PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

On Finite Amplitude Convection in a Rotating Magnetic System

P. H. Roberts and K. Stewartson

Phil. Trans. R. Soc. Lond. A 1974 **277**, 287-315
doi: 10.1098/rsta.1974.0052

ON FINITE AMPLITUDE CONVECTION IN A ROTATING MAGNETIC SYSTEM

By P. H. ROBERTS
School of Mathematics, University of Newcastle upon Tyne

AND K. STEWARTSON, F.R.S.
Department of Mathematics, University College London

(*Received* 11 *February* 1974)

CONTENTS

The instabilities that can arise in a stratified, rapidly rotating, magnetohydrodynamic system such as the Earth's core are often thought to play a key role in dynamo theory – that is, in the study of how the magnetic field in the system is maintained in the face of ohmic dissipation. An account of such instabilities is to be found in the M.A.C.-wave theory of Braginsky (1967), who, however, laid his greatest emphasis on the dissipationless modes, an idealization which leads to difficulties described below. ohmic and thermal diffusion is therefore restored, and three key dimensionless parameters are isolated: q, the ratio of thermal to ohmic diffusivities; λ, a measure of the relative importance of Coriolis and magnetic forces; and R, a Rayleigh number, which is here the ratio of buoyancy to Coriolis forces.

This study concentrates on a particular M.A.C.-wave model originally proposed by Braginsky. It consists of a horizontal layer containing a uniform horizontal magnetic field, $\boldsymbol{B}_0$, and rotated about the vertial, an adverse temperature gradient being maintained on the horizontal boundaries to provide the unstable density stratification. In the rotationally dominant case of large λ, the principle of the exchange of stabilities holds, and the motions that arise in the marginal state are steady. The planform of the convection is in rolls orthogonal to $\boldsymbol{B}_0$. If q and λ are sufficiently small the principle of the exchange of stabilities remains valid, but the planform consists of one or other of two families of rolls oblique to $\boldsymbol{B}_0$, or a combination of each. If q is large but λq is small, the modes are again oblique, but overstability occurs, a type of oscillation which also arises when q is large and λ takes intermediate values, although the motion is then in rolls transverse to $\boldsymbol{B}_0$.

A theory is developed for the weakly nonlinear convection that arises when R exceeds only slightly the critical value $R_c(q, \lambda)$ at which marginal convection occurs.

A critical curve $q = q_D(\lambda)$ is located which roughly divides the (q, λ) plane into regions of small q and of large q, although when q is large it separates the large λq from the small. On the one side of the curve, where q or λq are sufficiently small, it is concluded that, starting from an arbitrary initial perturbation, the convection that arises when R exceeds R_c will ultimately become a completely regular tesselated pattern filling the horizontal plane. On the other side of the curve the situation is considerably more complicated but it is argued that, for sufficiently large q and λq, subcritical instabilities can occur and that supercritical bursting is likely; that is, the instability that arises from the general initial perturbation will focus into a small spot in a finite time. The relevance of the theory to sunspot formation is discussed.

In an appendix, the form of the weakly nonlinear convection that arises when q differs only slightly from q_D, and R only slightly from R_c, is considered in situations in which the exchange of stabilities holds.

1. Introduction

It is now generally believed that the existence of the geomagnetic field is a manifestation of a finite amplitude instability of the Earth's core. It is thought that some agency, perhaps thermal or non-thermal convection, or perhaps precessional forces, drives the core into motion. A purely hydrodynamic flow, unaccompanied by magnetic field, could result. In fact, however, this solution is magnetically unstable, in the sense that any stray field is amplified by the induced e.m.fs associated with the motion of the electrically conducting fluid in the core. As this field increases, the Lorentz force it creates modifies and reduces the flow velocities until field growth ceases, and a different state of balance with the driving forces is struck, one in which the magnetic field is non-zero. This state is presumed to be quasi-stable, except possibly to a sufficiently large perturbation that causes the entire field, $\boldsymbol{B}$, to reverse ($-\boldsymbol{B}$ being as valid a solution to the magnetohydrodynamic equations as $+\boldsymbol{B}$).

Clearly, several questions deserve answers. First, it is necessary to explain why the non-magnetic state should be magnetically unstable. This is the so-called 'kinematic dynamo problem', which has received considerable attention in recent years. It is now known, in a general way, what the strength and character of the fluid motions must be if they are to provide sufficient dynamo action. It is necessary that the magnetic Reynolds number

$$R_m = UL/\eta = \mu\sigma UL \tag{1.1}$$

should be 'sufficiently large'. Here L is a typical length, U a typical relative velocity, μ is the permeability of the fluid, σ is its conductivity, and η is its magnetic diffusivity. A significant theoretical milestone was reached when it was realized that L is not necessarily the length scale, L_V, of the motions, but is related to the length scale, L_B, of the field created. Not only could a plausible theory be developed for cases in which $L_B \gg L_V$, but also astrophysical systems could be observed, particularly the solar convection zone, which are magnetic and in which $L_B \gg L_V$. A mathematical apparatus could be most readily built for cases in which the microscale Reynolds number $R_V = UL_V/\eta$ is small, the macroscale Reynolds number, $R_B = UL_B/\eta$ is large, and their product is of order unity. In this way, Steenbeck, Krause & Rädler (1966) could develop a theory adumbrated by Parker (1955*b*) to show that such flows could regenerate field if they possessed sufficient helicity, that is if their velocity and vorticity were sufficiently correlated; they also suggested mechanisms through which helicity might arise in rotating fluids. Their model of the induction process was one in which the small-scale motions are turbulent in character, but this

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

stochastic element is not necessary, as was realized by Childress (1967) who, by examining laminar tessalated flows, could provide a firm mathematical foundation to their theory; see also Roberts (1970). The technique of expansion in powers of L_V/L_B as a small parameter is now usually called 'the two-scale method'.

Variants of the two-scale method exist in which large-scale motions are present, which assist the dynamo to function. For example, in the so-called $\alpha\omega$-dynamo process, a toroidal flow creates a strong alined field by shearing the poloidal field lines (Elsasser 1947). Nevertheless, this model, and all others referred to as two-scale dynamos, rely on the existence of induction on the microscale, in this case to generate poloidal field from the toroidal.

Another successful approach to dynamo theory, one which proves that not all dynamos need be of two-scales, is the asymptotic theory of induction at large magnetic Reynolds number initiated by Braginsky (1964*a*) and developed by Soward (1972) and Gubbins (1973). Induction by small-scale motions is not invoked: L_V and L_B are comparable with the dimensions of the fluid container, L, and the asymptotic theory for $R_m \to \infty$ is successfully developed to the point where dynamo action can be demonstrated in models (see, for example, Braginsky 1964*b*; Roberts 1972*a*).

Despite the growing successes of the kinematic theory, it can only provide a part – and probably the smaller part – of an understanding of the existence and stability of the magnetohydrodynamic solution to which we referred in our opening paragraph. Plausible order of magnitude arguments indicate that the Earth's core is in a state of approximate magnetogeostrophic balance – that is, the Lorentz, Coriolis and pressure forces balance† in the primary motion. The reason for this is unknown. And, in particular, it is not understood why such a balance should be struck in the Earth when, in another highly rotating system, the solar convection zone, the magnetic forces are apparently not great enough to affect the primary flow seriously.

