Thermal and magnetically driven convection in a rapidly rotating fluid layer

By A. M. SOWARD

Institute of Geophysics and Planetary Physics, University of California, Los Angeles[†]

(Received 30 November 1977 and in revised form 26 April 1978)

The stability of an electrically conducting Boussinesq fluid which is confined between two horizontal planes a distance d apart is investigated. The fluid is heated from below, cooled from above and the whole system rotates rapidly with angular velocity Ω_c about a vertical axis. A weak non-uniform horizontal magnetic field, whose strength is measured by the Alfvén angular velocity $\Omega_M [\ll \Omega_c$, see (1.2)] permeates the fluid and corresponds to the flow of a uniform electric current parallel to the rotation axis. When the modified Rayleigh number R [see (2.1)] is greater than zero and $q = \kappa/\lambda < 1$, where κ and λ are the thermal and magnetic diffusivities respectively, instability sets in as a westward-propagating wave with a low frequency of order κ/d^2 .

When R = 0 and $\Omega_M > 2(\nu/\lambda)^{\frac{1}{2}} \Omega_c$, where ν is the viscosity, Roberts & Loper (1979) have isolated an exceptional class of unstable fast inertial waves which grow on the magnetic diffusion time scale $\tau_{\lambda} = d^2/\lambda$. When R < 0 and $\Gamma = \tau_{\lambda} \Omega_M^2/\Omega_c$ exceeds some value dependent upon q, a class of unstable slow waves also exists for a range of negative values of R. These waves propagate eastwards (westwards) when q is less (greater) than unity. In this case the fluid is stably stratified and the energy for the disturbance is taken from the magnetic field. The resulting description of the stability boundary for R < 0 in the Γ, R plane extends and clarifies the results of Roberts & Loper (1979), which are valid when both Γ and q are large.

1. Introduction

The fluid motions responsible for driving the geodynamo are generally believed to be the result of convection, thermal or otherwise. The only serious alternative is precession. Though a large amount of energy can be transmitted into the fluid by forces acting at the core surface, the recent investigations of Loper (1975) and Rochester *et al.* (1975) suggest that most of the available precessional energy is dissipated in boundary layers, leaving an inadequate supply to drive the dynamo. On the other hand, convection maintained directly by buoyancy forces leads to motions in the main body of the core, which are likely to be sufficiently complex to regenerate magnetic field. Furthermore the wastage of energy in boundary layers is likely to be less severe. Out of all the possible convective processes, thermal convection is generally adopted in theoretical dynamo models (e.g. see Busse 1975; Childress & Soward 1972; Soward 1974), owing to its relative simplicity.

There have been a number of studies of hydromagnetic convection in rapidly

[†] Permanent address: School of Mathematics, University of Newcastle upon Tyne, England.

rotating Boussinesq fluid layers of density ρ_0 (e.g. see Chandrasekhar 1961 chap. 5). A more recent comprehensive study has been made by Eltayeb (1972, 1975). He considered a horizontal fluid layer of thickness d rotating with angular velocity Ω_c and permeated by a uniform magnetic field \mathbf{B}_0^* . By heating the fluid from below and cooling it from above an adverse temperature gradient is maintained. The stability of the system was investigated for a number of different orientations of Ω_c and \mathbf{B}_0^* together with a variety of boundary conditions. In all cases the results indicate that instability occurs most readily when the Coriolis and Lorentz forces are comparable. Whether instability sets in as steady convection or overstability depends on the ratio

$$q = \kappa / \lambda \tag{1.1}$$

of the thermal diffusivity κ to the magnetic diffusivity λ . Furthermore when $q \leq O(1)$, the time scale associated with marginal convection is never shorter than the magnetic diffusion time scale

$$\tau_{\lambda} = d^2 / \lambda. \tag{1.2a}$$

For this reason, the ratio of the sizes of the Lorentz and Coriolis forces is measured conveniently by

$$\Gamma = \tau_{\lambda} \Omega_{M}^{2} / \Omega_{c}, \qquad (1.2b)$$

where Ω_M is the Alfvén angular velocity,

$$\Omega_M = B_0^* / d(\mu \rho_0)^{\frac{1}{2}}, \tag{1.2c}$$

and μ is the magnetic permeability.

For the particular case of rotation about a vertical axis z^* and a uniform horizontal magnetic field \mathbf{B}_0^* , Roberts & Stewartson (1974) have isolated the various regions in the q, Γ^{-1} plane in which overstability can occur and where it is preferred. In the case of geophysical interest, for which q < 1, overstability can never occur. It is widely believed, however, that the secular variation of the earth's magnetic field may be the manifestation of the westward propagation of slow hydromagnetic waves (e.g. see Hide 1966). If such waves are driven by buoyancy forces, this is not apparent from Eltayeb's (1972) model and so some ingredient must be added in order that travelling waves will ensure at the onset of instability. The required ingredient is certainly present in the full sphere model considered by Eltayeb & Kumar (1977). They investigate convection in a rapidly rotating self-gravitating fluid sphere containing heat sources and a uniform electric current $J^* = 2B_0^*/\mu d$ flowing through the sphere parallel to the rotation axis. Relative to cylindrical polar co-ordinates (ϖ^*, ϕ, z^*) the corresponding magnetic field has components

$$\mathbf{B}^* = (0, B_0^* \, \varpi^* / d, 0) \tag{1.3}$$

and its strength increases linearly with distance ϖ^* from the z^* rotation axis. The model combines the non-magnetic convection problem considered previously by Roberts (1968) and Busse (1970) with the non-dissipative magnetic problem considered by Malkus (1967), in which buoyancy forces were omitted. Eltayeb & Kumar (1977) show that when Γ is small and q is less than some order-one value, the waves propagate eastwards. In this limit Busse (1976) notes that the primary force balance is geostrophic, i.e. the Coriolis forces are balanced by the pressure forces alone. Since geostrophy imposes a severe constraint on the fluid motion Busse (1976) is able to

671

demonstrate the key interactions in the sphere through a particularly simple annulus model. Here a slight tilt of the top boundary mimics the geometrical structure of the spherical container. The resulting geostrophic constraint is broken by the combined effect of the Lorentz force and axial torques. The latter are maintained by the interaction of a radial gravitational field and an adverse density gradient and are responsible for driving the motion. As one would expect, in the case q < 1 he also finds that the waves necessarily propagate eastwards. When Γ is of order one, on the other hand, the Coriolis and Lorentz forces are comparable and Eltayeb & Kumar (1977) find from their numerical calculations that the waves propagate westwards. In view of the complexity of their model it would be helpful to have a simple model similar in spirit to Busse's (1976) annulus model which isolates the crucial mechanism leading to the westward propagation of waves. This is the prime objective of this paper.

