J. Fluid Mech. (2009), *vol.* 634, *pp.* 291–319. © 2009 Cambridge University Press
doi:10.1017/S0022112009007253 Printed in the United Kingdom

Linear theory of compressible convection in rapidly rotating spherical shells, using the anelastic approximation

C. A. JONES[1]†, K. M. KUZANYAN[1] AND R. H. MITCHELL[2]

[1]School of Mathematics, University of Leeds, Leeds LS2 9JT, UK

[2]School of Engineering, Computing and Mathematics, University of Exeter, North Park Road, Exeter EX4 4QF, UK

(Received 4 June 2008 and in revised form 24 March 2009)

The onset of compressible convection in rapidly rotating spherical shells is studied in the anelastic approximation. An asymptotic theory valid at low Ekman number is developed and compared with numerical solutions of the full equations. Compressibility is measured by the number of density scale heights in the shell. In the Boussinesq problem, the location of the onset of convection is close to the tangent cylinder when there is no internal heating only a heat flux emerging from below. Compressibility strongly affects this result. With only a few scale heights or more of density present, there is onset of convection near the outer shell. Compressibility also strongly affects the frequencies and preferred azimuthal wavenumbers at onset. Compressible convection, like Boussinesq convection, shows strong spiralling in the equatorial plane at low Prandtl number. We also explore how higher-order linear modes penetrate inside the tangent cylinder at higher Rayleigh numbers and compare modes both symmetric and antisymmetric about the equator.

1. Introduction

Convection occurs in many stars and planets and is ultimately responsible for generating their winds and magnetic fields. The rotation rate of most of these objects is much faster than their diffusive time scales, even if these are enhanced by turbulence. Consequently, rapidly rotating convection in spherical shell geometry is a key problem in astrophysical and planetary fluid dynamics. In most applications, the convection is in the strongly nonlinear regime rather than close to onset as assumed here. Nevertheless, an understanding of the linear problem is an essential prerequisite to understanding the very complex behaviour found in the nonlinear regime. At a very basic level, comparison with a well-established linear theory is an important validation for the sophisticated nonlinear computational fluid dynamics (CFD) codes used to develop insight into rotating convecting fluids. Many features observed in such simulations can be understood in terms of linear results, and linear theory allows a reasonably complete coverage of the large parameter space. Fully three-dimensional high-resolution simulations (e.g. Miesch *et al.* 2000; Heimpel, Aurnou & Wicht 2005) are too computationally expensive to allow coverage of the multi-dimensional parameter space, so the linear theory provides an invaluable guide as to the most suitable areas for nonlinear exploration.

† Email address for correspondence: c.a.jones@maths.leeds.ac.uk

The Boussinesq theory of the onset of rapidly rotating convection is now fairly well understood, but the rapidly rotating compressible case has received much less attention. The compressible problem is however more relevant to the problem of rapidly rotating convection in stars and planets, because convection occurs over many scale heights of density in all stars and in giant planets (Guillot 1999*a*, *b*).

The asymptotic theory of the onset of rapidly rotating convection in a Boussinesq sphere was developed by Roberts (1968) and Busse (1970). These papers established the local theory of convection. However, although this local theory has many points of contact with experiments and numerical calculations and forms a useful simple picture of rotating convection, it became clear (e.g. Zhang 1992) that the predicted critical Rayleigh numbers were incorrect except at very large Prandtl number. This problem was resolved with the development of the global theory of convection by Jones, Soward & Mussa 2000 and Dormy *et al.* 2004. In this paper we extend this global asymptotic theory to the compressible case. Busse & Hood (1982) and Zhang (1992) also established the spiralling nature of rotating convection, that is in the equatorial section the columnar structures spiral in a prograde direction or in other words in cylindrical coordinates (s, ϕ, z) in the positive ϕ direction as s increases. This spiralling is important in driving strong zonal flows from the Reynolds stresses (e.g. Jones, Rotvig & Abdulrahman 2003; Rotvig & Jones 2006). If there is no spiralling, any zonal flow has to be in the form of thermal or magnetic winds.

In seminal papers on rotating compressible convection, Gilman & Glatzmaier (1981) and Glatzmaier & Gilman (1981*a*, *b*) showed that compressibility made the convection occur preferentially nearer the outer boundary. Their model differed in some respects from ours, but nevertheless we also find this to be a strong effect, as did Drew, Jones & Zhang (1995). Drew *et al.* (1995) also found the surprising result that convection could occur for negative Rayleigh number Ra in some parameter regimes. Here we show that this cannot be the case if entropy diffusion dominates over thermal diffusion, as is likely to be the case in most astrophysical applications.

The structure of the paper is as follows: first we formulate the problem, noting some differences between our model and some previous models. We then establish rigorously that for our current model instability is not possible at negative Rayleigh number, in distinction to the previous model of Drew *et al.* (1995). Next we develop the asymptotic theory as the Ekman number $E \to 0$ for compressible convection. At fixed compressibility and moderate Prandtl number the convection takes the form of tall thin columns, as does Boussinesq convection. Then we discuss the results from a numerical eigenvalue code for the full equations, comparing them with the results of our asymptotic theory. Further we study some of the higher-Ra modes, i.e. beyond the lowest mode, which form the high latitudinal structures. This sheds some light on to how convection penetrates inside the tangent cylinder as the Rayleigh number is increased.

2. Governing equations

The geometry of the problem is a spherical shell lying between $r = r_i$ and $r = r_o$, where $r_o - r_i = d$. The radius ratio $r_i/r_o = \eta$. We consider a polytropic basic state of the atmosphere with radial gravity, the effective mass M being entirely within the shell so that gravity $\boldsymbol{g}$ falls off as $1/r^2$ within the shell.

$$\frac{\mathrm{d}p}{\mathrm{d}r} = -\frac{GM\rho}{r^2}, \quad \frac{p}{p_0} = \left(\frac{\rho}{\rho_0}\right)^{1+\frac{1}{n}}, \tag{2.1a, b}$$

where p_0 and ρ_0 are the reference values of pressure and density at the mid-point $r=(r_i+r_o)/2$ in the layer and n is the polytropic index. We non-dimensionalize on the length unit d, so that from now on r is the dimensionless radial coordinate. Assuming the perfect gas law, the solution of these equations is then written in the form (Gilman & Glatzmaier 1981)

$$\rho=\rho_0\left(\frac{GM\rho_0}{rp_0(n+1)d}+c_0\right)^n=\rho_0\zeta^n,\quad \zeta=\frac{c_1}{r}+c_0,\quad c_1=\frac{GM\rho_0}{p_0(n+1)d},\quad (2.2a,b,c)$$

$$p=p_0\zeta^{n+1},\qquad T=T_0\zeta,\qquad (2.2d,e)$$

T being the temperature and T_0 its value at the mid-point. We suppose the polytrope has N_ρ density scale heights, that is $\rho(r_i)/\rho(r_o)=\exp N_\rho$, leading to

$$\zeta_o=\frac{\eta+1}{\eta\exp\left(N_\rho/n\right)+1},\quad c_0=\frac{2\zeta_o-\eta-1}{1-\eta},\quad c_1=\frac{(1+\eta)(1-\zeta_o)}{(1-\eta)^2},\qquad (2.3a,b,c)$$

ζ_o being the value of ζ at $r=r_o$, showing that the dimensionless polytrope is completely determined once n, η, N_ρ are specified. Note that the dimensionless r satisfies

$$r_i=\frac{\eta}{1-\eta}<r<r_o=\frac{1}{1-\eta}.\qquad (2.4)$$

At the mid-point of the layer, where ρ, p and T have their reference values, $\zeta=1$. The departure from the Boussinesq case, which is the limit $N_\rho\to 0$ (see § A) is measured by N_ρ. In the case $N_\rho=5$, which we use below to illustrate a strongly compressible case, the density ratio between the inner and outer boundaries is about 150.

The equations are formulated in terms of entropy, which for a perfect gas is given by

$$S=c_p\left(\frac{1}{\gamma}\ln p-\ln\rho\right),\qquad (2.5)$$

where c_p is the specific heat at constant pressure and γ is the ratio of the specific heats c_p/c_v. In the anelastic approximation, we assume the convection is driven by small disturbances p_c and ρ_c to the reference state pressure $p(r)$ and density $\rho(r)$ (small meaning $p_c\ll p$ and $\rho_c\ll\rho$). The corresponding entropy perturbation is

$$S_c=\frac{c_p}{\gamma}\left(\frac{p_c}{p}-\frac{\gamma\rho_c}{\rho}\right).\qquad (2.6)$$

Lantz (1992) and Braginsky & Roberts (1995) independently discovered that when the basic reference state is close to an adiabatic state, the nonlinear momentum equation can be written

$$\frac{\partial\boldsymbol{u}}{\partial t}=\boldsymbol{u}\times\boldsymbol{\omega}-2\boldsymbol{\Omega}\times\boldsymbol{u}-\nabla\left(\frac{p_c}{\rho}+\frac{1}{2}\boldsymbol{u}^2\right)+\nu\boldsymbol{F}_v-\boldsymbol{g}\frac{S_c}{c_p}.\qquad (2.7)$$

Here $\boldsymbol{u}$ is the velocity, $\boldsymbol{\Omega}$ the angular rotation vector, $\boldsymbol{\omega}$ the vorticity and

$$\boldsymbol{F}_v=\frac{1}{\rho}\frac{\partial}{\partial x_j}\rho\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)-\frac{2}{3\rho}\frac{\partial}{\partial x_i}\rho\frac{\partial u_j}{\partial x_j}.\qquad (2.8)$$

This form of the viscous force corresponds to a constant kinematic viscosity ν (Batchelor 1967, p. 164 and p. 175). In many applications, the viscosity will be a turbulent viscosity, whose precise form is uncertain, but the present form has

the advantage of simplicity. An alternative, not explored here, would be to choose constant dynamic viscosity $\mu = \rho\nu$.

The significant difference between (2.7) and the more general compressible equation of motion (see e.g. Chandrasekhar 1961, equation (18)) is that use has been made of the relation

$$-\frac{1}{\rho}\nabla p_c + \boldsymbol{g}\frac{\rho_c}{\rho} = -\nabla\left(\frac{p_c}{\rho}\right) - \boldsymbol{g}\frac{S_c}{c_p} + \frac{p_c}{\rho}\left\{\frac{1}{\gamma p}\frac{\mathrm{d}p}{\mathrm{d}r} - \frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r}\right\}\hat{\boldsymbol{r}} \approx -\nabla\left(\frac{p_c}{\rho}\right) - \boldsymbol{g}\frac{S_c}{c_p}. \quad (2.9)$$

For a fuller discussion, see §4.2 of Braginsky & Roberts (1995), in particular their equation (4.19). If $\gamma = 1 + 1/n$, using (2.2*a*, *d*) the term in the curly braces is zero, and provided γ is close to this value and any departure of p and ρ from the exact polytrope is small, this approximation is valid. The great advantage of this representation is that when the curl and double curl of (2.7) are taken, the only thermodynamic convective variable left is the entropy; so we avoid having to solve a separate Poisson equation for the pressure perturbation (see e.g. Clune *et al.* 1999), which would be required to evaluate ρ_c. Since it is often the case that convection in planets and stars leads to a reference state that is close to adiabatic, (2.9) is a useful approximation. Note that if the layer were an exact polytrope with $\gamma = 1 + 1/n$, the entropy drop across the layer would be zero, and no convection would occur; so there must be small departures of p and ρ from the exact polytropic values which give rise to a small but finite entropy drop ΔS across the layer which is the same order of magnitude as S_c.

The continuity equation has the anelastic form (e.g. Gough 1966)

$$\nabla \cdot \rho\,\boldsymbol{u} = 0. \quad (2.10)$$

To derive the entropy equation, we start with the dimensional equation of heat transfer in the absence of external heat sources (Landau & Lifshitz 1959, equation (49.4)),

$$\rho T\left(\frac{\partial S}{\partial t} + (\boldsymbol{u}\cdot\nabla)S\right) = \nabla\cdot\rho c_p\kappa_m\nabla T + \rho\nu Q, \quad (2.11)$$

where κ_m is the thermal diffusivity due to molecular processes (thermal conductivity and radiative conductivity) and

$$Q = 2\left[e_{ij}e_{ij} - \frac{1}{3}(\nabla\cdot\boldsymbol{u})^2\right], \quad \text{with} \quad e_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right), \quad (2.12)$$

is the viscous heating. The linearized form of this equation was used in the previous studies of linear compressible convection (Glatzmaier & Gilman 1981*a*; Drew *et al.* 1995). However, in planets and stars the turbulence will give rise to a diffusion of entropy which will normally be much larger than the molecular conductivity term. Furthermore, in compressible flow, turbulent elements preserve their entropy, not their temperature, when the conductivity is small. Prandtl's mixing-length ideas suggest that turbulent elements will move a certain distance and then release their entropy content into their surroundings. This suggests that the turbulent entropy flux is proportional to the entropy gradient, not the temperature gradient. While in Boussinesq convection, eddy thermal diffusion can take a form similar to the molecular thermal diffusion, but with a much larger diffusivity, in compressible flow this is no longer the case. This was recognized by Gilman & Glatzmaier (1981), who included a diffusive flux proportional to $\nabla T - \nabla T_{ad}$, i.e. proportional to potential temperature rather than actual temperature. We must now decide whether to model the effect of turbulence, bearing in mind that there is no universally agreed theory of how this should be

done, or ignore it. The model used here was developed by Braginsky & Roberts (1995) in the context of Earth's core convection (see their §§4.3 and 4.4, particularly their equations (4.22), (4.30) and (4.38)) and was used (independently) in the stellar convection context by Clune *et al.* (1999, equation (3)). The essential assumption is that there is a turbulent velocity $\boldsymbol{u}^t$ which gives rise to a turbulent entropy fluctuation S^t, and these can be averaged over a short length scale so that $\overline{\boldsymbol{u}^t} = \overline{S^t} = 0$, but $\overline{\rho \boldsymbol{u}^t S^t} = \boldsymbol{I}^t$, a non-zero entropy flux. We now adopt the well-known gradient–diffusion model (see e.g. Davidson 2004, p. 165),

$$\boldsymbol{I}^t = -\rho\kappa\nabla S. \tag{2.13}$$

If we view the small-scale velocity $\boldsymbol{u}^t$ as prescribed, the small-scale turbulent entropy fluctuation is forced by the term $\rho(\boldsymbol{u}^t \cdot \nabla)S$ and so is linearly proportional to ∇S. The most general form for the turbulent entropy flux is then

$$I_i^t = -\rho\kappa_{ij}\frac{\partial S}{\partial x_j}, \tag{2.14}$$

where κ_{ij} is an anisotropic eddy diffusivity. This argument relies on the effect of turbulence being local; so the correlation length of the turbulence must be much less than any radius of curvature length of the entropy profile S. Furthermore, as pointed out by Braginsky & Roberts (1995), it is not clear that the isotropic form $\kappa_{ij} = \kappa\delta_{ij}$ is always appropriate in a rotating system, but one might hope that if the unresolved turbulent velocity has sufficiently small length and time scales, it will be unaffected by the rotation. We adopt here the isotropic form (2.13), which has been the most popular choice in simulations, e.g. Clune *et al.* (1999), and which forms the basis of the much used nonlinear anelastic spherical harmonic (ASH) code for anelastic convection.

