J. Fluid Mech. (2000), *vol.* 405, *pp.* 157–179. Printed in the United Kingdom

The onset of thermal convection in a rapidly rotating sphere

By CHRIS A. JONES, ANDREW M. SOWARD AND ALI I. MUSSA

School of Mathematical Sciences, University of Exeter, North Park Road, Exeter, Devon EX4 4QE, UK
e-mail: C.A.Jones@ex.ac.uk; A.M.Soward@ex.ac.uk; A.I.A.Mussa@ex.ac.uk

(Received 1 March 1999 and in revised form 8 October 1999)

The linear stability of convection in a rapidly rotating sphere studied here builds on well established relationships between local and global theories appropriate to the small Ekman number limit. Soward (1977) showed that a disturbance marginal on local theory necessarily decays with time due to the process of phase mixing (where the spatial gradient of the frequency is non-zero). By implication, the local critical Rayleigh number is smaller than the true global value by an $O(1)$ amount. The complementary view that the local marginal mode cannot be embedded in a consistent spatial WKBJ solution was expressed by Yano (1992). He explained that the criterion for the onset of global instability is found by extending the solution onto the complex s-plane, where s is the distance from the rotation axis, and locating the double turning point at which phase mixing occurs. He implemented the global criterion on a related two-parameter family of models, which includes the spherical convection problem for particular $O(1)$ values of his parameters. Since he used one of them as the basis of a small-parameter expansion, his results are necessarily approximate for our problem.

Here the asymptotic theory for the sphere is developed along lines parallel to Yano and hinges on the construction of a dispersion relation. Whereas Yano's relation is algebraic as a consequence of his approximations, ours is given by the solution of a second-order ODE, in which the axial coordinate z is the independent variable. Our main goal is the determination of the leading-order value of the critical Rayleigh number together with its first-order correction for various values of the Prandtl number.

Numerical solutions of the relevant PDEs have also been found, for values of the Ekman number down to 10^{-6}; these are in good agreement with the asymptotic theory. The results are also compared with those of Yano, which are surprisingly good in view of their approximate nature.

1. Introduction

The onset of thermal convection in a rapidly rotating sphere has been the subject of many papers. Partly this is due to the relevance of the problem for geophysics and astrophysics; in many planetary cores, and also in many stars, convection is believed to occur. The pattern of this convection will in many cases be dominated by rotational effects, while viscous effects play a less important role. In astrophysical and geophysical problems, the Rayleigh number R for convection is usually well beyond the value at the onset of convection; nevertheless, without an understanding

of the linear problem, it is hard to put nonlinear calculations on a firm foundation. Furthermore, perhaps the most important aspect of the linear theory is the form of convection selected at onset. This may be modified at high Rayleigh number, but nevertheless the form at onset gives the framework of discussion for nonlinear effects.

The second reason why this problem has attracted attention is that it has interesting mathematical properties, so that although the problem seems at first sight similar to many classical stability problems that were solved in the late 19th and early 20th centuries, in fact it has distinctive features that have hitherto prevented its full solution.

The classical problem is defined by a sphere filled with Boussinesq fluid; it is uniformly heated throughout and maintained at constant temperature at the surface of the sphere. We will consider the case of stress-free boundary conditions. Gravity is assumed to act radially inward, with the local acceleration due to gravity proportional to distance from the centre, as appropriate for self-gravitating uniform-density fluid. Rapid rotation means that the Taylor number, Ta, is large, or, equivalently, the Ekman number, E, is small. In this limit it is convenient to refer the system to cylindrical polar coordinates in which s measures distance from the rotation axis and z measures the axial distance from the equatorial plane. The mathematical problem is to find the critical Rayleigh number, and elucidate the form of the convection, in the asymptotic limit of large Ta. This is the problem we shall address in this paper (see § 2).

There are many variations on this classical problem. One of the most important is the case where the fluid lies between concentric spherical shells, rather than within a single sphere. We shall see that in many cases, though not all, the present asymptotic theory can be extended to cover the concentric shell problem too. Other variants can be formulated by changing the internal heating, to (say) a fixed flux of heat emerging from an inner core, or changing the thermal boundary conditions. The mechanical boundary conditions can be varied too, but in the limit of small E, there is no leading-order distinction between rigid boundaries and stress-free boundaries (Zhang & Jones 1993), as the Ekman layers make only a higher-order correction to the critical Rayleigh number. Yet other variants can be constructed, where, for example, the forcing is not gravitational but centrifugal, as is often the case in laboratory experiments (Busse & Carrigan 1976).

Early attempts at the problem concentrated on axisymmetric modes (see e.g. Chandrasekhar 1961). An important advance was made by Roberts (1968), and extended by Busse (1970), who were the first to recognize that the important modes in the rapidly rotating system are non-axisymmetric with the convection characterized by a 'cartridge belt' pattern when the Prandtl number P is order unity (see § 3.1). That is to say, the convective rolls have comparatively small azimuthal and radial extent and are localized in the vicinity of a circular cylinder $s = s_L$ (say), whose radius is roughly one half the radius of the whole sphere. These discoveries were nicely illustrated by figure 1 of Busse (1970), which has been reproduced on many occasions.

The basis of the Roberts–Busse local theory is the assumption that the azimuthal dependence is sinusoidal and that convection onsets in the neighbourhood of a critical cylinder $s = s_L$ where the radial dependence is sinusoidal. Thus the axial structure of the convective rolls is governed by a second-order differential equation in the axial coordinate z. It is solved subject to the condition that the radial flow normal to the boundary is zero, or equivalently (within the local theory) that the temperature perturbation there vanishes. The solution leads to a local dispersion relation for the complex growth rate $-\mathrm{i}\omega$. The marginal solutions ($\mathrm{Im}\{\omega\} = 0$) corresponding to the

lowest critical Rayleigh number were listed in table 1 of Busse (1970), and correspond to a wavenumber k in the radial direction whose critical value is zero (see § 3.2).

Yano (1992) noted that it was not possible to embed the Roberts–Busse local theory into a WKBJ solution which decayed exponentially in the radial s-direction on both sides of the critical cylinder $s = s_L$ (see also Kida 1994). This notion builds on the view expressed by Soward (1977) that small disturbances at values of the Rayleigh number just above those given by the local theory can grow initially, but cannot be sustained; the reason may be traced to the non-zero radial frequency gradient $\partial\omega/\partial s$, which leads to phase mixing. Essentially, disturbances continually decrease their radial s-length scale, and the resulting enhanced dissipation ultimately leads to their decay. Put another way, in *Fourier transform space* the region occupied by the initially growing transformed disturbance is advected out of that unstable region and subsequently decays. This is analogous to the effect of a non-zero group velocity $\partial\omega/\partial k$, which also convects growing disturbances out of their unstable regions but here in *real space*. Though this latter effect occurs in many stability problems with spatially varying backgrounds (see e.g. Huerre & Monkewitz 1990), it is not the prime consideration in our problem because here the local theory gives zero group velocity.

A similar phase mixing problem occurs for the linear stability of flow in the narrow gap between two rotating concentric spheres, the spherical Taylor–Couette problem; Soward & Jones (1983) found the critical Taylor number for the onset of instability, and again it is an $O(1)$ amount greater than that predicted by local theory. Yano (1992) applied the Soward & Jones technique to a modified version of the rotating convection problem. In order to make progress he considered a model with flattened geometry. At any radial distance s from the rotation axis, the local configuration closely resembled Busse's (1970) annulus with cylindrically radial gravity and tilted boundaries of small inclination. However, by considering radial variations, such as curvature of the boundaries, Yano was able to capture the essential features of the true spherical geometry. His analysis involved two parameters, one of which measuring boundary tilt needed to be small to justify his perturbation method. However, since the critical cylinder has radius roughly half that of the sphere, the boundaries are actually inclined at roughly 45° to the equatorial plane, which clearly does not define a small tilt angle.

The key step in Yano's (1992) implementation of the ideas was to recognize that to find a solution of WKBJ type, which decays exponentially on both sides of the critical cylinder, it is necessary to find a double turning point in the complex s-plane. At this double turning point $s = s_c$ both the complex phase mixing and the complex group velocity vanish. As in the local theory, the vanishing of the group velocity is a trivial matter, which is met by $k = k_c = 0$. So in the neighbourhood of s_c the associated s-length scale is long. The central issue is to find the double turning point in the complex s-plane located where phase mixing vanishes. A single turning point suffers from the undesirable feature of the Stokes phenomenon, namely that as WKBJ solutions pass from one Stokes domain to another, an unwanted second solution is picked up, whose magnitude is proportional to the Stokes constant. The second solution is 'unwanted' in the sense that it does not decay and meet the right boundary conditions. It was shown in Soward & Jones (1983) that the existence of a double turning point ensures that by fine tuning the appropriate parameters the Stokes constant can be adjusted to zero, so allowing a uniformly valid WKBJ solution to exist. The critical Rayleigh number required is therefore that at which the double turning point comes into existence (see § 3.3).

The advantage gained from Yano's (1992) small-tilt expansion is that the dispersion relation $\omega = \omega(s,k)$ is simply an algebraic expression, just as it is for the Busse annulus (Busse 1986). In contrast, in our approach to the full spherical problem we must first solve the second-order ordinary differential equation, mentioned above, governing the z-structure. This must be done numerically for various complex values of s and k to extract the required details of the dispersion relation, including the important complex location of the double turning point. Interestingly, that is not a straightforward task. So like Yano, we first solve a model problem with two adjustable parameters, but then vary them until we obtain solutions to the actual spherical problem (see § 3.4).

Surprisingly, despite the limitations of Yano's (1992) small-tilt approximation, he obtained critical values R_{crit} of the Rayleigh number quite close to those which we find here (see table 2).

