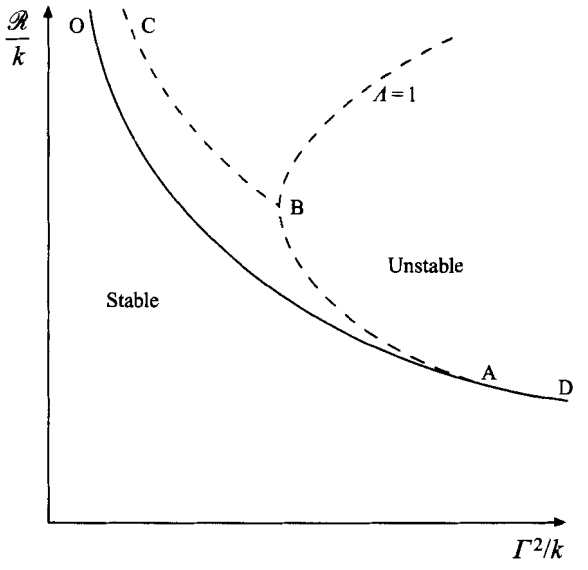

FIGURE 1. The stability boundary $q = 0$ in the $(\Gamma^2/k, \mathscr{R}/k)$-plane for the static state ($\mathscr{U} = 0$) indicated by the continuous line OAD. Stationary convection occurs on the $\Lambda = 1$ curve (appropriate to the $\beta = 0$ case and adopted for consistency with figure 2), which passes through the points B, A and D. Neutral unmodulated ($\alpha = 0$) oscillatory ($\omega \neq 0$) convection occurs on the curve CB. The frequency vanishes ($\omega = 0$) at B, where there is a Takens–Bogdanov bifurcation. The curve OA is the envelope of the corresponding modulated ($\alpha \neq 0$) oscillatory modes. The bifurcation on OA is a Hopf one and on AD a pitchfork.

3.2. *Unmodulated disturbances* $\alpha = 0$

When there is no spatial modulation in the azimuthal direction, α (and $\mathscr{A}$) vanishes. Then (3.5) reduces to

$$k^2\mathscr{P}^2 - \mathscr{R}\Gamma^2\mathscr{P} + \mathscr{R}^2 = 0. \tag{3.6}$$

Oscillatory solutions ($\omega \neq 0$) only occur when $\Gamma^2 < 2k$. Overstability occurs when

$$\frac{2(\beta^2 + k^2)}{k\mathscr{R}} \leqslant \frac{\Gamma^2}{k} < 2. \tag{3.7}$$

The overstability boundary $\mathscr{R}\Gamma^2 = 2(\beta^2 + k^2)$, illustrated by the broken curve CB on figure 1, terminates at B $\left(2, (\beta/k)^2 + 1\right)$, where the frequency vanishes ($\omega = 0$) and a direct mode of instability sets in. Since $\mathscr{R}$ increases with β, the minimum (over β) occurs at $\beta = 0$. When the geometry permits, however, the modulated waves discussed in §3.3 below are the preferred mode of instability. Nevertheless, in the geophysical application, the long azimuthal length scale necessary for the modulations might not exist. Then all that can be inferred is that the critical Rayleigh number lies between the two stability boundaries OA and CB.

Stationary convection ($p = 0$) occurs when $\Gamma = \Gamma_s(\mathscr{R})$, where

$$\frac{\Gamma_s^2}{k} = \frac{(\beta^2 + k^2)^2 + k^2\mathscr{R}^2}{(\beta^2 + k^2)k\mathscr{R}}. \tag{3.8a}$$

For the special case $\beta = 0$, this neutral curve is illustrated by the $\Lambda = 1$ curve on figure 1. When $\Gamma^2 \geqslant 2k$, its lower branch DAB provides the stability boundary to the

direct mode of convection with the convenient parametric representation

$$\frac{\Gamma_s^2}{k} := \frac{2}{\sin 2\varepsilon}, \qquad \frac{\mathscr{R}_s}{k} := \frac{\beta^2 + k^2}{k^2} \tan\varepsilon \qquad (0 < \varepsilon \leqslant \tfrac{1}{4}\pi) \tag{3.8b,c}$$

for Γ^2/k and $\mathscr{R}/k$.

From (3.4*a*) marginal convection everywhere on the neutral curve ($0 < \varepsilon < \frac{1}{2}\pi$) is given by

$$\hat{\theta}_\pm = -(\hat{U}/k\mathscr{R}) \sin\varepsilon e^{\mp i\varepsilon}. \tag{3.9}$$

So, when $\hat{U}$ is real, the actual geostrophic velocity and temperature fluctuations are

$$U = \hat{U}\cos\beta h, \qquad \theta = -(\hat{U}/k\mathscr{R}) \sin\varepsilon \cos\beta h \cos(ky - \varepsilon). \tag{3.10a,b}$$

Accordingly, we see that the phase lag ε increases from zero along the stability boundary as Γ^2 decreases from infinity beyond D, where $\mathscr{R}\Gamma^2 = \beta^2 + k^2$. It attains the value $\frac{1}{4}\pi$ at the minimum value of $\Gamma^2 (= 2k)$, where it meets the overstability branch at B. It continues to increase up to the value $\frac{1}{2}\pi$ as $\mathscr{R}$ increases indefinitely along the upper branch of the neutral curve beyond B. It is perhaps worth noting that stability is not regained above this upper branch, where instead there are two unstable modes.

3.3. *Modulated disturbances* $\alpha \neq 0$

The most unstable mode is found by locating that α, which maximizes the real part q of the complex growth rate p. The overstability boundary (the continuous curve OA on figure 1) is determined by the results of Appendix A with $\mathscr{Q} = (\beta/k)^2 + 1$, $A = \mathscr{R}/k$, $B = \mathscr{R}\Gamma^2/k^2$, $C = 0$. A convenient parametric representation of this boundary is obtained by setting $[(\beta/k)^2 + 1]/A = \tan 2\delta$. In this way (A2*a*) yields the values

$$\frac{\Gamma_o^2}{k} := 4\tan\delta, \qquad \frac{\mathscr{R}_o}{k} := \frac{\beta^2 + k^2}{k^2} \cot 2\delta \qquad (0 < \delta \leqslant \tfrac{1}{6}\pi) \tag{3.11a,b}$$

for Γ^2/k and $\mathscr{R}/k$. The corresponding values of the modulation wavenumber α, phase velocity $c (:= \omega/\alpha)$ and real group velocity $c_g (= \mathrm{d}\omega/\mathrm{d}\alpha$; since $\mathrm{d}q/\mathrm{d}\alpha = 0)$ determined from (A2*b–e*) are given by

$$\alpha = \frac{\beta^2}{8k}(2\cos 2\delta - 1)^{1/2}, \tag{3.11c}$$

$$\left(\frac{\beta^2}{2k^2\Gamma_o^2}\right) c = \frac{\beta^2 + k^2}{k^2(1 - \cos 2\delta)} \; (> 0), \qquad \left(\frac{\beta^2}{2k^2\Gamma_o^2}\right) c_g = \frac{\beta^2 + k^2}{k^2 \cos 2\delta} \; (> 0). \tag{3.11d,e}$$

Substitution of the results into (3.4) gives

$$\hat{\theta}_\pm = -(\hat{U}/4k\mathscr{R}\sin\delta)\left[e^{\mp i\delta} \pm (8k\alpha/\beta^2)e^{\pm i\delta}\right]. \tag{3.12}$$

Hence with $\hat{U}$ real, (3.1) and (2.11) give

$$U = 2\hat{U}_0 \sin\delta\cos\varphi \qquad (\hat{U}_0 := \tfrac{1}{2}\hat{U}/\sin\delta), \tag{3.13a}$$

$$\theta = -\tfrac{1}{2}(\hat{U}_0/k\mathscr{R})\left[\cos\varphi\cos(ky - \delta) - (8k\alpha/\beta^2)\sin\varphi\sin(ky + \delta)\right]. \tag{3.13b}$$

By taking the negative value of the square roots in (3.11*c*) a second solution is

obtained. It can be combined with (3.13) to give the guided wave

$$U = 2\widehat{U}_0 \sin\delta \cos\beta h \cos[\alpha(\zeta - ct)], \tag{3.14a}$$

$$\theta = -\tfrac{1}{2}(\widehat{U}_0/k\mathscr{R})\cos\beta h\{\cos[\alpha(\zeta - ct)]\cos(ky - \delta) - (8k\alpha/\beta^2)\sin[\alpha(\zeta - ct)]\sin(ky + \delta)\} \tag{3.14b}$$

travelling with positive phase velocity ($c > 0$). Like the steady solution (3.10) it vanishes on the lateral boundaries $\beta h = \pm\frac{1}{2}\pi$.

The overstability boundary OA meets the steady neutral curve at A $\left(4/\sqrt{3}, [(\beta/k^2)+1]/\sqrt{3}\right)$, when $\delta = \varepsilon = \frac{1}{6}\pi$. There the solutions (3.14) and (3.10) coincide ($\sin\delta = \sin\varepsilon = \frac{1}{2}, \alpha = 0$). As δ decreases from $\frac{1}{6}\pi$ at A to zero beyond O, so does Γ, while $\mathscr{R}$ increases to infinity. Note, however, that on this boundary we have $\mathscr{R}\Gamma^2 = 2(\beta^2 + k^2)(1 - \tan^2\delta)$ which is less than the value $2(\beta^2 + k^2)$ for the unmodulated oscillatory mode (see below (3.7)). That is why the modulated mode is preferred.

4. Stationary convection on long radial length scales

4.1. *Finite-amplitude uniform states*

When there is no azimuthal spatial modulation the governing equations (2.12), (2.16), (2.17) reduce to

$$\left(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2} - \mathscr{R}^2\Lambda^2\right)\tilde{\theta}_r - k(U + \mathscr{R})\tilde{\theta}_i = 0, \tag{4.1a}$$

$$\left(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2} + k^2\right)\tilde{\theta}_i + k(U + \mathscr{R})\tilde{\theta}_r = 0, \tag{4.1b}$$

where

$$U = -\mathscr{R}\Gamma^2\tilde{\theta}_r, \tag{4.1c}$$

$$\Lambda^2 := (\mathscr{R}\Gamma^2 - k^2)/\mathscr{R}^2, \tag{4.2}$$

and $\tilde{\theta}_r$ and $\tilde{\theta}_i$ are the real and imaginary parts of $\tilde{\theta}(:= \tilde{\theta}_r + \mathrm{i}\tilde{\theta}_i)$. A generalized version of the earlier parametric representation (3.8) of Γ and $\mathscr{R}$ is given by (C1) and (C2) of Appendix C. The particular special case useful here is

$$\frac{\Gamma^2}{k} := \frac{2\Lambda}{\sin 2\varepsilon_\Lambda}, \qquad \frac{\mathscr{R}}{k} := \frac{1}{\Lambda}\tan\varepsilon_\Lambda. \tag{4.3a,b}$$

In the case of the marginal modes, considered in the previous section, the duct solutions (3.10) have the property that

$$U = \tilde{\psi} = \tilde{\theta} = 0 \qquad \text{on} \qquad \pm 2h = \Lambda = \pi/\beta. \tag{4.4}$$

According to (3.8*a*), the most unstable modes occur in the wide-duct limit ($\Lambda \uparrow \infty$ $\beta \downarrow 0$) for which the neutral curve (3.8*b*,*c*) is defined by $\Lambda = 1$. This linear result implies that convection attempts to occupy the longest available radial length scale.