Since the kinematic induction problem is a necessary ingredient, it is natural that the initial attacks on the magnetohydrodynamic problem have been either through two-scale models or via Braginsky's asymptotic method. Busse (1973) has constructed a dynamo in a Bénard layer in which the convective motions are supplemented by a unidirectional plane Poiseuille flow set up by an externally applied pressure gradient. The scale of the magnetic field generated is large compared with the dimensions of the convection cells and the induction equation is therefore amenable to the two-scale method. Busse showed that, if the Rayleigh number, R, sufficiently exceeds the critical value, R_c, necessary for thermal instability, amplification of field will occur, and he determined the amplitude of the small steady-state field which could exist when $R-R_c$ is small. Childress & Soward (1972) and Soward (1974) examined a model which is perhaps closer to the solar dynamo. The motions are set up in a highly rotating Bénard layer, in which (cf. Chandrasekhar 1961, ch. 3) the horizontal size of the most unstable convection cells is a small fraction, of under $T^{-\frac{1}{6}}$, where T is the Taylor number, of the depth of the layer. This difference in dimension allows a variant of the two-scale method to be invoked in a particularly natural way. Again magnetic instability was established for sufficiently large $R-R_c$,

† The inertial forces are, however, weak. If, following the review by Roberts & Soward (1972), we take $V \simeq 10^{-4}$ m/s, $B \simeq 0.04$ tesla, $\rho \simeq 10^5$ kg/m³ and $L = 3.5 \times 10^6$ m = the core radius, we find that the Lorentz force $B^2/\mu L$ is about 4×10^{-4} k/m² s², while the Coriolis force $\rho\Omega V$ is about 7×10^{-4} kg/m² s², where Ω is the angular velocity. The Rossby number $V/\Omega L$ is, however, only 4×10^{-7}, and the inertial forces $\rho V^2/L$ only 3×10^{-10} kg/m² s². For future reference, we may note that the magnetic and thermal diffusivities are $\eta \simeq 3$ m²/s and $\kappa \simeq 10^{-5}$ m²/s, so that $q = \kappa/\eta \simeq 3 \times 10^{-6}$. Owing to the importance of the radiative conductivity, however, q might be of order 10^5 in a star. Consideration of this case is deferred to §7 below.

and again a theory for small steady-state fields was given for small $R-R_c$. In both these models magnetic fields had only small dynamical effects, and viscosity played as significant a part as buoyancy in determining the motions.

Quite a different model was considered by Braginsky (1964*c*, 1967) in relation to the Earth. His starting point is a primary state of toroidal axisymmetric flow and magnetic field, the latter being created from the former by shearing of the poloidal field. His asymptotic theory of induction at large magnetic Reynolds numbers, R_m, showed that asymmetric motions, with amplitude only a small fraction, of order $R_m^{-\frac{1}{2}}$, of the primary zonal flow, could regenerate the poloidal field from the toroidal. It was, then, necessary to find a dynamical reason for such motions. Braginsky pictured these as magnetohydrodynamic instabilities of the primary state, and he made the point, particularly significant in view of the well-known theorem of Cowling (1933), that such instabilities would be expected to arise even when the state was stable to axisymmetric motions. Unlike the constructs of Busse, Childress and Soward, the magnetic field of the Braginsky model is not small, and the *M*agnetic (Lorentz) force is as potent as the *A*rchimedean (buoyancy) force driving the instability and as the *C*oriolis forces in determining the course of the instabilities, a fact which led Braginsky to christen them 'M.A.C.-waves'. Also, unlike the models of Busse, Childress and Soward, the viscous forces have a negligible effect.

Braginsky developed an idealized model of M.A.C.-waves (1964*c*) and later gave a general theory for them (1967). The large values of R_m inferred for the Earth, together with the success of his large R_m induction theory, led him to discount the effects of ohmic and thermal† diffusion in his theory of M.A.C.-waves. This may be objected to on general grounds, as it alters the physical characters of the phenomenon from one in which persistent motions are possible to one in which any instability results in a single convulsion of overturning, together with associated transients: contrast, for example, the behaviour of the Bénard layer, in which diffusion of density occurs and in which viscosity provides a counter to buoyancy, to Rayleigh–Taylor instability, in which motion occurs to bring about a stable density stratification when one does not initially exist, a diffusion of vorticity and density playing no essential role.

A more mathematical objection may also be voiced, which for simplicity we will level against his first simple model (Braginsky 1964*c*, § 4). This consists of a plane horizontal layer, $0 \leqslant z \leqslant d$, rotating about the vertical with uniform angular velocity, Ω, and containing a uniform primary flow, V_0, and an alined field, B_0, both in the y-direction, which is taken to correspond to that of increasing longitude, ϕ, in the Earth's core. A top-heavy gradient, $\rho\alpha\beta$, of density, ρ, is created by an applied thermal gradient β, where α is the coefficient of volume expansion. Perturbation solutions are sought proportional to $\exp[\mathrm{i}(lx+my+nz+\omega t)]$. Clearly boundary conditions impose a discrete spectrum on n; for example, the vanishing of the normal velocity on $z=0$ and $z=d$ requires that n is an integral multiple of π/d. Recalling that ϕ is a periodic coordinate in the core, it is also reasonable to suppose that m is an integral multiple for some basic quantity πD^{-1}, where D has the dimensions of length. The asymmetric wave-motions yield the dispersion relation

$$\omega^2 = \frac{m^2B_0^2}{4\mu\rho\Omega^2}\,\frac{(l^2+m^2+n^2)}{n^2}\left[\frac{m^2B_0^2}{\mu\rho}-\frac{g\alpha\beta(l^2+m^2)}{(l^2+m^2+n^2)}\right], \tag{1.2}$$

† Braginsky has argued that the Earth's core is driven by non-thermal buoyancy of light silicates released when the ferrous components of the core fluid, in contact with the inner body, crystallizes onto its surface. This view has been prompted by consideration of the relative inefficiency of thermal convection as a driving engine, a topic outside the scope of this paper. We therefore adopt the simpler picture of thermal convection, and use 'thermal diffusion' rather than the coefficient of diffusion for a solute.

which shows that the wave becomes unstable once

$$\frac{g\alpha\beta\mu\rho}{B_0^2} > m^2 + \frac{m^2n^2}{(l^2+m^2+n^2)}. \tag{1.3}$$

It is clear that the most unstable mode is given by $m = \pi/D$ and $l = \infty$, irrespective of the value of n, i.e. there is an infinity of values of l and n for which unstable modes exist, no matter by how little $g\alpha\beta\mu\rho D^2/B_0^2$† exceeds unity. Further, if one wishes to study the evolution of the initial state of the fluid if unstable, then the mechanisms considered by Braginsky are insufficient since the growth involves very rapid oscillations in the x-direction and diffusive effects cannot then be neglected. It seems likely that of these effects viscosity plays a subservient role and only thermal and ohmic diffusivities are of significance in studies of relevance to the Earth's core. The dominant part played by magnetic forces in Braginsky's theory must also be retained, in contrast to the theories of Busse, Childress and Soward, since this too is the case in the Earth's core.