Eltayeb & Kumar's (1977) convection model must be interpreted cautiously. For even in the absence of buoyancy forces, Malkus (1967) has shown that instability is possible. In particular, in the absence of dissipation, a mode proportional to $e^{i\phi}$ is unstable for sufficiently strong magnetic fields ($\Omega_M \sim \Omega_c$). Since geophysically relevant models have $\Omega_M \ll \Omega_c$, this instability is not present and its existence for larger values of Ω_M does not restrict the usefulness of the model. On the other hand, Roberts & Loper (1979) have isolated a number of diffusive instabilities, which operate at moderate values of Γ . First, they have shown that in the absence of buoyancy forces and viscosity a fast inertial $e^{i\phi}$ mode with a frequency of order $\Omega_c \ (\gg \Omega_M)$ can grow on the magnetic diffusion time scale τ_{λ} for all values of Γ . Second, when the container boundary has finite electrical conductivity and $\Gamma \gg 1$, they have shown that there exist unstable slow waves with frequencies of order $\Omega_M^2/\Omega_c (\gg \tau_\lambda^{-1})$ which can grow on the time scale τ_{λ} . In addition, when the container is perfectly conducting and $\Gamma \gg 1$, they have shown that a slow wave instability is possible only in the presence of buoyancy forces. That such modes of instability exist for sufficiently large adverse density gradients is, of course, to be expected. The surprising feature, however, is that the instability also exists when the fluid is sufficiently bottom heavy ! In other words, the stratification acts only as a catalyst, while the motion is driven by the Lorentz forces.

It is made clear in Busse's (1976) annulus model that, when $\Gamma \leq 1$, the geometrical shape of the container is responsible for the eastward propagation of the marginal modes. For the order-one values of Γ investigated by Eltayeb & Kumar (1977) the geometrical shape plays a less crucial role. Instead, the sense of rotation and magnetic field line curvature alone are responsible for the westward propagation of the waves. To isolate this key interaction in its simplest form, Eltayeb's (1972) plane-layer model is considered with one slight modification. Instead of taking a uniform applied magnetic field, the fluid is supposed to be permeated by the magnetic field (1.3) adopted by Malkus (1967), Eltayeb & Kumar (1977) and Roberts & Loper (1979).

The outline of the paper is as follows. In §2 the governing equations and boundary conditions are described. The normal modes of the system are obtained and the dispersion relation (2.8) for the complex frequency s is derived. A key parameter in (2.8) is the magnetic Ekman number

$$\epsilon = (\tau_{\lambda} \Omega_c)^{-1} = \lambda/d^2 \Omega_c, \qquad (1.4a)$$

A. M. Soward

which measures the magnetic diffusion decay rate in units of the rotation frequency. Since it is assumed throughout that

$$\epsilon \ll 1, \quad \Omega_M / \Omega_c = (\epsilon \Gamma)^{\frac{1}{2}} \ll 1,$$
 (1.4b, c)

the quintic equation (2.8), in general, has three small and two large roots. The former three are considered in §3. They are associated with slow wave instabilities and, when both q and Γ are of order one, have frequencies of order τ_{λ}^{-1} When q and Γ take their limiting values the order of magnitude of each of the three frequencies is modified and a number of new values may be distinguished; these are considered in §§3.1-3.3. For completeness a brief discussion of the remaining two roots of (2.8), which correspond to fast inertial waves, is included but since few of the results are new, it is relegated to an appendix. Here the fast instability found by Malkus (1967), which can occur only when Ω_M and Ω_c are of comparable size, is ruled out by (1.4c). On the other hand, though $\epsilon\Gamma$ is small, (1.4b) allows for the possibility that Γ itself may be large as in §3.2 and in the case discussed by Roberts & Loper (1979). In §4, the geophysical relevance of the results obtained in §3.3 for $q \ll 1$ are discussed.

2. The governing equations

An electrically conducting Boussinesq fluid confined between two horizontal planes a distance d apart is considered. The system rotates rapidly with angular velocity Ω_c about the vertical axis and the fluid is permeated by the magnetic field **B*** defined by (1.3). The boundaries are both perfect electrical and thermal conductors and a temperature difference ΔT is maintained between the boundaries, with the bottom boundary the warmer. A convenient measure of the buoyancy forces is the modified Rayleigh number

$$R = \alpha g(\Delta T) d / \Omega_c \kappa, \qquad (2.1)$$

where α is the coefficient of thermal expansion and g is the acceleration due to gravity.

When R is sufficiently large the state of rigid-body rotation becomes unstable to infinitesimal perturbations and convection occurs. Adopting d for the unit of length, τ_{λ} [see (1.2a)] for the unit of time and letting U be the magnitude of a typical perturbation velocity, the dimensionless variables

$$\mathbf{x}^*/d = \mathbf{x}, \quad t/\tau_{\lambda} = t, \quad \mathbf{u}^*/U = \mathbf{u},$$
 (2.1*a*, *b*, *c*)

$$\mathbf{b}^*/B_0^* = \varpi \hat{\mathbf{\Phi}} + (U\tau_\lambda/d) \,\mathbf{b}, \quad T^*/\Delta T = -z + q^{-1}(U\tau_\lambda/d) \,\theta \qquad (2.1d, e)$$

are introduced, where T^* is the temperature. The system is referred to the cylindrical polar co-ordinates (ϖ, ϕ, z) and $\hat{\varpi}$, $\hat{\phi}$ and \hat{z} are used to denote the unit vectors in the ϖ , ϕ and z directions respectively. Marginal convection is governed by the linearized equations

$$\epsilon \,\partial \mathbf{u}/\partial t + 2\mathbf{\hat{z}} \times \mathbf{u} = -\nabla p + R\theta\mathbf{\hat{z}} + \Gamma(\partial_1\mathbf{b}/\partial\phi + 2\mathbf{\hat{z}} \times \mathbf{b}) + q\sigma\epsilon\nabla^2\mathbf{u}, \qquad (2.2a)$$

$$\partial \mathbf{b}/\partial t = \partial_1 \mathbf{u}/\partial \phi + \nabla^2 \mathbf{b},$$
 (2.2*b*)

$$q^{-1}\partial\theta/\partial t = \hat{\mathbf{z}} \cdot \mathbf{u} + \nabla^2\theta, \qquad (2.2c)$$

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{b} = 0, \qquad (2.2d, e)$$

 $\mathbf{672}$

where $\partial_1/\partial\phi$ denotes ϕ differentiation keeping unit vectors fixed in direction, q, Γ and ϵ are defined by (1.1), (1.2b) and (1.4c) respectively, while

$$\sigma = \nu/\kappa \tag{2.3}$$

is the usual Prandtl number.