In the long-length-scale context, it is of interest to note that the system (4.1) admits two constant amplitude solutions

$$U = \mathscr{U} := \mathscr{R}(\Lambda - 1) \qquad (\Lambda = \pm|\Lambda|), \tag{4.5a}$$

$$\tilde{\theta} = \Theta := \Theta_r + \mathrm{i}\Theta_i := -\frac{\mathscr{U}}{\mathscr{R}\Gamma^2}\left(1 - \mathrm{i}\frac{\mathscr{R}\Lambda}{k}\right) \equiv -\frac{\mathscr{U}\sin\varepsilon_\Lambda}{k\mathscr{R}\Lambda}\mathrm{e}^{-\mathrm{i}\varepsilon_\Lambda} \tag{4.5b}$$

provided $\mathscr{R}\Gamma^2 > k^2$. By analogy with (3.10*b*), they define the temperature distribution

$$\theta = -(\mathscr{U}/k\mathscr{R}\Lambda)\sin\varepsilon_\Lambda \cos(ky - \varepsilon_\Lambda). \tag{4.6}$$

According to (2.18*b*) it leads to the mean radial heat flux

$$\mathscr{H} = \tfrac{1}{2}\mathscr{U}^2/\mathscr{R}\Gamma^2. \tag{4.7}$$

So at each point in the $(\Gamma^2/k, \mathscr{R}/k)$-plane to the right of the continuous $\Lambda = 0$ curve $\mathscr{R}\Gamma^2 = k^2$ illustrated on figure 2, (4.2) has two distinct solutions $\pm|\Lambda|$. The positive (negative) value (see figure 2*a* (figure 2*b*)) defines a small (large)-amplitude solution with positive (negative) phase lag ε_Λ. According to (4.2), the curves Λ =constant shift to the left (right) as $|\Lambda|$ decreases (increases). They are bounded on the left by the $\Lambda = 0$ curve on which the small- and large-amplitude solutions coincide. Rather than to refer to $\pm|\Lambda|$-solutions, we will parameterize them individually by Λ in the entire range $-\infty < \Lambda < \infty$, while the specification is completed by the value of ε_Λ in the range $-\frac{1}{2}\pi < \varepsilon_\Lambda < \frac{1}{2}\pi$.

In view of (2.15), the temperature distribution (4.6) defines the small ageostrophic contribution

$$\psi - \overline{\psi} = -\gamma\mathscr{R}\theta \tag{4.8}$$

to the streamfunction (the 'Sverdrup relation' (2.16)). Since it is independent of h, it defines a net inflow or outflow at each value of y, reminiscent of salt fingering found in thermohaline convection. Indeed, the stability calculations for modulated disturbances to our finite-amplitude state in §5 below are reminiscent of Holyer's (1981) analysis of the stability of salt fingers to long-length-scale internal wave perturbations. Clearly that net mass flux must eventually be returned, which is why we prefer when possible to apply the duct boundary conditions (4.4). Nevertheless, it is possible that, when Λ is large, (4.5) provides a valid approximation to the solution of (4.1) subject to (4.4) throughout a large part of the flow region, with the return flow being limited to boundary layer regions at the edges. We explore this possibility for small-amplitude solutions in the following subsection.

4.2. *Weakly nonlinear theory*

Taking into account the nature of the steady finite-amplitude solution (4.5), we make the preliminary change of variables

$$X + \mathrm{i}Y := \widetilde{\theta}/\Theta. \tag{4.9}$$

Accordingly the governing equations (4.1*a–c*) become

$$\left[\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2}\right] X - k\left[2\mathscr{R}\Lambda + (U - \mathscr{U})\right] Y = 0, \tag{4.10a}$$

$$\left[\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2} - (\mathscr{R}^2\Lambda^2 - k^2)\right] Y + k(U - \mathscr{U})X = 0 \tag{4.10b}$$

with

$$U = \mathscr{U}\left(X + \frac{\mathscr{R}\Lambda}{k}Y\right). \tag{4.10c}$$

The finite-amplitude equilibrium (4.5) is simply $(X, Y) = (1, 0)$.

We restrict attention to small-amplitude solutions close to critical

$$|\Lambda - 1| \ll 1, \qquad |\mathscr{U}| \ll \mathscr{R}, \tag{4.11a,b}$$

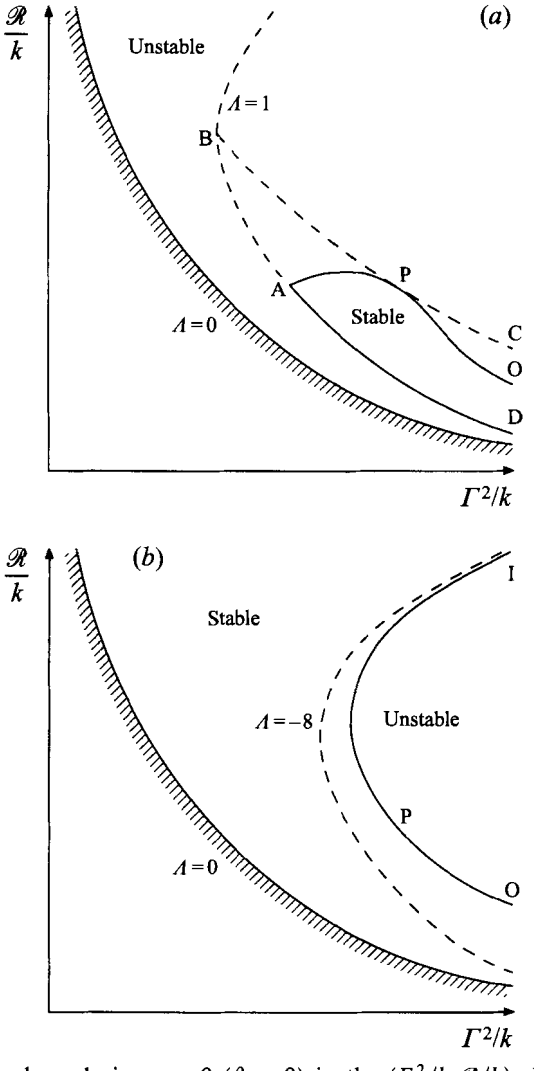

FIGURE 2. The stability boundaries $q = 0$ ($\beta = 0$) in the $(\Gamma^2/k, \mathscr{R}/k)$-plane for finite-amplitude geostrophic flow ($\mathscr{U} \neq 0$), which exists when $\mathscr{R}\Gamma^2 \geqslant k^2$. Equality defines the boundary of this region indicated by the continuous $\Lambda = 0$ curve on which $\mathscr{U} = -\mathscr{R}$. It is also a stability boundary. (*a*) The case $\Lambda > 0$ for which $\mathscr{U} > -\mathscr{R}$. This state is unstable everywhere except within the tongue bounded by the continuous curve OPAD. The curves CB, OA, AD have the same significance as in figure 1. (*b*) The case $\Lambda < 0$ for which $\mathscr{U} < -\mathscr{R}$. Instability sets in as a Hopf bifurcation across the continuous curve IO. It is bounded on its left by the broken curve $\Lambda = -8$.

which vary on slow time and long length scales. In this limit, we have $|Y| \ll |X|$ with (4.10) reducing to

$$\left[\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2}\right] X = 2k\mathscr{R}Y, \tag{4.12a}$$

$$\left[\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2} - (\mathscr{R}^2 - k^2)\right] Y = k\mathscr{U}(1 - X)X, \qquad U = \mathscr{U}X. \tag{4.12b,c}$$

Provided that $\mathscr{R}$ is not close to k (equivalently ε_A is not close to $\frac{1}{4}\pi$), we may neglect

the partial derivatives in (4.12*b*) and obtain

$$\left(\frac{\partial}{\partial t} - \frac{\partial^2}{\partial h^2}\right) X \mp (2K_\pm)^2(1-X)X = 0, \tag{4.13a}$$

where

$$(2K_\pm)^2 := \pm\frac{2k^2\mathscr{R}}{k^2 - \mathscr{R}^2}\mathscr{U} \equiv \pm(\tan 2\varepsilon_\Lambda)\mathscr{U} > 0 \tag{4.13b}$$

with the sign selected to ensure the reality of the appropriate K_+ or K_-. In fact, since $\mathscr{U}$ is positive (negative) on the right (left) of the stability boundary $\Lambda = 1$, K_+ (K_-) is appropriate above (below) both the lower $\mathscr{R} < k$ and upper $\mathscr{R} > k$ stability boundaries (illustrated by DAB and beyond B on figure 2*a*) respectively.

Steady solutions of (4.13) can be constructed in terms of Jacobian Elliptic functions. Consider, for example, the finite-amplitude solution $U = \mathscr{U}$ just above the lower stability boundary DAB, where K_+ is applicable; it can be terminated by the boundary layer

$$U = \tfrac{1}{2}\mathscr{U}[-1 + 3\tanh^2(K_+ h)] \tag{4.14}$$

relative to a suitably chosen h-origin. It is also clear by inspection of (4.13) that the constant finite-amplitude state is stable to small perturbations. On the other hand, the finite-amplitude solution which occurs below the lower stability boundary DAB, where K_- is applicable, cannot be terminated by a boundary layer. Again by inspection this state is unstable.

The above weakly nonlinear results are applicable when the duct width is large compared with the boundary layer width which, in turn, must be large compared with the bump width:

$$k\Lambda \gg K_\pm \Lambda \gg 1. \tag{4.15}$$

The results imply that not all finite-amplitude states can be terminated by boundary layers. They do, however, suggest that their existence is closely linked to the temporal stability. This observation must be interpreted cautiously as the amplitude modulation equations (4.13) are of lower order than the complete system (4.10).

4.3 Temporal stability

To test the stability of the uniform state (4.5), we simply consider perturbations proportional to $\exp(pt)$. Upon substitution into (4.1), or equivalently (4.10), linearization yields the quadratic dispersion relation

$$p^2 + (k^2 - \mathscr{R}^2\Lambda)p + 2k^2\mathscr{R}^2\Lambda(\Lambda - 1) = 0 \tag{4.16a}$$

for p. The roots are complex when and only when Λ satisfies

$$\left|\frac{k}{\Lambda\mathscr{R}} + \frac{3\mathscr{R}}{k}\right| < \left(8\left[1 + \left(\frac{\mathscr{R}}{k}\right)^2\right]\right)^{1/2} \qquad (\mathrm{Im}\, p \neq 0). \tag{4.16b}$$

When $0 < \Lambda < 1$, condition (4.16*b*) is not met, the product of the roots is negative, both are real and one is positive. At $\Lambda = 0$ and 1, one root vanishes, but importantly both roots vanish at the Takens–Bogdanov point B (2, 1), where $\Lambda = (k/\mathscr{R})^2 = 1$ ($\varepsilon_\Lambda = \frac{1}{4}\pi$). Outside the range $0 \leqslant \Lambda \leqslant 1$, the product of the roots is positive, either both roots form a complex conjugate pair or are real with the same sign (depending on whether or not (4.16*b*) is satisfied); when $\Lambda > (<)(k/\mathscr{R})^2$, the sum of the roots is positive (negative) and both have positive (negative) real part; when $\Lambda = (k/\mathscr{R})^2$, the

sum of the roots vanishes and p is purely imaginary, tending to zero as Λ decreases to unity at the Takens–Bogdanov point B.