The generalization of Braginsky's linear theory to include thermal and ohmic dissipation was carried out by Eltayeb (1972), whose results are remarkable and stand in contrast to those of the diffusionless theory. For example, rotation plays no part in the stability criterion (1.3). When diffusion is restored the criterion for instability for sufficiently strong fields and sufficiently small values of $q = \kappa/\eta$ (the ratio of the thermal and magnetic diffusivities) is that

$$R \equiv \frac{g\alpha\beta d^2}{2\Omega\kappa\pi^2} > 3\sqrt{3}, \tag{1.4}$$

a condition which is independent of the field strength. The most unstable mode is associated with finite values of l^2+m^2, which suggests that the forces neglected in deriving (1.4) are unlikely to be significant. Moreover when the instability is marginal the growth rate of the critical mode of disturbance is

$$\frac{g\alpha\beta}{2\Omega\sqrt{3}} - \frac{3\kappa\pi^2}{d^2}, \tag{1.5}$$

which is also independent of the field. Other properties are set out in § 3 of the present paper.

Our intention here is to begin a study of the evolutionary properties of the instabilities discovered by Eltayeb when the disturbances become sufficiently large for nonlinear effects to be important. Our ultimate aim is, however, that of examining the prospects for regenerating the applied magnetic field. As in Busse's model, a sufficiently slow variation of field in the horizontal will allow two-scale methods to be adopted, both for spatial modulation of the nonlinear convection, and to the dynamo process to which it gives rise.

After formulating our problem mathematically in § 2, we summarize the theory of its linear stability given by Eltayeb (1972) and develop his theory farther (§ 3). The evolution of weakly nonlinear motions is described by the analysis of § 4, the results of which are divided, for clarity, into the cases in which instability arises as aperiodic growth (§ 5) and those in which it arises as an oscillation of increasing amplitude (§ 6). The results are discussed in the geophysical and astrophysical contexts in § 7.

2. Statement of the problem

We consider an inviscid fluid of constant conductivity, σ, and thermal conductivity, κ, at rest between two rigid horizontal plates which are perfectly conducting in both the thermal and electrical senses. We assume that the whole system is in uniform rotation, with angular

† See note added in proof on page 315.

velocity, $\boldsymbol{\Omega}$, about the vertical. The plates are maintained at uniform temperatures $\theta_0 \pm \frac{1}{2}\beta d$, where $\beta d > 0$, with the hotter plate beneath the colder. A uniform magnetic field, $\boldsymbol{B}_0$, acts in a horizontal direction which is fixed relative to the rotating frame. (Such a field could, for example, be generated by co-rotating external coils.) The fluid is Newtonian with equation of state

$$\rho^* = \rho_0[1-\alpha(\theta^*-\theta_0)], \tag{2.1}$$

where ρ^* is the density, θ^* is the temperature, and ρ_0, θ_0 and α are positive constants, the coefficient of volume expansion, α, being such that $\alpha\beta d \ll 1$ so that the Oberbeck–Boussinesq approximation may be adopted. The temperature of the fluid at a distance z^* below the plane midway between the plates is

$$\eta^* = \theta_0 + \beta z^*, \tag{2.2}$$

the plates themselves being $z^* = \pm\frac{1}{2}d$. This static state of the fluid is subjected at time $t = 0$ to an infinitesimal perturbation. We wish to examine the conditions under which this disturbance is amplified, and to decide its ultimate fate.

Define an orthogonal set of axes $Ox^*y^*z^*$ relative to the rotating frame, with Oz^* downwards, Oy^* in the direction of the applied field, and O in the mid-plane. Let the fluid velocity be $\boldsymbol{V}^*$ and the magnetic field be $\boldsymbol{B}^*$. Introduce dimensionless variables by the transformations

$$\left.\begin{aligned} \boldsymbol{x}^* &= \frac{d}{n\pi}\boldsymbol{x}, \quad t^* = \frac{d^2}{\kappa n^2\pi^2}t,\\ \boldsymbol{V}^* &= \frac{n\pi\kappa}{d}\epsilon \boldsymbol{V}(\boldsymbol{x},t),\\ \boldsymbol{B}^* &= B_0[\hat{\boldsymbol{y}} + \mu\sigma\kappa\epsilon \boldsymbol{B}(\boldsymbol{x},t)],\\ \theta^* &= \theta_0 + \frac{\beta d}{n\pi}[z+\epsilon\theta(\boldsymbol{x},t)],\end{aligned}\right\} \tag{2.3}$$

where σ is the electrical conductivity of the fluid, μ is its permeability, ϵ is a small parameter representing the magnitude of the disturbance applied at $t = 0$, and n is an integer which we will temporarily leave undetermined; later, it will emerge that the case $n = 1$ is of greatest interest.

The equations governing the motion of the fluid now assume the forms

$$\operatorname{div}\boldsymbol{V} = \operatorname{div}\boldsymbol{B} = 0, \tag{2.4a}$$

$$\nabla^2\theta - V_z - \partial\theta/\partial t = \epsilon(\boldsymbol{V}\cdot\nabla\theta), \tag{2.4b}$$

$$\partial\boldsymbol{V}/\partial y + \nabla^2\boldsymbol{B} - q\,\partial\boldsymbol{B}/\partial t = -\epsilon q\operatorname{curl}(\boldsymbol{V}\times\boldsymbol{B}), \tag{2.4c}$$

$$\partial\boldsymbol{B}/\partial y - \lambda\hat{\boldsymbol{z}}\times\boldsymbol{V} - \lambda R\theta\hat{\boldsymbol{z}} - \operatorname{grad}\Pi + \epsilon q(\operatorname{curl}\boldsymbol{B})\times\boldsymbol{B} = n^2\delta^2(\partial\boldsymbol{V}/\partial t + \epsilon\boldsymbol{V}\cdot\nabla\boldsymbol{V}). \tag{2.4d}$$

Here
$$\lambda = \frac{2\Omega\rho_0}{\sigma B_0^2}, \quad R = \frac{g\alpha\beta d^2}{2\Omega\kappa\pi^2n^2}, \quad q = \mu\sigma\kappa, \quad \delta^2 = \frac{\lambda\pi^2\kappa}{2\Omega d^2}, \tag{2.5}$$

$\hat{\boldsymbol{z}}$ is a unit vector in the direction of the acceleration due to gravity, $\boldsymbol{g}$, and

$$\Pi = [p^* - (\Omega^2 d^2\rho_0/2\pi^2n^2)(x^2+y^2)]/qB_0^2 \tag{2.6}$$

is the dimensionless reduced pressure. These are the standard equations of magnetohydrodynamics including buoyancy in the sense of Oberbeck & Boussinesq, but by implying that the Ekman number $n^2\pi^2\nu/\Omega d^2$, where ν is the kinematic viscosity, is negligible, all viscous effects have been excluded.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The boundary conditions associated with (2.4) are that, at both $z = -\frac{1}{2}n\pi$ and $z = \frac{1}{2}n\pi$,

$$V_z = 0, \tag{2.7a}$$

since the velocity of the fluid normal to the plates must be zero at the plate;

$$\theta = 0, \tag{2.7b}$$

since the plates are perfect thermal conductors; and,

$$\partial B_x/\partial z = \partial B_y/\partial z = B_z = 0, \tag{2.7c}$$

since the plates are perfect electrical conductors. The conditions on $\boldsymbol{V}$, $\boldsymbol{B}$ and θ for large x and y are not so readily decided by the geophysical motivations of the study, but it is reasonable to suppose that they should be periodic in x and y, or at least remain bounded as $|x|$ and $|y|$ become infinite. It is perhaps worth noticing that, by using (2.7c) and the equivalence of the divergence of the Maxwell stress to the Lorentz force, we can then show that

$$\hat{\boldsymbol{z}} \times \int_{-\infty}^{\infty} \mathrm{d}x^* \int_{-\infty}^{\infty} \mathrm{d}y^* \int_{-\frac{1}{2}d}^{\frac{1}{2}d} \mathrm{d}z^* (\mathrm{curl}^* \boldsymbol{B}^*) \times \boldsymbol{B}^* = 0, \tag{2.8}$$

for all t.