The fluid velocity **u** and the magnetic field **b** are split into their toroidal and poloidal parts

$$\mathbf{\mu} = \nabla \times \Psi \hat{\mathbf{z}} + \nabla \times (\nabla \times \Phi \hat{\mathbf{z}}), \qquad (2.4a)$$

$$\mathbf{b} = \nabla \times f \,\mathbf{\hat{z}} + \nabla \times (\nabla \times g \,\mathbf{\hat{z}}) \tag{2.4b}$$

so that (2.2e, f) are automatically satisfied. For simplicity, it is supposed that the upper and lower boundaries are both stress free and so the boundary conditions satisfied by Ψ and Φ are

$$\partial^2 \Phi / \partial z^2 = \Phi = \partial \Psi' / \partial z = 0 \tag{2.5a}$$

on z = 0, 1. Similar conditions apply to g and f, namely

$$\dot{g} = \partial f / \partial z = 0, \qquad (2.5b)$$

while the perturbation temperature θ satisfies

$$\theta = 0. \tag{2.5c}$$

With the decomposition (2.4) and the boundary conditions (2.5), there exist separable solutions

$$(\Phi, g, \theta) = (\Phi, \tilde{g}, \tilde{\theta}) J_m(k\varpi) \sin n\pi z \, e^{i(m\phi - st)}, \tag{2.6a}$$

$$(\Psi, f, p) = (\tilde{\Psi}, \tilde{f}, \tilde{p}) J_m(k\varpi) \cos n\pi z \, e^{i(m\phi - st)}, \tag{2.6b}$$

where n is a non-zero integer. Upon substitution of (2.6) into (2.2), it follows that the quantities with a tilde satisfy the algebraic equations

$$\epsilon(-is + \sigma q a^2) n\pi \tilde{\Phi} + 2(\tilde{\Psi} - \Gamma \tilde{f}) = -\tilde{p} + \Gamma i m n\pi \tilde{g}, \qquad (2.7a)$$

$$-\epsilon(-is+\sigma qa^2)\Psi + 2n\pi(\Phi - \Gamma\tilde{g}) = -\Gamma imf, \qquad (2.7b)$$

$$\epsilon(-is + \sigma q a^2) k^2 \Phi = n\pi \tilde{p} + R\tilde{\theta} + \Gamma i m k^2 \tilde{g}, \qquad (2.7c)$$

$$(-is+a^2)\tilde{f} = im\tilde{\Psi}, \quad (-is+a^2)\tilde{g} = im\Phi, \quad (2.7d, e)$$

$$(-iq^{-1}s+a^2)\tilde{\theta} = k^2\tilde{\Phi}, \qquad (2.7f)$$

where

$$a^2 = k^2 + n^2 \pi^2. \tag{2.7g}$$

Here (2.7a, b) describes the horizontal momentum balance, (2.7c) is the vertical component of the momentum equation, (2.7d, e) are the equations of magnetic induction and (2.7f) is the heat-conduction equation. The system of algebraic equations (2.7) has a solution provided that

$$\begin{aligned} \frac{-is+a^2}{-iq^{-1}s+a^2} R\delta^2 &= 4 \frac{\{(-is+a^2)-im\Gamma\}^2}{\Gamma m^2 + \epsilon(-is+\sigma q a^2) (-is+a^2)} \\ &+ (1+\delta^2) \{\Gamma m^2 + \epsilon(-is+\sigma q a^2) (-is+a^2)\}, \end{aligned} \tag{2.8a}$$

where

$$\delta = k/n\pi. \tag{2.8b}$$

† To compare (2.8) with the results of Roberts & Loper (1979) it is necessary to note that they use the notation $\tau = \Omega_M$, $\Gamma = \epsilon$, $P_m = \sigma q$ and $\omega = s/\tau_{\lambda}$, while for the special case of the plane layer $k\lambda = a^2/\tau_{\lambda}$.

FLM 90

When m = 0, there are no magnetic field perturbations induced by the motion so the Lorentz force vanishes. Since the influence of the magnetic field is absent, this axisymmetric case is not discussed here and only non-zero integer values of m are considered in detail in the following sections. It must be emphasized that only one new term is introduced by the magnetic field line curvature, which is absent in Eltayeb's (1972) model with uniform magnetic field. It is $-im\Gamma$ in the first term on the right of (2.8*a*) which originates from the magnetic hoop stress in the horizontal momentum balance. In addition it may be noted that though the magnetic field strength is not uniform there are no instabilities of the tearing-mode type (e.g. see Furth, Killeen & Rosenbluth 1963). The reason is that in our case curvature exactly compensates the field gradient and as a result the equations (2.7) governing the perturbations (2.6) to the equilibrium state have constant coefficients. Consequently there can be no critical radius and corresponding resonant surface on which a tearingmode instability could occur.

3. MAC waves and related instabilities

Unlike the case of the fast waves considered in the appendix, the primary force balance for the slow waves includes in addition to the Coriolis force both the magnetic and the buoyancy forces. When $1 \ll \Gamma \ll e^{-1}$, there is a particular class of waves with frequencies of order Γ^{-1} which are subject to only slight dissipation. These waves, which are considered in §3.2 below, are the so-called MAC waves discussed first by Braginskii (1967). Elsewhere in this section, a number of other wave motions are considered which may justifiably be called MAC waves also. On the other hand, since their frequencies depend upon the thermal and/or magnetic diffusivities, they are readily distinguished from Braginskii's MAC waves and following popular usage the term MAC wave will not be used to describe them.