Zhang (1992) performed numerical calculations on the problem of rotating convection between concentric spherical shells at large but finite Taylor number. Because the solutions are concentrated on the critical cylinder at large Ta, we would expect that these solutions can be converted into approximate equivalent solutions of the full sphere problem. These calculations indicated that the Roberts–Busse local theory was only approximately correct, with a suggested error of the order of 25% for the critical Rayleigh number (see tables 1 and 2). For the purpose of comparing further our asymptotic results with numerical computations at large but finite Ta, we have repeated these calculations, and extended them to the case of whole sphere convection (see § 5). Our computational results confirm Zhang's numerical values (see table 4). A most encouraging finding is that the numerical results are in excellent agreement with our asymptotic theory, and clearly show that the local theory results do not give the correct values of the Rayleigh number, azimuthal wavenumber or frequency at onset (see also tables 5 and 6). Significantly, our asymptotic theory predicts that the critical cylinder $s = s_M$, upon which the convection amplitude is maximized, is significantly further from the axis than the local theory value $s = s_L$ (see § 4 and tables 1 and 3). On the critical cylinder, the convection rolls are tilted, a feature that has led to the term columnar-spiralling modes being used to describe them. Our asymptotic determination of the critical cylinder location and roll tilt angle is again both in agreement with the results of Zhang (1992) and our new numerical results.

The critical Taylor numbers predicted by Soward & Jones (1983) for the spherical Taylor–Couette problem turned out to be only a few percent larger than those predicted by the local theory, so that direct numerical methods would find it very difficult to discriminate between the two theories. In the current convection problem, the difference between the two theories is much larger, and the direct numerical calculations come down decisively in favour of the Soward & Jones (1983) theory. Interestingly the disparity between local and global theory was noted in a different context by Killworth (1980), whose analysis also relied on the isolation of a double turning point in the complex plane.

Before continuing, we emphasize that our paper does not address the issue of wall-attached modes (Zhang & Busse 1987), which are relatively high-frequency inertial waves attached to the equatorial region of the outer boundary (Zhang 1993). This raises issues beyond the scope of the present analysis. At fixed large Taylor number, the wall-attached mode may be preferred at sufficiently small Prandtl number. This feature is consistent with our numerical results, which show the increasing importance of the outer boundary condition as the Prandtl number is decreased. The competition between the wall-attached modes and the internal columnar-spiralling modes has been addressed by Hirsching & Yano (1994).

2. The governing equations

We consider a sphere of radius r_0 rotating about an axis with angular velocity Ω. Coordinates are chosen so that the axis of rotation is the z-axis. For the asymptotic theory cylindrical polar coordinates (s, ϕ, z) are used, while for direct numerical calculations spherical polar coordinates (r, ϑ, ϕ) are more convenient. The sphere is filled with Boussinesq fluid with coefficient of thermal expansion α, kinematic diffusivity ν and thermal diffusivity κ. Since the fluid contains a homogeneous heat source, the static temperature gradient in the absence of convection is $-\beta \boldsymbol{r}$. The gravity field is assumed to be $-g\boldsymbol{r}$. The linearized equation of motion is then

$$E\frac{\partial \boldsymbol{u}}{\partial \boldsymbol{t}} + \hat{\boldsymbol{z}} \times \boldsymbol{u} = -\nabla p + ER\theta \boldsymbol{r} + E\nabla^2 \boldsymbol{u}, \tag{2.1}$$

where the unit of time is r_0^2/ν, the viscous time, the unit of length is the sphere radius r_0 and the unit of temperature is $\beta \nu r_0^2/\kappa$. The linearized temperature equation is

$$P\frac{\partial \theta}{\partial t} = \boldsymbol{u} \cdot \boldsymbol{r} + \nabla^2 \theta, \tag{2.2}$$

and the equation of continuity is

$$\nabla \cdot \boldsymbol{u} = 0. \tag{2.3}$$

The dimensionless parameters are the Ekman number E, the Rayleigh number R and the Prandtl number P defined by

$$E = \frac{\nu}{2\Omega r_0^2}, \qquad R = \frac{g\alpha\beta r_0^6}{\kappa\nu}, \qquad P = \frac{\nu}{\kappa}.$$

Note that this definition of R corresponds to that used by Chandrasekhar (1961). The other commonly employed parameter, the Taylor number $Ta = 4\Omega^2 r_0^4/\nu^2$, is related to the Ekman number by $Ta = E^{-2}$.

If we set

$$\boldsymbol{u} = \nabla \times \Psi \hat{\boldsymbol{z}} + \nabla \times \nabla \times \xi \hat{\boldsymbol{z}} = \boldsymbol{u}_H + W\hat{\boldsymbol{z}},$$

the z-component of the curl of the momentum equation becomes

$$E\left(\frac{\partial}{\partial t} - \nabla^2\right)\nabla_H^2 \Psi + \frac{\partial W}{\partial z} = ER\frac{\partial \theta}{\partial \phi} \tag{2.4}$$

and the z-component of the double curl of the momentum equation is

$$E\left(\frac{\partial}{\partial t} - \nabla^2\right)\nabla^2 W - \frac{\partial}{\partial z}(\nabla_H^2 \Psi) = ER\left[z\nabla_H^2 \theta - \frac{1}{s}\frac{\partial}{\partial s}\left(s^2\frac{\partial \theta}{\partial z}\right)\right], \tag{2.5}$$

where

$$\nabla_H^2 \equiv \frac{1}{s}\frac{\partial}{\partial s}\left(s\frac{\partial}{\partial s}\right) + \frac{1}{s^2}\frac{\partial^2}{\partial \phi^2}.$$

The temperature equation (2.2) becomes

$$\left(P\frac{\partial}{\partial t} - \nabla^2\right)\theta = \frac{\partial \Psi}{\partial \phi} + zW + s\frac{\partial^2 \xi}{\partial s \partial z}, \qquad \nabla_H^2 \xi = -W. \tag{2.6}$$

While the form (2.4)–(2.6) of the equations is useful for the asymptotic analysis, for

direct numerical simulation at finite E, a decomposition based on spherical, rather than cylindrical, polar coordinates is required. Accordingly, following Roberts (1968), we employ the toroidal–poloidal representation

$$\boldsymbol{u} = \nabla \times \mathscr{T}\boldsymbol{r} + \nabla \times \nabla \times \mathscr{S}\boldsymbol{r}. \tag{2.7}$$

We insert this expansion into the radial components of the curl and double curl of (2.1) to obtain

$$E\left(\frac{\partial}{\partial t} - \nabla^2\right) Ł^2 \mathscr{T} - \frac{\partial \mathscr{T}}{\partial \phi} + Q\mathscr{S} = 0, \tag{2.8}$$

$$E\left(\frac{\partial}{\partial t} - \nabla^2\right) Ł^2 \nabla^2 \mathscr{S} - \frac{\partial}{\partial \phi}\nabla^2 \mathscr{S} - Q\mathscr{T} + ER\, Ł^2 \theta = 0, \tag{2.9}$$

where

$$Ł^2 = -\frac{1}{\sin\vartheta}\frac{\partial}{\partial \vartheta}\left(\sin\vartheta \frac{\partial}{\partial \vartheta}\right) - \frac{1}{\sin^2\vartheta}\frac{\partial^2}{\partial \phi^2} \qquad \text{and} \qquad Q = \hat{\boldsymbol{z}} \cdot \nabla - \tfrac{1}{2}(Ł^2 \hat{\boldsymbol{z}} \cdot \nabla + \hat{\boldsymbol{z}} \cdot \nabla Ł^2).$$

The temperature equation becomes

$$P\frac{\partial \theta}{\partial t} = Ł^2 \mathscr{S} + \nabla^2 \theta. \tag{2.10}$$

We use stress-free boundary conditions, as at finite E we expect these to reach the asymptotic limit more rapidly than in the rigid boundary case. The boundary conditions are then

$$\mathscr{S} = \theta = \frac{\partial^2 \mathscr{S}}{\partial r^2} = \frac{\partial}{\partial r}\left(\frac{\mathscr{T}}{r}\right) = 0, \tag{2.11}$$

all applied at $r = r_0$. At $r = 0$, regularity conditions apply, so that only solutions finite as $r \to 0$ are admitted. In practice, these conditions are enforced by the choice of expansion functions.

The system (2.8)–(2.11) formulated by Roberts (1968) is solved numerically in § 5 below.

3. The asymptotic analysis

3.1. *The Roberts–Busse formulation*

Starting with the equations in the form (2.4)–(2.6), we are guided by Roberts (1968) and Busse's (1970) results concerning the appropriate scaling and so make the ansatz

$$\frac{1}{s}\frac{\partial}{\partial \phi} \sim \frac{\partial}{\partial s} \sim O\left(E^{-1/3}\right), \qquad \frac{\partial}{\partial z} \sim O(1) \quad \text{as } E \downarrow 0,$$

which is appropriate to the mainstream exterior to any boundary layers. To obtain the dispersion relation, we assume that disturbances are of the form

$$\exp\left[\mathrm{i}(Ks + M\phi - \Omega t)\right],$$

where Ω is in general complex. At leading order, this implies that both

$$\nabla^2,\ \nabla_H^2 \ \to\ -K^2 - \frac{M^2}{s^2} = -A^2$$

so that using (2.4)–(2.5) the z-dependence of the amplitudes is governed by

$$E(A^2 - \mathrm{i}\Omega)(A^2\Psi) - \frac{\mathrm{d}W}{\mathrm{d}z} = -\mathrm{i}MER\theta, \tag{3.1}$$

$$E(A^2 - \mathrm{i}\Omega)W - \frac{\mathrm{d}\Psi}{\mathrm{d}z} = zER\theta \tag{3.2}$$

and (2.6) becomes

$$(A^2 - \mathrm{i}P\Omega)\theta = \mathrm{i}M\Psi + zW. \tag{3.3}$$

From (3.3) we can deduce that the boundary conditions $\boldsymbol{u}\cdot\boldsymbol{r} = 0$ and $\theta = 0$ on the surface of the sphere are equivalent. This is remarkably fortuitous, because it means that the thermal boundary layer is only triggered by the small z-derivatives of the mainstream solution that we construct. It is consequently very weak and does not even influence the $O(E^{1/3})$ correction terms (3.16*b*) which we consider later. Equations (3.1)–(3.3), together with this boundary condition define the dispersion relation giving Ω as a function of all the parameters in the problem. We can now scale E out of the dispersion relation, as expected for our $E \downarrow 0$ limit. The scalings are

$$\left.\begin{array}{l} R = E^{-4/3}\mathscr{R}, \quad \Omega = E^{-2/3}\omega, \quad M = E^{-1/3}m, \quad K = E^{-1/3}k, \\ A = E^{-1/3}a, \quad \theta = \theta, \quad \Psi = E^{-1/3}\psi, \quad W = E^{-2/3}w. \end{array}\right\} \tag{3.4}$$