Equations (2.4) and (2.7) form a predictive set from which the evolution of $\boldsymbol{V}$, $\boldsymbol{B}$ and θ can be followed from an arbitrarily assigned critical state. We shall now, however, make an additional assumption, namely that

$$\delta \ll 1, \tag{2.9}$$

and that the inertia terms on the right-hand side of (2.4d) are therefore negligible in comparison with the Coriolis and Lorentz forces on the left. The geophysical plausibility of this assumption was discussed briefly in § 1, but the more immediate question is whether, after the neglect of $\partial \boldsymbol{V}/\partial t$ in (2.4d), the set (2.4) and (2.7) remains predictive. More precisely, assuming that (2.4d) is obeyed at time $t = 0$, does the set determine $\boldsymbol{V}$, $\boldsymbol{B}$ and θ uniquely at later times?

It may be seen that, when (2.4d) lacks its right-hand side, it implies

$$\hat{\boldsymbol{z}} \cdot \int_{-\frac{1}{2}d}^{\frac{1}{2}d} \mathrm{d}z^* \,\mathrm{curl}^* \left[(\mathrm{curl}^* \boldsymbol{B}^*) \times \boldsymbol{B}^*\right] = 0, \tag{2.10}$$

for all t. In particular, (2.4) and (2.7) cannot determine the evolution of arbitrarily assigned $\boldsymbol{V}$, $\boldsymbol{B}$ and θ, but only those obeying (2.10). We will refer to (2.10) as 'Taylor's condition', although Taylor (1963) was primarily concerned with motions in a spherical container for which a restriction less severe than (2.10) is necessary. Condition (2.10) may be contrasted with (2.8). The latter must be obeyed by all initial centred disturbances, and is then automatically satisfied at all later times. On the other hand, it is possible to envisage initial states which do not obey (2.10), and whose evolution would then require the presence of the terms on the right-hand side of (2.4d). We may expect, however, that in a time of order $\delta\sqrt{q}$, Alfvén waves would travel from the centre of disturbance leaving (2.10) obeyed in their wake. Thereafter (2.4) and (2.7) would be a predictive set, even though the inertial terms in (2.4d) were neglected.

We should emphasize that (2.10) is automatically obeyed for the types of disturbance we envisage below. The authors hope to consider more general solutions in which (2.10) is not automatically satisfied, in a subsequent publication.

3. Linearized theory

The general aim of the investigation is to determine the fate of a small centred disturbance made to the static equilibrium at a value of R slightly in excess of the critical value, $R_c(q, \lambda)$, below which all infinitesimal disturbances die out, and above which small perturbations in certain ranges of wavelength and wave direction are amplified. We wish to trace the evolution of this small disturbance beyond the range of validity of the linear equations. We also hope to decide whether a mean component of magnetic field will develop in the same direction as the applied field, $B_0\hat{\mathbf{y}}$ which might help to regenerate that field against ohmic dissipation if the external coils were removed.

The first step in this programme is that of investigating the linear stability properties of the static state. A number of such analyses in related geometries have been carried out by a number of authors, but we wish to call particular attention to the study made by Eltayeb (1972) of the model considered here; his paper also includes a useful list of references. To examine the stability problem, we neglect ϵ in (2.4) and assume that all components of θ, $\boldsymbol{V}$, $\boldsymbol{B}$ and Π are functions of z multiplied by

$$\mathrm{e}^{\mathrm{i}\omega t} E_1 \equiv \mathrm{e}^{\mathrm{i}lx+\mathrm{i}my+\mathrm{i}\omega t}, \tag{3.1}$$

where l and m are assigned constants and ω is, it transpires, determined by the dispersion relation

$$\lambda R m^2 k^2(k^2+1+\mathrm{i}\omega q) = m^4(k^2+1)(k^2+1+\mathrm{i}\omega)+\lambda^2(k^2+1+\mathrm{i}\omega q)^2(k^2+1+\mathrm{i}\omega), \tag{3.2}$$

where $k^2 = l^2+m^2$. Marginal stability for the given l and m is obtained by setting $\mathrm{Im}(\omega)$ zero. With q and λ fixed, this condition defines a discrete set of values of R of which we consider only the one $(n = 1)$ for which the corresponding value of $g\alpha\beta d^2/2\Omega\kappa\pi^2$ is least, for this mode is the most readily excited to convection as β is increased. This R depends on l and m, and we vary these until it takes it minimum value, $R_c(q, \lambda)$, which we call the critical value. If R is increased from zero with q and λ fixed, some disturbances will amplify in time when R exceeds R_c, and the static state is therefore necessarily unstable for $R > R_c$. It is found that there are two distinct cases to be considered.

In the first, the value of ω at marginal stability is zero, so that the instability at R_c $(= R_{ce}$, say) arises in the manner conventionally known as the exchange of stabilities. Although we also will use this terminology, it is here open to misinterpretation, for it will appear that the static state can (for some q and λ) lose its stability to finite amplitude disturbances at values of R smaller than R_c. Similar behaviour is known in examples of penetrative convection (see, for example, Moore & Weiss 1973). We find that there are two possibilities: first,

$$\text{if} \quad \lambda \leqslant 2/\sqrt{3}, \quad R_{ce} = 3\sqrt{3}, \tag{3.3a}$$

independently of λ. The corresponding values of l and m,

$$l^2 = 2-\lambda\sqrt{3}, \quad m^2 = \lambda\sqrt{3}, \tag{3.3b}$$

are non-zero, and for this reason we will refer to the convection as occurring in oblique rolls. For small λ, the rolls are almost parallel to the applied field. Secondly,

$$\text{if} \quad \lambda \geqslant 2/\sqrt{3}, \quad R_{ce} = (m^2+1)(m^2+2)/2\lambda. \tag{3.4a}$$

The corresponding value of l is zero, and for this reason we will refer to the convection as occurring in cross-waves, the word 'cross' emphasizing that the axes of the convection rolls are perpendicular to the applied field. Their wave-number, m, is the only positive root of

$$m^6 = 2\lambda^2(m^2+1). \tag{3.4b}$$

In the second main case, ω is non-zero in the marginal state, so that the instability at R_c ($= R_{co}$, say) arises through the mechanism conventionally known as overstability. Again there are two possibilities: first,

$$\text{if} \quad \lambda \leqslant 2/[(1+q)\sqrt{3}], \quad R_{co} = 6\sqrt{3}/q, \tag{3.5a}$$

again independently of λ. The corresponding values of l and m

$$l^2 = 2-\lambda(1+q)\sqrt{3}, \quad m^2 = \lambda(1+q)\sqrt{3}, \quad \omega^2 = 9(q^2-2)/q^2, \tag{3.5b}$$

are again non-zero, and the description 'oblique' is once more appropriate. As for the corresponding exchange mode (3.3b), k is $\sqrt{2}$. Secondly, overstable cross-waves ($l = 0$) can occur for some values of λ when $q > 1$. Defining

$$\lambda_0(q) = 2(q^2-1)^{\frac{3}{2}}(q+1)^{-\frac{1}{2}}(2q^2-1)^{-\frac{1}{2}}, \tag{3.6}$$

we find that

$$\text{if} \quad \lambda_0 \geqslant \lambda \geqslant 2/[(1+q)\sqrt{3}], \quad R_{co} = (m^2+1)(m^2+2)/[\lambda q(1+q)], \tag{3.7a}$$

the corresponding values of m and ω being given by

$$m^6 = 2\lambda^2(1+q)^2(m^2+1), \quad \omega^2 = (m^2+1)^2(2q^2-2-m^2)/(qm)^2. \tag{3.7b}$$

The upper bound λ_0 on λ is determined by (3.7b) and marks the vanishing of ω^2. Overstable modes do not exist for $\lambda > \lambda_0$ since, by (3.7b), the corresponding ω^2 are negative.