Just as molecular diffusion gives a source term creating entropy, so turbulent diffusion also gives rise to a source term σ^t; so

$$\rho\left(\frac{\partial S}{\partial t} + (\boldsymbol{u}\cdot\nabla)S\right) = -\nabla\cdot\boldsymbol{I}^t + \sigma^t + \frac{1}{T}\nabla\cdot\rho c_p\kappa_m\nabla T + \frac{\rho\nu Q}{T}, \quad \sigma^t = -\frac{1}{T}(\boldsymbol{I}^t\cdot\nabla)T, \tag{2.15}$$

the expression for σ^t being given by Braginsky & Roberts (1995, equations (4.37) and (3.7*b*)). So the nonlinear heat transport equation is

$$\rho T\left(\frac{\partial S}{\partial t} + (\boldsymbol{u}\cdot\nabla)S\right) = \nabla\cdot\rho T\kappa\nabla S + \nabla\cdot\rho c_p\kappa_m\nabla T + \rho\nu Q. \tag{2.16}$$

Note that this form of σ^t ensures that the turbulent diffusion appears only as a divergence in (2.16); so there is no source of energy arising from the turbulence, only a source of entropy, consistent with the first law of thermodynamics. Consistency with the second law requires that the entropy source term is positive, which from (2.13) and (2.15) requires $\nabla S\cdot\nabla T \geqslant 0$. Since the layer is unstably stratified, this will normally be the case. In this paper, we adopt the opposite extreme of that taken in Drew *et al.* (1995) and ignore the molecular κ_m in comparison with the turbulent κ. Since the reference state is assumed to be close to adiabatic, we can take $S = S_c$ in the nonlinear heat transport equation. We can always add an arbitrary constant to entropy; so we can take the entropy S_c as zero at $r = r_o$ and ΔS at $r = r_i$.

We non-dimensionalize our equations, using the length scale d, time scale d^2/ν, (where ν is the constant kinematic viscosity), mass $\rho_0 d^3$ and unit of entropy $Pr\,\Delta S$ (where Prandtl number $Pr = \nu/\kappa$, κ is the constant entropy diffusion coefficient and

ΔS the entropy drop across the layer); so the dimensionless entropy S satisfies the boundary conditions

$$S = Pr^{-1} \text{ on } r = r_i, \quad S = 0 \text{ on } r = r_o. \tag{2.17a, b}$$

The six dimensionless parameters that govern anelastic compressible convection are

$$\left.\begin{aligned} Ra = \frac{GMd\Delta S}{\nu\kappa c_p}, \quad Pr = \frac{\nu}{\kappa}, \quad E = \frac{\nu}{\Omega d^2} \\ N_\rho = \ln\left(\frac{\rho(r_i)}{\rho(r_o)}\right), \quad n, \quad \eta = \frac{r_i}{r_o}, \end{aligned}\right\} \tag{2.18a–f}$$

Ra being the Rayleigh number and E the Ekman number.

We now linearize the entropy equation and put it in dimensionless form. We linearize about a state with $\boldsymbol{u} = 0$ and no time dependence. The basic state entropy $\bar{S}(r)$ is determined by the nonlinear entropy equation (2.16) together with the boundary conditions (2.17). Since we neglect κ_m, (2.16) becomes

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r} r^2 \zeta^{n+1} \frac{\mathrm{d}\bar{S}}{\mathrm{d}r} = 0, \tag{2.19}$$

using (2.2a, d). The dimensionless solution, using boundary conditions (2.17) is

$$\bar{S} = \frac{Pr^{-1}\left(\zeta_o^{-n} - \zeta^{-n}\right)}{\zeta_o^{-n} - \zeta_i^{-n}}, \tag{2.20}$$

where

$$\zeta_i = c_0 + \frac{c_1}{r_i}, \quad \zeta_o = c_0 + \frac{c_1}{r_o}. \tag{2.21}$$

We now assume $\boldsymbol{u}$ and S' are small and ignore second-order quantities. The basic state entropy $\bar{S}(r)$ can be balanced by a (small) static pressure in (2.7), so we can replace S_c in (2.7) by S'. We obtain

$$Pr\frac{\partial S'}{\partial t} = -Pr\,\boldsymbol{u}\cdot\nabla\bar{S} + \zeta^{-n-1}\nabla\cdot\zeta^{n+1}\nabla S'. \tag{2.22}$$

Note that a term from the advection down the mean entropy gradient, $\nabla\bar{S}$, has now appeared. The boundary conditions on S' are taken as fixed entropy conditions, so

$$S' = 0 \quad \text{on} \quad r = r_i,\ r_o\,. \tag{2.23}$$

The linearized, dimensionless form of (2.7) used in this paper is

$$\frac{\partial \boldsymbol{u}}{\partial t} = -2E^{-1}\hat{\boldsymbol{z}}\times\boldsymbol{u} - \nabla\left(\frac{p'}{\rho}\right) + \boldsymbol{F}_v + \frac{RaS'}{r^2}\hat{\boldsymbol{r}}. \tag{2.24}$$

Equations (2.10), (2.22) and (2.24) form the basis of the remainder of this paper. It should be noted that we used the turbulent diffusion of entropy to determine the basic static entropy state about which we linearized. This may appear surprising because the static state cannot be turbulent. However, we view the linear theory as the small-amplitude limit of the nonlinear problem, and even at very low amplitudes turbulent diffusion will dominate molecular diffusion, and so the state described by (2.20) will be approached as the Rayleigh number is reduced towards critical. Ultimately, as the amplitude falls further, convection will be so slow that the turbulent diffusion will fall below even the molecular diffusion, and a new basic state determined by (2.16) with only molecular diffusion will be approached. We do not consider such extremely small-amplitude convection here, as we are primarily interested in the limit

$Ra \to Ra_{crit}$ from above. The Boussinesq limit of these equations is discussed in § A.1. Equations (2.22) and (2.24) are complemented by the boundary conditions (2.23) and either no-slip or stress-free boundary conditions for the velocity $\boldsymbol{u}$. The numerical implementation of the velocity boundary conditions is discussed in § 5 below.

3. Impossibility of convection at negative *Ra*

A surprising result of Drew *et al.* (1995) was the existence of growing modes even at negative *Ra* at some parameter values. In that paper it was conjectured that this anomalous behaviour was due to the presence of temperature diffusion rather than entropy diffusion in the entropy equation (2.16), i.e. $\kappa = 0$ but $\kappa_m \neq 0$. Using entropy diffusion only, (2.22), one might expect that negative Rayleigh number always gives stability, and we now prove that this is indeed the case.

We multiply (2.24) by $\rho\boldsymbol{u}$ and integrate over the whole shell to get

$$\frac{\partial}{\partial t}\int \frac{1}{2}\zeta^n \boldsymbol{u}^2 \,\mathrm{d}v = \int \zeta^n \boldsymbol{u}\cdot\boldsymbol{F}_v \,\mathrm{d}v + Ra\int \zeta^n S'\frac{u_r}{r^2}\,\mathrm{d}v, \tag{3.1}$$

where the pressure term is removed using the divergence theorem and (2.10). Now multiply (2.22) by $RaPr^{-1}\rho(\mathrm{d}\bar{S}/\mathrm{d}r)^{-1}S'/r^2$ and integrate over the shell to get

$$\frac{RaPr}{2nc_1}(\zeta_o^{-n}-\zeta_i^{-n})\frac{\partial}{\partial t}\int \zeta^{2n+1}(S')^2\,\mathrm{d}v = Ra\int \zeta^n S'\frac{u_r}{r^2}\,\mathrm{d}v + \frac{Ra(\zeta_o^{-n}-\zeta_i^{-n})}{nc_1}\int \zeta^n S'\nabla\cdot\zeta^{n+1}\nabla S'\,\mathrm{d}v, \tag{3.2}$$

Subtracting (3.2) from (3.1) we obtain

$$\frac{\partial}{\partial t}\int \frac{1}{2}\zeta^n\boldsymbol{u}^2 - \frac{RaPr}{2nc_1}(\zeta_o^{-n}-\zeta_i^{-n})\zeta^{2n+1}(S')^2\,\mathrm{d}v = \int \zeta^n\boldsymbol{u}\cdot\boldsymbol{F}_v\,\mathrm{d}v - \frac{Ra(\zeta_o^{-n}-\zeta_i^{-n})}{nc_1}\int \zeta^n S'\nabla\cdot\zeta^{n+1}\nabla S'\,\mathrm{d}v. \tag{3.3}$$

Now if $Ra < 0$, the integral on the left-hand side of (3.3) is positive; so for growing modes the left-hand side must be positive. However, we show below that for negative *Ra* both integrals on the right-hand side are non-positive. It is therefore impossible to have growing modes at negative Rayleigh number.

3.1. *Viscous term*

Using (2.8) with the summation convention,

$$V = \int \zeta^n\boldsymbol{u}\cdot\boldsymbol{F}_v\,\mathrm{d}v = \int \frac{\partial}{\partial x_j}\left\{u_i\rho\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)-\frac{2}{3}u_j\rho\frac{\partial u_i}{\partial x_i}\right\}\mathrm{d}v - \int \frac{\partial u_i}{\partial x_j}\rho\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\mathrm{d}v + \frac{2}{3}\int \rho\frac{\partial u_i}{\partial x_i}\frac{\partial u_j}{\partial x_j}\,\mathrm{d}v. \tag{3.4}$$

The divergence term in (3.4) vanishes if either no-slip or stress-free boundary conditions apply; so

$$\begin{aligned} V &= -\frac{1}{2}\int \rho\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)\mathrm{d}v + \frac{2}{3}\int \rho\left(\frac{\partial u_i}{\partial x_i}\right)^2\mathrm{d}v \\ &= -\frac{4}{3}\int \rho\left(\frac{\partial u_i}{\partial x_i}\right)^2\mathrm{d}v - \frac{1}{2}\int \rho\sum_{i\neq j}\left(\frac{\partial u_i}{\partial x_j}+\frac{\partial u_j}{\partial x_i}\right)^2\mathrm{d}v \leqslant 0, \end{aligned} \tag{3.5}$$

establishing that for either no-slip or stress-free boundary the viscous term is always negative.

3.2. *Entropy term*

Define

$$H = \int \zeta^n S' \nabla \cdot \zeta^{n+1} \nabla S' \, \mathrm{d}v = \int \nabla \cdot (\zeta^{2n+1} S' \nabla S') \, \mathrm{d}v - \int \zeta^{n+1} \nabla (\zeta^n S') \cdot \nabla S' \, \mathrm{d}v. \quad (3.6)$$

The divergence term vanishes if $S' = 0$ on the boundaries, the case studied here, and also if the normal derivative of S' vanishes on the boundaries. It is not immediately apparent that the second term on the right of (3.6) has definite sign, as it is not a square. However, we can rewrite this term so that

$$H = -\int \zeta \, [\nabla (\zeta^n S')]^2 \, \mathrm{d}v + \int (\zeta^{1-n} \nabla \zeta^n) \cdot \nabla \frac{1}{2} (\zeta^n S')^2 \, \mathrm{d}v \quad (3.7)$$

The first term is clearly now negative definite, and the second term vanishes if $S' = 0$ on the boundaries. To see this, using (2.2*b*),

$$\int (\zeta^{1-n} \nabla \zeta^n) \cdot \nabla \frac{1}{2} (\zeta^n S')^2 \, \mathrm{d}v = -\frac{n c_1}{2} \int \int \int \frac{\partial}{\partial r} (\zeta^n S')^2 \, \sin\theta \, \mathrm{d}r \, \mathrm{d}\theta \, \mathrm{d}\phi = 0, \quad (3.8)$$

provided the boundary conditions are $S' = 0$ at the surface.

This establishes that both V and H are non-positive; so there cannot be growing modes if $Ra < 0$. In the case $Ra = 0$, it is also easy to establish there can be no growing modes; so unlike the Drew *et al.* (1995) problem, we only have linear growth for $Ra > 0$. The method can be generalized for any thermal diffusivity of the form $\kappa = \kappa_0 \rho^\alpha$, which includes the case of constant thermal conductivity $k = \kappa\rho$, $\alpha = -1$. The reference state is now

$$\bar{S} = \frac{\zeta_o^{-n(1+\alpha)} - \zeta^{-n(1+\alpha)}}{Pr \left[\zeta_o^{-n(1+\alpha)} - \zeta_i^{-n(1+\alpha)} \right]} \ (\alpha \neq -1), \quad \bar{S} = \frac{\ln \zeta_o - \ln \zeta}{Pr \, [\ln \zeta_o - \ln \zeta_i]} \ (\alpha = -1). \quad (3.9)$$

The method of proof to establish that growing modes cannot occur for negative Ra is the same as above.

Of course, this proof does not rule out the possibility of subcritical instability to nonlinear disturbances. Note that the proof fails if we use temperature diffusion, that is ignoring the turbulent term in (2.16), rather than entropy diffusion, which ignores the κ_m term. Then the integrals corresponding to (3.6) contain products of $\nabla S'$ and $\nabla T'$, and no definite conclusions about sign can be deduced. There is therefore no contradiction here with the results of Drew *et al.* (1995).

Note also that our proof does depend on specific assumptions about the equilibrium model. If for example there were internal heating, then the form of $\mathrm{d}\bar{S}/\mathrm{d}r$ would change, and our method might no longer apply. Also, we required the boundary condition $S' = 0$ on the boundaries to establish (3.8); it remains possible that with other boundary conditions negative Rayleigh number instability can occur even with entropy diffusion.