The system (3.1)–(3.3) can now be written as a single second-order ordinary differential equation in z:

$$\frac{\mathrm{d}^2w}{\mathrm{d}z^2} + \left[\frac{\mathscr{R}(a^2 - \mathrm{i}\omega)(m^2 + a^2z^2)}{a^2 - \mathrm{i}P\omega} - \frac{\mathrm{i}\mathscr{R}m}{a^2 - \mathrm{i}P\omega} - (a^2 - \mathrm{i}\omega)^2a^2\right] w = 0, \tag{3.5}$$

where $a^2 = k^2 + m^2/s^2$ (Roberts 1968, equations (7.10)–(7.12); Busse 1970, equation (5.2)) with boundary conditions

$$\mathrm{i}m\frac{\mathrm{d}w}{\mathrm{d}z} + (a^2 - \mathrm{i}\omega)a^2zw = 0 \quad \text{at } z = \pm(1 - s^2)^{1/2} = \pm h(s) \tag{3.6}$$

(Roberts 1968, (7.13); Busse 1970, (5.3)). Busse noted that the relevant solutions are antisymmetric about the equator $z = 0$, so we can replace the boundary conditions (3.6) by

$$\mathrm{i}m\frac{\mathrm{d}w}{\mathrm{d}z} + (a^2 - \mathrm{i}\omega)a^2zw = 0 \quad \text{at } z = h(s), \qquad w = 0 \quad \text{at } z = 0. \tag{3.7}$$

Note that $w = w(s, z)$ remains a function of both s and z through the s-dependence of a and through the boundary conditions. Nevertheless, neither the equation (3.5) nor the boundary conditions (3.6) contain s-derivatives at leading order; in consequence, the asymptotic theory requires only ordinary differential equations in z to be solved at each value of s.

3.2. *Local stability criteria*

To construct the dispersion relation for our problem, we solve the two-point boundary value problem (3.5)–(3.7) for the complex eigenvalue ω. This then defines ω as a function of

$$\omega = \omega(s, k, m, \mathscr{R}) \tag{3.8}$$

(following the scaling, m is no longer restricted to being an integer) and it is straightforward to find ω for any value of the parameters using standard numerical techniques. The objective of the local theory employed by Roberts and Busse is to maximize the

P	0.01	0.1	1	10	100
$\mathscr{R}_L$	0.01566	0.3130 (0.3151)	3.3814 (3.382)	6.2639 (6.265)	6.5674
m_L	0.08511	0.1762 (0.176)	0.3004 (0.3003)	0.4329 (0.432)	0.4660
s_L	0.5174	0.5135 (0.513)	0.5004 (0.5004)	0.5508 (0.550)	0.5664
ω_L	3.8257	1.6595 (1.65)	0.4362 (0.4362)	0.04237 (0.0423)	0.003975

TABLE 1. Local theory results; values in brackets are taken from Busse (1970)

growth rate $\mathrm{Im}\{\omega\}$ over all admissible real s and real k. The marginal modes are then identified by the minimum value of the Rayleigh number $\mathscr{R}$ which gives steady-state solutions. Evidently this local disturbance is neutral neither growing nor decaying in space, $\mathrm{Im}\{k\} = 0$, or time. Thus the latter condition

$$\mathrm{Im}\{\omega\} = 0 \tag{3.9}$$

can be thought of as the equation determining the Rayleigh number $\mathscr{R}$ and frequency $\mathrm{Re}\{\omega\}$ in terms of s, k and m. Let us first minimize $\mathscr{R}$ over m at fixed s and k. Then for an increment $\mathrm{d}m$ we have

$$\mathrm{Im}\{\mathrm{d}\omega\} = \mathrm{Im}\left\{\frac{\partial\omega}{\partial m}\right\}\mathrm{d}m + \mathrm{Im}\left\{\frac{\partial\omega}{\partial\mathscr{R}}\right\}\mathrm{d}\mathscr{R}. \tag{3.10}$$

Since the minimum is characterized by $\mathrm{d}\mathscr{R} = 0$ subject to the constraint (3.9), $\mathrm{Im}\{\mathrm{d}\omega\} = 0$, (3.10) implies that

$$\mathrm{Im}\left\{\frac{\partial\omega}{\partial m}\right\} = 0. \tag{3.11}$$

We stress that up to this point in the argument the local theory and the global theory, which we discuss below, are in total agreement.

Consider now minimization over k and s. By arguments similar to those used for (3.10) but involving increments $\mathrm{d}k$ and $\mathrm{d}s$ instead, we obtain the conditions

$$\mathrm{Im}\left\{\left(\frac{\partial\omega}{\partial k}\right)_L\right\} = 0 \quad \text{and} \quad \mathrm{Im}\left\{\left(\frac{\partial\omega}{\partial s}\right)_L\right\} = 0, \tag{3.12a,b}$$

which complete the determination of the minimum $\mathscr{R}_L$ of $\mathscr{R}$ over all real values of m, k and s. The solution of the system of equations (3.9) and (3.11)–(3.12) provides the local stability criteria, values for which are identified by the subscript L. It is important to note that, because the eigenvalue problem (3.5), (3.7) depends on k only through k^2, so does the eigenvalue ω itself. Accordingly, (3.12a) is met trivially by $k_L = 0$, as does the group velocity in the global case (see (3.22) below).

The four equations just mentioned were used by Roberts and Busse to determine the four real unknowns, $\mathscr{R}_L$, m_L, k_L, and s_L. To solve these four nonlinear equations numerically, an eigenvalue solver is used as the inner loop of an iterative nonlinear equation solver. We used the routine C05NBF of the Numerical Algorithm Group (NAG) library. The partial derivatives in equations (3.11)–(3.12) were evaluated using centred finite differences. In the course of the iterative solution of the four nonlinear equations, the second-order eigenvalue problem for ω must be solved many times, but this causes no difficulty to modern computers. Our results, at various values of the Prandtl number, can be found in table 1; they are very close to those found by Busse (1970).

3.3. *Global stability criteria*

It is important to appreciate that the theory outlined above is strictly local both in space and time. From the initial value point of view, the description is valid for a limited period of time but does not necessarily determine the long time evolution of small disturbances (Soward 1977). Linked is the fact that, with the time dependence separated out, the local spatial structure cannot necessarily be extended throughout the spatial domain and still meet the boundary conditions (Yano 1992). Since similar local theories have been proposed for other problems (see e.g. the historical survey section of Soward & Jones 1982), it is important to understand that the local (in space) theory does not necessarily provide an approximation to the complete (global) eigenvalue problem. In fact for our convection problem, the local results do not even provide approximations of the true global solution in the limit $E \downarrow 0$; the local and global results differ by an $O(1)$ amount, as explained by Soward (1977) and Yano (1992). This means that previous local solutions do not provide approximate eigensolutions of the problem they purport to solve. In view of the previous literature, we summarize very briefly the reason why the local theory is inadequate, and what needs to be done to find the correct value of $\mathscr{R}$ for global instability. For more details the reader is referred to Soward & Jones (1982), Huerre & Monkewitz (1990), Soward (1992) and Yano (1992).

Our primary objective is to construct a valid WKBJ solution of (2.4)–(2.6), which satisfies all appropriate boundary conditions, in the form

$$E^{2/3} W \sim w(s,z) \exp\left[\frac{\mathrm{i}}{E^{1/3}}\left(\int k(s)\mathrm{d}s + m\phi - \frac{\omega t}{E^{1/3}}\right)\right], \tag{3.13}$$

where guided by the results of § 3.1 above it is anticipated that w and k both vary on the $O(1)$ length scale of the sphere. In addition, for given m, $\mathscr{R}$ and ω the function $k = k(s)$ is determined by the dispersion relation (3.8), while the function $w(s,z)$ is uniquely defined by the corresponding solution of (3.5), (3.7) up to an arbitrary multiplicative function of s, which like w is assumed to vary on an $O(1)$ length scale. Since k only appears in the dispersion relation for ω in the combination k^2, the local condition (3.12*a*) is satisfied simply when $k = 0$ and then both the real and imaginary parts of $\partial\omega/\partial k$ vanish. The point s_L at which $\partial\omega/\partial k = 0$ corresponds to a double root of the the dispersion relation, where two distinct WKBJ solutions (here $k(s)$ and $-k(s)$) have coalesced. The central issue of the eigenvalue problem is to connect those WKBJ modes which decay exponentially to zero on both sides of $s = s_L$.

The connection problem outlined above is resolved locally in the neighbourhood of s_L by considering the amplitude equation given by Yano (1992), equation (5.3), which is similar to (3.16*a*) below. The only significant difference is that the term proportional to $(s - s_L)^2$ in (3.16*a*) is neglected in comparison with a larger term proportional to $s - s_L$. Yano's term linear in $s - s_L$ in his (5.3) is present because there is non-zero phase mixing:

$$\mathrm{Re}\left\{\left(\frac{\partial\omega}{\partial s}\right)_L\right\} \neq 0$$

(see (3.12*b*)). As Yano points out, the resulting Airy equation has no solutions that decay to zero on both sides of s_L. By implication it is impossible to connect the two appropriate WKBJ solutions across $s = s_L$. Indeed, there is no value of the Prandtl number at which $\mathrm{Re}\{(\partial\omega/\partial s)_L\} = 0$. The point is simply that with $\mathscr{R}$ slightly above $\mathscr{R}_L$, phase mixing shortens the radial length of any initial disturbance causing it to ultimately decay.

To find a satisfactory solution, we must follow Soward & Jones (1983) and consider the analytic extension of the exact solution separable in ϕ and t

$$W(s,z)\,\exp\,[\mathrm{i}(M\phi-\Omega t)]$$

of the entire convection problem into the complex s-plane. For real s, that solution has the asymptotic WKBJ representation (3.13). Clearly, it also provides the asymptotic WKBJ representation of the analytic extension provided that w solves (3.5), (3.7) albeit for now complex s. Tackling partial differential equations with two essential variables, here s and z, by techniques whose origin lies with one-dimensional problems was the main thrust of the Soward & Jones development. Once this point is appreciated, the key step is to determine the corresponding analytic extension of the dispersion relation (3.8) by the recipe described. Except for this added complication (ω is the eigenvalue of an ODE rather than the solution of algebraic equations), our analysis parallels Yano's development.