Anticipating that the inertia and viscous forces are negligible, it is assumed that both s and a are of order one. Consequently (2.8) is linearized on the basis of $\epsilon \ll 1$ to give

$$\frac{-is+a^2}{-iq^{-1}s+a^2}R\delta^2 = 4\{(-is+a^2)-im\Gamma\}^2/m^2\Gamma + (1+\delta^2)m^2\Gamma.$$
(3.1)

The value of R at which marginal convection occurs is now obtained by supposing that s is real. In particular, the real and imaginary parts yield the two separate identities

$$F\left(s,a^2,\frac{2q}{1+q}\,m\Gamma\right) = \Theta,\tag{3.2a}$$

$$R\delta^{2} = \frac{-8}{q(1-q)\,m^{2}\Gamma}\,F(s,qa^{2},m\Gamma),$$
(3.2b)

where

$$\Theta = \frac{1-q}{1+q} m^2 \Gamma^2 \left\{ 1 - \frac{m^2}{4} \left(1 + \delta^2 \right) \right\}$$
(3.2c)

and

$$F(\xi, \alpha, \beta) = (\xi^2 + \alpha^2) \left(\xi + \beta\right) / \xi. \tag{3.2d}$$

When the values of δ , a, m and Γ are given, (3.2a) is a cubic in s which for certain values of the constants a^2 , $2q(1+q)^{-1}m\Gamma$ and Θ has three real roots. For each value



FIGURE 1. The function $F(\xi, \alpha, \beta)$ is plotted against ξ for two cases: (a) $\beta < 3^{\frac{3}{2}}\alpha$ and (b) $\beta > 3^{\frac{3}{2}}\alpha$. In (a) the dashed curves I and II are the asymptotes $F_{\rm I} = \alpha(\xi + \beta)/\xi$ and $F_{\rm II} = \xi^2 + \alpha^2$ valid for $\xi \ll \alpha$ and $\xi \gg \beta$ respectively. When $\beta \ll \alpha$, F is approximated uniformly for all values of ξ by either $F_{\rm I}$ or $F_{\rm II}$. This is of particular importance for the cases $\Gamma \ll 1$ and $q \ll 1$ considered in §§3.1 and 3.3 respectively. In (b) the dashed curves III and IV are the asymptotes $F_{\rm III} = \beta(\xi^2 + \alpha^2)/\xi$ and $F_{\rm IV} = \xi(\xi + \beta)$ valid for $\xi \ll \beta$ and $\xi \gg \alpha$ respectively. When $\alpha \ll \beta$, F is approximated uniformly for all values of ξ by either $F_{\rm III} = \beta(\xi^2 + \alpha^2)/\xi$ and $F_{\rm IV} = \xi(\xi + \beta)$ valid for $\xi \ll \beta$ and $\xi \gg \alpha$ respectively. When $\alpha \ll \beta$, F is approximated uniformly for all values of ξ by either $F_{\rm III}$ or $F_{\rm IV}$. This is of particular importance for the case $\Gamma \gg 1$ considered in §3.2.

of s there is a corresponding value R(s) of the Rayleigh number. Though these values are difficult to obtain in general there are a few observations that are readily made.

To begin with, the three roots s_1 , s_2 and s_3 of (3.2a) satisfy

$$s_1 + s_2 + s_3 = s_1 s_2 s_3 / a^4 = \frac{-2q}{1+q} m\Gamma, \qquad (3.3a)$$

so (3.2b) implies that

$$\frac{R(s_i) - R(s_j)}{s_i - s_j} = \frac{8}{qm^2\Gamma} \left\{ \frac{2+q}{2} s_k - \frac{m\Gamma}{1+q} \right\},$$
(3.3b)

23-2

A. M. Soward

where (i, j, k) is any permutation of (1, 2, 3). The result indicates that whenever

$$s_k < 2m\Gamma/(2+q)(1+q) \quad (s_i < s_j)$$
 (3.3c)

the value of $R(s_j)$ is less than $R(s_i)$. The approximate location of the roots is determined by inspection of the graph of $F(\xi, \alpha, \beta)$ vs. ξ , which is indicated in figure 1. The character of the graph of F depends upon whether β is less than or greater than $3^{\frac{3}{4}}\alpha$, while the realized values of s and the resulting values of R depend critically on the sign of Θ .

When q < 1 and $m \ge 2$ (or m = 1 and $\delta > \sqrt{3}$) the constant Θ is negative and (3.2a) has one, two or three roots all lying in the interval

$$-2q(1+q)^{-1}m\Gamma \leq s \leq 0.$$

Here R is positive and the result (3.3) is applicable. Therefore the smallest value of R corresponds to the mode with the smallest frequency |s|. When q < 1, m = 1 and $\delta < \sqrt{3}$, the constant Θ is positive and (3.2*a*) has one root in the interval

$$-\Gamma \leq s \leq -2q(1+q)^{-1}\Gamma$$

 $\partial F/\partial s$

The reason is that in this interval

$$< 0$$
 (3.4*a*)

and

$$\frac{1+q}{1-q}F\left(-\Gamma,a^2,\frac{2q}{1+q}\Gamma\right) = \Gamma^2 + a^4 > \frac{1+q}{1-q}\Theta > 0. \tag{3.4b}$$

The corresponding value of R(s) is positive. On the other hand, when Γ exceeds some value Γ_M (say) which depends on q, two additional positive roots s_2 and s_3 of (3.2a) are also possible. The corresponding values of R are negative and, in view of (3.3),

$$0 > R(s_2) > R(s_3) \quad (0 < s_2 < s_3). \tag{3.5}$$

To summarize, the left-hand branch of $F(s, a^2, 2q(1+q)^{-1}m\Gamma)$ corresponds to the usual thermal instability driven by buoyancy forces, which in all cases leads to west-ward-propagating waves. The right-hand branch, which is of interest only when m = 1, corresponds to a magnetic diffusion instability in which the energy for the disturbance is taken from the magnetic field. Thus for given Γ , as the bottom-heavy stratification is increased, instability first sets in as a slow wave, when $R = R(s_2)$. As the stratification is increased the fluid remains unstable until $R = R(s_3)$. At this stage the marginal wave is relatively fast and any further increase in R renders the system stable again. The marginal waves with frequencies s_2 and s_3 both propagate eastwards.