4. Small E asymptotic theory

The onset of Boussinesq convection in a rapidly rotating sphere has been solved in the asymptotic limit $E \to 0$ (Jones *et al.* 2000) and in the spherical shell case (Dormy *et al.* 2004). Here we extend this Wentzel-Kramers-Brillouin (WKB) theory from Boussinesq to anelastic compressible convection.

In cylindrical polar coordinates (s, ϕ, z) we can satisfy the continuity equation (2.10) setting

$$\zeta^n \boldsymbol{u} = \nabla \times \Psi \,\widehat{\boldsymbol{z}} + \nabla \times \nabla \times \Xi \,\widehat{\boldsymbol{z}}. \tag{4.1}$$

In the limit $E \to 0$, the numerical solutions discussed below indicate that the convection at onset takes the form of tall thin columns in compressible convection as well as in Boussinesq convection. Following Jones *et al.* (2000) we adopt the scalings

$$\frac{1}{s}\frac{\partial}{\partial \phi} \sim \frac{\partial}{\partial s} \sim O(E^{-1/3}), \quad \frac{\partial}{\partial z} \sim O(1), \tag{4.2}$$

when acting on perturbed quantities. Since we are seeking WKB solutions, we assume disturbances are proportional to

$$\exp[\mathrm{i}(ks + m\phi - \omega t)], \tag{4.3}$$

where ω is in general complex. So at leading order,

$$\nabla^2,\ \nabla_H^2 \to -\left(k^2 + \frac{m^2}{s^2}\right) = -a^2. \tag{4.4}$$

At the boundaries, $u_r = 0$, so in general u_z and u_s are of the same order; therefore from (4.1) $\Xi \sim E^{1/3}\Psi$. It follows that

$$\zeta^n u_s = \frac{\partial^2 \Xi}{\partial z \partial s} + \frac{1}{s}\frac{\partial \Psi}{\partial \phi} \sim \frac{1}{s}\frac{\partial \Psi}{\partial \phi}, \quad \zeta^n u_\phi = \frac{1}{s}\frac{\partial^2 \Xi}{\partial z \partial \phi} - \frac{\partial \Psi}{\partial s} \sim -\frac{\partial \Psi}{\partial s}. \tag{4.5}$$

The appropriate asymptotic scalings for the variables are (Jones *et al.* 2000)

$$\left.\begin{aligned} Ra = E^{-4/3}\mathscr{R}, \quad \omega = E^{-2/3}\hat{\omega}, \quad m = E^{-1/3}\hat{m}, \quad k = E^{-1/3}\hat{k}, \\ a = E^{-1/3}\hat{a}, \quad S' = S', \quad \Psi = E^{-1/3}\psi, \quad u_z = E^{-2/3}w. \end{aligned}\right\} \tag{4.6}$$

4.1. *Equations for the z structure and the local dispersion relation*

We insert these expressions into the z component of the curl of the momentum equation (2.24) and the z component of the double curl of the momentum equation, retaining only leading-order terms. A considerable simplification results because gradients of the density are only $O(1)$, whereas horizontal derivatives of perturbed quantities are larger at $O(E^{-1/3})$. Together with the entropy equation, we obtain

$$\frac{1}{\zeta^n}\frac{\mathrm{d}\psi}{\mathrm{d}z} = \frac{1}{2}(\hat{a}^2 - \mathrm{i}\hat{\omega})w - \frac{\mathscr{R} z S'}{2r^3} + \frac{nz\psi}{r\zeta^{n+1}}\frac{\mathrm{d}\zeta}{\mathrm{d}r}, \tag{4.7}$$

$$\frac{\mathrm{d}w}{\mathrm{d}z} = \left(\frac{\hat{a}^2}{2\zeta^n}(\hat{a}^2 - \mathrm{i}\hat{\omega}) - \frac{\mathrm{i}\hat{m}n}{r\zeta^{n+1}}\frac{\mathrm{d}\zeta}{\mathrm{d}r}\right)\psi + \frac{\mathrm{i}\hat{m}\mathscr{R}S'}{2r^3} - \frac{nzw}{\zeta r}\frac{\mathrm{d}\zeta}{\mathrm{d}r}, \tag{4.8}$$

$$S' = \frac{1}{(\mathrm{i}\hat{\omega}Pr - \hat{a}^2)(\zeta_o^{-n} - \zeta_i^{-n})}\frac{\mathrm{d}\zeta}{\mathrm{d}r}\left(\frac{\mathrm{i}\hat{m}n\psi}{r\zeta^{2n+1}} + \frac{nzw}{\zeta^{n+1}r}\right). \tag{4.9}$$

On eliminating S' we obtain a second-order two-point boundary value problem in z with eigenvalue $\hat{\omega}$. The boundary conditions are

$$\mathrm{i}\hat{m}\psi + zw\zeta^n = 0, \quad \text{on} \quad z = \pm\left(\frac{1}{(1-\eta)^2} - s^2\right)^{1/2}. \tag{4.10}$$

This system is the compressible equivalent of the Roberts–Busse equations (see equation (3.5) of Jones *et al.* 2000 and equation (3.11) of Dormy *et al.* 2004 for the

Boussinesq equivalents). It defines the local dispersion relation for $\hat{\omega}$ in terms of the parameters. There are solutions both symmetric and antisymmetric about the equator, but as in the Boussinesq problem, the first modes to onset always appear to be those symmetric in ψ and antisymmetric in w. The system has to be solved numerically, but of course this is a very simple one-dimensional problem compared to the task of solving the full system numerically, which involves the inversion of very large matrices (see § A.2).

4.2. *Local analysis*

The first step in analysing the dispersion relation is to require that the growth rate $Im\{\hat{\omega}\}=0$ and also that $\partial Im\{\hat{\omega}\}/\partial\hat{m}=0$ and $\partial Im\{\hat{\omega}\}/\partial s=0$. As in the Boussinesq case, $\hat{\omega}$ is a function of $\hat{k}^2$ rather than $\hat{k}$ alone; so minimizing over $\hat{k}$ results in $\hat{k}=0$. We hold Pr, N_ρ and n fixed and find the values of $\hat{m}$, $\mathscr{R}$ and s that satisfy these three conditions. The critical value of s found may lie in (i) $0<s<\eta/(1-\eta)$ or (ii) $\eta/(1-\eta)<s<1/(1-\eta)$, depending on the parameters. The system is singular at $s=1/(1-\eta)$, and critical s cannot exceed that value. As explained in Dormy *et al.* (2004), these two cases must be treated differently. In case (i) the local maximum of $Im\{\hat{\omega}\}$ does not lie inside the fluid. In consequence, the minimum critical Rayleigh number is achieved at the tangent cylinder, and convection will occur there first as $\mathscr{R}$ is increased. The leading-order value of critical $\mathscr{R}$ is given by setting $s=s_i$, its value at the tangent cylinder, and solving $Im\{\hat{\omega}\}=0$ and $\partial Im\{\hat{\omega}\}/\partial\hat{m}=0$ for $\mathscr{R}$ and $\hat{m}$. The width of the convective region near onset is $O(E^{2/9})$, and the (non-zero) frequency is given by $Re\{\hat{\omega}\}$.

4.3. *Global analysis*

In case (ii), the global theory of instability must be used. Now the WKB theory predicts that onset lies in the neighbourhood of some $s=s_M$ inside the fluid; so a solution is required for which the amplitude decays to zero as $(s-s_M)/E^{1/3}\to\pm\infty$. For such solutions to exist, there must be a value of s in the complex plane at which both the real and imaginary parts of $\partial\hat{\omega}/\partial s=0$. So these two conditions, together with $Im\{\hat{\omega}\}=0$ and $\partial Im\{\hat{\omega}\}/\partial\hat{m}=0$, give four equations for four unknowns, $\mathscr{R}$, $\hat{m}$ and $s=s_r+\mathrm{i}s_i$. The value of $\mathscr{R}_c$ that emerges from global theory is larger by an order one amount from the local value. Once the turning point $s_c=s_r+\mathrm{i}s_i$ has been identified in the complex plane, $\mathscr{R}_c$, $\hat{\omega}_c$ and $\hat{m}_c$ are determined; so the dispersion relation $\hat{\omega}(\hat{k},s)=\hat{\omega}_c$ becomes an equation that determines complex $\hat{k}$ as a function of complex s, and at s_c, $\hat{k}$ has a double zero. If we insert real values s and find the value of $s=s_M$ at which $\hat{k}$ is purely real, then the solution has s dependence in the neighbourhood of $s=s_M$,

$$\sim\exp\left[\mathrm{i}\hat{k}_M\frac{s-s_M}{E^{1/3}}+\frac{\mathrm{i}\hat{k}'_M}{2}\frac{(s-s_M)^2}{E^{1/3}}\right],\quad\text{with}\quad\hat{k}'_M=\left(\frac{\mathrm{d}\hat{k}}{\mathrm{d}s}\right)_M,\tag{4.11}$$

provided the $\hat{k}$ root with $Im\{\mathrm{d}\hat{k}/\mathrm{d}s\}>0$ is chosen. It follows that $s=s_M$ is where the convection has maximum amplitude in case (ii). The quantities $\hat{k}_M$ and $Im\{\mathrm{d}\hat{k}/\mathrm{d}s\}_M$ give the radial wavenumber near the onset of convection and the inverse width of the convecting region respectively; $\hat{k}_M s_M/\hat{m}$ measures the amount of spiralling in the solution, discussed further in § 6. Other important quantities are the values of s at which the anti-Stokes lines cut the real s-axis, s_- and s_+. The anti-Stokes lines are

defined by the path

$$Im\left\{\int_{s_c}^{s}\hat{k}(s)\,\mathrm{d}s\right\}=0, \tag{4.12}$$

s_c being the double turning point in the complex plane. As discussed in Jones *et al.* (2000), the global asymptotic theory is only valid when the interval (s_-, s_+) lies entirely within the fluid. If $s_- < s_i$ or $s_+ > s_o$, then the boundary conditions at $s = s_i$ or $s = s_o$ matter, and the asymptotic theory becomes considerably more complicated.

4.4. *Algorithms required to evaluate asymptotic results*

To study the asymptotic theory, a suite of five programs is required. First, for any set of parameters η, N_ρ, n and Pr the dispersion relation given by solving (4.7)–(4.10) is used to minimize $\mathscr{R}$ over s and $\hat{m}$ (with $\hat{k}=0$). If the minimizing value of s satisfies $s \leqslant \eta/(1-\eta)$, then we must fix $s=\eta/(1-\eta)$ and use a second program that minimizes $\mathscr{R}$ over $\hat{m}$ only, to get the correct value of $\mathscr{R}_c$. If on the other hand $s > \eta/(1-\eta)$, then the global theory must be applied. We use a third program that solves complex $\mathrm{d}\hat{\omega}/\mathrm{d}s=0$, $Im\{\hat{\omega}\}=0$ and $Im\{\partial\hat{\omega}/\partial\hat{m}\}=0$ for the four unknowns $\mathscr{R}$, $\hat{m}$, s_r and s_i. Then a fourth program which inputs the values of $\mathscr{R}_c$, $\hat{m}$ and $\hat{\omega}_c$ from the global theory program is used to find $\hat{k}(s)$ from the relation $\hat{\omega}_c=\hat{\omega}(\hat{k}, s)$. This program must find the real value of $s=s_M$ at which $Im\{\hat{k}\}=0$ and also find the corresponding values of $\hat{k}=\hat{k}_M$ and $Im\{\mathrm{d}\hat{k}/\mathrm{d}s\}_M$, being careful to select the sign of $\hat{k}_M$ such that $Im\{\mathrm{d}\hat{k}/\mathrm{d}s\}_M > 0$. The fifth program, which finds s_- and s_+, evaluates the complex path integral

$$Im\left\{\int_{s_c}^{s}\hat{k}(s)\,\mathrm{d}s\right\}$$

from the double turning point to any real s, again calculating $\hat{k}$ from $\hat{\omega}_c=\hat{\omega}(\hat{k}, s)$. The program must then find the zeros of this real integral as s varies to obtain s_- and s_+. Since the dispersion relation $\hat{\omega}_c=\hat{\omega}(\hat{k}, s)$ always has two roots $\hat{k}$ with opposite signs, care must be taken to ensure that the same root is taken as the complex integral is evaluated.

5. Numerical formulation for finite E

5.1. *Toroidal and poloidal decomposition*

From the continuity equation (2.10) we can set

$$\boldsymbol{u}=\frac{1}{\rho}\nabla\times\nabla\times\hat{\boldsymbol{r}}f\rho+\frac{1}{\rho}\nabla\times\hat{\boldsymbol{r}}e\rho, \tag{5.1}$$

where e and f are the toroidal and poloidal velocity component potentials, respectively. Note that here we use Chandrasekhar's (1961, Appendix III) definition of toroidal and poloidal scalars based on the unit vector $\hat{\boldsymbol{r}}$. The anelastic continuity equation (2.10) is automatically satisfied for this decomposition. The boundary conditions can be taken to be either no-slip or stress-free,

$$f=e=\frac{\partial f}{\partial r}=0, \quad \text{no-slip}, \tag{5.2}$$

$$f=r\frac{\partial e}{\partial r}-2e=r\frac{\partial^2 f}{\partial r^2}-2\frac{\partial f}{\partial r}+\frac{r}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r}\frac{\partial f}{\partial r}=0, \quad \text{stress-free}. \tag{5.3}$$

Stress-free boundary conditions were used in all the numerical results presented here. At low E, there is not a great deal of difference in the linear theory between stress-free and no-slip boundary conditions, as the leading-order asymptotic results are the same in both cases (see e.g. Dormy *et al.* 2004). It is only when zonal flow generation is considered that the difference between stress-free and no-slip conditions becomes crucial (see e.g. Gillet & Jones 2006).

We are looking for the solutions for velocity and entropy in the form of azimuthal waves

$$e = e(r,\theta)\,\mathrm{e}^{\mathrm{i}(m\phi-\omega t)}, \quad f = f(r,\theta)\,\mathrm{e}^{\mathrm{i}(m\phi-\omega t)}, \quad S' = S'(r,\theta)\,\mathrm{e}^{\mathrm{i}(m\phi-\omega t)}, \tag{5.4}$$

where m is the azimuthal wavenumber, and positive values of the frequency ω correspond to prograde motion and negative values to the retrograde one. At low E prograde modes appear to be the most unstable.

Further details of the toroidal–poloidal equations, together with a brief description of the numerical method used, are given in § B.