Close to the stability boundary $\Lambda = 1$ of the static state but away from the Takens–Bogdanov point B, where $\mathscr{R} = k$, the assumptions (4.11) and (4.15) imply that the two real roots of (4.16*a*) are

$$\frac{p}{k^2} = \begin{cases} \mp(2K_\pm)^2\mathscr{U} + O(\mathscr{U}^2), \\ -1 + (\mathscr{R}/k)^2 + O(\mathscr{U}). \end{cases} \tag{4.17a,b}$$

The former small-p case (4.17*a*), obtained by neglecting p^2 in (4.16*a*), corresponds to the slow temporal evolution of small $X - 1$ as governed by (4.13) and discussed in the previous subsection. The latter relatively large-p case (4.17*b*), obtained by neglecting the term proportional to $\Lambda - 1$ in (4.16*a*), defines another direct mode which is stable (unstable) when $\mathscr{R} < (>)k$.

When $\mathscr{R} < k$, the small finite-amplitude state $0 < \Lambda < 1$, which bifurcates subcritically from the lower branch of the stability boundary DAB, is unstable. This direct mode of instability (p real) switches on at $\Lambda = 1$ (4.17*a*) and off at $\Lambda = 0$ (see figure 2*a*). The supercritical bifurcation from the lower boundary DAB, on the other hand, is stable over the range $1 < \Lambda \leqslant (k/\mathscr{R})^2$, specifically

$$\left(\frac{\mathscr{R}}{k}\right)^2 < \frac{\mathscr{R}\Gamma^2}{k^2} - 1 \leqslant \left(\frac{k}{\mathscr{R}}\right)^2. \tag{4.18}$$

Overstability sets in (p imaginary) at the stability boundary $\Lambda = (k/\mathscr{R})^2$ (illustrated by the broken curve BPC) and persists ($\mathrm{Re}\, p > 0$) for all larger values $\Lambda > (k/\mathscr{R})^2$.

When $\mathscr{R} > k$, the bifurcation from the upper branch of the neutral curve (the broken curve beyond B) is unstable for all positive values of Λ; the range of stability defined by (4.18) no longer exists.

From a general point of view, the bifurcation from the neutral curve $\Lambda = 1$ is best understood by considering (4.10) with the spatial derivatives omitted. Evidently, there is a Takens–Bogdanov bifurcation at B, where $\mathscr{R} = k$, $\Lambda = 1$. Our further reduction to (4.12) is appropriate for unfolding that bifurcation (but see, Guckenheimer & Holmes, p. 365). From these general considerations, we see that, following the Hopf bifurcation on BPC, a stable limit cycle emerges. That eventually loses stability via a global bifurcation which occurs when the limit cycle becomes the homoclinic trajectory of infinite period through the unstable zero-amplitude equilibrium. Subsequently, all solutions converge on the large-amplitude equilibrium states.

When $\Lambda \leqslant 0$, the corresponding large-amplitude states are stable, $\mathrm{Re}\, p \leqslant 0$, for all values of $\mathscr{R}$. Indeed, we may speculate that this is generally the preferred equilibrium state and it is characterized by a negative value of the geostrophic velocity

$$\mathscr{U} \leqslant -\mathscr{R}. \tag{4.19}$$

5. Modular stability of the uniform finite amplitude states

5.1. *The dispersion relation*

When the finite-amplitude uniform states (4.5) can be supported over a substantial radial distance, it is of interest to investigate their stability to additional perturbations $\tilde{\theta} - \Theta$ and $U - \mathscr{U}$ of the form (3.1*a*) and (3.1*b*) respectively. The linear equations

(2.10) and (2.14) yield (3.3) as before, while linearization of (2.16) and (2.17) yields

$$(k^2\mathscr{P} \pm \mathrm{i}k\mathscr{R}\Lambda)\hat{\theta}_{\pm} = -(1 - k\Theta_i \pm \mathrm{i}k\Theta_r)\hat{U} \tag{5.1}$$

in place of (3.4*a*). Together (3.3) and (5.1) lead to the quadratic dispersion relation

$$1 + \mathrm{i}\mathscr{A} = \left(\frac{\mathscr{R}\Gamma^2}{2k^2}\right)\left(\frac{1 - k\Theta_i + \mathrm{i}k\Theta_r}{\mathscr{P} + \mathrm{i}\Lambda(\mathscr{R}/k)} + \frac{1 - k\Theta_i - \mathrm{i}k\Theta_r}{\mathscr{P} - \mathrm{i}\Lambda(\mathscr{R}/k)}\right), \tag{5.2}$$

for $\mathscr{P}$ as functions of Γ^2/k, $\mathscr{R}/k$ and $\mathscr{A}$.

As in §3.2 we consider the most unstable modes obtained by choosing $\mathscr{A}$ to maximize the growth rate q for some frequency ω with the other parameters fixed. By inspection of (5.2), the choice of the coefficients in Appendix A is $A = \Lambda(\mathscr{R}/k)$, $B = \Lambda(\mathscr{R}/k)^2 + 1$, $C/A = (-\Lambda + 1)/\Lambda$. The subsequent analysis of the results is cumbersome and distracts from our main theme. We, therefore, relegate the details to Appendix B and simply extract the main results here.

Of course, the fate of unmodulated disturbances (α (or $\mathscr{A}$) $= 0$) can be understood within the framework of the dispersion relation (4.16*a*) with p replaced by $p + \beta^2$ and does not involve the elaborate minimization just described. In that case the most unstable modes are of long length scale, $\beta \to 0$. This continues to be the case for our modulated disturbances. It should be noted, however, that with $\mathscr{A}$ fixed α tends to zero in concert with β. It means that the azimuthal modulation length scale must increase with radial scale. Bearing this limitation in mind, attention is focused below on the marginal modes

$$q = 0 \qquad \text{and} \qquad \beta \to 0. \tag{5.3}$$

We also argue in Appendix B that the neutral curve for our minimizing modes provides us with the stability boundary.

There are two stability boundaries, one for positive ($\mathscr{U} > 0$) and the other for negative ($\mathscr{U} < 0$) geostrophic flows. They are determined remarkably simply by (B8), which yields

$$\Lambda = \frac{2\sin^2\varepsilon_\Lambda}{\sin^2 2\varepsilon_\Lambda - \frac{1}{4}}. \tag{5.4}$$

It leads immediately via (4.3) to the parametric representation

$$\frac{\Gamma^2}{k} = \frac{2\tan\varepsilon_\Lambda}{\sin^2 2\varepsilon_\Lambda - \frac{1}{4}}, \qquad \frac{\mathscr{R}}{k} = \frac{\sin^2 2\varepsilon_\Lambda - \frac{1}{4}}{\sin 2\varepsilon_\Lambda} \tag{5.5a,b}$$

of the stability boundaries in terms of the phase angle ε_Λ. The corresponding mean geostrophic velocity and temperature fields, upon which the instability rides, are given by (4.5) which reduce to

$$\frac{\mathscr{U}}{k} = \frac{(\frac{1}{2} - \cos 2\varepsilon_\Lambda)^2}{\sin 2\varepsilon_\Lambda}, \qquad k\Theta = \frac{(\frac{1}{2} - \cos 2\varepsilon_\Lambda)^2}{2\sin\varepsilon_\Lambda}\mathrm{e}^{-\mathrm{i}\varepsilon_\Lambda}. \tag{5.6a,b}$$

The frequency of the marginal modes are given by (B4), which using (B2) yields

$$\frac{\omega}{k^2} \equiv \mathscr{F} = \frac{(\frac{1}{2} - \cos 2\varepsilon_\Lambda)^{1/2}}{\cos\varepsilon_\Lambda}. \tag{5.7a}$$

Their corresponding phase velocity and real (since $\mathrm{d}q/\mathrm{d}\alpha = 0$) group velocity deter-

mined by (B5) are

$$\left(\frac{\beta^2}{2k^2\Gamma^2}\right)c \equiv \mathscr{C} = \frac{4}{1+2\cos 2\varepsilon_\Lambda}, \qquad \left(\frac{\beta^2}{2k^2\Gamma^2}\right)c_g \equiv \mathscr{C}_g = \frac{4}{3-2\cos 2\varepsilon_\Lambda} \ (>0). \tag{5.7b,c}$$

From these results we may determine the modulation wavenumber

$$2(\Gamma/\beta)^2\alpha \equiv \mathscr{A} := \mathscr{F}/\mathscr{C}. \tag{5.7d}$$

The appropriate ranges of ε_Λ are fixed by the obvious requirements that Γ^2/k, $\mathscr{R}/k$ and $\mathscr{F}^2$ are all positive. The two distinct cases of positive and negative geostrophic flows are discussed in detail in the two subsections which follow. It is perhaps significant that the group velocity defined by (5.7*c*) is always positive.

5.2. *Positive geostrophic flows, $\mathscr{U} > 0$*

We noted in §4.3 that, when $\mathscr{R} < k$, the positive geostrophic flow $\mathscr{U}(>0)$, which sets in following the supercritical instability at $\Lambda = 1$, is stable to unmodulated disturbances on the range (4.18), namely $1 < \Lambda < (k/\mathscr{R})^2$ (see figure 2*a*). A direct mode of instability sets in to the left of the curve BAD ($\Lambda = 1$), while overstability sets in above the curve BPC ($\Lambda = (k/\mathscr{R})^2$). The lines intersect at the Takens–Bogdanov point B (2,1). The nature of the solutions on the line $\Lambda = 1$ is given by the results of §3.2 for vanishing geostrophic flow, $\mathscr{U} = 0$.