In this paper, we are mainly concerned with the nonlinear aspects of the marginal stability problem, and it is important to decide in the linear range whether, as R is increased, instability arises first through the exchange of stabilities or by overstability. It may be seen at once that, since ω^2 vanishes at $\lambda = \lambda_0$ for the overstable modes, and since R_{ce} is a strict minimum over all l and m for all modes with $\omega^2 = 0$, R_{co} cannot be less than R_{ce} when $\lambda = \lambda_0$. Comparison of (3.3a) and (3.5a) shows that $R_{ce} < R_{co}$ for all $q < 2$ for which oblique overstability occurs, and the same result in fact follows generally. Even when $q > 2$, there is a range, $\lambda > \lambda_E(q)$, in which $R_{ce} < R_{co}$. The bound $\lambda_E(q)$ may be determined by equating R_{ce} and R_{co}, and (3.4a) and (3.7a) show readily that $\lambda_E \to 3.273\ldots$ as $q \to \infty$, the corresponding value of R_c being $6.546\ldots$.

The salient features of the linear stability analysis are summarized in figure 1. It is interesting to note that the manifestation of the instability is not a continuous function of q and λ. At the boundary $\lambda = \lambda_E(q)$ between the overstable and exchange regions of figure 1, there are two neutrally stable modes for one of which $\omega = 0$ while for the other $\omega \neq 0$; the associated values of m^2 are different. At neighbouring points on either side of the curve, one or other of these modes becomes unstable first as R increases, the other remaining neutral or decaying.

No matter which mode is operative at the values of q, λ and R selected, the linearized solution of the governing equation takes the form

$$\boldsymbol{V}_{11} = \left[-\frac{iA}{k^2}\left\{l+\frac{\lambda}{m}(k^2+1+i\omega q)\right\}\sin z,\ -\frac{iA}{k^2}\left\{m-\frac{\lambda l}{m^2}(k^2+1+i\omega q)\right\}\sin z,\ A\cos z\right]E_1\,e^{i\omega t}, \tag{3.8a}$$

$$\boldsymbol{B}_{11} = \frac{im\,\boldsymbol{V}_{11}}{k^2+1+i\omega q}, \tag{3.8b}$$

$$\theta_{11} = -\frac{A\cos z}{k^2+1+i\omega}E_1\,e^{i\omega t}, \tag{3.8c}$$

where A is a constant at present undetermined, and the suffices 11 have been added to the variables in anticipation of the developments of the theory which follow.

Solutions for other values of n are similar to (3.8); for example, the $\cos z$ appearing in the expression for θ_{11} is replaced by $\cos nz$ if n is odd, or by $\sin nz$ if n is even. It is now clear that n is a number of free oscillations of the solution between the plates. Since, however, the critical value of $g\alpha\beta d^2/2\Omega\kappa\pi^2$ for given l, m, q and λ increases with increasing n, it is not necessary to consider modes for which $n \geqslant 2$.

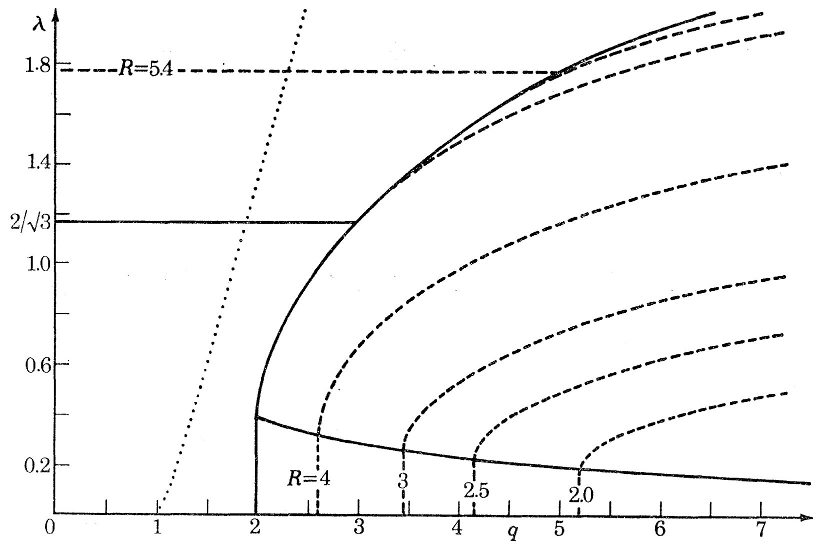

FIGURE 1. A summary of the linear stability results. Overstability is possible only to the right of the dotted curve, but is preferred only to the right of the full curve which starts at (2, 0) and asymptotes to $\lambda = 3.273...$ as $q \to \infty$. On the left of this curve, the preferred mode is transverse to $\boldsymbol{B}_0$ if $\lambda > 2/\sqrt{3}$ and is oblique otherwise. On the right of the curve, transverse modes are preferred above the nearly horizontal line that asymptotes to $\lambda = 0$ for $q \to \infty$; otherwise oblique modes are preferred. Some curves of constant Rayleigh number, R, are given (dashed).

4. Finite amplitude convection in the form of simple rolls

The analysis presented in the last section shows that, as R increases for fixed q and λ, the stability of the static conduction solution is lost at a critical value, R_c, of R which is given by either (3.3a), (3.4a), (3.5a) or (3.7a), depending on which of the four domains shown in figure 1 the point (q, λ) lies. At this value of R, the most unstable mode is given by (3.8). From a number of studies of the evolution of small centred disturbances in a marginally unstable state (see, for example, Stewartson & Stuart, 1971), we may expect a filtering process to ensue in which all waves, except those that are marginally unstable, decay leaving at large times a wave-packet consisting of the most unstable mode together with modes of neighbouring wave numbers. This wave-packet will travel with the appropriate group velocity, and will grow exponentially in central amplitude while simultaneously spreading out horizontally to distances proportional to $\sqrt{t}$. Nonlinear effects must eventually modify this picture and, although one can anticipate the general form which the disturbance is then likely to take, its detailed properties depend crucially on the values of certain coefficients which must be evaluated before further progress can be made. Our aim in this section is to derive some fundamental results of the nonlinear theory which are basic to the understanding of the evolution of the centred disturbance when nonlinear effects become significant. For this purpose, we temporarily abandon the notion of a centred disturbance, and suppose that the linear disturbance is given by (3.8), added to its complex conjugate.