6. Results

6.1. *The Boussinesq case*

Before discussing the effects of compressibility, we first recall some of the known features of Boussinesq rapidly rotating convection, gleaned from Busse (1970), Zhang (1992), Jones *et al.* (2000), Al-Shamali, Heimpel & Aurnou (2004) and Dormy *et al.* (2004). In spherical shell models at low E, convection first occurs outside the tangent cylinder surrounding the inner core. In the case of differential heating, that is no internal heating and a prescribed temperature drop across the shell, the onset of convection occurs in the neighbourhood of the tangent cylinder, whereas with internal heating onset can occur in the interior of the shell. In this paper we only consider the differential heating case. With slow rotation, convection may be axisymmetric at onset (Geiger & Busse 1981), but in rapidly rotating systems, convection always takes a non-axisymmetric columnar form except at very low Prandtl number (Zhang 1994) at which inertial modes may occur first. In this paper, we have not explored these very low-Pr cases.

The preferred azimuthal wavenumber at onset increases with radius ratio. At $Pr = 1$ and $E = 2 \times 10^{-4}$, $m = 6$ is preferred at $\eta = 0.35$, but onset with $m = 61$ is first at $\eta = 0.85$. Asymptotically, the Rayleigh number scales as $Ra \sim E^{-4/3}$ as $E \to 0$, $m \sim E^{-1/3}$, and the frequency of the most unstable mode scales as $\omega \sim E^{-2/3}$, although this asymptotic dependence is only completely established at very low E (Dormy *et al.* 2004), and for $E \approx 10^{-4}$ Al-Shamali *et al.* (2004) found that for the Rayleigh number dependence an exponent $E^{-1.16}$ fitted the numerical data best. The frequency first increases and then decreases as η is increased. If the Prandtl number is decreased, the critical azimuthal wavenumber is somewhat reduced, and the frequency is increased. Thus at $Pr = 0.1$, $E = 2 \times 10^{-4}$ and $\eta = 0.35$, $m = 5$ is preferred, while at $\eta = 0.85$, $m_{crit} = 43$ with the frequency being about three times the $Pr = 1$ value.

In all cases, the mode with z vorticity symmetric about the equator (and the z velocity antisymmetric) is preferred over the modes with the opposite symmetry (Busse 1970).

6.2. *Compressible results*

Dormy *et al.* (2004) noted that if the heat flux driving the convection were applied at the inner boundary and if there were no internal heating (the case there called

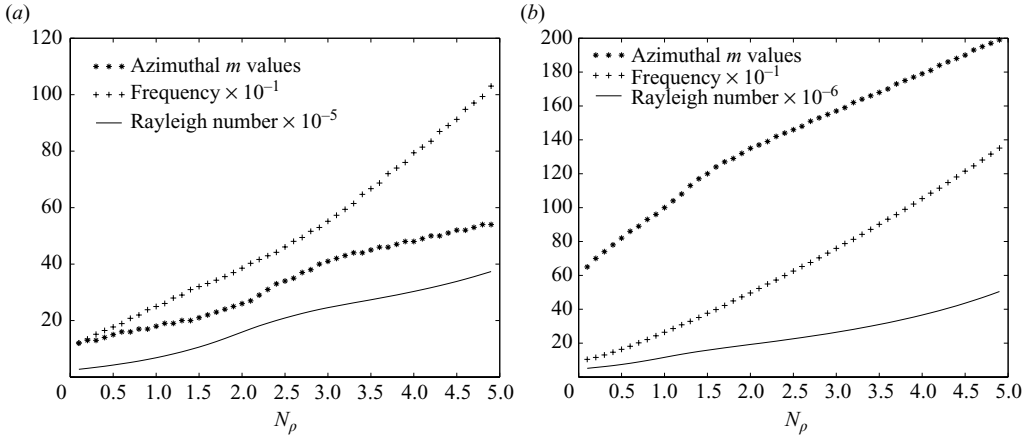

FIGURE 1. Critical Ra, ω and m as functions of N_ρ ($E = 2 \times 10^{-4}$, $Pr = 1$, $n = 2$): (*a*) $\eta = 0.5$; (*b*) $\eta = 0.85$.

differential heating), then there would be onset of convection at the tangent cylinder; i.e. case (i) of § 4.2 would apply irrespective of the size of the inner core. Since this is the form of heating adopted here, we might expect case (i) to occur always. Furthermore, the form of gravity adopted here $g \sim 1/r^2$ favours convection closer to the tangent cylinder than the $g \sim r$ used in Dormy *et al.* (2004). Remarkably, even a rather modest amount of compressibility completely overcomes these effects, and convection frequently occurs in the interior, case (ii). The tendency for compressibility to push the convection towards the outer boundary was noted in Glatzmaier & Gilman (1981*a*) and Drew *et al.* (1995). It is clearly a powerful effect, overcoming the rapid diminution of gravity as r increases.

In figures 1(*a*) and 1(*b*) we show the variation of Ra_{crit}, m_{crit} and ω_{crit} as a function of N_ρ, which increases with increasing density variation. Here $Pr = 1$, $n = 2$ and $E = 2 \times 10^{-4}$. The variation with E may be approximately inferred from the scaling laws, and we consider the effect of varying Pr below. We fix $n = 2$ throughout the paper. This value was motivated by models of Jupiter's atmosphere, where n ranges from approximately 1 near the transition region to around 3 near the surface. We see that the critical Rayleigh number, azimuthal wavenumber and frequency all increase with N_ρ. The increase in Ra and m appears to be due to the tendency for the convection to move outward towards the low-density region when the shell contains many density scale heights. In consequence, the effect of increasing N_ρ is similar to that of increasing η in Boussinesq convection, where as mentioned above Ra_{crit} and m_{crit} increase with η. The increase in frequency cannot be attributed to this effect because frequency does not vary consistently with η. The increase in frequency is actually due to the local vortex-stretching mechanism in compressible fluids, discussed by Evonuk & Glatzmaier (2004) and Evonuk (2008). In Boussinesq convection, the convection columns form Rossby waves in which the restoring force is due to the vortex stretching associated with the sloping boundary as a column moves towards or away from the rotation axis. In compressible convection, as a fluid element moves into less dense (denser) surroundings it expands (contracts), and its vorticity is stretched (reduced) as it does so. In consequence, there is an additional restoring force acting on the columns, which reinforces the restoring force due to the sloping boundaries. This compressible effect therefore increases the wave frequency,

and the greater the compressibility N_ρ, the faster the wave travels. This effect can also be understood in terms of the Proudman–Taylor theorem. When the vorticity equation is formed, the Coriolis term on taking the curl of $2\boldsymbol{\Omega} \times \boldsymbol{u}$ no longer simply becomes $-2\boldsymbol{\Omega} \cdot \nabla \boldsymbol{u}$ because this result relies on $\nabla \cdot \boldsymbol{u} = 0$, which is no longer true in our compressible situation. In consequence an additional term is present in the vorticity equation, corresponding to the physical mechanism discussed above.

In figures 2(*a*)–2(*c*) we show how the eigenfunctions change as N_ρ increases. The parameters are $E = 2 \times 10^{-5}$, $\eta = 0.5$, $Pr = 1$. The Boussinesq case (figure 2*a*) shows the onset of convection occurring near the tangent cylinder as expected. The value of E is sufficiently small for the asymptotic behaviour to be apparent, and the meridional cross-section, taken at a ϕ value indicated by the horizontal line in figure 2(*a*), shows the 'tall thin column' behaviour. At moderate compressibility, $N_\rho = 2$ in figure 2(*b*), the convection has moved out into the fluid interior, still maintaining the columnar structure. Note that in figure 2(*b*) there is strong prograde spiralling; that is each convective column is located at increasing ϕ as s increases (Zhang 1992).

At large compressibility, onset of the convection is close to the outer shell, and despite the very strong curvature of the outer boundary here the columnar structure is still evident. The physical reason why the outer boundary near the equator is the preferred location in compressible convection appears to lie primarily in the form of the reference state entropy gradient, which from (2.20) is

$$\frac{\mathrm{d}\bar{S}}{\mathrm{d}r} = -\frac{c_1 Pr^{-1}}{\zeta_o^{-n} - \zeta_i^{-n}} \frac{n}{\zeta^{n+1} r^2}; \tag{6.1}$$

it becomes large and negative as ζ becomes small near the outer boundary. In consequence, for a given velocity the corresponding entropy fluctation is much larger near the outer boundary (see (4.9)), and this strongly enhances the buoyancy terms in (4.7) and (4.8), enabling them to overcome the dissipation. Only near the equator is it possible for convection columns to exist predominantly in the low-density region. Far from the equator, columns must extend into the interior, with consequently larger damping relative to only a thin region of strong driving. In Boussinesq convection the large slope of the boundaries as the equatorial region is approached makes the flow strongly ageostrophic; so Boussinesq convection avoids this region, but the strong driving in this region in the compressible case overcomes this disadvantage. The only other place ζ terms enter (4.7)–(4.9) is in the additional terms proportional to $\mathrm{d}\zeta/\mathrm{d}r$ in (4.7) and (4.8), consequences of the compressible continuity equation (2.10), but these do not appear to be responsible for the strong preference of convection to occur at the outer boundary. This explanation is supported by the results of Glatzmaier & Gilman (1981*a*) and Drew *et al.* (1995) who also found convection occurring primarily near the outer boundary in the vicinity of the equator in their constant thermal diffusivity case. With constant thermal diffusivity the model based on molecular thermal conductivity also has a much stronger temperature gradient near the outer boundary. On the other hand, Glatzmaier & Gilman (1981*b*) found that with constant conductivity, i.e. $\kappa \sim 1/\rho$, where the reference state temperature gradient is much more uniform, there was no such tendency for convection to move towards the equatorial region.

The asymptotic regime we consider here is $E \to 0$ with N_ρ fixed. Figure 2(*c*) suggests that another asymptotic regime may develop, in which $E \to 0$ and $N_\rho \to \infty$ and where the modes are trapped at the equator. This double limit has not been explored here in detail. Generally, the frequency of the modes is increasing with N_ρ, and this raises the possibility that these modes may be connected with equatorially

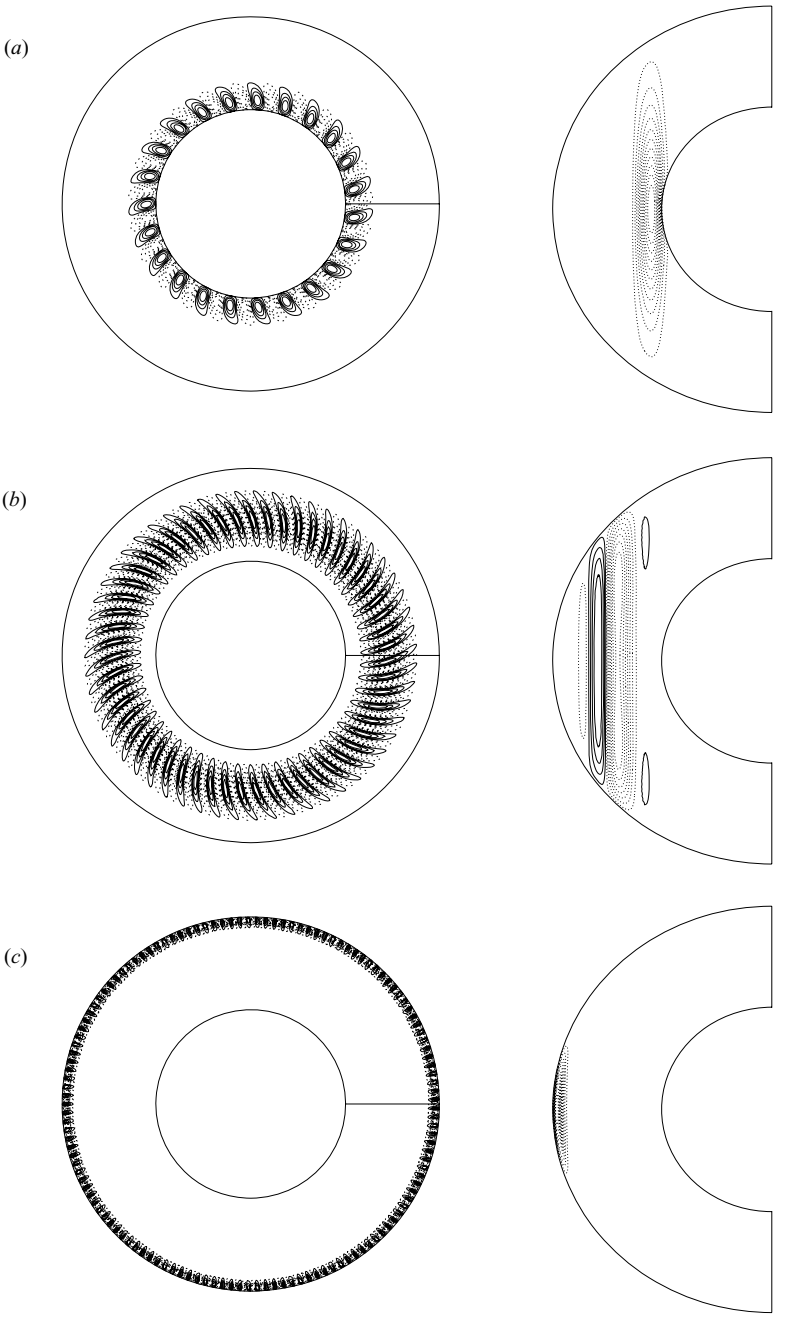

FIGURE 2. Left panel: equatorial section of entropy fluctuation S'. The horizontal radius marks the ϕ location at which the right-hand side meridional section is taken. Right panel: meridional section at the longitude marked in the left panel ($E = 2\times10^{-5}$, $\eta = 0.5$, $Pr = 1$, $n = 2$): (*a*) Boussinesq, $N_\rho = 0$, $Ra_{crit} = 3.2280 \times 10^6$, $\omega_{crit} = 534.36$, $m_{crit} = 23$; (*b*) $N_\rho = 2.0$, $Ra_{crit} = 3.3258 \times 10^7$, $\omega_{crit} = 1844.42$, $m_{crit} = 55$; (*c*) $N_\rho = 5.0$, $Ra_{crit} = 6.6570 \times 10^7$, $\omega_{crit} = 5614.99$, $m_{crit} = 133$.

trapped inertial modes (see equation (4.1) of Zhang, Liao & Busse 2007, where the anelastic form of the inertial wave equation is given). An essential difference between the columnar modes described here and inertial modes are that the frequency of

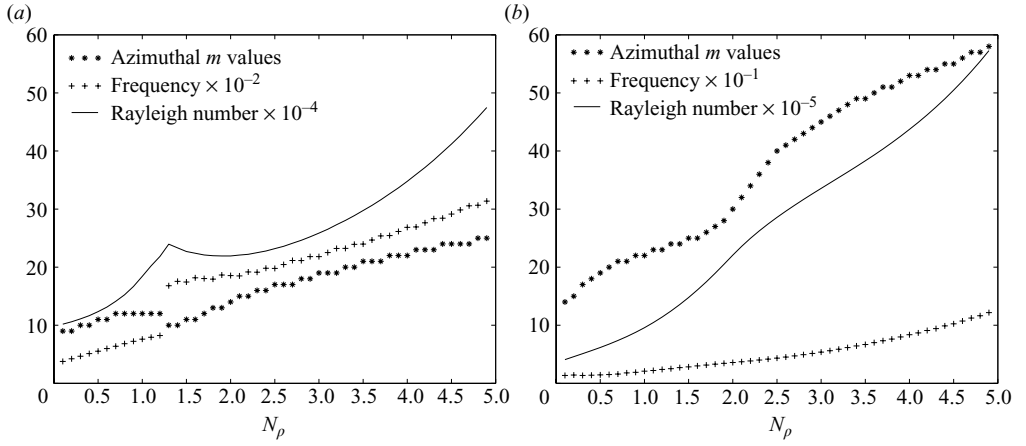

FIGURE 3. Critical Ra, ω and m as functions of N_ρ ($E = 2 \times 10^{-4}$, $\eta = 0.5$, $n = 2$): (a) $Pr = 0.1$; (b) $Pr = 10$.

inertial modes is $O(\Omega)$, whereas the frequency of columnar modes is smaller, only $O(\Omega E^{1/3})$. The increase in frequency as N_ρ increases may represent a move to a more inertial character of the convection in the double limit $E \to 0$ and $N_\rho \to \infty$.