To begin we locate a point $s=s_c$ and wavenumber $k=k_c$, at which both the complex group velocity and complex phase mixing vanish:

$$\left(\frac{\partial\omega}{\partial k}\right)_c=0\quad\text{and}\quad\left(\frac{\partial\omega}{\partial s}\right)_c=0 \tag{3.14a,b}$$

(see Huerre & Monkewitz 1990). In general, this means that both s_c and k_c are complex. For our problem, however, though the location s_c is complex, the wavenumber is real and simply vanishes, $k_c=0$, just as it does in the case of the local analysis. Our extra requirement that the phase mixing vanishes enables the connection problem, which occurs when the wavenumber has a repeated root, to be resolved. It is accomplished by considering the local modulated representation

$$E^{2/3}W\sim\mathscr{W}(x)w(s,z)\,\exp\left[\frac{\mathrm{i}}{E^{1/3}}\left(m_c\phi-\frac{\omega_c t}{E^{1/3}}\right)\right]\quad\text{with } x=(s-s_c)/E^{1/6}, \tag{3.15a}$$

which is valid in the neighbourhood of s_c. A match is obtained with the WKBJ solution (3.13) provided that

$$\mathscr{W}\sim\exp\left[\frac{\mathrm{i}}{2}k_c'x^2\right]\quad\text{in which } k_c'=\lim_{s\to s_c}\frac{\mathrm{d}k}{\mathrm{d}s}. \tag{3.15b}$$

Since $k_c=0$ is a repeated root of the dispersion relation (3.8), there is a second neighbouring root $k_1(s)=-k(s)$ in the vicinity of $s=s_c$ for which

$$\lim_{s\to s_c}\frac{\mathrm{d}k_1}{\mathrm{d}s}=-k_c'\qquad(k_1(s_c)=k_c=0)\,.$$

Our limiting process points to the fact that $\mathrm{d}k/\mathrm{d}s$ is not uniquely defined at $s=s_c$ and highlights the importance of the matching condition (3.15b).

Notice that the radial length in the vicinity of $s=s_c$ is relatively long, of $O(E^{1/6})$. The systematic expansion of the governing equations (2.8)–(2.10) is particularly sensitive to the relative sizes of the partial derivatives in the ∇^2 operator: $\partial^2/\partial\phi^2=O(E^{-2/3})$, $\partial^2/\partial s^2=O(E^{-1/3})$ and $\partial^2/\partial z^2=O(1)$. Since $k_c=0$, only the ϕ-derivatives influence the lowest-order result $\omega_c=\omega(s_c,0,m_c,\mathscr{R}_c)$ at say $O(1)$. At $O(E^{1/3})$, we begin to capture the influence of the the radial s-derivatives but the z-derivatives remain small

and it is at this level we obtain the amplitude equation

$$-\frac{1}{2}\left(\frac{\partial^2\omega}{\partial k^2}\right)_c \frac{\mathrm{d}^2\mathscr{W}}{\mathrm{d}x^2} + \left\{\frac{1}{2}\left(\frac{\partial^2\omega}{\partial s^2}\right)_c x^2 + \left(\frac{\partial\omega}{\partial\mathscr{R}}\right)_c \mathscr{R}_1 - \left[\omega_1 - \left(\frac{\partial\omega}{\partial m}\right)_c m_1\right]\right\}\mathscr{W} = 0, \tag{3.16a}$$

with

$$\mathscr{W} \to 0 \quad \text{as } x \to \pm\infty$$

(more precisely, $\mathscr{W} \to 0$ in the appropriate sector of the complex plane, see below (3.17*c*)), where

$$\mathscr{R} = \mathscr{R}_c + E^{1/3}\mathscr{R}_1, \qquad m = m_c + E^{1/3}m_1, \qquad \omega = \omega_c + E^{1/3}\omega_1. \tag{3.16b}$$

A similar equation, but for the case of a real double turning point, was derived by Connor, Hastie & Taylor (1979), their equation (36).

The simple representation (3.16*a*) stems directly from the small-k expansion of the dispersion relation (3.8) using the asymptotic result

$$\frac{E^{1/3}}{\mathscr{W}}\frac{\mathrm{d}^2\mathscr{W}}{\mathrm{d}x^2} \sim -k^2$$

which follows from the functional dependence of $w(s,z)$ in (3.15*a*) on $k(s)$ and the consistency of the two representations (3.13) and (3.15*a*). By this device (3.16*a*) captures all the appropriate terms at $O(E^{1/3})$. Interestingly at the higher $O(E^{1/6})$ level, we might have expected large terms of the form $\mathrm{d}\mathscr{W}/\mathrm{d}x$ and $x\mathscr{W}$ to appear in (3.16*a*) but these terms (linear in k and x) are absent because of our ansatz (3.14). In their absence the structure of the amplitude equation is controlled by the second partial derivatives of ω with respect to both k and s, which are the coefficients of the quadratic terms in the dispersion relation expansion. Significantly, in addition to the ∇^2 scalings already mentioned, $\partial/\partial s$ and s commute to the level of approximation taken and this leads to the absence of a term $x\mathrm{d}\mathscr{W}/\mathrm{d}x$, which is linked to the result $(\partial^2\omega/\partial k\,\partial s)_c = 0$, a consequence of ω being a function of k^2 and $k_c = 0$. Finally, we emphasize that the only way z-derivatives influence the problem is through their connection with the Coriolis acceleration and the process is encapsulated by the equation (3.5) for $w(s,z)$.

The locations at which the coefficient of $\mathscr{W}$ in (3.16*a*) vanishes are called turning points; being quadratic in x, there are in general two. Loosely speaking, when a WKBJ solution arrives in the vicinity of a turning point a second WKBJ solution is triggered, as in the case of Airy's equation. Essentially our global stability criteria (3.14) isolate, correct to lowest order, a double turning point s_c. The required small separation, $O(E^{1/6})$, is only apparent in the higher-order theory (3.16). Thus by fine tuning the values of $\mathscr{R}_1$ and ω_1, the locations of the turning points can be adjusted so as to eliminate the unwanted second WKBJ solution, which we identified by a second wavenumber $k_1(s) = -k(s)$ above. By this device, we obtain the desired WKBJ solution tending to zero both as $x \to -\infty$ and $x \to +\infty$. This procedure identifies a sequence of eigenfunctions, the first one of which is

$$\mathscr{W} = \mathscr{W}_0 \exp\left[-\frac{1}{2}\left(\frac{\partial^2\omega/\partial s^2}{\partial^2\omega/\partial k^2}\right)_c^{1/2} x^2\right] \tag{3.17a}$$

provided that

$$\left(\frac{\partial\omega}{\partial\mathscr{R}}\right)_c \mathscr{R}_1 - \left[\omega_1 - \left(\frac{\partial\omega}{\partial m}\right)_c m_1\right] = -\frac{1}{2}\left(\frac{\partial^2\omega}{\partial k^2}\frac{\partial^2\omega}{\partial s^2}\right)_c^{1/2}. \tag{3.17b}$$

Clearly the eigenfunction (3.17*a*) coincides with the asymptotic representation (3.15*b*) and so fixes the sign to be taken by the square roots; the opposite and inappropriate signs are linked to the second unwanted WKBJ solution identified by the second wave number k_1. Since $(\partial\omega/\partial m)_c$ is real (see (3.11)), (3.17*b*) is an equation that fixes the real values of $\mathscr{R}_1$ and $\omega_1 - (\partial\omega/\partial m)_c m_1$. In view of the fact that m_1 determines an $O(1)$ correction to the wavenumber M, it is reasonable for the purposes of comparison with the numerics to define

$$m_1 = M_N - E^{-1/3} m_c, \tag{3.17c}$$

where M_N is the value of M predicted by the numerical solution. It should be emphasized that though this choice influences the realized frequency correction ω_1 the realized value of $\mathscr{R}_1$ remains uniquely defined by the imaginary part of formula (3.17*b*).

Of course, the double turning point analysis is natural when s_c is real. Nevertheless, it is equally applicable when s_c is complex; but then to complete the specification we need to ensure that (3.17*a*) merges with the correct WKBJ solution in the appropriate Stokes sectors in the complex s-plane (see §4 below). When this is accomplished the WKBJ solution (3.13) gives a uniformly valid approximation of the solution for real s. Though the procedure outlined is well known within the framework of ordinary differential equations (see, for example, Wasow 1985), the notion of analytic continuation of solutions of partial differential equations, as required here, is less intuitive and the interesting aspect of Soward & Jones' (1983) earlier study.

3.4. *Numerical implementation of the global criteria*

Since we need to extend our problem onto the complex s-plane to find the double turning point, it is more convenient to formulate (3.5) and (3.7) as an ODE on the fixed interval [0,1], rather than allowing the boundary points to be possibly complex. We extend the definition of

$$h(s) = (1 - s^2)^{1/2} \tag{3.18}$$

into the complex plane, keeping to the branch corresponding to the positive square root. We can then rescale the equation and boundary condition using

$$\left.\begin{array}{l} \mathscr{R} \to h^{-10/3}\widehat{R}, \quad \omega \to h^{-2/3}\widehat{\omega}, \quad m \to h^{2/3}\widehat{m}, \quad k \to h^{-1/3}\widehat{k}, \\ a \to h^{-1/3}\widehat{a}, \quad w \to w, \quad s \to h\widehat{s}, \quad z \to h\widehat{z}. \end{array}\right\} \tag{3.19}$$

Note that this implies $\widehat{a}^2 = \widehat{k}^2 + \widehat{m}^2/\widehat{s}^2$. The form of the system (3.5) and (3.7) is left invariant by this transformation and so

$$\frac{\mathrm{d}^2 w}{\mathrm{d}\widehat{z}^2} + \left[\frac{\widehat{R}(\widehat{a}^2 - \mathrm{i}\widehat{\omega})(\widehat{m}^2 + \widehat{a}^2\widehat{z}^2)}{\widehat{a}^2 - \mathrm{i}P\widehat{\omega}} - \frac{\mathrm{i}\widehat{R}\widehat{m}}{\widehat{a}^2 - \mathrm{i}P\widehat{\omega}} - (\widehat{a}^2 - \mathrm{i}\widehat{\omega})^2\widehat{a}^2\right] w = 0, \tag{3.20}$$

$$\mathrm{i}\widehat{m}\frac{\mathrm{d}w}{\mathrm{d}\widehat{z}} + (\widehat{a}^2 - \mathrm{i}\widehat{\omega})\widehat{a}^2\widehat{z}w = 0 \quad \text{at } \widehat{z} = 1, \qquad w = 0 \quad \text{at } \widehat{z} = 0. \tag{3.21}$$

Note that if s becomes complex, h becomes complex, so it follows that in system (3.20)–(3.21) $\widehat{R}$, $\widehat{a}$ and $\widehat{m}$ are now complex, as well as $\widehat{\omega}$.