When q > 1 and $m \ge 2$ (or m = 1 and $\delta > \sqrt{3}$), the constant Θ is positive and (3.2a) always has one negative root s, which lies in the interval

$$-\infty < s \leq -2q(1+q)^{-1}m\Gamma.$$

According to (3.2b) the corresponding value of R is positive. For sufficiently large values of Θ there are also two positive roots s_2 and s_3 . In both cases R is positive and according to (3.3) $R(s_2) > R(s_3) > 0 \quad (0 < s_2 < s_3).$ (3.6)

Unfortunately, since the inequality (3.3c) with k = 2 may or may not be satisfied, there is no indication as to which of $R(s_1)$ and $R(s_3)$ is smaller. When q > 1, m = 1 and $\delta < \sqrt{3}$, the constant Θ is negative and (3.2a) has one root in the interval

$$-2q(1+q)^{-1}\Gamma \leqslant s \leqslant -\Gamma,$$

 $\mathbf{676}$

since here both the inequalities (3.4a, b) hold. The corresponding value of R is positive. When $2q(1+q)^{-1} \Gamma \ge 3^{\frac{3}{2}}a^2$, (3.2a) may have an additional two roots lying in the interval $-\Gamma \le s \le 0$.

The corresponding values of R are negative and according to (3.3)

$$0 > R(s_2) > R(s_3) \quad (-\Gamma < s_2 < s_3 < 0). \tag{3.7}$$

In order for these roots to exist Θ must take a sufficiently large negative value. This imposes a lower limit Γ_M (say) at which this magnetic instability can take place. Unlike the case q < 1, instability first sets in with increasing -R as a fast[†] wave and is ultimately suppressed as a slow wave. In both cases the marginal waves propagate westwards.

In the following subsections various limiting cases are discussed in more detail.

3.1. The weak field case, $\Gamma \ll 1$

When Γ is small instability never sets in for negative R. As R is increased from zero, marginal convection occurs first in a localized region at a large distance from the rotation axis where Lorentz and Coriolis forces are comparable. There the magnetic field is almost uniform and the results of Eltayeb (1972) and Roberts & Stewartson (1974) are applicable to lowest order. The main interest in this section is the small modifications which are induced by curvature.

When q < 2, the most unstable mode has

$$n = 1, \quad m^2 \Gamma = 2 \times 3^{\frac{1}{2}} \pi^2, \quad \delta^2 = 2.$$
 (3.8*a*)

The corresponding value of the critical Rayleigh number is

$$R_c = 2 \times 3^{\frac{3}{2}} \pi^2, \tag{3.8b}$$

while the frequency is

$$s_c = -qm\Gamma = -2^{\frac{1}{2}} \times 3^{\frac{1}{2}}\pi q\Gamma^{\frac{1}{2}}.$$
(3.8c)

For a uniform magnetic field this mode describes steady convection (s = 0). It follows that the effect of the field line curvature is to induce a slow westward drift of the convection pattern.

When q > 2, the most unstable mode has

$$n = 1, \quad m^2 \Gamma = 2 \times 3^{\frac{1}{2}} \pi^2 (1+q), \quad \delta^2 = 2.$$
 (3.9*a*)

The corresponding value of the critical Rayleigh number is

$$R_c = 4 \times 3^{\frac{3}{2}} \pi^2 / q. \tag{3.9b}$$

To lowest order the frequency may take either of the values

$$s = \pm 3\pi^2 (q^2 - 2)^{\frac{1}{2}}, \tag{3.9c}$$

but to determine which is preferred it is necessary to proceed to the next approximation. This procedure may, however, be bypassed if the third slow mode corresponding to the values (3.9a) is calculated from (3.2a). This is

$$s = \frac{2qm\Gamma}{(1+q)(q^2-2)} > \frac{2m\Gamma}{(1+q)(2+q)},$$
(3.10)

so according to (3.3) the negative value of s in (3.9c) minimizes the Rayleigh number.

[†] Here and elsewhere in this section the terms fast and slow are used to distinguish the relative sizes of the frequencies s_1 , s_2 and s_3 . In §1, on the other hand, all these waves were described as slow in comparison with the inertial waves, which are very fast.

The above results indicate that the onset of instability takes the form of westwardpropagating waves for all values of q. The slow wave is preferred when q < 2. Furthermore, though the azimuthal wavenumber m is large the ϕ length scale π/m of the waves in the convection region is of order one.

3.2. The strong magnetic field case

When the magnetic field is strong, or more precisely when

$$(1+q)/q \ll \Gamma \ll e^{-1}, \tag{3.11}$$

the function $F(s, a^2, 2q(1+q)^{-1}m\Gamma)$ in (3.2a) is generally characterized by

$$2q(1+q)^{-1}m\Gamma \gg a^2.$$

In this case F is approximated for all values of s by either F_{III} or F_{IV} (defined and illustrated in figure 2b). Furthermore, except possibly when m = 1, Θ is large, of order Γ^2 . Consequently (3.2a) has two large roots of order Γ and one small root of order Γ^{-1} . The two fast waves are the MAC waves discussed previously by Eltayeb & Kumar (1977) and Roberts & Loper (1979). The slow wave is new but is closely related to the mode of steady convection discussed by Eltayeb (1972) and Roberts & Stewartson (1974) for a uniform magnetic field.

Rather than determine the marginal modes from (3.2), it is more instructive to return to (3.1). In the case of the slow modes, for which s is of order Γ^{-1} , a quasi-steady state is achieved when the Rayleigh number is

$$\Gamma R_0 = \Gamma \{-4 + m^2 (1 + \delta^2)\} / \delta^2. \tag{3.12a}$$

If instead the Rayleigh number takes the value

$$R = \Gamma R_0 + R_1 \tag{3.12b}$$

it is clear from (3.1) that s is complex and is given approximately by

$$-is = \frac{qa^2}{(1-q)\Gamma} \left(\frac{R_1}{R_0} + i \frac{8a^2}{mR_0 \delta^2} \right).$$
(3.12c)

Evidently instability occurs whenever

$$R_1/R_0(1-q) > 0. (3.13)$$

In the case of fast modes, for which s is of order Γ , non-dissipative MAC waves are possible when their frequencies

$$\Gamma s_{0\pm} = \Gamma \{ -m \pm \frac{1}{2} m [(1+\delta^2) m^2 - qR\delta^2 / \Gamma] \}, \qquad (3.14a)$$

determined from (3.1), are real. Once the effect of dissipation has been taken into account the frequencies have order-one corrections $s_{1\pm}$. They are determined by substituting

$$s = \Gamma s_{0\pm} + s_{1\pm} \tag{3.14b}$$

into (3.1) and yield the growth rates

$$-is_{1\pm} = \frac{1+q}{2} a^2 \left\{ \frac{\Theta - F(\Gamma s_{0\pm}, a^2, 2q(1+q)^{-1}m\Gamma)}{\Gamma^2 s_{0\pm}(s_{0\pm}+m)} \right\} + O(\Gamma^{-1}).$$
(3.14c)

The value at which $s_{\pm 1}$ is zero, of course, determines the marginal modes.