There are two points that should be borne in mind at this juncture. First, although the preferred direction introduced by the field has destroyed the degeneracy of the field-free case, in which R_c depends on $k = \sqrt{(l^2+m^2)}$ but not on l/m, some indeterminancy remains for the oblique modes: there are two families of rolls for which R_c is the same, one in which $lm > 0$ and one in which $lm < 0$. Similarly, for every overstable oblique and cross-mode there are apparently two possible directions of wave motion, depending on the sign of (say) ωm. Thus, for example, we might expect that the temperature variation in the linear solution is given by

$$\theta_{11} = -\frac{2|A|\cos z}{[(k^2+1)^2+\omega^2]^{\frac{1}{2}}}\cos\left[\omega t + lx + my + \alpha_1\right], \tag{4.1}$$

with additional similar terms in which the sign of l and/or ω is opposite. (Here α_1 is a constant phase.) It may be seen, however, that the centred disturbance will give rise to four wave-packets, in the case of the oblique modes, or two in the case of the cross-modes, and that, by the time their amplitude has become so large that nonlinear terms are important, they have become so spatially separated that they may be treated independently. Thus we consider solutions of the form (4.1) in isolation. It should be recognized that this, in the oblique exchange case, represents a definite restriction on the class of initial disturbances considered.

Second, if $R-R_c$ is small, equations (3.8) formally rule out the slow rate of growth of the disturbance. This difficulty is avoided if we assume that A is a slowly varying function of t. Hitherto ϵ has been introduced only as a small parameter, but we now define it precisely by writing

$$\epsilon^2 = R - R_c, \tag{4.2}$$

and introduce a slow time-scale

$$\tau = \epsilon^2 t \tag{4.3}$$

for variations of A.

Following the method of Stuart (1958), we now consider how the nonlinear terms affect the evolution of A at large times, t. We assume that all dependent variables, for instance θ, can be expanded in the form

$$\theta = \theta_0(z,\tau,\epsilon) + \{\theta_1(z,\tau,\epsilon)\, E_1\, e^{i\omega t} + \theta_2(z,\tau,\epsilon)\, E_1^2\, e^{2i\omega t} + \ldots + \text{complex conjugate}\}. \tag{4.4}$$

We also assume that θ_n, and like components of the other dependent variables, can be expanded in powers of ϵ in the form

$$\left.\begin{aligned} \theta_0 &= \epsilon^2\theta_{02}(z,\tau) + \epsilon^3\theta_{03}(z,\tau) + \ldots, \\ \theta_1 &= \epsilon\theta_{11}(z,\tau) + \epsilon^2\theta_{12}(z,\tau) + \epsilon^3\theta_{13}(z,\tau) + \ldots, \\ \theta_2 &= \epsilon^2\theta_{22}(z,\tau) + \epsilon^3\theta_{23}(z,\tau) + \ldots. \end{aligned}\right\} \tag{4.5}$$

These forms, are now substituted into the governing equations (2.4) and the coefficients of the terms $\epsilon^n E_1^m$ ($n = 1, 2, \ldots$; $m = 0, 1, 2, \ldots$) are successively equated to zero. There results a set of equations for the unknown functions (θ_{mn}, for example) which may be solved seriatim. The full details of the analysis are too tedious for inclusion here; only the principal results are recorded.

The first equations arise from the coefficients of ϵE_1, and are identical to those of the linear theory of § 3. The appropriate solution is given by (3.8), with conjugate complex expressions added, and with A a function of τ alone. Continuing the expansion, the next equations to arise stem from the coefficients of $\epsilon^2 E_1^2$, and are merely algebraic leading to the solution

$$\left.\begin{aligned} \boldsymbol{V}_{22} &= -\frac{\mathrm{i}}{m}(2k^2+\mathrm{i}\omega q)\,\boldsymbol{B}_{22}, \\ \boldsymbol{B}_{22} &= \frac{\lambda q A^2\,\mathrm{e}^{2\mathrm{i}\omega t}}{2mk^2(k^2+1+\mathrm{i}\omega q)}\,(-m, l, 0), \\ \theta_{22} &= 0. \end{aligned}\right\} \tag{4.6}$$

From the coefficients of $\epsilon^2 E_1$ we simply obtain a solution proportional to that already derived from the coefficient of ϵE_1, the proportionality factor being a slowly varying function of time. We may, without loss of generality, absorb these $\epsilon^2 E_1$ terms into those of ϵE_1, and therefore set to zero all terms involving the pair 12 of suffices. From the coefficients of $\epsilon^2 E_1^0$ we obtain

$$\left.\begin{aligned} \boldsymbol{V}_{02} &= -\frac{2q\omega|A|^2\cos 2z}{k^2[(k^2+1)^2+\omega^2q^2]}(l, m, 0), \\ \boldsymbol{B}_{02} &= -\frac{mq(k^2+1)\,|A|^2\cos 2z}{k^2[(k^2+1)^2+\omega^2q^2]}\left[l+\frac{\lambda}{m}(k^2+1),\ m-\frac{\lambda l}{m^2}(k^2+1),\ 0\right], \\ \theta_{02} &= -\frac{(k^2+1)\,|A|^2\sin 2z}{2[(k^2+1)^2+\omega^2]}. \end{aligned}\right\} \tag{4.7}$$

It is now seen that both the condition (2.8) on the Lorentz force and the Taylor condition (2.10) are satisfied by (3.8).

To this point, the derivation of the various terms of (4.5) has been straightforward, but when we consider the coefficients of $\epsilon^3 E_1$, we obtain a set of equations in which the homogeneous terms (proportional to $\boldsymbol{V}_{13}$, $\boldsymbol{B}_{13}$ or θ_{13}) are linearly dependent, because they are identical to those used to obtain the linear solution (3.8). We obtain, in fact

$$\left.\begin{aligned} &\mathrm{i}m(\mathrm{curl}\,\boldsymbol{B}_{13})_z+\lambda\,\partial V_{13z}/\partial z = Q_1\,\mathrm{e}^{\mathrm{i}\omega t}\sin z+\tilde{Q}_1\,\mathrm{e}^{\mathrm{i}\omega t}\sin 3z, \\ &\mathrm{i}m(k^2+1)\,B_{13z}+\lambda\,\partial(\mathrm{curl}\,\boldsymbol{V}_{13})_z/\partial z-\lambda k^2 R_c\,\theta_{13} = Q_2\,\mathrm{e}^{\mathrm{i}\omega t}\cos z+\tilde{Q}_2\,\mathrm{e}^{\mathrm{i}\omega t}\cos 3z, \\ &(k^2+1+\mathrm{i}\omega)\,\theta_{13}+V_{13z} = Q_3\,\mathrm{e}^{\mathrm{i}\omega t}\cos z+\tilde{Q}_3\,\mathrm{e}^{\mathrm{i}\omega t}\cos 3z, \\ &(k^2+1+\mathrm{i}\omega q)\,B_{13z}-\mathrm{i}mV_{13z} = \mathrm{i}Q_4\,\mathrm{e}^{\mathrm{i}\omega t}\cos z+\mathrm{i}\tilde{Q}_4\,\mathrm{e}^{\mathrm{i}\omega t}\cos 3z, \\ &(k^2+1+\mathrm{i}\omega q)\,(\mathrm{curl}\,\boldsymbol{B}_{13})_z-\mathrm{i}m(\mathrm{curl}\,\boldsymbol{V}_{13})_z = \mathrm{i}Q_5\,\mathrm{e}^{\mathrm{i}\omega t}\sin z+\mathrm{i}\tilde{Q}_5\,\mathrm{e}^{\mathrm{i}\omega t}\sin 3z, \end{aligned}\right\} \tag{4.8}$$