Essentially, the results of Appendix B show that the domain of stability (4.18) is eroded when modulated disturbances are considered. Specifically, instability sets in once the stability boundary APO on figure 2(*a*), defined parametrically by (5.5) on the range

$$\tfrac{1}{6}\pi \leqslant \varepsilon_\Lambda < \tfrac{5}{12}\pi \qquad (1 \leqslant \Lambda < \infty), \tag{5.8}$$

is crossed. It begins at A $(4/\sqrt{3}, 1/\sqrt{3})$ $((\varepsilon_\Lambda, \Lambda) = (\frac{1}{6}\pi, 1))$. As Γ^2/k increases so does $\mathscr{R}/k$ until it reaches a maximum at $(\frac{8}{3}, \frac{3}{4})[(\varepsilon_\Lambda, \Lambda) = (\frac{1}{4}\pi, \frac{4}{3})]$. With further increase of Γ^2/k, the value of $\mathscr{R}/k$ decreases indefinitely. The boundary touches and is tangent to the unmodulated stability boundary BPC at P $(4\sqrt{3}, 1/\sqrt{3})$ $((\varepsilon_\Lambda, \Lambda) = (\frac{1}{3}\pi, 3))$. Finally it continues on to infinity beyond O, where

$$\mathscr{R}\Gamma^2/k^2 \sim 2(1+\sqrt{3})^2 \qquad \text{as} \qquad (\varepsilon_\Lambda, \Lambda) \to (\tfrac{5}{12}\pi, \infty) \tag{5.9}$$

in contrast to $\mathscr{R}\Gamma^2/k^2 \sim 1$ on $\Lambda = 1$ beyond D.

On the stability boundary geostrophic flow and wave characteristics are as follows. At A the disturbance is degenerate in as much as both the frequency ω and modulation wavenumber α vanish; yet the phase and group velocities c and c_g are finite:

$$\mathscr{U}/k = \mathscr{F} = \mathscr{A} = 0, \qquad \mathscr{C} = \mathscr{C}_g = 2 \qquad (\Lambda = 1). \tag{5.10}$$

At P the modulation wavenumber again vanishes but the phase velocity is infinite:

$$\mathscr{U}/k = 2/\sqrt{3}, \qquad \mathscr{F} = 2, \qquad \mathscr{A} = 0, \qquad \mathscr{C}_g = 1 \qquad (\Lambda = 3). \tag{5.11}$$

Beyond O we have the asymptotic results

$$\left.\begin{aligned} \mathscr{U}/k &\sim (2+\sqrt{3}), \qquad \mathscr{F} \sim (1+\sqrt{3})^{3/2}, \qquad \mathscr{A} \sim -\tfrac{1}{2}(1+\sqrt{3})^{1/2}, \\ \mathscr{C} &\sim -2(1+\sqrt{3}), \qquad \mathscr{C}_g \sim 2(\sqrt{3}-1)/\sqrt{3} \qquad (\Lambda \uparrow \infty). \end{aligned}\right\} \tag{5.12}$$

The main point is that the phase velocity is positive (negative) on AP (PO) changing sign through infinity at P, while the group velocity is positive everywhere.

5.3. *Negative geostrophic flows, $\mathscr{U} < 0$*

In §4.3 we noted that the negative large-amplitude geostrophic flows $\mathscr{U}$, characterized by $\Lambda < 0$, were always stable to unmodulated disturbances. According to the results of Appendix B, however, instability to modulated disturbances sets in as the stability boundary IO on figure 2(*b*), defined parametrically by (5.5) on the range

$$-\tfrac{1}{2}\pi < \varepsilon_\Lambda < -\tfrac{5}{12}\pi \qquad (-8 > \Lambda > -\infty), \tag{5.13}$$

is crossed to the right. The condition $\Lambda < -8$ with (4.2) gives the crude lower bounds

$$\Gamma^2 > 64\mathscr{R} + k^2/\mathscr{R}, \qquad -\mathscr{U} > 9\mathscr{R} \tag{5.14}$$

on the bump size Γ and the magnitude of the geostrophic flow for instability. The curve $\Lambda = -8$ is illustrated broken on figure 2(*b*).

The stability boundary begins on the upper curve beyond I on figure 2(*b*), where

$$\Gamma^2 \sim 64\mathscr{R} \qquad \text{as} \qquad (\epsilon_\Lambda, \Lambda) \to (-\tfrac{1}{2}\pi, -8). \tag{5.15}$$

There the geostrophic velocity, frequency and modulation wavenumber tend to infinity, while the phase and group velocities remain finite:

$$\left.\begin{array}{lll} \mathscr{U}/k \sim -9\mathscr{R}/k, & \mathscr{F} \sim 4\sqrt{6\mathscr{R}/k}, & \mathscr{A} \sim -\sqrt{6\mathscr{R}/k}, \\ \mathscr{C} \sim -4, & \mathscr{C}_g \sim 4/5 & (\mathscr{R} \uparrow \infty). \end{array}\right\} \tag{5.16}$$

As ε_Λ increases, $-\Lambda$ increases and $\mathscr{R}$ decreases in concert. Simultaneously, Γ decreases to a minimum above that predicted by (5.14) and again increases to infinity. The lower curve beyond O is again given asymptotically by (5.9) as $(\epsilon_\Lambda, \Lambda) \to (-\frac{5}{12}\pi, -\infty)$, while the corresponding wave characteristics continue to be given by (5.12) but the geostrophic velocity changes sign:

$$\mathscr{U}/k \sim -(2+\sqrt{3}) \qquad (\mathscr{R} \downarrow 0). \tag{5.17}$$

The phase velocity c remains negative and the group velocity c_g remains positive everywhere on the stability boundary IO.

Curiously, our results show that, at fixed bump height, stability is regained at sufficiently large Rayleigh number. For that reason we have been careful in Appendix B to establish the coincidence of the neutral curve and the stability boundary. Presumably, the large geostrophic flows, associated with increasing $\mathscr{R}$, are eventually responsible for the stabilization. It is nevertheless important to remember the limitations of the present theory in which many terms (particularly nonlinear) have been ignored and must eventually be important. We comment further on this matter in the concluding remarks of §7.

6. Travelling jets

6.1. *Wave trains*

Here we continue our investigation of the negative uniform geostrophic flows $\mathscr{U}$ initiated in §5.3. Following the Hopf bifurcation across the stability boundary in the $(\Gamma^2/k, \mathscr{R}/k)$-plane a new finite-amplitude state is achieved, which is sensitive to the boundary conditions chosen. For simplicity, we assume that the disturbance is localized on a length scale β^{-1} small compared to the channel width Δ. Far from the lateral boundaries, we may reasonably expect travelling wave solutions dependent only on the variable φ defined by (3.2*a*) to be preferred, which are independent of the remote boundary conditions. Of course, other types of solution are possible

controlled by the boundary conditions. Our objective here, however, is to construct the simplest possible finite-amplitude solution of permanent form.

Under our simple wave assumption the vorticity equation (2.14) reduces to

$$\frac{\partial U}{\partial \varphi} - \mathscr{A}(U - \langle U \rangle) = -\frac{\mathscr{R}\Gamma^2}{2}\left(\frac{\partial \tilde{\theta}}{\partial \varphi} + \frac{\partial \tilde{\theta}^*}{\partial \varphi}\right), \tag{6.1a}$$

while Taylor's condition (2.12) provides the boundary condition

$$\langle U \rangle = -\tfrac{1}{2}\mathscr{R}\Gamma^2(\langle \tilde{\theta} \rangle + \langle \tilde{\theta}^* \rangle), \tag{6.1b}$$

where, since wave train solutions are sought periodic in φ over the 2π-interval, the cylinder average (1.4*b*) reduces to the mean value

$$\langle \ldots \rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ldots \mathrm{d}\varphi. \tag{6.2}$$

If the fluctuating geostrophic flow is expressed in terms of a fluctuating streamfunction

$$U - \langle U \rangle = -\beta \frac{\partial \Psi}{\partial \varphi} \qquad (\langle \Psi \rangle = 0), \tag{6.3a}$$

then integration of (6.1*a*) subject to (6.1*b*) yields

$$U + \mathscr{A}\beta\Psi = -\tfrac{1}{2}\mathscr{R}\Gamma^2(\tilde{\theta} + \tilde{\theta}^*). \tag{6.3b}$$

The vorticity equation (6.1) is supplemented by the heat conduction equation (2.17), which becomes

$$\left(-\beta^2 \frac{\partial^2}{\partial \varphi^2} - \omega \frac{\partial}{\partial \varphi} + k^2\right)\tilde{\theta} + \mathrm{i}k(U + \mathscr{R})\tilde{\theta} = -U. \tag{6.4}$$

Note that k can be scaled out of the governing equations and we take advantage of this property in our discussion of the numerical solution in the subsections below.

The energetics of the system (6.1) and (6.4) helps us to understand the nature of their solution. Upon averaging the product of (6.1a) and Ψ, we obtain with (6.4) the energy balance

$$\mathscr{D}_v = \mathscr{W} = \mathscr{D}_\kappa, \tag{6.5a}$$

where

$$\mathscr{D}_v = \frac{1}{\mathscr{R}\Gamma^2}\langle U^2 \rangle, \qquad \mathscr{W} = -\tfrac{1}{2}\langle U(\tilde{\theta} + \tilde{\theta}^*) \rangle, \qquad \mathscr{D}_\kappa = \left\langle \beta^2 \left| \frac{\partial \tilde{\theta}}{\partial \varphi} \right|^2 + k^2 |\tilde{\theta}|^2 \right\rangle. \tag{6.5b–d}$$

It says that the viscous dissipation $\mathscr{D}_v$ in the Ekman layers balances the rate of working of the buoyancy force $\mathscr{W}$ which equals the the thermal dissipation $\mathscr{D}_\kappa$.

6.2. *Numerical solution*

Periodic solutions of (6.1) and (6.4) were found numerically using a Fourier series representation. Essentially, the linearized marginal solutions existing on the stability boundary were extended to their finite-amplitude states located in the region of instability. Some care is needed in the choice of modulation wavenumber α. In the marginal case it is selected by our earlier linear theory to maximize the growth rate. Then, if Γ is increased with α and $\mathscr{R}$ fixed, the nature of the finite-amplitude propagating wave evolves and its frequency ω changes. Nevertheless, eventually at some second critical value of Γ the wave evaporates. This simply means that, for

larger values of Γ, modes with the prescribed α and $\mathscr{R}$ decay. Nevertheless, at this point in the $(\Gamma^2/k, \mathscr{R}/k)$-plane the system is unstable to modes with different α.

To avoid this difficulty, we selected α/k to take roughly the value that maximized the linear growth rate, rather than fixing it arbitrarily to take a value appropriate to some point on the stability boundary. Of course, this recipe does not necessarily select the preferred mode nor for that matter guarantee stability of the mode selected to other forms of disturbance. Nevertheless it does ensure that, as we increase Γ^2/k at fixed $\mathscr{R}/k$, a finite-amplitude solution is located. During this increase we find that the geostrophic velocity profile sharpens up into a jet-like structure. This is illustrated in figure 3(*a*) for the case

$$\Gamma/k^{1/2} = 192.3\ldots \tag{6.6}$$

with

$$\mathscr{R}/k = 1.377\ldots \qquad (\Lambda = -163.9\ldots), \tag{6.7a}$$

$$\beta/k = 1.0, \qquad \alpha/k = 0.00007541\ldots \qquad (\mathscr{A} = 5.577\ldots). \tag{6.7b}$$

For this set of values the resulting frequency is given by

$$\omega/k^2 \equiv \mathscr{F} = -164.9\ldots. \tag{6.7c}$$

Given Γ^2/k and $\mathscr{R}/k$ the mechanism for determining the maximum linear growth rate is explained in the paragraph below (B5) in Appendix B. The value of $\mathscr{A}$ obtained in this way differs slightly from the value chosen in (6.7*b*). Since there is nothing special about the solution with the maximal linear properties in the nonlinear regime, there is no reason to suppose that the nonlinear solution for maximizing $\mathscr{A}$ is qualitatively different to that for our actual choice.