In figures 3(a) and 3(b) we show how varying the Prandtl number affects the critical Rayleigh number, frequency and preferred wavenumber. The general trend is for all these quantities to increase with N_ρ, but the discontinuous frequency near $N_\rho = 1.3$ indicates that mode crossing has occurred. The wave speeds are much higher at low Prandtl number, and the preferred m values are lower.

In figures 4(a)–4(c) we show the eigenfunctions at a low Prandtl number of 0.1. This can be compared with figure 2, which was for Prandtl number equal to 1. The general trend of the convection moving outward as N_ρ increases is found here also, but note that the convecting region is much broader than at $Pr = 1$, and there is very strong prograde spiralling. We found that this strong spiralling only occurs as the Ekman number is reduced below $\sim 10^{-5}$. At $E = 2 \times 10^{-4}$ there is only weak spiralling. The amount of spiralling is important, as zonal flows are only generated if the convection spirals outward. Stronger zonal flows will therefore be generated at lower Prandtl number.

6.3. *Results from the asymptotic theory*

In table 1, results from (a) a wide-gap shell, $\eta = 0.5$, with compressibility $N_\rho = 2$ and (b) a narrow-gap shell, $\eta = 0.8$, at compressibility $N_\rho = 1$ are shown. Both cases were run at $Pr = 0.1$, 1 and 10. In all these cases, onset of instability is in the interior of the fluid, which in case (a) is $1 < s < 2$ and in case (b) is $4 < s < 5$ in the dimensionless unit. The frequency $\hat{\omega}_c$ at onset decreases strongly with Prandtl number. The azimuthal wavenumber $\hat{m}_c$ increases strongly with η. The turning point in the complex plane is at $s_r + \mathrm{i}s_i$, and s_M is where onset of the convection is first; $\hat{k}_M$ is the radial wavenumber at onset, and $\hat{k}_M s_M/\hat{m}_c$ measures the ratio of radial to azimuthal wavenumber. A large value means the convection spirals strongly in the equatorial s–ϕ plane; $\hat{k}_M$ is generally negative, which means the spiralling is prograde, as in figures 2(b) and 4(b); that is the spiral has increasing ϕ as s increases. At high Prandtl number, $\hat{k}_M s_M/\hat{m}_c$ is small, and there is very little spiralling, the convection pattern taking the form of radial spokes. At low Prandtl number $\hat{k}_M s_M/\hat{m}_c$ is large, and there is strong spiralling as noted in figure 4(b). A curiosity is that the spiralling can actually be retrograde

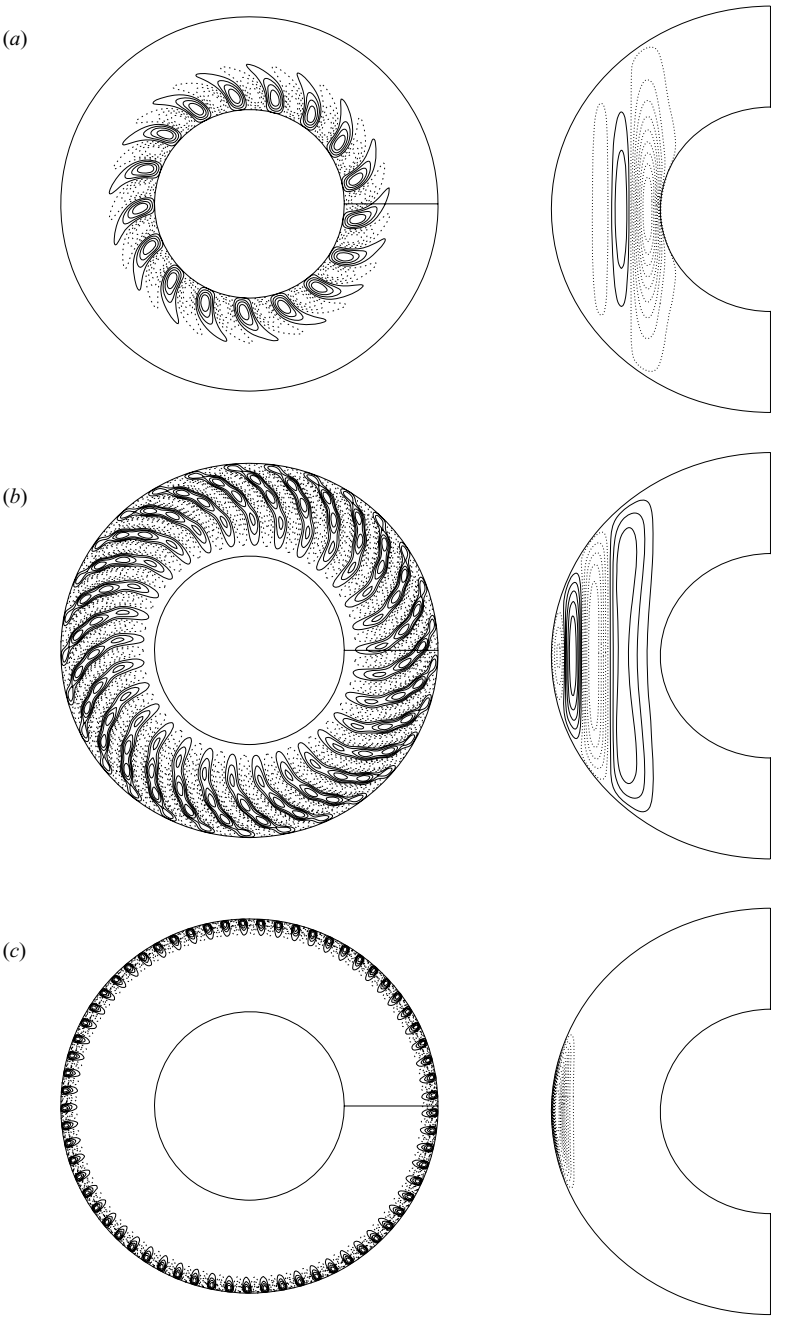

FIGURE 4. Left panel: equatorial section of entropy fluctuation S'. Right panel: meridional section at the longitude marked in the left panel ($E = 2 \times 10^{-5}$, $\eta = 0.5$, $Pr = 0.1$, $n = 2$): (*a*) Boussinesq, $Ra_{crit} = 1.0381 \times 10^6$, $\omega_{crit} = 1606.12$, $m_{crit} = 17$; (*b*) $N_\rho = 2.0$, $Ra_{crit} = 4.6853 \times 10^6$, $\omega_{crit} = 6900.95$, $m_{crit} = 29$; (*c*) $N_\rho = 5.0$, $Ra_{crit} = 7.6817 \times 10^6$, $\omega_{crit} = 18\,933.3$, $m_{crit} = 65$.

at high Prandtl number, as in the case $\eta = 0.8$, $Pr = 10$. In Boussinesq spherical convection, only prograde spiralling has been found (Zhang 1992; Jones *et al.* 2000). The quantity $Im(\mathrm{d}\hat{k}/\mathrm{d}s)_M$ measures how confined the convection is. The low value

η	0.5			0.8		
Pr	0.1	1.0	10.0	0.1	1.0	10.0
N_ρ	2.0	2.0	2.0	1.0	1.0	1.0
$\mathscr{R}_c$	3.2374	17.6348	24.4301	16.6877	73.4440	98.9767
$\hat{\omega}_c$	4.5882	1.3780	0.1260	2.9716	0.9147	0.0853
$\hat{m}_c$	0.7945	1.5310	1.8190	2.5281	4.1368	4.7366
s_r	1.0557	1.3328	1.3670	4.3778	4.2931	4.3310
s_i	−0.2811	−0.1518	−0.0057	−0.7233	−0.2137	0.0466
s_M	1.7438	1.4210	1.3683	4.7044	4.3537	4.3290
$\hat{k}_M$	−0.4974	−0.3247	−0.0150	−0.4623	−0.2141	0.0587
$Im(\mathrm{d}\hat{k}/\mathrm{d}s)_M$	0.2368	1.3348	2.4801	0.4927	0.8675	1.2003
s_-	0.9610	1.2345	1.3624	3.8290	4.1176	4.2794
s_+	—	1.6260	1.3741	—	4.5723	4.3770

TABLE 1. Global bifurcation data.

η	0.5			0.8		
Pr	0.1	1.0	10.0	0.1	1.0	10.0
N_ρ^{LG}	1.7432	1.7706	1.8101	0.8917	0.9064	0.8344
$\mathscr{R}_c$	1.2618	13.5384	20.7268	5.9826	63.2909	81.6280
$\hat{\omega}_c$	5.2824	1.3205	0.1206	3.5073	0.8596	0.0745
$\hat{m}_c$	0.6123	1.1261	1.3150	1.9957	3.6949	4.0979
N_ρ^{AS}	2.0745	1.8154	1.8092	1.0672	0.9458	0.8546
$\mathscr{R}_c$	3.5165	14.9529	20.7292	17.9730	69.3181	83.9908
$\hat{\omega}_c$	4.6658	1.3255	0.1208	3.1356	0.8786	0.0759
$\hat{m}_c$	0.8336	1.2155	1.3181	2.6740	3.9455	4.1839
s_r	1.1024	1.0833	1.0103	4.5077	4.1948	4.0327
s_i	−0.2982	−0.1438	−0.0150	−0.6171	−0.2426	0.0295
s_+	—	1.4379	1.0328	—	4.5227	4.0593

TABLE 2. Local and global bifurcation data; N_ρ^{LG} is the value of N_ρ at which the s value which gives minimum critical Rayleigh number on local theory coincides with the inner sphere. For $N_\rho < N_\rho^{LG}$ onset of convection is first at the tangent cylinder. For $N_\rho > N_\rho^{LG}$ onset of convection is in the interior of the gap, and global theory must be used; N_ρ^{AS} is the value of N_ρ at which the lower anti-Stokes point, s_-, lies on the tangent cylinder; s_+ is the value of the upper anti-Stokes point at $N_\rho = N_\rho^{AS}$, and R_c, ω_c, m_c, s_r and s_i are evaluated using global bifurcation theory at the same point.

at low Prandtl number means the disturbance spreads over a large fraction of the cell, and this can be observed by comparing figure 2 with figure 4. The higher values at high Prandtl number mean that at onset the convection is tightly constrained to values of s close to s_M, decaying rapidly in both s directions.

Also given in table 1 are the values of s_- and s_+, where the anti-Stokes lines cut the real axis. The global theory is only strictly valid if this interval lies entirely inside the fluid. This causes no difficulty at $Pr = 1$ and $Pr = 10$, but at $Pr = 0.1$ the interval does not lie within the fluid. The values of s_- are outside the interval, and there is no value of s_+, and the $\hat{k}$ emerging from the dispersion relation becomes singular as $s \to 1/(1-\eta)$. This means that at low Prandtl number, because the convection is so spread out at onset, the solution is not sufficiently localized to be completely independent of the boundary conditions at the tangent cylinder and at $s = 1/1-\eta$.