where

$$\left.\begin{aligned} Q_1 &= \frac{\lambda q^2|A|^2 A}{[(k^2+1)^2+\omega^2q^2]}\left[1-\frac{(k^2+1)\,(3k^2+3+\mathrm{i}\omega q)}{2(k^2+1+\mathrm{i}\omega q)}\right], \\ Q_2 &= -\frac{\lambda k^2 A}{(k^2+1+\mathrm{i}\omega)}-\frac{q^2m^2(k^2+1)\,(k^2-3)\,|A|^2A}{2[k^2+1+\mathrm{i}\omega q]\,[(k^2+1)^2+\omega^2q^2]}, \\ Q_3 &= \frac{1}{(k^2+1+\mathrm{i}\omega)}\frac{\mathrm{d}A}{\mathrm{d}\tau}+\left[\frac{k^2+1}{2[(k^2+1)^2+\omega^2]}-\frac{\mathrm{i}\omega q^2}{(k^2+1+\mathrm{i}\omega)\,[(k^2+1)^2+\omega^2q^2]}\right]|A|^2A, \\ Q_4 &= -\frac{mq}{(k^2+1+\mathrm{i}\omega q)}\frac{\mathrm{d}A}{\mathrm{d}\tau}-\frac{q^2m[(k^2+1)^2+\mathrm{i}\omega q(k^2-1)]}{2(k^2+1+\mathrm{i}\omega q)\,[(k^2+1)^2+\omega^2q^2]}|A|^2A, \\ Q_5 &= \frac{\lambda q}{m}\frac{\mathrm{d}A}{\mathrm{d}\tau}+\frac{\lambda q^2[k^4+8k^2+3-\mathrm{i}\omega q(k^2-1)]}{2m[(k^2+1)^2+\omega^2q^2]}|A|^2A, \end{aligned}\right\} \tag{4.9}$$

with similar expressions for $\tilde{Q}_1$–$\tilde{Q}_5$, which, however, are not required, and are therefore not given.

By linearly combining the terms on the left-hand sides of equations (4.8), we see that they are soluble only if

$$\frac{\lambda}{m^2}(k^2+1+\mathrm{i}\omega q)\,Q_1+Q_2+\frac{\lambda k^2 R_c\,Q_3}{(k^2+1+\mathrm{i}\omega)}+\frac{m(k^2+1)\,Q_4}{(k^2+1+\mathrm{i}\omega q)}+\frac{\lambda}{m}Q_5=0. \tag{4.10}$$

This consistency relation implies that A must satisfy a first-order differential equation, which therefore controls the evolution of the nonlinear modal disturbance. Such equations have, of course, been of common occurrence since the advent of the Landau–Stuart theory (see, for example Stuart 1958). The coefficients of the various terms depend on which of the four regions of figure 1 in which (q, λ) lies. We shall consider these four cases in the next two sections.

5. Exchange instability

The linear theory of § 2 has shown that, when $q < 2$ or when $q > 2$ and $\lambda > \lambda_E(q)$, marginally unstable disturbances are non-oscillating, or – as it is often said – the principle of the exchange of stabilities is valid. Now (4.10) can be written as

$$\mathrm{d}A/\mathrm{d}\tau = d_1 A + k_1|A|^2 A, \tag{5.1}$$

where we have used the notation for the coefficients of d_1 and k_1 which is usually adopted. In the exchange case, these quantities are real, and

$$\text{if} \quad \lambda \leqslant 2/\sqrt{3}, \quad d_1 = 1/\sqrt{3}, \quad k_1 = (2q^2-9)/18, \tag{5.2a}$$

$$\text{if} \quad \lambda \geqslant 2/\sqrt{3} \quad \left\{\begin{aligned} d_1 &= 2\lambda[m^2+2+q(m^2-2)]^{-1},\\ k_1 &= \frac{2q^2(m^6-m^2-2)-(m^2+1)^2(m^2+2)}{2(m^2+1)^2[m^2+2+q(m^2-2)]}, \end{aligned}\right\} \tag{5.2b}$$

where m is the only positive root of (3.4b).

In order that the solution to (5.1), valid for $\tau > 0$, should match with the linearized solution valid for $t \to \infty$, we require that A should be equal to some constant A_0 (say) at $\tau = 0$. The differential equation is readily solved and, provided $k_1 < 0$, we find that

$$|A_1| \to (-d_1/k_1)^{\frac{1}{2}}, \quad \tau \to \infty, \tag{5.3}$$

irrespective of the value of A_0. This behaviour is often described by saying that the static state is unstable to spectral evolution. The alternative behaviour, that arises when $k_1 > 0$, leads to an infinite $|A|$ at a finite value of t, and is often called catastrophic instability. According to (5.2), the dividing case, $k_1 = 0$, is given by $q = 3/\sqrt{2}$ if $\lambda < 2/\sqrt{3}$, and by $\lambda = \lambda_D(q)$ if $\lambda > 2/\sqrt{3}$, where $\lambda_D(q)$ is obtained by eliminating m between (3.4b) and

$$q^2 = \frac{(m^2+1)^2(m^2+2)}{2(m^6-m^2-2)}. \tag{5.4}$$

The curve $\lambda = \lambda_D(q)$ is shown in figure 2. To its left, spectral evolution occurs, while to its right the instability is catastrophic.

When $k_1 > 0$, finite amplitude subcritical solutions occur, which are unstable in the sense that an initial disturbance of smaller amplitude reverts to the static conduction solution, while one of slightly larger amplitude amplifies catastrophically. This is often taken to mean that, at even larger amplitudes, stable steady nonlinear solutions exist. If we tentatively accept the existence of these solutions, we see from figure 2 that they may well exist in regions of the $q\lambda$-plane;

for instance, those of large q and $\lambda < 3.273$, in which the linear theory indicates that over-stability is preferred. We will encounter a similar paradox when studying the overstable modes in § 6. Clearly, then, the initial state may, at sufficiently large q, follow very different courses depending on whether it is of small or large amplitude. There is, of course, no possibility within the framework of the present theory of discovering the fate of large amplitude disturbances, or even locating the conjectured stable subcritical finite amplitude solutions. We will, therefore, consistently take the view that the initial disturbance is small, while recognizing that this assumption, though leading to a consistent picture, may not be representative of the behaviour of the system when $k_1 > 0$ if the initial disturbance is of arbitrary amplitude.

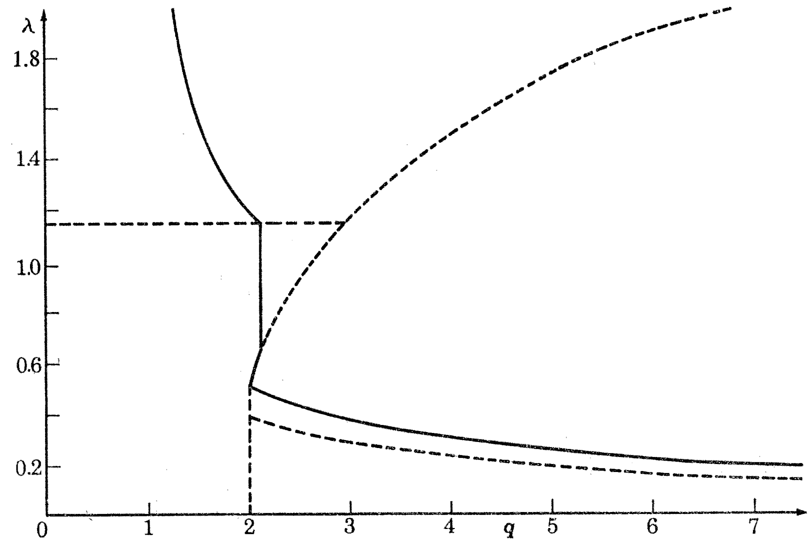

FIGURE 2. A summary of some nonlinear stability results. The full line $\lambda = \lambda_D(q)$ is that on which either $k_1 = 0$ or Re $k_1 = 0$, depending on whether the curve lies in the stable or overstable regions (shown dashed for ease of reference with figure 1).