We now derive the conditions for the existence of a double turning point in these new variables. The condition $(\partial\omega/\partial k)_c = 0$ leads to

$$\left(\widehat{k}\frac{\partial\widehat{\omega}}{\partial\widehat{a}}\right)_c = 0. \tag{3.22}$$

As indicated by Yano (1992), the desired solution of this is $\widehat{k}_c = 0$; so $\widehat{a}_c = \widehat{m}_c/\widehat{s}_c$. Note that although this implies $k_c = 0$ at the complex turning point, when the solution is continued down to the real axis, k is $O(1)$. With $\widehat{k}$ now eliminated, $\widehat{\omega}$ can be viewed as a function of $\widehat{a}$, $\widehat{R}$, and $\widehat{m}$. The condition $(\partial\omega/\partial s)_c = 0$ then leads to

$$\left(\frac{1}{3}+\frac{\widehat{a}^2}{\widehat{m}^2}\right)_c\left(\widehat{a}\frac{\partial\widehat{\omega}}{\partial\widehat{a}}\right)_c + \frac{10}{3}\left(\widehat{R}\frac{\partial\widehat{\omega}}{\partial\widehat{R}}\right)_c - \frac{2}{3}\left(\widehat{m}\frac{\partial\widehat{\omega}}{\partial\widehat{m}}\right)_c - \frac{2\widehat{\omega}_c}{3} = 0, \tag{3.23}$$

the real and imaginary parts of which provide two equations, while (3.9) and (3.11) give another two equations,

$$\mathrm{Im}\,\{(1+\widehat{s}_c^2)^{1/3}\widehat{\omega}_c\} = 0, \qquad \mathrm{Im}\left\{(1+\widehat{s}_c^2)^{2/3}\left(\frac{\partial\widehat{\omega}}{\partial\widehat{m}} + \frac{\widehat{a}}{\widehat{m}}\frac{\partial\widehat{\omega}}{\partial\widehat{a}}\right)_c\right\} = 0, \tag{3.24a,b}$$

and $\mathrm{Im}\,\{\mathscr{R}_c\} = 0$ and $\mathrm{Im}\,\{m_c\} = 0$ give the last two equations

$$\mathrm{Im}\,\{(1+\widehat{s}_c^2)^{5/3}\widehat{R}_c\} = 0, \qquad \mathrm{Im}\,\{(1+\widehat{s}_c^2)^{-1/3}\widehat{m}_c\} = 0. \tag{3.25a,b}$$

Equations (3.23)–(3.25) comprise six (real) equations for three complex unknowns, $\widehat{a}_c$, $\widehat{R}_c$ and $\widehat{m}_c$. This defines the numerical problem we have to solve.

The procedure adopted was similar to that used to find the local dispersion relation. A two-point boundary value solver was used to find the eigenvalue $\widehat{\omega}$ for any given complex $\widehat{a}$, $\widehat{R}$ and $\widehat{m}$. This was achieved using a finite difference scheme; the truncation error was monitored by doubling the number of points used until satisfactory convergence was obtained. This eigenvalue solver then forms the inner loop of an iteration procedure to solve the six nonlinear equations (3.23)–(3.25) for the six unknowns, the real and imaginary parts of $\widehat{a}_c$, $\widehat{R}_c$ and $\widehat{m}_c$. The Numerical Algorithm Group (NAG) routine C05NBF was used to perform this outer loop iteration.

However, solving nonlinear equations of this type, when the system contains as many as six unknowns, requires some care with the initial estimate used. Random initial estimates did not lead to satisfactory convergence. The procedure adopted was to use a comparison problem, defined by

$$\frac{\mathrm{d}^2w}{\mathrm{d}\widehat{z}} + \left[\frac{\widehat{R}(\widehat{a}^2-\mathrm{i}\widehat{\omega})(\widehat{m}^2+\eta_2\widehat{a}^2\widehat{z}^2)}{\widehat{a}^2-\mathrm{i}P\widehat{\omega}} - \frac{\mathrm{i}\eta_1\widehat{R}\widehat{m}}{\widehat{a}^2-\mathrm{i}P\widehat{\omega}} - (\widehat{a}^2-\mathrm{i}\widehat{\omega})^2\widehat{a}^2\right]w = 0, \tag{3.26}$$

$$\mathrm{i}\widehat{m}\frac{dw}{d\widehat{z}} + \eta_2(\widehat{a}^2-\mathrm{i}\widehat{\omega})\widehat{a}^2\widehat{z}w = 0 \quad \text{at } \widehat{z}=1, \qquad w = 0 \quad \text{at } \widehat{z}=0, \tag{3.27}$$

which differs from (3.20)–(3.21) only by the additional factors of η_1 and η_2. When $\eta_1 = \eta_2 = 1$ we recover the desired problem, but when $\eta_1 = \eta_2 = 0$ we have a much simpler problem which can be solved by elementary methods. For $\eta_1 = \eta_2 = 0$ the eigenfunction $w = \sin(\pi\widehat{z}/2)$, and for $P = 1$ the solution to (3.23)–(3.25) is $\widehat{R} = \pi^{4/3}/2$, $\widehat{m} = \pi^{1/3}3^{1/2}/2$ and $\widehat{a} = \pi^{1/3}/2^{1/2}$ with $\widehat{\omega} = 0$.

P	0.01	0.1	1	10	100
			Absolute instability results		
$\mathscr{R}_c$	0.4762	1.1295	4.1173	8.0470	8.5954
m_c	0.08536	0.1770	0.3029	0.4848	0.5363
s_r	0.4928	0.4895	0.5342	0.6805	0.7114
s_i	−0.1647	−0.1479	−0.09667	−0.08546	−0.08782
ω_c	2.6668	1.2386	0.4715	0.06638	0.006433
$\mathscr{R}_c/\mathscr{R}_L$	30.406	3.608	1.218	1.285	1.372
			Yano's (1992) results		
$\mathscr{R}_c/\mathscr{R}_L$		3.58	1.24	1.29	
m_c		0.169	0.297	0.498	
ω_c		1.261	0.475	0.0763	

First-order correction coefficients

P	0.1	1	10
$(\partial\omega/\partial\mathscr{R})_c$	0.05756 + 0.5453 i	0.03658 + 0.1062 i	0.006633 + 0.006709 i
$(\partial^2\omega/\partial k^2)_c$	−3.3041 − 3.7099 i	−0.9652 − 0.9895 i	−0.2028 − 0.08236 i
$(\partial^2\omega/\partial s^2)_c$	−1.4271 − 27.6405 i	−5.8492 − 13.7241 i	−2.0611 − 0.6950 i
$(\partial\omega/\partial m)_c$	3.1556	0.5186	−0.04156
$\mathscr{R}_1$	9.9468	17.7815	17.8976
ω_1-$(\partial\omega/\partial m)_c m_1$	−1.6543	−0.6089	−0.2047

TABLE 2. Absolute instability results, with $s_c = s_r + \mathrm{i}s_i$. $\mathscr{R}_c/\mathscr{R}_L$ is ratio of critical Rayleigh number of absolute instability to critical Rayleigh number for local instability. Also given are Yano's (1992) small inclination results and the first order correction coefficients.

The simplest procedure would be to increase η_1 and η_2 together until they reached unity, following the nonlinear solution. In practice, this was not successful, as the solution is lost at $\eta_1 = \eta_2 \sim 0.25$. The successful technique was to first increase η_1 from zero to unity in steps of 0.1, keeping $\eta_2 = 0$, and then to increase η_2 from zero to unity in steps of 0.1, keeping $\eta_1 = 1$. The first part of this operation, with $\eta_2 = 0$, is simplified by the fact that the eigenfunction remains at $w = \sin(\pi\hat{z}/2)$, so the dispersion relation can be written down explicitly, which makes it possible to find the conditions (3.23)–(3.25) as explicit formulae. These nonlinear equations can then be solved without the need to use a boundary value solver for the eigenvalue ω, but they still have to be solved numerically. The second part of the operation, increasing η_2 from zero to unity at $\eta_1 = 1$ requires a routine evaluating ω as a function of the parameters, which is then used to evaluate the derivatives needed to solve (3.22)–(3.25) numerically.