The numerically calculated mean values

$$\langle U\rangle/k = -107.1\ldots, \qquad \mathscr{D}_v = <U^2>/(\mathscr{R}\Gamma^2) = 1.0038\ldots, \tag{6.8a}$$

$$k\langle\tilde{\theta}_r\rangle = 0.002103\ldots, \qquad k\langle\tilde{\theta}_i\rangle = 1.0009\ldots, \qquad \langle\mathscr{H}\rangle = 0.5019\ldots \tag{6.8b}$$

for our solution may be compared with

$$\mathscr{U}/k = -227.0\ldots, \qquad \mathscr{U}^2/(\mathscr{R}\Gamma^2) = 1.0122\ldots, \tag{6.8c}$$

$$k\Theta_r = 0.00459\ldots, \qquad k\Theta_i = 1.006\ldots, \qquad \mathscr{H} = 0.5061\ldots \tag{6.8d}$$

for the undisturbed uniform state upon which the wave rides. Our numerical evidence at other parameter values suggests that the geostrophic jet illustrated on figure 3(*a*) continues to exist with indefinite increase of Γ and is typical for large-amplitude solutions. On figure 3(*b*) we plot the real and imaginary parts of $k(\tilde{\theta} - \langle\tilde{\theta}\rangle)$ together with the scaled streamfunction $-(\mathscr{A}k\beta/\mathscr{R}\Gamma^2)\Psi$ $(= k\tilde{\theta}_r + kU/\mathscr{R}\Gamma^2)$ for the geostrophic flow. To simplify the notation the vertical axes in figures 3(*a*) and 3(*b*) have been plotted with $k = 1$ but, of course, under the scalings introduced here those results renormalize to all values of k.

It is of interest to compare our results with an asymptotic analysis based on large Γ (see (6.6)), or more precisely

$$\Gamma/k^{1/2} \gg 1, \tag{6.9}$$

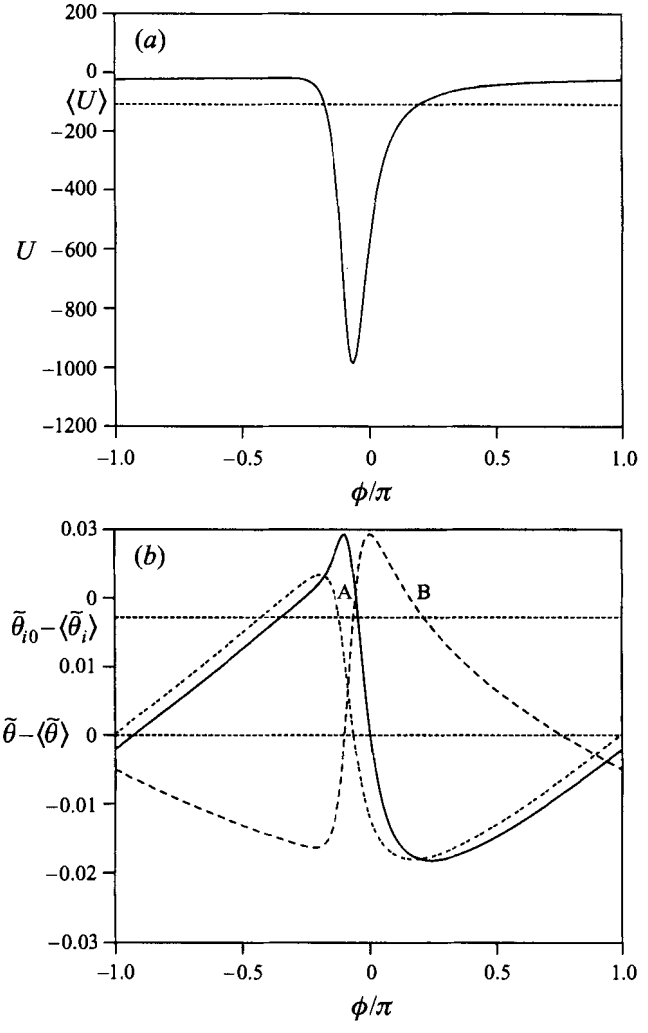

FIGURE 3. The large-amplitude wave train solution characterized by the data (6.6), (6.7) with $k = 1$. (*a*) The geostrophic velocity U plotted versus φ. Its mean value $\langle U \rangle$ is shown by the broken line. The corresponding tranverse velocity is proportional to $U - \langle U \rangle$. (*b*) The fluctuating temperature $\tilde{\theta} - \langle \tilde{\theta} \rangle$ plotted versus φ. The real and imaginary parts are distinguished by the continuous and evenly broken lines respectively. The straight line $\tilde{\theta}_{i0} - \langle \tilde{\theta}_i \rangle$ is drawn broken. It intersects the curve $\tilde{\theta}_i - \langle \tilde{\theta}_i \rangle$ at the points A and B. The scaled streamfunction $\tilde{\theta}_i + (\mathscr{R}\Gamma^2)^{-1}U$ is given by the unevenly broken line.

while the magnitude of the parameters (6.7) suggest the estimates

$$\mathscr{R}/k = O(1) \qquad (\Lambda = O(\Gamma/k^{1/2})), \tag{6.10a}$$

$$\beta/k = O(1), \qquad \alpha/k = O(k/\Gamma^2) \qquad (\mathscr{A} = O(1)), \tag{6.10b}$$

$$\omega/k^2 \equiv \mathscr{F} = O(\Gamma/k^{1/2}). \tag{6.10c}$$

If we set

$$k\tilde{\theta} = \mathrm{i} - \mathscr{F}^{-1}\vartheta, \tag{6.11a}$$

then the solution portrayed in figure 3 is largely consistent with the assumption that

$$\vartheta = O(1), \qquad U/k = O(\Gamma/k^{1/2}) \tag{6.11b}$$

with the φ-scale of order unity. Accordingly, the energy balance (6.5) together with (4.2) and (4.5*a*) gives

$$\langle U^2 \rangle \sim \mathscr{R}\Gamma^2 \sim \mathscr{U}^2 \qquad (\mathscr{D}_v = \mathscr{D}_\kappa \sim 1) \tag{6.12}$$

correct to leading order. Since $\langle U^2 \rangle = \langle U \rangle^2 + \langle (U - \langle U \rangle)^2 \rangle$, the result immediately explains why $\langle U \rangle$ (see (6.8*a*)) is smaller in magnitude than $\mathscr{U}$ (see (6.8*c*)).

Under the assumptions (6.9), (6.10) the leading-order terms in inverse powers of $\Gamma/k^{1/2}$ in the heat conduction equation (6.4) are

$$\frac{\partial \vartheta}{\partial \varphi} - \mathrm{i}\left(\frac{U}{k\mathscr{F}}\right)\vartheta = \frac{\mathscr{R}}{k} - \mathrm{i}. \tag{6.13}$$

All the terms in (6.1) are of comparable size. Nevertheless, upon eliminating the term $\partial\tilde{\theta}_r/\partial\varphi$ by use of the real part of (6.13), it reduces to

$$\frac{\partial U}{\partial \varphi} - \mathscr{A}\chi U = \mathscr{A}(U_0 - \langle U \rangle), \qquad \chi = 1 - \frac{\vartheta_i}{\vartheta_{i0}}, \tag{6.14a,b}$$

where

$$\vartheta_{i0} = \frac{k^2\mathscr{F}^2\mathscr{A}}{\mathscr{R}\Gamma^2}, \qquad \frac{U_0}{\mathscr{F}k} = \frac{\mathscr{R}^2\Gamma^2}{k^3\mathscr{F}^2\mathscr{A}}. \tag{6.14c,d}$$

Though we have dropped the highest derivatives in (6.4), the reduced equations (6.13), (6.14) appear to adequately describe our numerical solution. The higher derivatives are only required to describe internal boundary layers on length scales of order $k^{1/2}/\Gamma$. Even if they exist, they do not appear to control the leading-order solution. Further confirmation of this notion is provided by the energy balance (6.5). Our numerical results (6.8*a*,*b*) show that the high-order diffusive contribution made by $\langle \beta^2 |\partial\tilde{\theta}/\partial\varphi|^2 \rangle$ is negligible (see (6.12)). Therefore, upon assuming that our solution $\vartheta(\varphi), U(\varphi)/k\mathscr{F}, k\mathscr{F}/\mathscr{R}^{1/2}\Gamma$ is governed by (6.13) and (6.14), it is necessarily a function of $\mathscr{A}$ and $\mathscr{R}/k$ alone. Since $\mathscr{R}/k$ is evidently of order unity, the large-amplitude jet-like structure must, therefore, be traced to the moderately large size of $\mathscr{A}$ (see (6.7*b*)).

Corresponding to the data (6.6), (6.7), the formulae (6.14*c*,*d*) give

$$-\mathscr{F}^{-1}\vartheta_{i0} = 0.01806\ldots, \qquad U_0/k = -76.24\ldots. \tag{6.15a,b}$$

The value of $k\tilde{\theta}_{i0}$ $(= 1 - \mathscr{F}^{-1}\vartheta_{i0})$ is important because at that value of $k\tilde{\theta}_i$ the coefficient χ of U in (6.14*a*) vanishes (see (6.14*b*)). Below, we interpret the nature of our numerical solution on the basis that $\mathscr{A}$ is large.

Outside the jet, where $U/k\mathscr{F}$ is small, the advective term $-\mathrm{i}(U/k\mathscr{F})\vartheta$ can be neglected in (6.13), while the term $\partial U/\partial\varphi$ can be neglected in (6.14*a*) on the basis that $\mathscr{A}$ is large. It gives

$$-\omega\frac{\partial\tilde{\theta}}{\partial\varphi} \sim \mathscr{R} - \mathrm{i}k, \qquad U \sim \frac{-U_0 + \langle U \rangle}{\chi}. \tag{6.16a,b}$$

The former, (6.16*a*), says that $\tilde{\theta}$ varies linearly with φ, which is confirmed by figure 3(*b*). Of course, the latter, (6.16*b*), is singular when $\vartheta_i = \vartheta_{i0}$ towards the right-hand end of the jet, where χ vanishes. This means that (6.16*b*) can only be a reasonable approximation of the numerical solution to the right of the point B on figure 3(*b*). As φ

increases up to π, (6.16) indicates that χ and U are roughly proportional to φ and $1/\varphi$ respectively consistent with figure 3(*a*). Beyond $\varphi = \pi$(equivalently $-\pi$), a boundary layer, namely the jet, is eventually reached in which $\partial U/\partial\varphi$ becomes important. This occurs after φ has increased by nearly 2π, by which stage the coefficient $\mathscr{A}\chi$ of U has increased from roughly zero to roughly $2\mathscr{A}$ (see (6.14*b*) and figure 3*b*). This gives an e-folding distance on the left-hand side of the jet of $1/2\mathscr{A}$, which is roughly 0.1 and confirmed by the plot on figure 3(*a*). During this rise in amplitude, χ decreases and eventually vanishes (the point A on figure 3*b*) close to the peak amplitude of the jet. With further increase of φ, the sign of χ changes causing the jet amplitude to decay. While χ is negative (the interval AB on figure 3*b*), χ remains small compatible with the smaller slope on the right-hand side of the jet. As χ again passes through zero to positive values, the solution relaxes back to the exterior state described by (6.16).