The various types of instabilities that can arise may now be determined from the general theory. In particular, as |R| is increased from zero, instability first sets in at an order- Γ value either as a slow wave when q < 1 or as a fast MAC wave when q > 1. In the case q > 1 and $m \ge 2$ (or m = 1 and $s > \sqrt{3}$), for which there is the possibility of the excitation of two distinct MAC waves with positive R, (3.12c) and (3.3) indicate that the eastward-propagating (s > 0) mode is preferred. As q is increased, the size of the buoyancy forces required to excite the MAC waves decreases. Thus, when $q = O(\Gamma)$, the frequency s remains of order Γ but the term qR/Γ in (3.14a) becomes small, of order q^{-1} . The buoyancy force is no longer important in the primary force balance and the MC wave instabilities discussed by Roberts & Loper (1979) are recovered.

Some additional care must be taken in discussing the m = 1 mode, for in the limit $\delta \to \sqrt{3}$ the constant Θ in (3.2a) is no longer of order Γ^2 . Consequently the approximations leading to (3.11) and (3.12) are no longer applicable. Nevertheless the resulting critical Rayleigh number, either positive or negative, is readily determined from (3.2). For the appropriate values

$$m = 1, \quad \delta \doteq \sqrt{3}, \quad a \doteq 2\pi \quad (n = 1),$$

$$(3.15)$$

R and s are of order one and (3.2b) indicates that the Rayleigh number is given approximately by

$$R = \frac{-8}{3q(1-q)} \frac{s^2 + 16\pi^2}{s}.$$
(3.16)

For the case $\delta \downarrow \sqrt{3}$, R is positive and takes its critical value

$$R_c = 64\pi^2/3|1-q|, \qquad (3.17a)$$

when the frequency is

$$s_c = \pm 4\pi^2 q.$$
 (3.17b)

The positive (negative) sign is taken when q is greater (less) than one. The signs in (3.17a, b) are reversed when $\delta \uparrow \sqrt{3}$. It follows that the system is stable whenever

$$R_c > R > -R_c \tag{3.18}$$

and is unstable elsewhere. It should be emphasized, however, that these order-one values of R are possibly only in the case m = 1 so the instability is exceptional even for positive R. Thus if the m = 1 mode is excluded R_c will increase linearly with Γ as indicated by (3.12) and (3.13) above.

The fact that instability is possible for all values of $R < -R_c$ is a little surprising but the mathematical reason is simply that $R \to -\infty$ as $\delta \to 0$. The physical reason, on the other hand, is that the modes corresponding to $\delta \to 0$ (*n* fixed) have long horizontal length scales. This implies that the energy for the instability comes from the magnetic field at a great distance, where it is sufficiently large to overcome the static stability.



FIGURE 2. The regions of stability and instability in the Γ , R plane are indicated on the graph for the case $q \ll 1$. The smallest values of |R| for marginal stability always correspond to n = 1. To the right and below the boundary curve AC_{-} , the m = 1 mode is unstable. For a prescribed value of δ , however, the region of instability is confined within the curve $A_{\delta}BC_{\delta}$. The upper (lower) branch BC_{δ} (BA_{δ}) corresponds to the slow (fast) wave for which s is of order q(1) and R is of order $1(q^{-1})$. Both branches have R proportional to Γ as $\Gamma \to \infty$. The ordinate Γ_{δ} of the point B_{δ} increases with δ from Γ_{M} when $\delta = 0$ to infinity when $\delta = \sqrt{3}$. Once δ exceeds $\sqrt{3}$ the sign of R changes. The qualitative features of the stability boundaries for the modes m = 1 ($\delta > \sqrt{3}$), 2 and some other integer greater than 2 are indicated. For each individual mode $R \to \infty$ as either $\Gamma \to 0$ or ∞ . The exception is the m = 1 mode, for which $R \to \frac{64}{3}\pi^2$ as $\Gamma \rightarrow \infty$. The resulting stability boundary for every type of disturbance is continuous but only piecewise differentiable. It begins at C_+ with the mode $m = 1, \delta = \sqrt{3}$ at $\Gamma = \infty$ and ends at D, where $\Gamma \to 0$, $\delta \to \sqrt{2}$ and $m \to (12)^{\frac{1}{2}} \pi/\Gamma^{\frac{1}{2}}$. All modes resulting from bottom (top) heavy stratification R < 0 (> 0) propagate eastwards (westwards). The only region of stability lies between the curves DC_+ and AC_- . Notice that in the absence of stratification (R = 0) the system is stable.

3.3. Small thermal conductivity, $q \ll 1$

The section is concluded with a discussion of the geophysically interesting limit

$$q \ll 1, \quad \Gamma q \ll 1.$$
 (3.19)

In this case, approximate solutions of (3.2) can be obtained for order-one values of Γ , which are at any rate qualitatively applicable whenever q < 1.

The mode of instability which is excited for the smallest values of |R| is a slow mode, which oscillates on the thermal diffusion time scale [s = O(q)]. To isolate this mode the function F in (3.2a) is approximated by $F_{\rm I}$ (see figure 1a). In this way (3.2a) reduces to the equation

$$1 + \frac{2qm\Gamma}{s} = \frac{m^2\Gamma^2}{a^4} \left\{ 1 - \frac{m^2}{4} \left(1 + \delta^2 \right) \right\}$$
(3.20*a*)

for s, while the corresponding Rayleigh number determined by (3.2b) is

$$R = \frac{-8}{m\delta^2} \frac{s^2 + q^2 a^4}{sq}.$$
 (3.20b)

The positive and negative critical Rayleigh numbers are both obtained by minimizing (3.20b) subject to (3.20a). Various curves corresponding to marginal convection are illustrated in figure 2, together with the resulting stability boundary defined by the critical values of R.