Once the solution at $P = 1$ is obtained by this procedure, the solution was continued by increasing or decreasing P gradually. In this way, the conditions for absolute instability can be calculated. Once the 'hat' quantities are found, they can be converted back using (3.19) to give the real values of $\mathscr{R}_c$, m_c, ω_c; s_c is still complex. The values found, together with the small-inclination approximation results of Yano (1992), are given in table 2. Also given in table 2 are the coefficients

$$\left(\frac{\partial\omega}{\partial\mathscr{R}}\right)_c = \frac{1}{(1+\hat{s}_c^2)^{4/3}}\left(\frac{\partial\hat{\omega}}{\partial\hat{R}}\right)_c, \qquad \left(\frac{\partial^2\omega}{\partial k^2}\right)_c = \left(\frac{1}{\hat{a}}\frac{\partial\hat{\omega}}{\partial\hat{a}}\right)_c, \tag{3.28a}$$

$$\left(\frac{\partial\omega}{\partial m}\right)_c = (1+\widehat{s}_c^2)^{2/3}\left(\frac{\partial\widehat{\omega}}{\partial\widehat{m}} + \frac{\widehat{a}}{\widehat{m}}\frac{\partial\widehat{\omega}}{\partial\widehat{a}}\right)_c, \tag{3.28b}$$

$$\left(\frac{\partial^2\omega}{\partial s^2}\right)_c = \widehat{s}_c^2(1+\widehat{s}_c^2)^{4/3}\left[-\frac{4}{9}\widehat{\omega}_c + 2\frac{1+\widehat{s}_c^2}{\widehat{s}_c^4}\left(\widehat{a}\frac{\partial\widehat{\omega}}{\partial\widehat{a}}\right)_c + \left(\mathscr{D}^2\widehat{\omega}\right)_c\right], \tag{3.28c}$$

where the differential operator $\mathscr{D}$ is

$$\mathscr{D} = \left(\frac{1}{3}+\frac{1}{\widehat{s}_c^2}\right)\widehat{a}\frac{\partial}{\partial\widehat{a}} + \frac{10}{3}\widehat{R}\frac{\partial}{\partial\widehat{R}} - \frac{2}{3}\widehat{m}\frac{\partial}{\partial\widehat{m}}. \tag{3.28d}$$

In evaluating these quantities, we recall from (3.23) and (3.24) that

$$\widehat{a}_c = \frac{\widehat{m}_c}{\widehat{s}_c}, \qquad (\mathscr{D}\widehat{\omega})_c = \tfrac{2}{3}\widehat{\omega}_c. \tag{3.29}$$

The first-order corrections $\mathscr{R}_1$ and $\omega_1 - (\partial\omega/\partial m)_c m_1$ are determined from (3.17*b*) using the formula (3.28) and are also listed in table 2.

The local instability results are seen to be significantly different from the absolute instability results, particularly at lower Prandtl numbers. Also of note is how well Yano's (1992) results compare with the present results, despite the small-inclination approximation being used. This provides further evidence of how effective the annulus model is in describing the essential effects of the geometry of spherical convection.

4. Extending the solution to the real axis

Once the critical values of $\mathscr{R} = \mathscr{R}_c$, $m = m_c$ and $\omega = \omega_c$ are determined by finding the double turning point in the complex plane, the corresponding solution on the real axis can be constructed, recalling from (3.13) that the s,z-dependence has the form

$$\sim w(s,z)\,\exp\left[\frac{\mathrm{i}}{E^{1/3}}\int k(s)\,\mathrm{d}s\right].$$

Here the z-dependence of the eigenfunction $w(s,z)$ and the complex eigenvalue $k(s)$ at given s are determined by the solution of the eigenvalue problem (3.5)–(3.6), while the amplitude of that solution as function of s is determined by higher-order theory. In other words, the leading-order behaviour, even possibly at complex s, is controlled by the value of the complex wavenumber $k(s)$, which solves the dispersion relation (3.8) in the form

$$\omega(s,k,m_c,\mathscr{R}_c) = \omega_c \tag{4.1}$$

with the parameters m and $\mathscr{R}$ fixed at their critical values. By construction, at the double turning point $s = s_c$, the wavenumber has the double zero root $k = 0$. As we move away from this point in the complex plane, it splits into two non-zero values $\pm k$. So upon reaching the real s-axis, we must isolate which of the two gives the physically acceptable solution. To this end, we seek the point M, at $s = s_M$, on the real axis where $\mathrm{Im}\{k\} = 0$, and choose the k-root which has $\mathrm{Im}\{\mathrm{d}k/\mathrm{d}s\} > 0$. There the solution takes its maximum value, being approximated locally by

$$\sim w(s,z)\,\exp\left[\mathrm{i}k_M\frac{s-s_M}{E^{1/3}} + \frac{\mathrm{i}k'_M}{2}\frac{(s-s_M)^2}{E^{1/3}}\right] \quad \text{with } k'_M = \left(\frac{\mathrm{d}k}{\mathrm{d}s}\right)_M. \tag{4.2}$$

Since k_M and $\mathrm{Im}\{k'_M\}$ are both $O(1)$ at the point M, it follows that the radial wavenumber is $O(E^{1/3})$, the same as the azimuthal wavenumber, while the extent of the radial modulation of the disturbance is $O(E^{1/6})$. The location s_M, and the corresponding values of k_M and k'_M, are given in table 3 as a function of Prandtl number.

P	0.01	0.1	1	10	100
s_M	0.8692	0.7131	0.5915	0.6922	0.7195
k_M	−0.3325	−0.4243	−0.3486	−0.3007	−0.3154
$\mathrm{Re}\{(\partial k/\partial s)_M\}$	−1.336	−1.012	−1.595	−1.166	−1.046
$\mathrm{Im}\{(\partial k/\partial s)_M\}$	0.186	0.516	1.710	3.331	3.529
s_-	0.4042	0.3874	0.4634	0.6025	0.6286
s_+	1.0000	0.9926	0.7563	0.7772	0.8029

TABLE 3. Solution characteristics on the real axis.

Another question of interest is where the anti-Stokes lines emanating from the double turning point cut the real axis, say at the points A_- and A_+, with s-coordinates s_- and s_+ respectively. Since the two solutions for k are of the form $\pm k$, the anti-Stokes lines are defined by the path

$$\mathrm{Im}\left\{\int_{s_c}^{s} k(s)\,\mathrm{d}s\right\} = 0, \tag{4.3}$$

where as usual s_c is the double turning point. The significance of the anti-Stokes lines is that in the interval (s_-, s_+) the WKBJ solution based on the k with $\mathrm{Im}\{k'_M\} > 0$ is valid. Outside this interval, this is no longer the case. In consequence, our solution is only valid provided the interval (s_-, s_+) lies entirely in physical space, which for the whole sphere problem is $(0, 1)$. Fortunately, as we see from table 3, where s_- and s_+ are tabulated, this validity condition holds for all Prandtl numbers. It should be noted that at small P, s_+ becomes very close to unity. Indeed, it is only less than unity by a small amount at $P = 0.1$, and at $P = 0.01$ s_+ lies so close to unity that great accuracy is needed to resolve the difference. Although in theory the asymptotic solution is valid in the limit of small E at any fixed P, in practice extremely small values of E will be needed to get the behaviour near $s = 1$ correct when $P < 0.1$.

In the case of a spherical shell of finite gap, it is not necessarily the case that the interval (s_-, s_+) lies entirely within the fluid. It may happen that s_- is less than the s-value of the inner-sphere. In that case, the inner-sphere boundary condition at $s_i\,(< s_-)$ must be taken into account, and the critical Rayleigh number will be altered by an $O(1)$ amount. In some sense, this is rather remarkable, because the solution at the point A_- is exponentially small, so one might not think that setting it to zero could make much difference. Nevertheless, this naïve view is incorrect because, even though the WKBJ mode (k) is exponentially small, it cannot alone meet the homogeneous boundary condition on the inner sphere; the second WKBJ mode ($-k$) is subdominant and so cannot be invoked to resolve the difficulty. Evidently a new solution method is required when (s_-, s_+) does not lie entirely within the fluid; this issue will be addressed in a subsequent paper.

Another feature of interest in table 3 is the rather large values of s_M compared to the local theory values s_L given in table 1. Apart from P close to unity, it is no longer true that the disturbance is located at s roughly half the radius of the sphere; the disturbance is located considerably further from the axis. This is perhaps the most striking difference between the true global instability results and the local theory.

The radial extent of the disturbance is best measured by $\mathrm{Im}\{k'_M\}$, as can be seen from (4.2); this quantity determines the exponential cut-off. $\mathrm{Im}\{k'_M\}$ is much larger at high Prandtl number than at low Prandtl number, so the disturbance is much more

confined at high P than at low P. This is related to the fact that the interval (s_-, s_+) is much larger at low P. We therefore expect that the ratio of the radial extent of the rolls to azimuthal wavelength will increase significantly as the Prandtl number is lowered.

In the next section, we compare these expectations with numerical solutions of the partial differential equations.

5. Numerical solution of the PDEs

To solve the PDE eigenvalue problem at finite E, we use the toroidal–poloidal expansion (2.7), and spherical polar coordinates, so that the system becomes (2.8)–(2.11). We use expansion in spherical harmonics, and expansion in Chebyshev polynomials in the radial direction. Because of the azimuthal and temporal symmetry, the dependence $\exp \mathrm{i}(M\phi - \Omega t)$ can be assumed, but spherical harmonics of different degree are coupled. Following the work of Busse (1970) we only need to consider the symmetric modes, so the expansions take the form

$$\mathscr{S} = \sum_{n=1}^{N_x+2} \sum_{l=0}^{L} \mathscr{S}_{nl} r^{k_2} T_{2n-1}(r) P_{2l+m}^{m}(\cos\vartheta) \exp \mathrm{i}(M\phi - \Omega t), \tag{5.1}$$

$$\mathscr{T} = \sum_{n=1}^{N_x+1} \sum_{l=0}^{L} \mathscr{T}_{nl} r^{k_1} T_{2n-1}(r) P_{2l+m+1}^{m}(\cos\vartheta) \exp \mathrm{i}(M\phi - \Omega t), \tag{5.2}$$

where if m is even $k_1 = 2$ and $k_2 = 1$, if m is odd, $k_1 = 1$ and $k_2 = 2$ (see e.g. Jones, Longbottom & Hollerbach 1995), bearing in mind the need for the expansions to have the correct behaviour at the axis. The equations in the radial direction are derived by collocation at the positive zeros of $T_{2N_x}(r)$ together with the boundary conditions (2.11).

Significant simplification occurs because in the equation corresponding to degree $2l+m$ the interaction between harmonics is restricted to the adjacent modes of degrees $2l+m-1$ and $2l+m+1$. In consequence the corresponding matrices have a banded structure, aiding their inversion. An LU decomposition was applied to the banded matrix, and inverse iteration was used to find the (complex) eigenvalue Ω. Iteration on the Rayleigh number then reduces the imaginary part of Ω to zero.

The results for $P = 10$, $P = 1$ and $P = 0.1$ are given in tables 4, 5 and 6. The truncations N_x and L are also listed, together with the (integer) value of M which gives minimum critical Rayleigh number, along with the corresponding Rayleigh number and frequency. For comparison, the leading-order asymptotic results at the given Taylor number (recall $Ta = E^{-2}$) are listed. Also tabulated are the asymptotic results for R and Ω including the first-order corrections using equations (3.16*b*), (3.17*b*) and (3.17*c*), with the numerical results listed in table 2. The value of Ω given for the first-order correction assumes that the critical value of M is that given by the numerical results; this is necessary as the asymptotic theory does not restrict M to integer values.