Inside the jet boundary layer, the geostrophic velocity is large and the dominant balance in (6.13) is

$$-\omega\frac{\partial\tilde{\theta}}{\partial\varphi} \sim -\mathrm{i}kU(\tilde{\theta} - \mathrm{i}k^{-1}). \tag{6.17a}$$

Its solution has the property that

$$|\tilde{\theta} - \mathrm{i}k^{-1}| = \text{constant}. \tag{6.17b}$$

From a more general point of view, the nonlinear term causes the complex vector $\tilde{\theta} - \mathrm{i}k^{-1}$ to rotate anti-clockwise about the origin with increasing φ, while the complex constant on the right-hand side of (6.13) encourages the linear behaviour exhibited outside the jet.

It must be stressed that the focusing of the jet depends on the large size of $\mathscr{A}$. For smaller values of $\mathscr{A}$ this might not occur. It is not our intention to undertake a comprehensive study of the parameter space in which finite-amplitude solutions exist. We only wish to show by example their existence and to interpret the nature of the particular solution reported here. We believe that the case considered is particularly illuminating for an understanding of the underlying structure of solutions.

6.3. *Physical interpretation*

When the geostrophic flow velocity is large, the isotherms and streamlines almost coincide as evinced by the leading-order balance

$$\mathrm{i}kU\tilde{\theta} = -U \tag{6.18}$$

in the heat conduction equation (2.17), which yields the leading-order result (6.11*a*). Essentially the resulting sinusoidal temperature distribution characterized by $\tilde{\theta}_i$ ($\sim k^{-1}$) is thus forced by the corresponding sinusoidal shape of the geostrophic contours. This feature holds for both the uniform basic state, which exists at large $|A|$, and for the finite-amplitude wave trains, which ride upon it. The smaller-order fine structure of the two types of solution is generally different. To understand the nature of the solutions it is important to remember that motion across the geostrophic contours is of two types. The larger ageostrophic part fluctuates on the length scale of the bumps, while the smaller part is a modulated departure from the geostrophic flow. They are identified by

$$\mathrm{i}k\gamma\tilde{\psi} \qquad \text{and} \qquad \alpha\gamma^2\frac{\partial\Psi}{\partial\varphi} = -\frac{\alpha}{\beta}\gamma^2(U - \langle U\rangle) \tag{6.19}$$

respectively.

In the case of the uniform basic state, a small out-of-phase component of the temperature distribution, characterized by $\tilde{\theta}_r$, is induced by thermal diffusion and corresponds to the balance

$$kU\tilde{\theta}_r = -k^2\tilde{\theta}_i \tag{6.20}$$

in (2.17). It is this out-of-phase component which provides the buoyancy force $\mathscr{R}\tilde{\theta}_r$ necessary to drive the flow against the drag induced by the Ekman boundary layers (see Taylor's condition (2.12)).

For the travelling waves, on the other hand, the situation is far more complicated. Since our wave is modulated on a long length scale, the plane wavefronts are tilted at a small angle $\gamma^2\alpha/\beta$ to the azimuthal y-direction in real space. Since the frequency ω is negative, they should be thought of as waves propagating predominantly radially inwards in the negative x-direction. In view of the small tilt, the effect is to produce a very large phase velocity in the negative y-direction.

Outside the jets, both the short-length-scale ageostrophic motion and the small departures from geostrophy are driven across the geostrophic contours by buoyancy forces. They are characterized by the balances

$$\tilde{\psi}_i \sim -\mathscr{R}\tilde{\theta}_i, \qquad -U + \langle U \rangle = \beta\frac{\partial \Psi}{\partial \varphi} \sim -\frac{\mathscr{R}\Gamma^2}{\mathscr{A}}\frac{\partial \tilde{\theta}_r}{\partial \varphi} \tag{6.21}$$

(see (2.16) and (6.1a) respectively: the 'Sverdrup relations'). The ageostrophic motion convects the basic temperature distribution causing the buoyancy force, responsible for modulating the geostrophic flow, to vary linearly with time:

$$-\omega\frac{\partial \tilde{\theta}_r}{\partial \varphi} \sim \mathscr{R}, \tag{6.22a}$$

while the in-phase component of temperature diffuses causing $\tilde{\theta}_i$, responsible for driving the ageostrophic motion, to decay linearly with time:

$$-\omega\frac{\partial \tilde{\theta}_i}{\partial \varphi} \sim -k \tag{6.22b}$$

(see (6.16a)).

Inside the jet, thermal diffusion is largely unimportant and the temperature readjusts with time in response to the large geostrophic flow (see (6.17a)). Here the ageostrophic flow is unimportant in modifying the temperature.

The energy balance (6.5) is interesting. The dominant contribution to the viscous dissipation $\mathscr{D}_v$ in Ekman layers occurs inside the jets, where U is large. On the other hand, since the temperature perturbation $\tilde{\theta}$ takes the almost constant value ik^{-1} everywhere (see (6.11a), thermal energy dissipation ($\mathscr{D}_\kappa \sim 1$) (see (6.12)) is dominated by losses outside the jet. According to our estimate (6.11b) the rate of working of the buoyancy force $-U\tilde{\theta}_r$ is of order unity. Nevertheless, inspection of figure 3(a,b) shows that the total work $\mathscr{W}(\sim 1)$ is dominated by the jet contribution. The jet is triggered at its forward (left) edge and is accompanied by useful (positive) work. After $\tilde{\theta}_r$ passes through zero the collapse of the jet follows with a smaller amount of destructive (negative) work such that the total work is unity. Though the jet propagates in the negative x-direction, it should be stressed that it is not the propagation direction that is responsible for this forward-edge boundary layer. Rather, it is due to the 'Stommel layer' balance of the two terms on the left-hand side of (6.1a). In contrast with that western coast boundary layer, ours has the character of an eastern boundary layer

(growing to the right). That is because the β-plane effect has the opposite sign in a full sphere to that in a spherical shell! Incidently, there is a comparable consistency about the group velocity of our linear waves. For all our maximizing solutions the group velocity is positive (eastward) for our sphere, just as the group velocity is always westward in a spherical shell. Of course, in application of our results to the region above the Earth's inner solid core, these sign changes are appropriate. Finally, it must not be overlooked that our unbounded duct geometry imposes no restrictions on the modulation length scale. The limitations imposed by finite geometry lead to an additional condition for the applicability of the results presented in this section.

7. Concluding remarks

A pervading motivation for studies of the type undertaken here is to determine the zonal flows set up in convecting rotating systems. Early investigations by Busse & Cuong (1977) with the annulus geometry, Gillman (1977) with spherical shells and Soward (1977) with the full sphere focused attention on the Reynolds stresses set up by the convection to see what differential rotation they drive. One particular application is to the atmospheres of major planets (Busse 1994). In a recent series of papers, which link closely with our present study, Busse & Or (1986), Or & Busse (1987), Schnaubelt & Busse (1992) and Or (1994) showed that mean flows could arise naturally as a secondary instability riding on the thermal Rossby waves (the primary instability of the static state). It is interesting to see that such mean flow arises in our system, albeit in a completely different parameter range, as the preferred primary instability when the radial extent is long enough to accommodate it. This observation is sufficiently important to warrant further comment.

From a general point of view, the issue of the instability leading almost directly to the generation of geostrophic flows has arisen previously with connection to Taylor's condition in the magnetohydrodynamic context. In particular, Soward (1980) considered the possibility that geostrophic motion could result as a secondary instability from its interaction with a familly of oblique convecting rolls (the primary instability). Their existence plays a role very similar to the sinusoidal bumps of our problem and the equations governing stability (cf. Soward 1980, equation (5.6), and (3.3), (3.4) above) have a close resemblance. The idea, therefore, is that the presence of the bumps introduces the crucial longitudinal variation necessary for the mean flow instability, which in the absence of bumps is provided by the thermal Rossby waves.

Or & Busse (1987) discuss the tertiary instability of their finite-amplitude mean flow and comment that the growth rate for unmodulated disturbances ($d = 0$ in their notation) differs very little from its maximum value for some non-zero d. This appears to tie in well with the picture which we developed here in §§4 and 5 for sufficiently large bumps. Specifically, in figure 2(a) at fixed Γ^2 to the right of A, we see that, as $\mathscr{R}$ is increased above the lower stability boundary AD, a stable positive geostrophic mean flow develops. Generally, it is first unstable to a modulated instability as it crosses the curve APO but that is quickly followed by an overstable unmodulated instability as it crosses the broken curve BPC. Indeed, in the geophysical application, the geometrical constraints on the modulation length scale may result in almost entirely emptying the small domains of modular instability bounded by the lines connecting PBA and CPO, as identified in §5. In Busse & Or's case overstability introduces a second frequency, in addition to the thermal Rossby wave frequency; that causes vacillation which leads to increasingly chaotic solutions. In contrast, our unmodulated disturbances are described by a single frequency. They have no radial spatial structure and are deter-

mined by the solutions of (4.1) independent of h. Their temporal evolution is easily understood by eliminating t from (4.1) and considering the phase-plane portraits of the integral curves in the $(\tilde{\theta}_r, \tilde{\theta}_i)$-plane. So on crossing BPC, the focal point $(\Theta_r, \Theta_i)(|\Lambda|)$ (corresponding to the uniform positive geostrophic flow) becomes unstable to a Hopf bifurcation and is then surrounded by a stable limit cycle. As $\mathscr{R}$ is increased the limit cycle expands and its period lengthens. This is eventually terminated by a homoclinic orbit which passes through the unstable static equilibrium (0, 0). A global bifurcation follows and all solutions now converge on the third remainining equilibrium point $(\Theta_r, \Theta_i)(-|\Lambda|)$ (corresponding to the large-amplitude negative geostrophic flow).

Our large-amplitude negative geostrophic flows are particularly interesting. They appear to have no counterpart in the Or & Busse (1987) results and so are peculiar to our bump geometry. It should also be remembered that they can only be reached by the finite jump in states described above. The apparent continuous connection through a subcritical transition of unstable states $1 > \Lambda \geqslant 0$ (or equivalently $0 > \mathscr{U} \geqslant -\mathscr{R}$) suggested by (4.5) is illusiory. When limited to finite radial extent, the weakly nonlinear theory of §4.2 suggests that these modes cannot accommodate the boundary conditions. This leaves our upper branch $0 > \Lambda$ portrayed in figure 2(*b*) unconnected. We may speculate that, as Λ increases through negative values to zero, the influence of remote boundaries increases in importance and the mode of convection approximated by the uniform geostrophic flow $\mathscr{U}(< -\mathscr{R})$ gains a structure which our present theory is inadequate to describe. This branch can now be continued, but presumably remains hanging disconnected from the static state.