In table 2, we again focus on $\eta = 0.5$ and $\eta = 0.8$, but we consider the critical values of N_ρ that can occur. Since in Boussinesq fluid, the onset of convection is

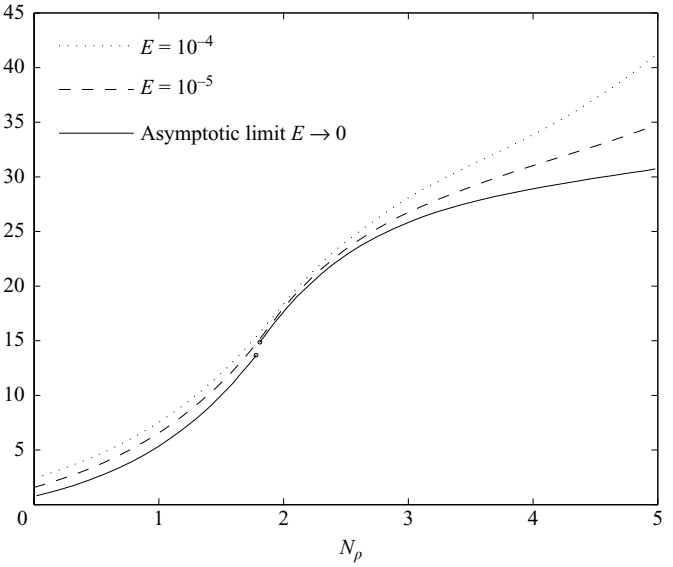

FIGURE 5. Asymptotic theory compared with full numerical code results ($\eta = 0.5$, $Pr = 1$, $n = 2$). Solid line: $\mathscr{R}$ against N_ρ from the asymptotic theory. Dashed line: $Ra \times E^{4/3}$ against N_ρ at $E = 10^{-5}$. Dotted line: $Ra \times E^{4/3}$ against N_ρ at $E = 10^{-4}$. The small dots showing the break in the solid line denote the asymptotic values at N_ρ^{LG} and N_ρ^{AS}. For $N_\rho < N_\rho^L G$ local theory at $s = s_i$ is used; for $N_\rho > N_\rho^{AS}$ the global theory is used.

always first at the tangent cylinder, whereas at large N_ρ onset is in the interior of the fluid, there must always be a critical value of N_ρ, denoted by N_ρ^{LG}, below which onset of convection is first at the tangent cylinder, and these are listed in table 2. For $N_\rho < N_\rho^{LG}$, local theory with $s = s_i$ must be used; for $N_\rho > N_\rho^{LG}$, global theory, with a complex s_c, is used; so $N_\rho = N_\rho^{LG}$ denotes the transition between local and global theories. As noted above, even quite modest amounts of compressibility push the convection away from the tangent cylinder, despite the gravity falling off steeply as the distance from the tangent cylinder increases. The corresponding values of $\mathscr{R}_c$, $\hat{\omega}_c$ and $\hat{m}_c$ are given at this critical $N_\rho = N_\rho^{LG}$; N_ρ^{AS} is the value of N_ρ at which the smaller of the two anti-Stokes points, s_-, lies on the tangent cylinder; s_+ is the value of the larger anti-Stokes point at $N_\rho = N_\rho^{AS}$, and R_c, ω_c, m_c, s_r and s_i are evaluated using global bifurcation theory at the same point. For $N_\rho > N_\rho^{AS}$ and $s_+ < 1/(1-\eta)$ the entire interval (s_-, s_+) lies in the fluid, and so global bifurcation theory gives the correct asymptotic $E \to 0$ limit. For values of N_ρ in $N_\rho^{LG} < N_\rho < N_\rho^{AS}$ neither the global nor the local asymptotic theory applies.

Most of the features seen in figures 1–4 can be inferred from the asymptotic theory, though values of $E < 10^{-4}$ are needed before the asymptotic regime is unambiguously reached. Thus the higher values of m, the lower frequencies and the higher Rayleigh numbers at higher Prandtl number can all be deduced from the asymptotic theory.

In figure 5, the approach to the asymptotic limit is examined, for $Pr = 1$ and $\eta = 0.5$. We show $Ra \times E^{4/3}$ as a function of N_ρ at $E = 10^{-4}$ and $E = 10^{-5}$ and also show the asymptotic value of $\mathscr{R}$. In the limit $E \to 0$ the full numerical simulations should approach the asymptotic curve, and it appears that they do. Note that below $N_\rho^{LG} = 1.771$ the local asymptotic theory is used, and above $N_\rho^{AS} = 1.815$ the global asymptotic theory is used. Between the two is a gap, as the the global asymptotic

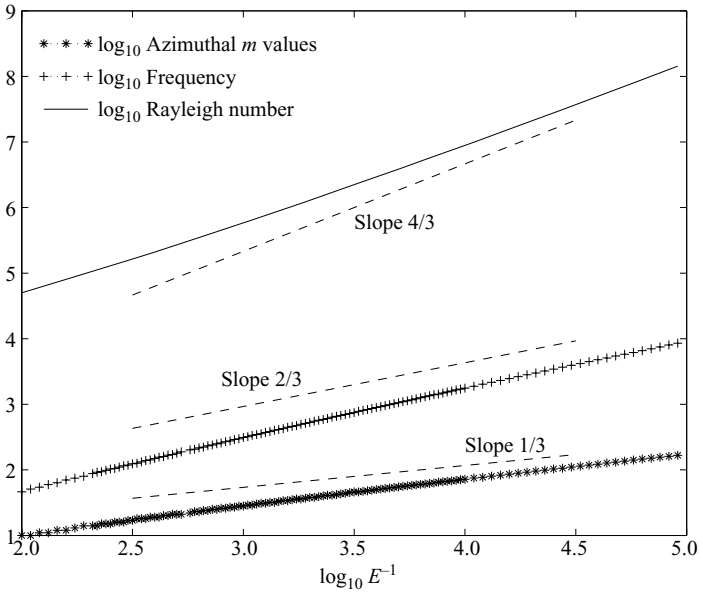

FIGURE 6. Plot of $\log_{10} Ra_{crit}$, $\log_{10} \omega_{crit}$ and $\log_{10} m_{crit}$ as functions of $-\log_{10} E$ ($N_\rho = 5$, $\eta = 0.5$, $Pr = 1.0$, $n = 2$). Dashed lines with the slopes corresponding to the expected asymptotic power laws are drawn for comparison.

theory is not strictly valid until s_- lies inside the fluid. The asymptotic limit is approached most rapidly when the convection lies well inside the fluid interior. This is not too surprising, as when the convection lies at the tangent cylinder, the errors are $O(E^{2/9})$ (Dormy *et al.* 2004), whereas in the interior they are $O(E^{1/3})$. When the convection occurs near the outer shell, it is restricted in the z direction by the geometry (see figure 2*c*); so for the tall thin columns necessary for the asymptotic theory to be a good approximation, naturally the radial extent must be correspondingly less, and this requires even smaller E.

In figure 6 we show the variation of Ra, ω and m as E varies for a strongly stratified case, $N_\rho = 5$. Figure 6 is plotted on a logarithmic scale so that the power law variation with E can be discerned. We see that for the lowest values of E that could be comfortably achieved by the code at large N_ρ, about $E = 10^{-5}$, the slope of $\log_{10} Ra$ against $\log_{10} E^{-1}$ is 1.3, not far from the expected asymptotic value of 4/3. Similarly the slope of $\log_{10} \omega$ against $\log_{10} E^{-1}$ is 0.69, close to the expected value of 2/3, while the exponent of the azimuthal wavenumber dependence is $E^{-0.35}$ compared to the asymptotic value of $-1/3$.

6.4. *Solutions antisymmetric about the equator*

Busse (1970) noted that onset of solutions with ψ, u_r and u_ϕ symmetric about the equator is at lower critical Rayleigh number than the modes with the opposite parity in the Boussinesq case. Note that modes with this symmetry also have u_z antisymmetric about the equator. We explored whether modes with symmetric u_r are also preferred in compressible cases. In much the same way as for the case of symmetric modes about the equator, the critical Ra and azimuthal numbers m grow with N_ρ (see figure 7). If we compare figures 7(*a*) and 7(*b*), we see that the critical Rayleigh number is larger for the antisymmetric modes at all values of N_ρ; so there is no evidence that antisymmetric modes are preferred in compressible convection.

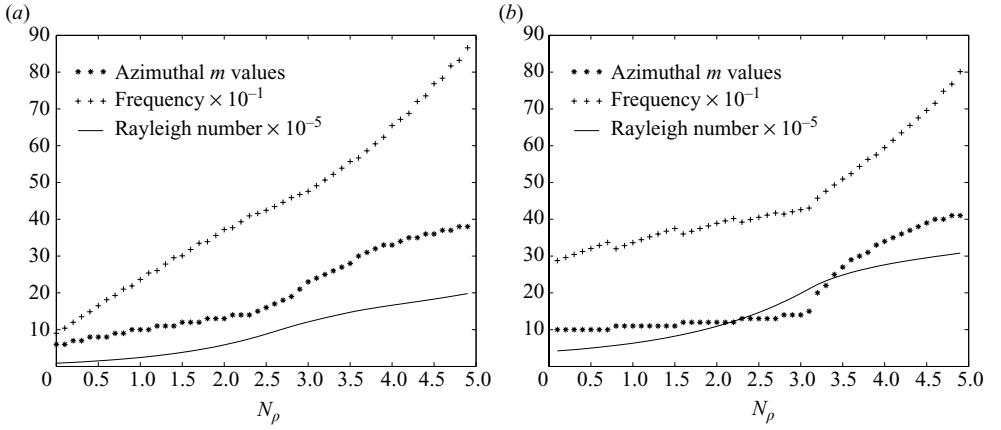

FIGURE 7. Critical Ra, ω and m as functions of N_ρ for (a) symmetric modes and (b) antisymmetric modes ($E = 2 \times 10^{-4}$, $\eta = 0.35$, $n = 2$).

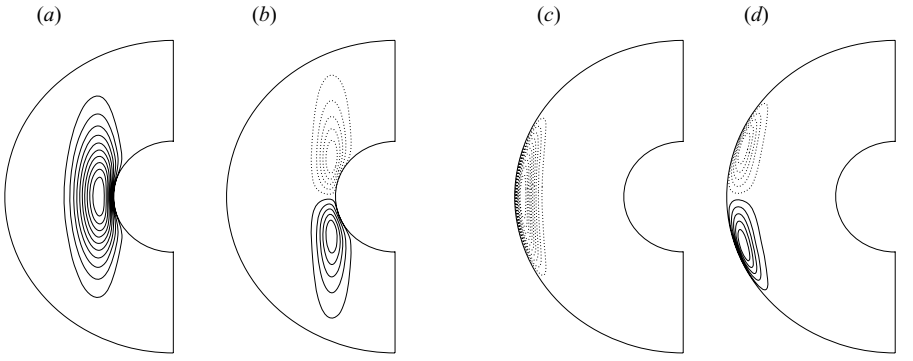

FIGURE 8. Meridional sections for $E = 2 \times 10^{-4}$, $\eta = 0.5$, $Pr = 1$, $n = 2$: (a) symmetric Boussinesq mode, $N_\rho = 0$, $Ra_{crit} = 8.7529 \times 10^4$, $\omega_{crit} = 90.11$, $m_{crit} = 6$; (b) antisymmetric mode with $Ra = 4.0712 \times 10^5$, $\omega = 279.59$, $m = 10$; (c) symmetric mode with $N_\rho = 5.0$, $Ra_{crit} = 2.02198 \times 10^5$, $\omega_{crit} = 881.28$, $m_{crit} = 39$; (d) antisymmetric mode with $Ra = 3.11521 \times 10^6$, $\omega = 820.29$, $m = 42$.

The structure of the solutions is shown in figure 8, for both Boussinesq and strongly compressible cases. Apart from the equatorial parity, the behaviour is similar to that of the symmetric modes.

6.5. *Higher (polar) modes inside the tangent cylinder*

Convection always starts outside the tangent cylinder, but as the Rayleigh number is increased, onset of higher-Rayleigh-number modes takes place, and eventually convection occurs inside the tangent cylinder as well as outside it. Indeed, eventually more heat transport occurs inside the tangent cylinder than outside it (Tilgner & Busse 1997). Here we explore how convection extends into the tangent cylinder from a linear perspective. At a given m there is a two-parameter family of eigenfunctions, one corresponding to modes with additional zeros in the z-direction and another corresponding to higher modes in the s direction. The next mode after the fundamental in the z direction is just the antisymmetric mode discussed in the § 6.4, and the subsequent sequence has increasing structure in the columns. There are substantial

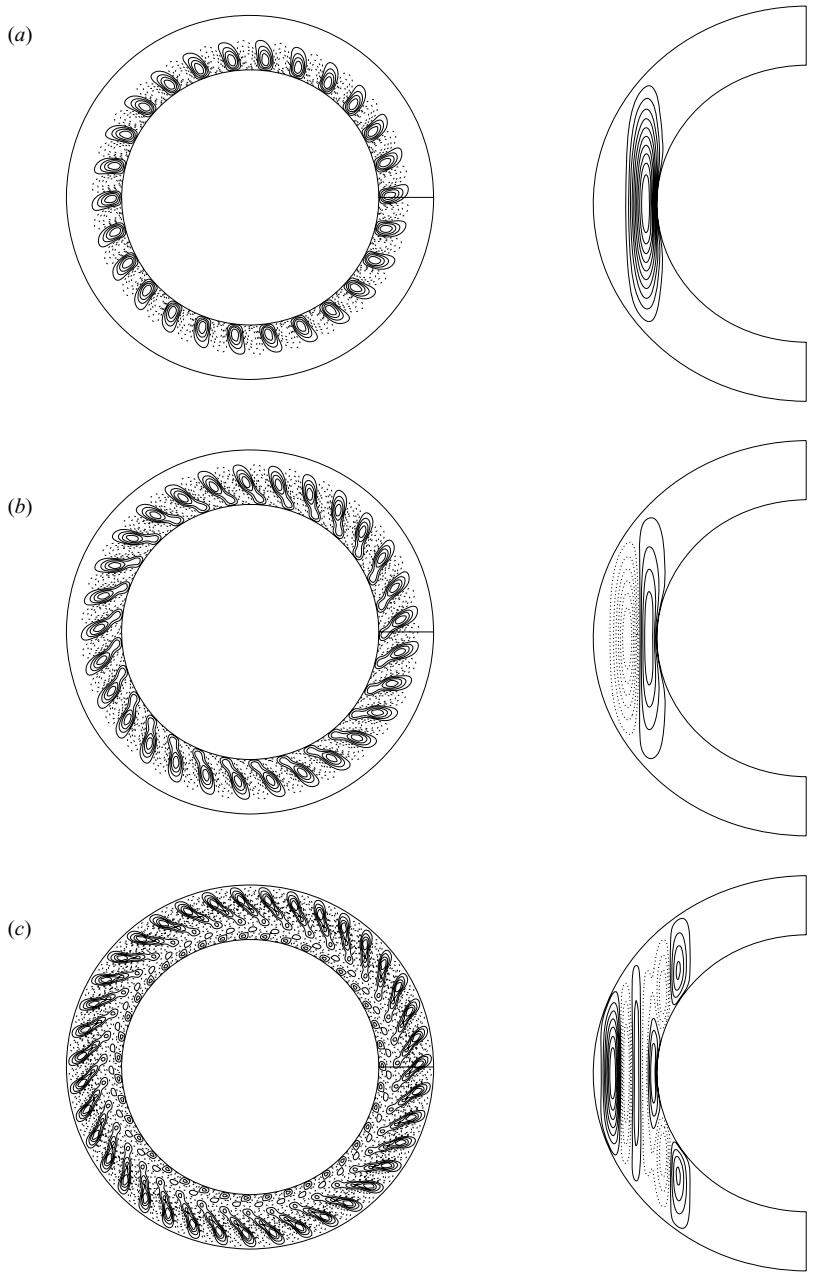

FIGURE 9. Some higher-Ra modes for Boussinesq convection ($E = 2 \times 10^{-4}$, $\eta = 0.7$, $Pr = 1.0$, $n = 2$): (*a*) $Ra_{crit} = 1.0536 \times 10^6$, $\omega_{crit} = 111.09$, $m_{crit} = 26$ (lowest mode); (*b*) $Ra_2 = 2.0979 \times 10^6$, $\omega_2 = 97.83$, $m_2 = 29$; (*c*) $Ra_5 = 8.6838 \times 10^6$, $\omega_5 = 58.69$, $m_5 = 39$. Left panel: equatorial section of radial velocity u_r. Right panel: meridional section at the longitude marked in the left panel.