The minimum value of Γ at which instability is possible for negative R is of some interest. It is

$$\Gamma_M = (2/3^{\frac{1}{2}})\pi^2 \tag{3.21}$$

and occurs as $\delta \to 0$, $s \to -\infty$ and $R\delta^2 \to -\infty$! To see how this limit is approached it is supposed that Γ is slightly in excess of Γ_M . Substitution of

 $\Gamma = \Gamma_M + \Delta \quad (q^{\frac{3}{5}} \ll \Delta \ll 1) \tag{3.22}$

into (3.20a) then gives

$$\Gamma_M^2 q/s = \Delta - \frac{7}{6} \Gamma_M \delta^2 \tag{3.23a}$$

so (3.20b) reduces to

$$R = -8\Gamma_M^2/\delta^2(\Delta - \frac{7}{6}\Gamma_M\delta^2). \qquad (3.23b)$$

The critical value of R is

$$R_c = -\frac{112}{3} \Gamma_M^3 / \Delta^2 \tag{3.24a}$$

and corresponds to the critical values

$$\delta_c^2 = \frac{3}{7} \Delta / \Gamma_M, \quad s_c = 2\Gamma_M^2 q / \Delta. \tag{3.24b, c}$$

It is readily seen that the approximations leading to (3.20) cease to be applicable once $s = O(q^{\frac{1}{2}})$ and so the result (3.24) is no longer valid when $\Delta = O(q^{\frac{3}{2}})$. In order to understand the significance of the order- $q^{\frac{1}{2}}$ value of s, the stability curve for m = 1, n = 1 with δ fixed is considered. In this particular case instability for negative values of R is not possible for values of Γ less than

$$\Gamma_0 = a^2 (1 - a^2 / 4\pi^2)^{-\frac{1}{2}} \quad (> \Gamma_M). \tag{3.25}$$

Upon writing

$$\Gamma = \Gamma_0 + q^{\frac{3}{2}} \Gamma_2, \quad s = q^{\frac{1}{2}} s, \tag{3.26a}$$

(3.2a-d) yield

$$s_1^3 - 2a^4(\Gamma_2/\Gamma_0)s_1 + 2\Gamma_0a^4 = 0 \tag{3.26b}$$

and

$$R = -\frac{8s_1}{\delta^2 q^{\frac{2}{3}}}.$$
 (3.26c)

For large negative Γ_2 , (3.26b) has only a single negative root s_1 . This is the slow wave which is excited by positive R. When

$$\Gamma_2 = \frac{3}{2} (\Gamma_0^5 / a^4)^{\frac{1}{3}} \tag{3.27a}$$

a second wave, with frequency

$s_1 = (a^4 \Gamma_0)^{\frac{1}{2}}, \tag{3.27b}$

is possible for

$$R = -(8/\delta^2) \left(a^4 \Gamma_0/q^2\right)^{\frac{1}{3}}.$$
 (3.27c)

For larger values of Γ_2 this wave splits into a slow wave with R negative and of order one and a fast wave [s = O(1)] with R negative and of order q^{-1} . Meanwhile, the original slow wave which occurred for negative Γ_2 changes to a fast wave when Γ_2 takes large positive values. It always corresponds to positive R.

4. Discussion

The commonly accepted values of the diffusivities in the earth's fluid core are $\kappa \neq 10^{-5} \text{ m}^2 \text{ s}^{-1}$ and $\lambda = 3 \text{ m}^2 \text{ s}^{-1}$, while ν is believed to be comparable to or smaller than κ . With a core radius of $d \neq 3.5 \times 10^6$ m, the geodynamo is characterized by

$$\sigma \lesssim 1, \quad q \approx 3 \times 10^{-6}, \quad \epsilon \approx 5 \times 10^{-9}.$$
 (4.1)

There is no general agreement, however, about the value of the magnetic field strength B_0^* (and Γ), which depends critically on the nature of the dynamo process. The traditional picture of the geodynamo is one of the Braginskii (1964) type, for which B_0^* is of the order of 400 gauss. On the other hand, in a more recent dynamo model proposed by Busse (1976) it is argued that B_0^* is of the order of 4 gauss, comparable to its surface value. If these two values are taken as upper and lower bounds on the value of B_0^* , the value of Γ realized lies in the interval

$$2 \times 10^{-2} \leqslant \Gamma \leqslant 2 \times 10^2. \tag{4.2}$$

An order-one value of Γ , which will lie inside this interval, is especially attractive, since all convective models indicate that the critical Rayleigh number is minimized in this case.

Though all values of Γ have been considered in this paper, only those results obtained for Γ of order one or larger are directly applicable to confined geometries. For this reason the westward propagation of waves isolated here is relevant to the geodynamo only when $\Gamma \gtrsim 1$ and supports Eltayeb & Kumar's (1977) full numerical calculations for the sphere. For smaller values of Γ the direction of wave propagation is likely to be eastward as predicted both by their model and Busse's (1976).

In addition to the usual thermal instabilities, some of the magnetic instabilities isolated here may also form an integral part of the dynamo process. In particular, they may provide a mechanism for limiting the growth of the magnetic field. Now the fast inertial wave instability described in the appendix is unlikely to occur for, unless σ is extremely small, the stability criterion (A 7) is met by the values of q, ϵ and Γ given by (4.1) and (4.2). On the other hand, the slow resistive modes of instability described in §3.3 are possible. They are of two types, which are most readily distinguished in the limit of large Γ discussed in §3.2. There they are typified by the two exceptional modes (3.15), which occur when R is either greater than or less than zero. It is, however, difficult to see how the system can evolve into a state in which the instability for R < 0 can operate. The reason is simply that no force is available to drive motions, which can first intensify the magnetic field in the required way. A further objection to this instability is that it is manifest by eastward-propagating waves in contrast with the westward-propagating waves which occur when q > 1. It would therefore appear reasonable to suppose that the only relevant magnetic instabilities are the exceptional westward-propagating waves, which closely resemble the usual thermal instability and occur when R > 0.

In their analysis of the spherical model, Eltayeb & Kumar (1977) isolated only the MAC wave instability when considering large Γ . Even with the exclusion of all the exceptional magnetic modes described in the previous paragraph, this mode of instability is likely to be preferred only when q exceeds some order-one value. In par-

ticular, the results of §3.2 indicate that for the plane layer the slow resistive mode with $s \sim \Gamma^{-1}$ is preferred when q < 1. Furthermore, in the geophysically interesting limit $q \ll 1, q\Gamma \ll 1$ discussed in §3.3, MAC wave marginal convection is non-existent.