It must be stressed that the only nonlinearity retained in our governing equations is the convection of heat by the geostrophic flow. Both the heat transport by the ageostrophic flow and the momentum transport by the Reynolds stresses have been ignored as a result of our asymptotic approximations. These approximations appear to limit the possible temporal structures. Indeed once the large-amplitude geostrophic flow is attained the only further transition found is the modulated wave identified in §5. The ensuing finite-amplitude wave riding on it appears to be stable within the framework of the equations we analysed. Or & Busse (1987) remark that the mean flow instability reduces the radial heat transport below that induced by the thermal Rossby waves. Though we cannot make that comparison, the numerical study reported in §6 shows that the mean heat transport $\langle \mathscr{H} \rangle$ by the modulated instability is also reduced below that produced by the uniform mean flow. Even for the large-amplitude jet the reduction is small (see (6.8*b*,*d*)) but nevertheless the trend was found consistently for all parameters that we investigated.

Finally, we remark briefly on the Bell's (1993) results reported in his thesis, which are not based on the present asymptotic reduction of the governing equations. Instead the entire Busse & Or (1986) quasi-geostrophic equations were employed. Conseqently, the Ekman number remains explicitly in the system and various small values of it were employed. He was thus able to identify the different regimes of linear instability with increasing bump size (at fixed Rayleigh number) from thermal Rossby to the modulated oscillatory and steady flow instabilities isolated here. Following the primary bifurcations to finite-amplitude states, some of his secondary instabilities produced structures very similar to those displayed by Or & Busse (1987) for their mean flow instabilities. His results suggest that, as the magnitude of the convection increases, the sensitivity of the motion to the bump structure decreases. This reinforces the self-evident fact that the large-amplitude results reported here are only applicable for sufficiently small Ekman number.

Peter Bell is grateful to the Science and Engineering Research Council for the award of a Research Studendship. Some of the research was accomplished by Andrew Soward during a one month visit (26 March to 23 April 1994) to the School of Mathematics and Statistics, The University of Sydney, NSW, Australia organized by Professor R. James. Their support and that of the British Council is gratefully acknowledged. Further stimulus for the ideas resulted from a three week visit (26 March to 16 April 1995) to the Courant Institute of Mathematical Sciences, New York, University, USA supported by Professor S. Childress under grant number NSF DMS 94044086. This support and that of SERC grant numbers GR/H03278 and GR/K58388 are all gratefully acknowledged. The study has benefited particularly from discussions with Carlo Barenghi, David Gubbins, Dominique Jault and KeKe Zhang. Comments made by Fritz Busse and two anonymous referees have been particularly helpful.

Appendix A

Many of the stability calculations involve finding the maximum value of the real part $\mathscr{Q}$ of the complex function $\mathscr{P}(\mathscr{A}) = \mathscr{Q} - \mathrm{i}\mathscr{F}$, which solves the quadratic equation

$$1 + \mathrm{i}\mathscr{A} = \tfrac{1}{2}\left(\frac{B + \mathrm{i}C}{\mathscr{P} + \mathrm{i}A} + \frac{B - \mathrm{i}C}{\mathscr{P} - \mathrm{i}A}\right), \tag{A 1}$$

where A, B, C are real constants, while $\mathscr{A}$ is the real parameter upon which $\mathscr{P}$ depends. The maximum is characterized by $\mathrm{d}\mathscr{Q}/\mathrm{d}\mathscr{A} = 0$ and given by the solution of

$$\left[(4A - C)\mathscr{Q} - AB\right]^2 = (C^2 + B^2)(\mathscr{Q}^2 + A^2). \tag{A 2a}$$

The corresponding imaginary part of $\mathscr{P}$ is given by

$$\mathscr{F}^2 = \tfrac{1}{2}\left[(4A - C)A + B\mathscr{Q}\right] - (\mathscr{Q}^2 + A^2). \tag{A 2b}$$

At this maximum, the scaled phase and real group velocities are

$$\mathscr{C} := \frac{\mathscr{F}}{\mathscr{A}}, \qquad \mathscr{C}_g := \frac{\mathrm{d}\mathscr{F}}{\mathrm{d}\mathscr{A}} = \frac{\mathscr{C}}{(\mathscr{C}/\mathscr{Q}) - 1} \tag{A 2c,d}$$

respectively, where

$$\frac{\mathscr{C}}{\mathscr{Q}} = \frac{(4A - C)A + B\mathscr{Q}}{AC + B\mathscr{Q}}. \tag{A 2e}$$

Appendix B

The objective is to find the most unstable modes of (5.2) obtained by choosing $\mathscr{A}$ to maximize the growth rate q, where

$$(q + \beta^2)/k^2 = \mathscr{Q} - 1. \tag{B 1}$$

In our analysis of it below, we find it convenient to introduce the parameter $\mathscr{S}$ defined by

$$\mathscr{S} := \frac{k^2}{\mathscr{R}^2 \Lambda} \equiv \frac{\Lambda}{\tan^2 \varepsilon_\Lambda}. \tag{B 2}$$

With the substitutions proposed in §5.1, the results of Appendix A are as follows.

The condition (A2*a*) characterizing the most unstable modes reduces to

$$\mathscr{G}(\Lambda, \mathscr{S}, \mathscr{Q}) = 0, \tag{B 3a}$$

where

$$\mathscr{G}(\Lambda, \mathscr{S}, \mathscr{Q}) := \left[\mathscr{S}^2 - 2(12\Lambda - 5)\mathscr{S} + 1\right]\mathscr{Q}^2 + 2(\mathscr{S} + 1)(5\Lambda - 1)\mathscr{Q} + (1 - \Lambda)^2. \tag{B 3b}$$

At fixed $\mathscr{Q}$, it defines $\mathscr{S}$ as a function of Λ, and consequently $\mathscr{R}/k$ as a function of Γ^2/k. The corresponding frequency defined by (A2*b*) is

$$\mathscr{F} = \left\{\tfrac{1}{2}\left[-(2\mathscr{Q} - 1)\mathscr{Q}\mathscr{S} + (3\Lambda + \mathscr{Q} - 1)\right]/\mathscr{S}\right\}^{1/2}, \tag{B 4}$$

while, using (A2*e*), the phase and group velocities determined by (A2*c*,*d*) are

$$\mathscr{C} = \frac{\mathscr{Q}(2\mathscr{S} + 5\Lambda + \mathscr{Q} - 1)}{2\mathscr{S} - \Lambda + \mathscr{Q} + 1}, \qquad \mathscr{C}_g = \frac{2\mathscr{C}}{\mathscr{C} - \mathscr{Q}}. \tag{B 5a,b}$$

When the values of Γ^2/k and $\mathscr{R}/k\,(\geqslant k/\Gamma^2)$ are specified the corresponding values of Λ and $\mathscr{S}$ are determined by (4.2) and (B2). Whence $\mathscr{Q}$ is the solution of the quadratic equation (B3). Of course, only those real roots $\mathscr{Q}$ of (B3), which also ensure that the frequency $\mathscr{F}$ of (B4) is real, are acceptable. In turn, $\mathscr{C}$, $\mathscr{C}_g$, $\mathscr{A}$ are determined by (B5) and (5.7*d*). The value of $\mathscr{Q}$ also fixes the magnitude of the maximized growth rate by the formula (B1) for q/k^2.

Throughout §5 it is assumed that the neutral curve $q = \beta = 0$, on which $\mathscr{Q} = 1$, coincides with the stability boundary. That is only the case if the curves with positive growth rate q, for which

$$\mathscr{Q} > 1, \tag{B 6}$$

do not cross it. To see that this does not happen, we note that the crossing of two neighbouring curves $\mathscr{Q}$ and $\mathscr{Q} + \delta\mathscr{Q}$ would lie on the envelope of the one-parameter familly of curves (B3) and be determined by the additional relation $\partial\mathscr{G}/\partial\mathscr{Q} = 0$. Together they determine the envelope

$$(\mathscr{S} + 1)^2/\mathscr{S} = -(1 - \Lambda)^2/\Lambda. \tag{B 7}$$

All points on the envelope lie in the quadrants where Λ and $\mathscr{S}$ have opposite signs. These envelopes are irrelevant because, according to (B2), the product $\Lambda\mathscr{S}$ must be positive.

On the stability boundary, our equation (B3), namely

$$\mathscr{G}(\Lambda, \mathscr{S}, \mathscr{Q}) = 0 \qquad \text{with} \qquad \mathscr{S} = \Lambda/\tan^2\varepsilon_\Lambda, \qquad \mathscr{Q} = 1, \tag{B 8}$$

yields the simple formula (5.4) for Λ in terms of ε_Λ. The other expressions (5.7) follow immediately from (B4), (B5). Note, however, that (B8) has two solutions; the second is $\Lambda = 0$ and, on the curve it defines, the frequency $\mathscr{F}$ is infinite. On neighbouring curves $\mathscr{Q}$ =constant, $\mathscr{F}$ is only real for $\mathscr{Q} < 1$. They correspond to decaying modes which are of no interest to us.

Appendix C

Main dimensionless parameters

Ekman number	$E := \nu/\Omega D^2$	$\ll 1$	(1.2)
Rayleigh number	$Ra := \alpha_T\beta_T g D^2/2\kappa\Omega$	$\ll 1$	(1.7)
Prandtl number	ν/κ		(2.7)

Aspect ratio	D/H	$\ll 1$	(1.15)
Boundary slope	η	$O(1)$	(1.15)
Geostrophic contour tilt	γ	$\ll 1$	(1.12)

Derived dimensionless parameters

Slope parameter	$S := (D/H)\eta$	$\ll 1$	(1.9)
Renormalized Rayleigh number	$\mathscr{R} := Ra/S$	$= O(1)$	(1.10)
Bump parameter	$\Gamma := E^{-1/4}\eta^{1/2}\gamma$	$= O(1)$	(1.13)