We note that although there are differences between the numerical results and the leading-order theory, the percentage difference between the asymptotic and numerical results steadily decreases as $Ta \uparrow \infty$ as we would hope. Furthermore, when the first-order corrections are included the agreement is significantly better, particularly in the cases $P = 10$ and $P = 1$. Indeed, the small residuals in these cases are consistent with the expansions (3.16*b*) being continued in the natural way to give second-order corrections $O(E^{2/3})$. In the case $P = 0.1$ the agreement is not so good,

Ta	10^9	10^{10}	10^{11}	3×10^{11}	10^{12}
			Numerical results		
R	8.761×10^6	3.945×10^7	1.797×10^8	3.713×10^8	8.240×10^8
Ω	61.748	136.49	297.23	430.06	649.58
M	14	21	32	39	47
N_x	24	26	34	36	42
L	16	18	20	22	26
			Zhang's (1992) finite-gap results converted using (5.3)		
R	8.6×10^6	3.9×10^7	1.7×10^8		8.0×10^8
Ω	62	1.3×10^2	3.3×10^2		6.2×10^2
M	16	21	34		52
			Leading-order asymptotic results		
R	8.047×10^6	3.735×10^7	1.734×10^8	3.606×10^8	8.047×10^8
Ω	66.38	143.02	308.12	444.38	663.82
M	15.33	22.50	33.03	39.67	48.48
			Asymptotic results including first-order corrections		
R	8.613×10^6	3.914×10^7	1.790×10^8	3.704×10^8	8.226×10^8
Ω	61.66	136.41	297.09	429.91	649.50

TABLE 4. Numerical and asymptotic solutions for $P = 10$.

Ta	10^9	10^{10}	10^{11}	3×10^{11}	10^{12}
			Numerical results		
R	4.761×10^6	2.105×10^7	9.466×10^7	1.947×10^8	4.302×10^8
Ω	4.428×10^2	9.849×10^2	2.124×10^3	3.073×10^3	4.638×10^3
M	9	14	20	24	30
N_x	24	26	34	36	40
L	16	16	20	22	28
			Leading-order asymptotic results		
R	4.117×10^6	1.911×10^7	8.870×10^7	1.845×10^8	4.117×10^8
Ω	4.715×10^2	1.0158×10^3	2.188×10^3	3.156×10^3	4.715×10^3
M	9.58	14.06	20.64	24.78	30.29
			Asymptotic results including first-order corrections		
R	4.680×10^6	2.089×10^7	9.433×10^7	1.943×10^8	4.295×10^8
Ω	4.427×10^2	9.861×10^2	2.124×10^3	3.073×10^3	4.639×10^3

TABLE 5. Numerical and asymptotic solutions for $P = 1$.

and the residuals cannot naturally be explained as being due to higher-order terms. We conclude that the anti-Stokes line coming so close to the outer boundary in the case $P = 0.1$ (see table 3) is affecting the accuracy of the asymptotic expansion at low Prandtl numbers. Note also that the numerical results cannot be reconciled with the local theory results, which lead to significantly smaller values of the critical Rayleigh number.

It is also possible to compare these results with the finite-gap numerical calculations of Zhang (1992). In that paper, the inner sphere is located at radius $\eta/(1-\eta)$ and

Ta	10^9	10^{10}	10^{11}	3×10^{11}
		Numerical results		
R	1.293×10^6	5.694×10^6	2.616×10^7	5.389×10^7
Ω	9.777×10^2	2.556×10^3	5.345×10^3	8.016×10^3
M	4	8	11	14
N_x	26	26	36	40
L	18	20	26	26
		Leading-order asymptotic results		
R	1.129×10^6	5.243×10^6	2.433×10^7	5.062×10^7
Ω	1.239×10^3	2.668×10^3	5.749×10^3	8.291×10^3
M	5.60	8.21	12.06	14.48
		Asymptotic results including first-order corrections		
R	1.444×10^6	6.237×10^6	2.748×10^7	5.607×10^7
Ω	1.027×10^3	2.560×10^3	5.408×10^3	8.032×10^3

TABLE 6. Numerical and asymptotic solutions for $P = 0.1$.

the concentric outer sphere at radius $1/(1-\eta)$, so that η is the radius ratio. Zhang's numerical results were for $\eta = 0.4$. To compare results we use the scalings

$$\left.\begin{array}{lll} R_g = (1-\eta)^{10/3}R, & M_g = (1-\eta)^{-2/3}M, & s_g = (1-\eta)^{-1}s, \\ \Omega_g = (1-\eta)^{2/3}\Omega, & K_g = (1-\eta)^{1/3}K, & z_g = (1-\eta)^{-1}z, \end{array}\right\} \tag{5.3}$$

where the quantities with g subscripts refer to the values found in the finite-gap calculations and the unsubscripted quantities to the whole sphere calculation. We see that the results of table 4 compare satisfactorily with table 2 of Zhang (1992); we have converted his results to the full sphere case using (5.3).

In figure 1 we show the eigenfunctions for the radial and azimuthal components of velocity, u_s and u_ϕ, on the equatorial plane. Both the ϕ-dependence and the radial dependence of these quantities are shown for Prandtl numbers in the range 0.1–10 for $Ta = 10^{11}$ and $Ta = 10^{12}$. The ϕ-dependence in figure 1 is shown for $s = s_M$ as defined in table 3, where the disturbance takes its maximum value according to asymptotic theory. The ϕ-value used for the radial dependence curves is one at which the ϕ-dependent curves have a maximum; since these ϕ-dependence curves are sinusoidal it does not matter which maximum is chosen. We see that in figures 1(a) and 1(b), which are for Prandtl number $P = 10$, the disturbance is already localized at $Ta = 10^{11}$, and becomes more localized at $Ta = 10^{12}$ in accord with the asymptotic theory. The excellent agreement between the asymptotic and numerical results is therefore not unexpected. In the case $P = 1$, figures 1(c), $Ta = 10^{11}$ and 1(d), $Ta = 10^{12}$, the disturbance is somewhat more spread out, but the amplitude does become very small before it reaches either the centre or the outer boundary, so again the good agreement between asymptotic theory and numerical results is no surprise. The case shown in figure 1(e), for $P = 0.1$, $Ta = 10^{11}$, is rather different; the amplitude has not dropped to very small values as $s \to 1$, the outer boundary. $u_r = 0$ at $s = 1$ is enforced by the boundary conditions, but u_ϕ is not constrained to be zero, because the outer boundary is stress-free, and it is in fact significantly non-zero. In consequence, we are not sufficiently far into the asymptotic regime in figure 1(e) for the boundary condition $\mathscr{W} \to 0$ as $x \to \infty$ in (3.16a) to be a good

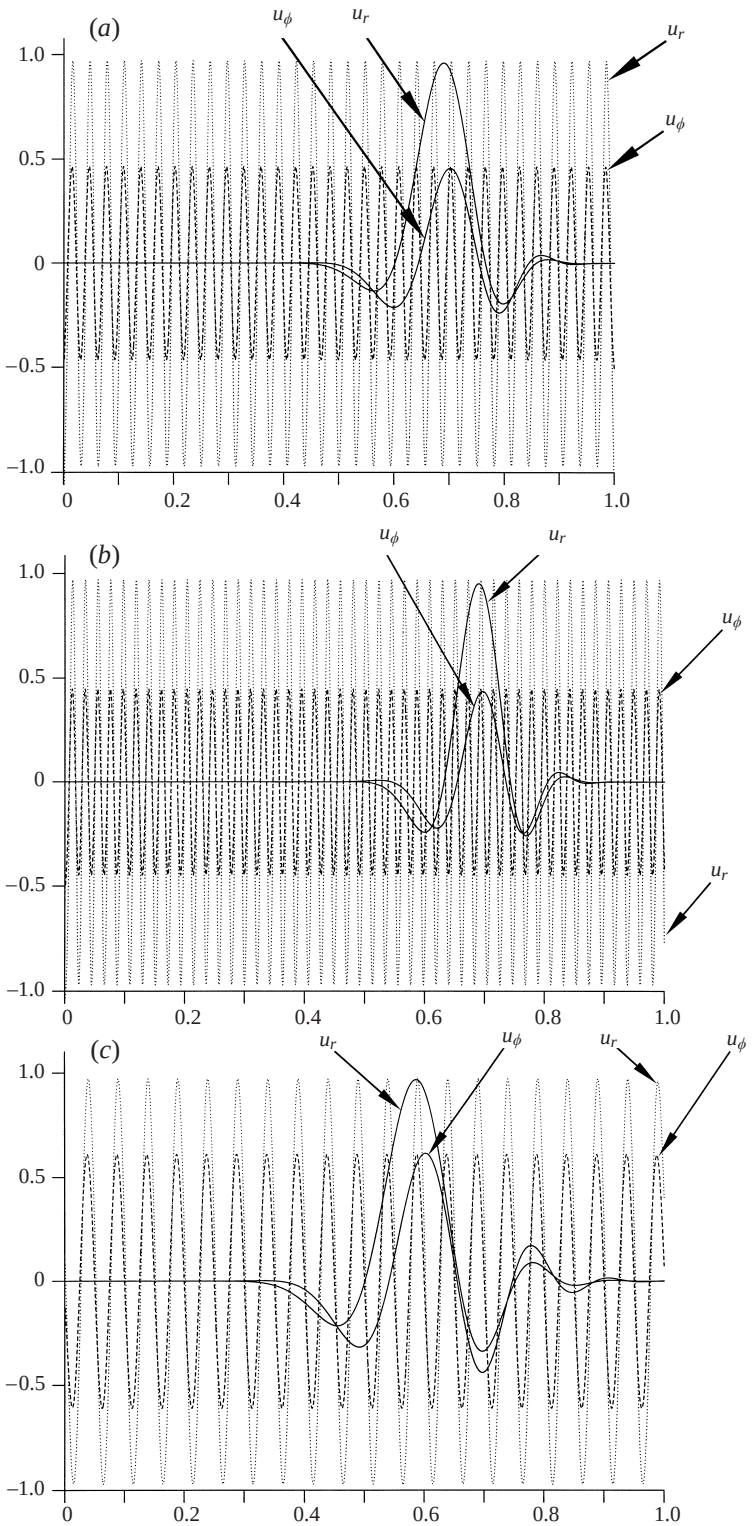

FIGURE 1 (*a–c*). For caption see facing page.