The author wishes to thank Professor P. H. Roberts for keeping him informed of his research with Professor D. E. Loper. Their discovery of magnetic instabilities for bottom-heavy fluids was a source of stimulation during the detailed study of this mechanism reported in § 3. The work was begun at the School of Mathematics, University of Newcastle upon Tyne and was completed under the support of National Science Foundation grant no. EAR 77-00145.

Appendix

Roberts & Loper (1979) have shown that in the absence of viscosity and buoyancy forces a fast ($cs \sim 1$) m = 1 mode is unstable at all values of Γ . Here the nature of this instability for the plane-layer model is described briefly.

Upon setting R = 0 (2.8) reduces to two separate quadratic equations

$$\mp i\{(-is+a^2)-im\Gamma\} = \frac{1}{2}(1+\delta^2)^{\frac{1}{2}}\{\Gamma m^2 + \epsilon(-is+\sigma q a^2)(-is+a^2)\}$$
(A 1)

for s with a total of four distinct solutions. When $\epsilon \ll 1$, two of them are of order one and the general theory of §3 shows that they correspond to damped waves. The other two solutions are large, of order ϵ^{-1} , and to a first approximation their values

$$e^{-1}s_{0\pm} = \pm e^{-1}2(1+\delta^2)^{-\frac{1}{2}}$$
 (A 2)

are the frequencies of two inertial waves. In order to determine whether these modes are growing or damped it is necessary to determine the complex frequency

$$s = e^{-1}s_{0\pm} + s_{1\pm} + O(\epsilon) \tag{A 3}$$

correct to order one. At this stage the key role of viscosity is isolated by supposing that

$$\Gamma = O(1), \quad \sigma q = O(\epsilon), \quad a^2 = O(\epsilon^{-1}). \tag{A 4}$$

It follows immediately from (3.1) that

$$-is_{1\pm} + \sigma q a^2 = -(m + s_{0\pm}) m \Gamma / (-is_{0\pm} + a^2 \epsilon)$$
(A 5a)

and so the order-one imaginary part of s is given by

$$\operatorname{Im} s_{1\pm} = -\sigma q a^2 + \frac{(\mp 1 - \frac{1}{2}(1 + \delta^2)^{\frac{1}{2}} m) m\Gamma}{\frac{1}{2}(1 + \delta^2)^{\frac{1}{2}}} \frac{a^2 \epsilon}{a^4 \epsilon^2 + 4/(1 + \delta^2)}.$$
 (A 5b)

When m = 1 and $\delta < \sqrt{3}$, the westward-propagating wave, $s = e^{-1}s_0$, may be unstable. In the absence of viscosity ($\sigma = 0$), the mode with the maximum growth rate is characterized by

$$k = O(1), \quad \delta = O(e^{\frac{1}{2}}), \quad a^2 \epsilon = 2$$
 (A 6*a*, *b*, *c*)

and has

$$\operatorname{Im} s_{1-} = \frac{1}{4}\Gamma. \tag{A 6d}$$

It has a short vertical length scale and grows on the magnetic diffusion time scale τ_{λ} based on the width of the layer. With viscosity included, the growth rate is reduced,

A. M. Soward

and if instead of (A 6c) one considers the limit $a^2 \epsilon \rightarrow 0$, it is readily seen that instability is totally eliminated when

$$4\sigma q > \epsilon \Gamma$$
 (i.e. $2(\nu/\lambda)^{\frac{1}{2}} \ge \Omega_M/\Omega_c$). (A 7)

Though (A 7) is implicit in the results of Roberts & Loper (1979), the result (A 6), which hinges on the improved formula (A 5), is new.

When $R \neq 0$ and the buoyancy forces are taken into account (A 5a) must be replaced by

$$-is_{1\pm} + \sigma q a^2 = \frac{-(m+s_{0\pm})m\Gamma}{-is_{0\pm} + \epsilon a^2} + \frac{R}{2} \left(\frac{\delta^2}{1+\delta^2}\right) \frac{1}{-iq^{-1}s_{0\pm} + \epsilon a^2}.$$
 (A 8)

Evidently, for the parameter range envisaged in (A 4) $[\sigma = O(1)]$, R must be extremely large before buoyancy forces have any significant effect on the fast modes with short vertical length scales.

REFERENCES

- BRAGINSKIĬ, S. I. 1964 Zh. Eksp. teor. Fiz. 47, 1084. [Trans. Sov. Phys. J. Exp. Theor. Phys. 20, 726 (1965).]
- BRAGINSKII, S. I. 1967 Geomag. Aero. (USSR) 7, 1050. [Trans. Geomag. Aero. 7, 851 (1967).]
- BUSSE, F. H. 1970 J. Fluid Mech. 44, 441.
- BUSSE, F. H. 1975 Geophys. J. Roy. Astr. Soc. 42, 437.
- BUSSE, F. H. 1976 Phys. Earth Planet. Int. 12, 292.
- CHANDRASEKHAR, S. 1961 Hydrodynamic and Hydromagnetic Stability. Oxford University Press.
- CHILDRESS, S. & SOWARD, A. M. 1972 Phys. Rev. Lett. 29, 837.
- ELTAYEB, I. A. 1972 Proc. Roy. Soc. A 326, 229.
- ELTAYEB, I. A. 1975 J. Fluid Mech. 71, 161.
- ELTAYEB, I. A. & KUMAR, S. 1977 Proc. Roy. Soc. A 353, 145.
- FURTH, H. P., KILLEEN, J. & ROSENBLUTH, M. N. 1963 Phys. Fluids 6, 459.
- HIDE, R. 1966 Phil. Trans. Roy. Soc. A 259, 615.
- LOPER, D. E. 1975 Phys. Earth Planet. Int. 11, 43.
- MALKUS, W. V. R. 1967 J. Fluid Mech. 28, 793.
- ROBERTS, P. H. 1968 Phil. Trans. Roy. Soc. A 263, 93.
- ROBERTS, P. H. & LOPER, D. E. 1979 J. Fluid Mech. 90, 641.
- ROBERTS, P. H. & STEWARTSON, K. 1974 Phil. Trans. Roy. Soc. A 277, 287.
- ROCHESTER, M. G., JACOBS, J. A., SMYLIE, D. E. & CHONG, K. F. 1975 Geophys. J. Roy. Astr. Soc. 43, 661.

Soward, A. M. 1974 Phil. Trans. Roy. Soc. A 275, 611.