jumps in Ra_{crit} between each mode in this sequence. In figure 9 we show some of the s sequence of modes for the Boussinesq case with $E = 2 \times 10^{-4}$, $\eta = 0.7$, $Pr = 1$, $n = 2$. The next symmetric mode in the z sequence has $Ra = 8.378 \times 10^6$ and vorticity which changes sign twice along one column. Figure 9(*a*) is the fundamental

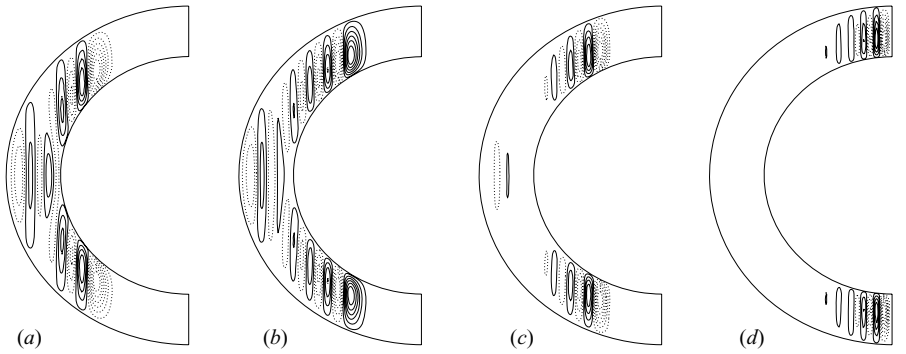

FIGURE 10. Meridional sections of the incompressible (Boussinesq) solutions for high latitudinal modes for $E = 2 \times 10^{-4}$, $\eta = 0.7$, $Pr = 1$, $n = 2$: (*a*) $Ra_6 = 1.0613 \times 10^7$, $\omega_6 = 58.09$, $m_6 = 40$; (*b*) $Ra_9 = 1.2618 \times 10^7$, $\omega_9 = 37.97$, $m_9 = 33$; (*c*) $Ra_{11} = 1.3512 \times 10^7$, $\omega_{11} = 37.89$, $m_{11} = 26$; (*d*) $Ra = 1.4797 \times 10^7$, $\omega = -3.052$, $m = 3$ (not optimized over m).

(lowest Ra_{crit}) mode, with a structure similar to those shown in figures 2(*a*) and 4(*a*). The next mode is shown in figure 9(*b*). In the meridional section we see a double row of vortices as s varies, but this is slightly deceptive, because the radial equatorial section in the left panel shows that there is in fact only one spiral, but constant ϕ lines cut both positive and negative spirals with substantial amplitudes. This can happen at low Prandtl number even for the lowest mode (see figures 4*a* and 4*b*). At very low Ekman number, the interval in Rayleigh number between adjacent modes becomes $O(RaE^{1/3})$ (Jones *et al.* 2000). We omit the third and fourth modes in the sequence, which exist at $Ra = 3.577 \times 10^6$ and $Ra = 5.514 \times 10^6$, and in figure 9(*c*) we show the fifth member of this eigenfunction sequence. Not only has the whole region outside the tangent cylinder been filled with convecting columns, but they have even penetrated inside the tangent cylinder, working in from the tangent cylinder itself. Nonlinear convection at this value of Ra will have significant amplitude inside the tangent cylinder. In figure 10, the s sequence continues, but we show just meridional sections. Figure 10(*a*) is the sixth in the s sequence and illustrates the rapid penetration of the convection into the tangent cylinder. Local convection is now possible inside the tangent cylinder, and so the problem starts to resemble the degenerate problem of convection between horizontal boundaries. In consequence there are now smaller gaps in Ra between the adjacent modes in the sequence, and it becomes harder to evaluate the mode sequence. The trend is that modes of smaller m start to appear, and the convection penetrates right up to the polar regions. The final picture (figure 10*d*) has a very low value of m. Unlike the other modes in this sequence it has not been optimized over m, but we include it to illustrate the existence of neutral polar modes at Rayleigh numbers of the same order as the other modes.

In figure 11 we show the equivalent s sequence for a strongly compressible case. As expected, all members of the sequence have convection starting close to the outer wall. The general trend of movement towards the polar regions occurs for these highly compressible modes also. However, the crossing of the tangent cylinder is much less of a barrier for the compressible modes, and the spacing in Rayleigh number is fairly regular. The sharp drop in the Rayleigh number intervals between adjacent modes that is such a clear feature in figures 9 and 10 does not occur.

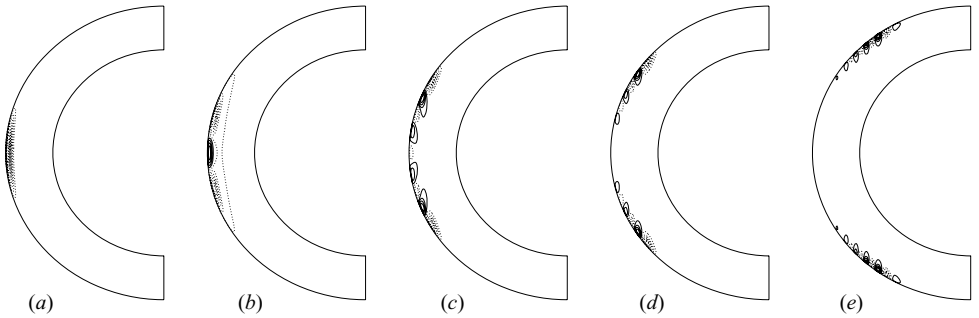

FIGURE 11. Meridional sections of the compressible ($N_\rho = 5.0$) solutions for high latitudinal modes for $E = 2 \times 10^{-4}$, $\eta = 0.7$, $Pr = 1$, $n = 2$: (*a*) the first unstable mode, $Ra_{crit} = 1.2087 \times 10^7$, $\omega_{crit} = 1249.7$, $m_{crit} = 98$; (*b*) the second mode, $Ra_2 = 1.6655 \times 10^7$, $\omega_2 = 1152.5$, $m_2 = 105$; (*c*) the fifth mode, $Ra_5 = 2.4845 \times 10^7$, $\omega_5 = 982.0$, $m_5 = 104$; (*d*) the eighth mode, $Ra_8 = 3.0735 \times 10^7$, $\omega_8 = 858.9$, $m_8 = 99$; (*e*) and a higher mode, $Ra_{high} = 4.1685 \times 10^7$, $\omega_{high} = 592.3$, $m_{high} = 74$;

7. Conclusion

Although our model has some differences between the earlier models of Glatzmaier & Gilman (1981*a*, *b*) and Drew *et al.* (1995), with the diffusion of entropy notably dominating in the heat equation in this work, many of the features found there also occur here. Compressibility moves the region of onset of convection away from the tangent cylinder towards the outer shell. This is a very powerful effect. Even a moderate degree of compressibility, with only a few scale heights in the layer, can overcome counteracting tendencies such as a gravity field and a basic state temperature gradient that drop off quickly as r increases. We also note that compressibility increases the critical azimuthal wavenumber at onset at a given Ekman number. This means that nonlinear compressible convection is considerably more challenging numerically than Boussinesq convection. To achieve the same Ekman number at $N_\rho = 5$ as in a Boussinesq code, the azimuthal resolution will need to be roughly a factor of five greater. Another notable feature is that the wave frequencies are substantially higher in strongly compressible cases. Wave frequency decreases rapidly with increasing Prandtl number, but compressibility increases frequencies at all Prandtl numbers. The strong spiralling effect seen particularly at low Prandtl numbers occurs also in compressible convection, and so strong zonal flows are more likely to result from low-Prandtl-number compressible convection, just as they are in Boussinesq convection.

We have also shed some light on the slightly mysterious result of convection at negative Rayleigh number reported in Drew *et al.* (1995). We have established this phenomenon cannot occur if entropy diffusion dominates over thermal diffusion in the heat equation. Since this is likely to be the case in even very weakly turbulent systems, it seem unlikely that the negative Rayleigh convection is of any great physical significance.

The asymptotic theory of convection developed in Jones *et al.* (2000) and Dormy *et al.* (2004) has been successfully extended to the compressible case in a natural way. Despite the fact that the Proudman–Taylor theorem does not strictly apply in a compressible fluid, the convection still forms tall thin columns in the limit $E \to 0$ if the other parameters remain fixed. This allows rapid exploration of the parameter space, and all the major effects reported here are found in the asymptotic theory, as well as by full numerical computation. At Pr about unity or greater there is excellent agreement between the asymptotic theory and the full numerics. However, there is

a difficulty at low Prandtl number with the asymptotic theory when onset of the convection is first in the interior of the fluid rather than at the tangent cylinder. In order to apply the asymptotic boundary condition that the disturbance vanishes as we both increase and decrease s from its critical location, it is necessary for the s interval between the points at which the anti-Stokes lines defined by (4.12) cut the real s axis to lie entirely within the fluid. At low Pr this is often not the case, and then the boundary conditions at $s = s_i, s_o$ matter even in the limit $E \to 0$. The correct procedure for obtaining the $E \to 0$ limit in this awkward case has not yet been fully worked out. Nevertheless, although precise values of Ra cannot be found at small E, the general trends predicted by the asymptotic theory are still found at $Pr = 0.1$.

We have considered the antisymmetric modes as well as the symmetric modes, but we find that the symmetric modes are preferred in compressible convection, just as they are in Boussinesq convection (Busse 1970). We also examined the higher-order modes, to see how nonlinear convection is likely to spread into the tangent cylinder. We find that at Rayleigh numbers a factor 10 higher than initial onset, modes both inside and outside the tangent cylinder are excited, in both Boussinesq and compressible convection.

Appendix A. Boussinesq limit $N_\rho \to 0$

We first consider the nonlinear equations (2.7) and (2.16) in the Boussinesq limit, $N_\rho \to 0$. From (2.3), $\zeta_0 \to 1$, $c_0 \to 1$ and $c_1 \to 0$. From (2.2c), this implies

$$\frac{g d\rho}{p} \to 0 \quad \text{and hence} \quad \frac{g d}{c_p T} \to 0, \tag{A 1a, b}$$

using the perfect gas law. In (2.7), the term $\nabla(p_c/\rho)$ has order of magnitude $p_c/\rho d$, while in the term

$$\frac{\boldsymbol{g} S_c}{c_p} = \frac{\boldsymbol{g}}{\gamma}\left(\frac{p_c}{p} - \frac{\gamma \rho_c}{\rho}\right) \tag{A 2}$$

the pressure perturbation part has order of magnitude $g p_c/p$ which by (A 1a) is negligibly small in comparison in the Boussinesq limit. The pressure fluctuation part in (A 2) is therefore negligible in comparison to the density fluctuation part. Then from the gas law, $\rho_c/\rho \to -T_c/T$; so

$$\frac{S_c}{c_p} = \frac{S}{c_p} \to \frac{T_c}{T} = \alpha T_c, \tag{A 3}$$

where $\alpha = 1/T$ is the coefficient of expansion. This can be extended to the case of a liquid, provided α is the appropriate coefficient of expansion (Anufriev, Jones & Soward 2005). The nonlinear momentum equation therefore reduces to

$$\frac{\partial \boldsymbol{u}}{\partial t} = \boldsymbol{u} \times \boldsymbol{\omega} - 2\boldsymbol{\Omega} \times \boldsymbol{u} - \frac{\nabla p_c}{\rho} - \nabla(\boldsymbol{u}^2/2) + \nu \boldsymbol{F}_v - \boldsymbol{g}\alpha T_c \tag{A 4}$$

in the Boussinesq limit.

Since in the reference state ρ, p and T all tend to constant values in the Boussinesq limit, (A 3) implies that the turbulent and molecular diffusion terms in the heat transport equation (2.16) have the same form as $N_\rho \to 0$. Furthermore, the viscous heating term involving Q is negligible in this limit, since from the momentum equation S_c/c_p has the order of magnitude $\nu u/g d^2$, which implies that the viscous heating term in (2.16) is smaller than the advection term $\rho T(\boldsymbol{u}\cdot\nabla)S_c$ by a factor $g d/c_p T$ which by

(A 1*b*) is very small in the Boussinesq limit. Since T_c and T differ only by a constant in the Boussinesq limit, the heat transport equation (2.16) becomes

$$\frac{\partial T}{\partial t} + (\boldsymbol{u}\cdot\nabla)T = \nabla\cdot\kappa\nabla T, \tag{A 5}$$

where κ is the sum of the turbulent and molecular diffusivities. The constant entropy boundary conditions reduce to constant temperature boundary conditions from (A 3).