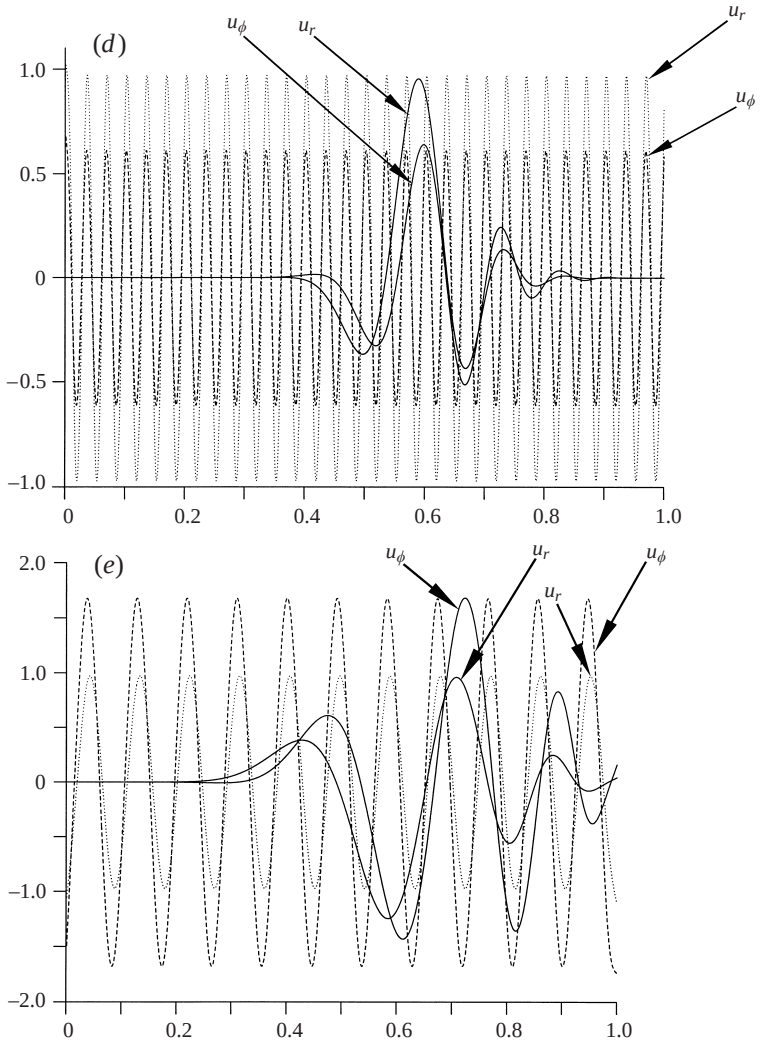

FIGURE 1. The azimuthal dependence and the radial dependence of the velocity components u_ϕ and u_r are shown in the equatorial plane. For the azimuthal plot, the dotted curve gives u_r and the dashed curve u_ϕ. The abscissa is $1/2\pi$ times the azimuthal angle ϕ, and the value of $r = s_M$, is taken from table 3. The ordinate is normalized so that the maximum value of u_r is unity, the relative magnitude of u_ϕ then being determined from the numerical calculation. For the radial plot, both u_r and u_ϕ are plotted as labelled solid lines. The abscissa is the r-coordinate, and the value of ϕ is chosen to be at a point where the azimuthal plot has a local maximum of u_r. The ordinate scaling is determined by the azimuthal plot normalization. (*a*) $P = 10$, $Ta = 10^{11}$ (azimuthal plot at $r = s_M = 0.6922$); (*b*) $P = 10$, $Ta = 10^{12}$; (*c*) $P = 1$, $Ta = 10^{11}$ (azimuthal plot at $r = s_M = 0.5915$); (*d*) $P = 1$, $Ta = 10^{12}$; (*e*) $P = 0.1$, $Ta = 10^{11}$ (azimuthal plot at $r = s_M = 0.7131$).

approximation, which is consistent with the comparatively poor agreement between asymptotic and numerical results at low P. At larger Ta, the disturbance is likely to become more localized, and the asymptotic theory to give better results, but unfortunately it becomes computationally very expensive to obtain numerical results at very large Ta and small P.

We also note that the location of the disturbance maximum, $s = s_M$, given in table

3 is in reasonable agreement with the graphs shown in figure 1: in particular, it is noticeable that the disturbance maximum is nearest the axis at $P = 1$, becoming significantly further out from the axis than local theory would indicate both at lower and higher Prandtl numbers. The location of the critical cylinder as determined in table 3 is also consistent with Zhang's (1992) figure 2.

Zhang also found that the rolls tilt at low Prandtl number, i.e. the local wave vector $(K_M, M/S_M) = (k_M, m/s_M)$ normal to the surface of constant phase changes its direction as P varies. The tilt angle, defined by $\tan^{-1}(-k_M s_M/m)$, varies from approximately 20° at $P = 10$ to just over 60° at $P = 0.1$, a similar range of variation to that found by Yano (1992) and Zhang (1992). Note that since k_M is negative (see table 3), the direction of the tilt is in the prograde sense as s increases, as in figure 3 of Zhang (1992).

The weak z-dependence, together with the continuity equation, implies that

$$\frac{u_s}{u_\phi} \sim \frac{\lambda_s}{\lambda_\phi}.$$

As noted above, the ratio $k_M s_M/m$ decreases as P increases, so the inverse ratio, λ_s/λ_ϕ, increases with P. We therefore expect u_s/u_ϕ to increase with P, and this trend can be seen in figure 1: at $P = 10$, u_s $(= u_r)$ is larger than u_ϕ, but at $P = 0.1$ this is reversed, and u_ϕ is larger than u_s.

6. Conclusions

The main result is that we have successfully applied the complex double turning point theory to find the condition for the absolute instability due to convection in a rapidly rotating sphere for the first time. It is found that the results are significantly different from the previous local theories. Although it is hard numerically to get into the asymptotic regime, it is possible to get far enough to show conclusively that the local theory is inadequate, and that the complex theory gives much better results. The critical azimuthal wavenumber, the frequency at onset as well as the critical Rayleigh number are all correctly predicted, as is the location of the critical cylinder on which the convection first onsets. The other major difficulty of the local theory, its failure to give the radial structure because it gives the critical radial wavenumber as $k \to 0$ in the limit $E \downarrow 0$, is also resolved. The predicted radial structure from the asymptotic theory is in accord with that derived from the numerical solutions.

Previous numerical calculations of Zhang (1992) are confirmed, and the small-inclination theory of Yano (1992) is shown to give a remarkably good approximation.

The location of the anti-Stokes lines for this problem indicates that the asymptotic theory will not be particularly useful at low Prandtl numbers; however, it is known that inertial modes (Zhang 1994) rather than Rossby modes are important at very low Prandtl number. We also note that this work can be extended to cover a large number of related instability problems in either whole sphere or spherical gap geometry.

We thank Dr J.-I. Yano and Professor K. Zhang for helpful comments, and we are grateful to Dr J. B. Taylor and Dr P. D. Killworth for drawing our attention to their related pioneering contributions in other research fields. C. A. J. and A. M. S. acknowledge support from the PPARC through research grant numbers GR/L22973 and GR/L40922.

REFERENCES

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

Busse, F. H. 1986 Asymptotic theory of convection in a rotating cylindrical annulus. *J. Fluid Mech.* **173**, 545–556.

Busse, F. H. & Carrigan, C. R. 1976 Laboratory simulation of thermal convection in rotating planets and stars. *Science* **191**, 81–83.

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.

Chandrasekhar, S. 1961 *Hydrodynamic and Hydromagnetic Stability*. Oxford University Press.

Connor, J. W., Hastie, R. J. & Taylor, J. B. 1979 High mode number stability of an axisymmetric toroidal plasma. *Proc. R. Soc. Lond.* A **365**, 1–17.

Hirsching, W. R. & Yano, J.-I. 1994 Metamorphosis of marginal thermal convection in rapidly rotating self-gravitating spherical shells. *Geophys. Astrophys. Fluid Dyn.* **74**, 143–179.

Huerre, P. & Monkewitz, P. A. 1990 Local and global instabilities in spatially developing flows. *Ann. Rev. Fluid Mech.* **22**, 473–537.

Jones, C. A., Longbottom, A. W. & Hollerbach, R. 1995 A self-consistent convection driven geodynamo model, using a mean field approximation. *Phys. Earth. Planet. Inter.* **92**, 119–144.

Kida, S. 1994 Stability of thermal convection in a rapidly rotating sphere. *J. Phys. Soc. Japan* **63**, 2964–2973.

Killworth, P. D. 1980 Barotropic and baroclinic instability in rotating stratified fluids. *Dyn. Atmos. Oceans* **4**, 143–184.

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.

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. 1992 Thin disc $\alpha\omega$-dynamo models II. Short length scale modes. *Geophys. Astrophys. Fluid Dyn.* **64**, 201–225.

Soward, A. M. & Jones, C. A. 1983 The linear stability of the flow in the narrow gap between two concentric rotating spheres. *Q. J. Mech. Appl. Maths* **36**, 19–42.

Wasow, W. R. 1985 *Linear Turning Point Theory*. Springer.

Yano, J.-I. 1992 Asymptotic theory of thermal convection in rapidly rotating systems. *J. Fluid Mech.* **243**, 103–131.

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

Zhang, K. 1993 On equatorially trapped boundary inertial waves. *J. Fluid Mech.* **248**, 203–217.

Zhang, K. 1994 On coupling between the Poincaré equation and the heat equation. *J. Fluid Mech.* **268**, 211–229.

Zhang, K. & Busse, F. H. 1987 On the onset of convection in rotating spherical shells. *Geophys. Astrophys. Fluid Dyn.* **39**, 119–147.

Zhang, K. & Jones, C. A. 1993 The influence of Ekman boundary layers on rotating convection. *Geophys. Astrophys. Fluid Dyn.* **71**, 145–163.