Dimensional characteristics

Units of length and time	D and D^2/κ		(2.1)
Bump height	$\eta D h$	$h := x + (\gamma/k)\sin(ky)$	(1.11*a*)
Long azimuthal length	$(D/\gamma^2)\zeta$	$\zeta \equiv \gamma^2 y$	(2.13)
Travelling wave coordinate	$\varphi := \beta h + \alpha\zeta - \omega t$		(3.2*a*)
Bump azimuthal wavenumber	$(1/D)k$	$k = O(1)$	(2.4*a*)
Radial wavenumber	$(1/D)\beta$		(3.2*a*)
Modulation wavenumber	$(\gamma^2/D)\alpha$		(3.2*a*)
or equivalently	$(E^{1/2}/2\eta D)\beta^2\mathscr{A}$	$\mathscr{A} \equiv 2(\Gamma/\beta)^2\alpha$	(3.3*b*)
Complex growth rate	$(\kappa/D^2)p$	$p := q - \mathrm{i}\omega$	(3.2*b*)
or equivalently	$(\kappa/D^2)[k^2\mathscr{P} - \beta^2 - k^2]$	$\mathscr{P} := \mathscr{Q} - \mathrm{i}\mathscr{F}$	(3.4*b*)
		$q + \beta^2 + k^2 \equiv k^2\mathscr{Q}$	(3.4*c*)
		$\omega \equiv k^2\mathscr{F}$	(3.4*d*)
Modulation phase velocity	$(D/\gamma^2\kappa)c$	$c := \omega/\alpha$	(3.2*a*)
or equivalently	$(2\eta D/E^{1/2}\kappa)(k/\beta)^2\mathscr{C}$	$\mathscr{C} := \mathscr{F}/\mathscr{A}$	(A2*c*)
		$c \equiv 2(\Gamma/\beta)^2k^2\mathscr{C}$	(5.7*b*)
Modulation group velocity	$(D/\gamma^2\kappa)c_g$	$c_g := \mathrm{d}\omega/\mathrm{d}\alpha$	(3.2*b*)
or equivalently	$(2\eta D/E^{1/2}\kappa)(k/\beta)^2\mathscr{C}_g$	$\mathscr{C}_g := \mathrm{d}\mathscr{F}/\mathrm{d}\mathscr{A}$	(A2*d*)
(Real case $\mathrm{d}q/\mathrm{d}\alpha = 0$)		$c_g \equiv 2(\Gamma/\beta)^2k^2\mathscr{C}_g$	(5.7*c*)

Parametric representation

Phase-lag angle function	$\varepsilon(\mathscr{Q})$	$\cos^2\varepsilon(\mathscr{Q}) := k^2\mathscr{Q}/(\Gamma^2\mathscr{R})$	(C1*a*)
Direct mode driving function	$\Lambda(\mathscr{Q}) := \frac{1}{2}(\Gamma^2/k)\sin(2\varepsilon(\mathscr{Q}))$		(C1*b*)
Inverse formula	$\Gamma^2/k = 2\Lambda(\mathscr{Q})/\sin(2\varepsilon(\mathscr{Q}))$		(C2*a*)
	$\mathscr{R}/k = (\mathscr{Q}\tan\varepsilon(\mathscr{Q}))/\Lambda(\mathscr{Q})$		(C2*b*)

Stationary convection boundary $p = 0$

Phase-lag angle	$\varepsilon := \varepsilon(\mathscr{Q})$	$\left(\mathscr{Q} = 1 + (\beta/k)^2\right)$	(3.8)
Stability boundary	$\Lambda(\mathscr{Q}) = 1$	$\left(\mathscr{Q} = 1 + (\beta/k)^2\right)$	(3.8)

Nonlinear usage

Phase-lag angle	$\varepsilon_\Lambda := \varepsilon(1)$		(4.3)
Direct mode driving parameter	$\Lambda := \Lambda(1)$	$\Lambda^2 \equiv (\mathscr{R}\Gamma^2 - k^2)/\mathscr{R}^2$	(4.2)

Finite-amplitude constants

Geostrophic velocity	$\mathscr{U} := \mathscr{R}(\Lambda - 1)$		(4.5*a*)
Temperature	$\Theta := \Theta_r + \mathrm{i}\Theta_i$		(4.5*b*)
	$\Theta_r := -\mathscr{U}/(\mathscr{R}\Gamma^2)$		(4.5*b*)
	$\Theta_i := \mathscr{U}\Lambda/(k\Gamma^2)$	$\equiv -\Theta_r \tan\varepsilon_\Lambda$	(4.5*b*)

REFERENCES

ANUFRIEV, A. P. & BRAGINSKY, S. I. 1975 The effect of a magnetic field on the stream of a rotating liquid at a rough surface. *Mag. Gidrod.* **4**, 62–68 (English transl. *Magnetohydrodynamics* **11**, 461–467 (1976)).

ANUFRIEV, A. P. & BRAGINSKY, S. I. 1977 Effect of irregularities of the boundary of the Earth's core on the speed of the fluid and on the magnetic field. III. *Geomag. Aeron.* **17**, 742–750 (English transl. pp. 492–496).

BELL, P. I. 1993 The effect of bumps on convection in the Earth's core. PhD dissertation, University of Newcastle upon Tyne.

BUSSE, F. H. 1970 Thermal instabilities in rapidly rotating systems. *J. Fluid Mech.* **44**, 441–460.

BUSSE, F. H. 1994 Convection driven zonal flows and vortices in the major planets. *Chaos* **4**, 123–134.

BUSSE, F. H. & CUONG, P. G. 1977 Convection in rapidly rotating spherical fluid shells. *Geophys. Astrophys. Fluid Dyn.* **8**, 17–44.

BUSSE, F. H. & HOOD, L. L. 1982 Differential rotation driven by convection in a rapidly rotating annulus. *Geophys. Astrophys. Fluid Dyn.* **21**, 59–74.

BUSSE, F. H. & OR, A. C. 1986 Convection in a rotating cylindrical annulus: thermal Rossby waves. *J. Fluid Mech.* **166**, 173–187.

BUSSE, F. H. & WICHT, J. 1992 A simple dynamo caused by conductivity variations. *Geophys. Astrophys. Fluid Dyn.* **64**, 135–144.

CARRIGAN, C. R. & BUSSE, F. H. 1983 An experimental and theoretical investigation of the onset of convection in rotating spherical shells. *J. Fluid Mech.* **126**, 287–305.

EWEN, S. A. & SOWARD, A. M. 1994 Phase mixed rotating magnetoconvection and Taylor's condition I. Amplitude equations. *Geophys. Astrophys. Fluid Dyn.* **77**, 209–230.

FEARN, D. R. & PROCTOR, M. R. E. 1992 Magnetostrophic balance in non-axisymmetric, non-standard dynamo models. *Geophys. Astrophys. Fluid Dyn.* **67**, 117–128.

GILLMAN, P. H. 1977 Non-linear dynamics of Boussinesq convection in a deep rotating spherical shell. *Geophys. Astrophys. Fluid Dyn.* **8**, 93–135.

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

GUBBINS, D. & RICHARDS, M. 1986 Coupling of the core dynamo and mantle: thermal or topographic? *Geophys. Res. Lett.* **13**, 1521–1524.

GUCKENHEIMER, J. & HOLMES, P. 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields.* Springer.

HIDE, R. 1967 Motions of the Earth's core and mantle and variations of the main geomagnetic field. *Science* **157**, 55–56.

HIDE, R. 1989 Fluctuations in the Earth's rotation and the topography of the core–mantle interface. *Phil. Trans. R. Soc. Lond.* A **328**, 351–363.

HIDE, R., CLAYTON, R. W., HAGER, B. H., SPIETH, M. A. & VOORHIES, C. V. 1993 Topographic core–mantle coupling and fluctuations in the Earth's rotation. In *Relating Geophysical Structures and Processes, The Jeffreys Volumes* (ed. K. Aki & R. Dmowska), vol. 76, pp. 107–120. AGU/IUGG.

HOLYER, J. Y. 1981 On the collective instability of salt fingers. *J. Fluid Mech.* **110**, 195–207.

JAULT, D. & LE MOUËL, J.-L. 1989 The topographic torque associated with a tangentially geostrophic motion at the core surface and inferences on he flow inside the core. *Geophys. Astrophys. Fluid Dyn.* **48**, 273–296.

JAULT, D. & LE MOUËL, J.-L. 1993 Circulation in the liquid core and coupling with the mantle. *Adv. Space Res.* **13**, 11, 221–233.

KUANG, W. & BLOXHAM, J. 1993 On the effect of boundary topography on flow in the Earth's core. *Geophys. Astrophys. Fluid Dyn.* **72**, 161–195.

MOFFATT, H. K. & DILLON, R. F. 1976 The correlation between gravitational and geomagnetic fields caused by interaction of the core fluid motion with a bumpy core–mantle interface. *Phys. Earth Planet. Inter.* **13**, 68–78.

MUNK, W. H. 1950 On the wind-driven ocean circulation. *J. Met.* **7**, 79–93.

OI, A. C. 1994 Chaotic transitions of convection rolls in a rapidly rotating annulus. *J. Fluid Mech.* **261**, 1–19.

OR, A. C. & BUSSE, F. H. 1987 Convection in a rotating cylindrical annulus. Part 2. Transitions to asymmetric and vascillating flow. *J. Fluid Mech.* **174**, 313–326.

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

ROBERTS, P. H. 1968 On the thermal instability of a rotating fluid sphere containing heat sources. *Phil. Trans. R. Soc. Lond.* A **263**, 93–117.

ROBERTS, P. H. 1988 On topographic core–mantle coupling. *Geophys. Astrophys. Fluid Dyn.* **44**, 181–187.

ROBERTS, P. H. & SOWARD, A. M. 1992 Dynamo theory. *Ann. Rev. Fluid Mech.* **24**, 459–512.

SCHNAUBELT, M. & BUSSE, F. H. 1990 Convection in a rotating cylindrical annulus with rigid boundaries. In *Nonlinear evolution of spatio–temporal structures in dissipative continuous systems* (ed. F. H. Busse & L. Kramer). NATO ASI Series B, vol. 225, pp. 67–72. Plenum.

SCHNAUBELT, M. & BUSSE, F. H. 1992 Convection in a rotating cylindrical annulus. Part 3. Vacillating and spatially modulated flows. *J. Fluid Mech.* **245**, 155–173.

SOWARD, A. M. 1977 On the finite amplitude thermal instability of a rapidly rotating fluid sphere. *Geophys. Astrophys. Fluid Dyn.* **9**, 19–74.

SOWARD, A. M.1980 Finite-amplitude thermal convection and geostrophic flow in a rotating magnetic system. *J. Fluid Mech.* **98**, 449–471.

STEWARTSON, K. 1957 On almost rigid rotations. *J. Fluid Mech.* **3**, 17–26.

STOMMEL, H. 1948 The westward intensification of wind-driven ocean currents. *Trans. Am. Geophys. Union* **99**, 202–206.

SUN, Z.-P., SCHUBERT, G. & GLATZMAIER, G. A. 1994 Numerical simulations of thermal convection in a rapidly rotating spherical shell cooled inhomogeneously from above. *Geophys. Astrophys. Fluid Dyn.* **75**, 199–226.

SVERDRUP, H. U. 1947 Wind-driven currents in a baroclinic ocean; with application to the equatorial currents of the eastern Pacific. *Proc. Natl Acad. Sci.* **33**, 318–326.

TAYLOR, J. B. 1963 The magneto-hydrodynamics of a rotating fluid and the Earth's dynamo problem. *Proc R. Soc. Lond.* A **274**, 274–283.

ZHANG, K. 1992 Spiralling columnar convection in rapidly rotating spherical fluid shells. *J. Fluid Mech.* **236**, 535–556.

ZHANG, K. & GUBBINS, D. 1992 On convection in the Earth's core driven by lateral temperature variations in the lower mantle. *Geophys. J. Intl* **108**, 247–255.

ZHANG, K. & GUBBINS, D. 1993 Convection in a rotating spherical shell with an inhomogenious temperature boundary at infinite Prandtl number. *J. Fluid Mech.* **250**, 209–232.