For numerical solution, it is convenient to rewrite the linearized dimensionless entropy equation (2.22) in the form

$$\left\{\nabla^2 - Pr\frac{\partial}{\partial t}\right\} S' = \frac{\xi}{\rho}\,\widehat{\boldsymbol{r}}\cdot\boldsymbol{u}\left(\frac{1}{\rho_o^{-1}-\rho_i^{-1}}\right) - \xi\frac{n+1}{n}\frac{\partial S'}{\partial r}. \tag{A 6}$$

Here n is the polytropic index; $\rho = \rho(r)$ is the density; $\rho_o = \rho(r_o)$; $\rho_i = \rho(r_i)$. Compressibility is involved in our equations mainly through the logarithmic derivative

$$\xi = \frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} = -c_1\frac{n}{\zeta}\frac{1}{r^2}, \tag{A 7}$$

using (2.2*a*, *b*). Now we show that the dimensionless linearized equations (2.22) and (2.24) reduce to the Boussinesq equations with the temperature fluctuation T' replacing the entropy fluctuation S' in the limit $N_\rho \to 0$. The following asymptotic formulae in this limit can be derived from (2.2) and (2.3):

$$\zeta_0 = 1 - \frac{\eta}{1+\eta}\frac{N_\rho}{n} + O\left({N_\rho}^2\right), \qquad c_0 = 1 - \frac{2\eta}{1-\eta^2}\frac{N_\rho}{n} + O\left({N_\rho}^2\right), \tag{A 8a, b}$$

$$c_1 = \frac{\eta}{(1-\eta)^2}\frac{N_\rho}{n} + O\left({N_\rho}^2\right), \ \zeta(r) = 1 + \frac{\eta N_\rho}{n(1-\eta)^2}\left[\frac{1}{r} - \frac{2(1-\eta)}{1+\eta}\right] + O\left(N_\rho^2\right), \tag{A 9a, b}$$

$$\xi(r) = -\frac{nc_1}{r^2\zeta} = -\frac{N_\rho}{r^2}\frac{\eta}{(1-\eta)^2} + O\left(N_\rho^2\right). \tag{A 10}$$

From (A 10) it is clear that the second term on the right-hand side of (A 6) becomes zero in the Boussinesq limit $N_\rho \to 0$. Using (A 8) and (A 9),

$$\zeta_o^{-n} = 1 + \frac{\eta N_\rho}{1+\eta} + O\left(N_\rho^2\right), \qquad \zeta_i^{-n} = 1 - \frac{N_\rho}{1+\eta} + O\left(N_\rho^2\right); \tag{A 11a, b}$$

so the entropy equation (A 6) reduces to

$$Pr\frac{\partial T'}{\partial t} = \frac{\eta(\widehat{\boldsymbol{r}}\cdot\boldsymbol{u})}{r^2(1-\eta)^2} + \nabla^2 T' \tag{A 12}$$

in the Boussinesq limit. Comparing with Dormy *et al.* (2004, equations (2.3*a*, *b*, *d*)), we see that the differences between their model and ours are that the gravity term in our equation of motion is taken as $\sim 1/r^2$ instead of $\sim r$, appropriate for a uniform sphere, and only the differential heating case is considered here. Also, the unit of entropy (temperature in the Boussinesq limit) has a factor of $\eta/(1-\eta)^2$ different in its definition; so this factor must be included in the Rayleigh number definition.

In principle the first term on the right-hand side of (A 12) can be written as

$$\frac{\eta}{(1-\eta)^2}\,\widehat{\boldsymbol{r}}\cdot\boldsymbol{u}\left(\frac{1-\lambda}{r^2} + \lambda r\right),$$

where the parameter $0 \leqslant \lambda \leqslant 1$ is the fraction of internal heating (cf. Dormy *et al.* 2004). However, we assume for our problem only differential heating, and so in this paper $\lambda = 0$.

Appendix B. The form of the numerical equations

The toroidal–poloidal decomposition (5.1) employed here implies that

$$\boldsymbol{u} = \left[\frac{\mathscr{L}^2\{f\}}{r^2}\right]\widehat{\boldsymbol{r}} + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial e}{\partial \phi} + \frac{\partial^2 f}{\partial r \partial \theta} + \frac{\partial f}{\partial \theta}\xi\right]\widehat{\boldsymbol{\theta}} + \frac{1}{r}\left[\frac{1}{\sin\theta}\frac{\partial^2 f}{\partial r \partial \phi} - \frac{\partial e}{\partial \theta} + \frac{1}{\sin\theta}\frac{\partial f}{\partial \phi}\xi\right]\widehat{\boldsymbol{\phi}}, \quad \text{(B 1)}$$

where the 'horizontal' operator

$$\mathscr{L}^2 \equiv -\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left\{\sin\theta\frac{\partial}{\partial \theta}\right\} - \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial \phi^2}. \quad \text{(B 2)}$$

Then we apply curl and double-curl operators to (2.24) and evaluate their radial parts by multiplying by $\widehat{\boldsymbol{r}}$. For evaluation of the Coriolis terms, as $\boldsymbol{\Omega} = \Omega\,\widehat{\boldsymbol{z}}$, it is useful to introduce the operator

$$\mathscr{Q} \equiv -\cos\theta\,\mathscr{L}^2\frac{\partial}{\partial r} - \sin\theta\frac{\partial}{\partial \theta}\frac{\partial}{\partial r} + \frac{\mathscr{L}^2}{r}\left\{\sin\theta\frac{\partial}{\partial \theta}\right\}. \quad \text{(B 3)}$$

Then we can write

$$\widehat{\boldsymbol{r}}\cdot\nabla\times(\widehat{\boldsymbol{z}}\times\boldsymbol{u}) = \frac{1}{r^2}\mathscr{Q}\{f\} - \frac{1}{r^2}\frac{\partial e}{\partial \phi} + \xi\left[-\frac{\cos\theta}{r^2}\mathscr{L}^2 - \frac{\sin\theta}{r^2}\frac{\partial}{\partial \theta}\right]\{f\} \quad \text{(B 4)}$$

and

$$\widehat{\boldsymbol{r}}\cdot\left[(\nabla\times)^2(\widehat{\boldsymbol{z}}\times\boldsymbol{u})\right] = \frac{1}{r^2}\mathscr{Q}\{e\} + \frac{1}{r^2}\mathscr{D}^2\frac{\partial f}{\partial \phi} - \xi\frac{\mathscr{L}^2}{r^3}\frac{\partial f}{\partial \phi} + \frac{1}{r^2}\frac{\partial}{\partial r}\left(\xi\frac{\partial f}{\partial \phi}\right), \quad \text{(B 5)}$$

where

$$\mathscr{D}^2 \equiv \frac{\partial^2}{\partial r^2} - \frac{\mathscr{L}^2}{r^2}. \quad \text{(B 6)}$$

Then $r^2\,\widehat{\boldsymbol{r}}\cdot$ curl of the momentum equation (2.24) is

$$\mathscr{L}^2\frac{\partial e}{\partial t} - \mathscr{L}^2\mathscr{D}^2\{e\} + 2E^{-1}\left[\mathscr{Q}\{f\} - \frac{\partial e}{\partial \phi} + \xi\left(-\cos\theta\,\mathscr{L}^2 - \sin\theta\frac{\partial}{\partial \theta}\right)\{f\}\right] = \xi\,r^2\frac{\partial}{\partial r}\left(\frac{\mathscr{L}^2}{r^2}\{e\}\right), \quad \text{(B 7)}$$

and applying $r^2\,\widehat{\boldsymbol{r}}\cdot\text{curl}^2$ gives

$$\begin{aligned} &-\mathscr{L}^2\mathscr{D}^2\frac{\partial f}{\partial t} - \mathscr{L}^2\frac{\partial}{\partial r}\left(\frac{\partial f}{\partial t}\xi\right) + \mathscr{L}^2\mathscr{D}^4\{f\} + \mathscr{L}^2\left\{\xi\frac{\partial^3 f}{\partial r^3} + 3\xi'\frac{\partial^2 f}{\partial r^2} + 3\xi''\frac{\partial f}{\partial r} + \xi'''f\right\} \\ &\quad - \frac{1}{r^2}\mathscr{L}^4\left\{\xi\frac{\partial f}{\partial r} + \xi' f\right\} + 2E^{-1}\left[\mathscr{Q}\{e\} + \mathscr{D}^2\frac{\partial f}{\partial \phi} - \frac{\xi}{r}\mathscr{L}^2\frac{\partial f}{\partial \phi} + \frac{\partial}{\partial r}\left(\xi\frac{\partial f}{\partial \phi}\right)\right] \\ &\quad = Ra\,\mathscr{L}^2\left\{\frac{S'}{r^2}\right\} + F_v, \end{aligned} \quad \text{(B 8)}$$

where the viscous term

$$F_v = \frac{\xi}{r^2}\frac{\partial}{\partial r}\left\{\mathscr{L}^4 f\right\} + \left(\frac{2}{3}\frac{\xi^2}{r^2} - \frac{2\xi}{r^3} - \frac{\xi'}{r^2}\right)\mathscr{L}^4 f - \xi\frac{\partial^3}{\partial r^3}\left\{\mathscr{L}^2 f\right\} + \left(\frac{2\xi}{r} - \xi^2 - \xi'\right)\frac{\partial^2}{\partial r^2}\left\{\mathscr{L}^2 f\right\} + \left(\frac{2\xi^2}{r} - \frac{2\xi}{r^2} + \frac{2\xi'}{r} - 3\xi\,\xi'\right)\frac{\partial}{\partial r}\left\{\mathscr{L}^2 f\right\} + \left(\frac{4}{r}\xi\,\xi' - \frac{2\xi^2}{r^2} - \xi'^2 - \xi\,\xi''\right)\mathscr{L}^2 f, \tag{B 9}$$

with

$$\xi' = \frac{\partial \xi}{\partial r}, \quad \xi'' = \frac{\partial^2 \xi}{\partial r^2}, \quad \xi''' = \frac{\partial^3 \xi}{\partial r^3}.$$

Finally, using expansion (5.1) for $\boldsymbol{u}$, the entropy equation (A 6) reads

$$\nabla^2 S' - Pr\frac{\partial S'}{\partial t} = \frac{\xi}{\rho}\left(\frac{1}{\rho_o^{-1} - \rho_i^{-1}}\right)\frac{\mathscr{L}^2}{r^2}\{f\} - \xi\,\frac{n+1}{n}\,\frac{\partial S'}{\partial r}. \tag{B 10}$$

We use the governing equations in the above forms for the full numerical computations. In our code, we solve the linear problem by a spectral collocation method, using a representation of our solution by Chebyshev polynomials in the radial coordinate and associated Legendre polynomials in the θ and ϕ directions. Crucially, the Coriolis term only couples adjacent spherical harmonics; so the matrices have a banded structure. We evaluate eigenvalues by inverse iteration. Typically the radial resolution was 30–50 polynomials in both directions, the higher N_ρ using the highest resolution. Our code can solve the eigenvalue problem either for symmetric or antisymmetric modes.

This research was supported by STFC, grant PP/E001092/1.

REFERENCES

Al-Shamali, F. M., Heimpel, M. H. & Aurnou, J. M. 2004 Varying the spherical shell geometry in rotating thermal convection. *Geophys. Astrophys. Fluid Dyn.* **98**, 153–169.

Anufriev, A. P., Jones, C. A. & Soward, A. M. 2005 The Boussinesq and anelastic liquid approximations for convection in the Earth's core. *Phys. Earth Planet. Inter.* **152**, 163–190.

Batchelor, G. K. 1967 *An Introduction to Fluid Dynamics*. Cambridge University Press.

Braginsky, S. I. & Roberts, P. H. 1995 Equations governing convection in earth's core and the geodynamo. *Geophys. Astrophys. Fluid Dyn.* **79**, 1–97.

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

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

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

Clune, T. C., Elliot, J. R., Miesch, M. S., Toomre, J. & Glatzmeier, G. A. 1999 Computational aspects of a code to study rotating turbulent convection in spherical shells. *Parallel Comput.* **25**, 361–380.

Davidson, P. A. 2004 *Turbulence: An Introduction for Scientists and Engineers*. Oxford University Press.

Dormy, E., Soward, A. M., Jones, C. A., Jault, D. & Cardin, P. 2004 The onset of thermal convection in rotating spherical shells. *J. Fluid Mech.* **501**, 43–70.

Drew, S. J., Jones, C. A. & Zhang, K. 1995 Onset of convection in a rapidly rotating compressible fluid spherical shell. *Geophys. Astrophys. Fluid Dyn.* **80**, 241–254.

Evonuk, M. 2008 The role of density stratification in generating zonal flow structures in a rotating fluid. *Astrophys. J.* **673**, 1154–1159.

Evonuk, M. & Glatzmaier, G. A. 2004 2D Studies of various approximations used for modeling convection in giant planets. *Geophys. Astrophys. Fluid Dyn.* **98**, 241–255.

Geiger, G. & Busse, F. H. 1981 On the onset of convection in slowly rotating fluid shells. *Geophys. Astrophys. Fluid Dyn.* **18**, 147–156.

Gillet, N. & Jones, C. A. 2006 The quasi-geostrophic model for rapidly rotating spherical convection outside the tangent cylinder. *J. Fluid Mech.* **554**, 343–369.

Gilman, P. A. & Glatzmaier, G. A. 1981 Compressible convection in a rotating spherical shell. Part 1. Anelastic equations. *Astrophys. J. Suppl. Ser.* **45**, 335–349.

Glatzmaier, G. A. & Gilman, P. A. 1981*a* Compressible convection in a rotating spherical shell. Part 2. A Linear Anelastic Model. *Astrophys. J. Suppl. Ser.* **45**, 351–380.

Glatzmaier, G. A. & Gilman, P. A. 1981*b* Compressible convection in a rotating spherical shell. Part 4. Effects of viscosity, conductivity, boundary conditions, and zone depth. *Astrophys. J. Suppl. Ser.* **47**, 103–115.

Gough, D. O. 1969 The anelastic approximation for thermal convection. *J. Atmos. Sci.* **26**, 448–456.

Guillot, T. 1999*a* A comparison of the interiors of Jupiter and Saturn. *Planet. Space Sci.* **47**, 1183–1200.

Guillot, T. 1999*b* Interior of giant planets inside and outside the solar system. *Science* **286**, 72–77.

Heimpel, M., Aurnou, J. & Wicht, J. 2005 Simulation of equatorial and high-latitude jets on Jupiter in a deep convection model. *Nature* **438**, 193–196.

Jones, C. A., Rotvig, J. & Abdulrahman, A. 2003 Multiple jets and zonal flow on Jupiter. *Geophys. Res. Lett.* **30**, 1–1.

Jones, C. A., Soward, A. M. & Mussa, A. I. 2000 The onset of thermal convection in a rapidly rotating sphere. *J. Fluid Mech.* **405**, 157–179.

Landau, L. D. & Lifshitz, E. M. 1959 *Fluid Mechanics*. Pergamon.

Lantz, S. R. 1992 *Dynamical Behavior of Magnetic Fields in a Stratified, Convecting Fluid Layer*, PhD Thesis, Cornell University.

Miesch, M. S., Elliott, J. R., Toomre, J., Clune, T. L., Glatzmaier, G. A. & Gilman, P. A. 2000 Three-dimensional spherical simulations of solar convection. Part 1. Differential rotation and pattern evolution achieved with laminar and turbulent states. *Astrophys. J.* **532**, 593–615.

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.

Rotvig, J. & Jones, C. A. 2006 Multiple jets and bursting in the rapidly rotating convecting two-dimensional annulus model with nearly plane-parallel boundaries. *J. Fluid Mech.* **567**, 117–140.

Tilgner, A. & Busse, F. H. 1997 Finite-amplitude convection in rotating fluid shells. *J. Fluid Mech.* **332**, 359–376.

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

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

Zhang, K., Liao, X. & Busse, F. H. 2007 Asymptotic solutions of convection in rapidly rotating non-slip spheres. *J. Fluid Mech.* **578**, 371–380.