The case $k_1 > 0$ also differs from the case $k_1 < 0$ in respect of the importance of spatial modulation. We consider the evolution of the infinitesimal centred disturbance at marginally unstable values of R. We suppose that we are in a domain of the $q\lambda$-plane in which linear instability first appears through the exchange of stabilities. An analysis parallel to that carried out by Stewartson & Stuart (1971) for plane Poiseuille flow shows that the temperature, θ, takes the form

$$\theta \sim \mathrm{Re}\frac{\delta_1 E_1}{t}\cos z \exp\left\{d_1(R-R_c)\,t - \frac{(bx^2-2hxy+ay^2)}{4(ab-h^2)\,t}\right\}[1+O(t^{-\frac{1}{2}})]$$
$$+\mathrm{Re}\frac{\delta_2 E_2}{t}\cos z \exp\left\{d_1(R-R_c)\,t - \frac{(bx^2+2hxy+ay^2)}{4(ab-h^2)\,t}\right\}[1+O(t^{-\frac{1}{2}})], \quad (5.5)$$

as $t \to \infty$, with $\boldsymbol{V}$ and $\boldsymbol{B}$ being given similarly. Here

$$E_2 = \mathrm{e}^{\mathrm{i}lx - \mathrm{i}my}, \quad (5.6)$$

l and m are real and positive, and are given by (3.3 b) when $\lambda < 2/\surd 3$; otherwise, $l = 0$ and m is given by (3.4 b). In this case, the second term of (5.5) does not appear. The constants a, b, h and d_1 are given by series expansion of $\mathrm{i}\omega$, as given by (3.2), about $R = R_c$, $l = l_c$ and $m = m_c$:

$$\mathrm{i}\omega = d_1(R-R_c) - a(l-l_c)^2 - b(m-m_c)^2 - 2h(l-l_c)(m-m_c). \quad (5.7)$$

Clearly $ab \geqslant h^2$, by the condition that $\mathrm{i}\omega$ is a maximum at the critical state. In the present case, we have

$$\text{if} \quad \lambda < 2/\sqrt{3} \quad \left\{\begin{aligned} a &= 2(2-\lambda\sqrt{3})/3, \\ b &= 2(\lambda^2-\lambda\sqrt{3}+3)/\lambda\sqrt{3}, \\ h &= (2-\lambda\sqrt{3})^{\frac{1}{2}}(2\lambda-\sqrt{3})/\lambda^{\frac{1}{2}}3^{\frac{1}{4}}, \end{aligned}\right\} \tag{5.8a}$$

$$\text{if} \quad \lambda > 2/\sqrt{3} \quad \left\{\begin{aligned} a &= (m^2+1)(m^2-2)/m^2[m^2+2+q(m^2-2)], \\ b &= 4(2m^2+3)/(m^2+1)[m^2+2+q(m^2-2)], \\ h &= 0, \end{aligned}\right\} \tag{5.8b}$$

where m in (5.8b) is given by (3.4b). Finally, δ_1 and δ_2 are constants of order 1 which depend on the precise specification of the initial centred disturbance.

The solution (5.5) clearly has the form of two wave-packets, each consisting of a family of simple rolls whose amplitude is a maximum at the origin and decreases over a distance of order $\sqrt{t}$ to exponentially smaller values. As we have already stated, we have elected to consider for $\lambda < 2/\sqrt{3}$, only those initial states for which $\delta_2 = 0$. We see that, although (2.8) is obeyed, Taylor's condition (2.10) is violated when $\delta_1\delta_2 \neq 0$. This means that the evolution of the double roll cannot, apparently, be treated in the framework of the present analysis. We will discuss this case in a subsequent paper.

It is possible to generalize (5.1) at once to the case of spatial modulation by following a line of reasoning similar to that taken by Stewartson & Stuart (1971) for plane Poiseuille flow. We introduce new variables

$$\xi = \epsilon x, \quad \eta = \epsilon y, \tag{5.9}$$

and allow A to be functions of ξ and η as well as τ. Then A satisfies the partial differential equation

$$\frac{\partial A}{\partial \tau} - a\frac{\partial^2 A}{\partial \xi^2} - b\frac{\partial^2 A}{\partial \eta^2} - 2h\frac{\partial^2 A}{\partial \xi\,\partial \eta} = d_1 A + k_1|A|^2 A. \tag{5.10}$$

As explained by Hocking, Stewartson & Stuart (1972), the extra terms in (5.10) arise because, while A is small, the nonlinear term can be neglected and A will therefore evolve into a packet composed of all waves in the neighbourhood of the critical rolls, specifically those in which $l-l_c$ and $m-m_c$ are both of order ϵ. For this to be possible, the linear terms of any differential equation satisfied by A must lead to a dispersion relationship identical with (5.7). The derivatives of A which must appear in (5.10) can therefore be written down at once. Higher derivatives with respect to ξ or η in (5.10) correspond to higher powers of $l-l_c$ and $m-m_c$ in the expansion of $\mathrm{i}\omega$ from (3.2) and hence have coefficients of order ϵ, which is negligible. The nonlinear term in (5.10) must be the same as that of (5.1), because it is already of the same order as the linear terms, and any differentiation with respect to ξ on η would make contributions of relative order ϵ, which could again be neglected. This argument would require modification if there were a mean flow of order ϵ^2 generated by the linear terms, i.e. if V_{02} contained a component whose mean value with respect to z were non-zero (see Davey, Hocking & Stewartson 1974). Here, however, $V_{02} = 0$ since $\omega = 0$, and the problem does not arise. We will also find (§ 6) that, even when $\omega \neq 0$, the mean value of V_{02} is zero and the difficulty is not encountered.

The appropriate initial and boundary conditions satisfied by A are

$$\left.\begin{aligned} A \sim \frac{\delta_3}{\tau}\exp\left[-\frac{(b\xi^2-2h\xi\eta+a\eta^2)}{4(ab-h^2)\,\tau}\right], \quad \text{as} \quad \tau \to 0, \\ |A| \to 0, \quad \xi^2+\eta^2 \to \infty, \end{aligned}\right\} \tag{5.11}$$

where δ_3 is a small constant.

Extensive studies have been made of solutions of (5.10) and (5.11); see Hocking *et al.* (1972), Hocking & Stewartson (1971, 1972), Zakharov & Shabat (1972). These have revealed a number of the properties of A. We infer, in the present case of real coefficients, that, if $k_1 < 0$, $|A|$ will tend to the constant limiting value (5.3) for all ξ and η as $\tau \to \infty$. This shows that the classic expectations of the principle of the exchange of stabilities are upheld when $k_1 < 0$, and that the stability of the static solutions is lost to that of a more complicated dynamical state in which A is constant, even though the initial disturbance was centred.

When $k_1 > 0$, the situation is very different, and the breakdown of the solution of (5.10) and (5.11) as $t \to \infty$ is not at all the same as that of the solution of (5.1). In particular, it is not true that $|A| \to \infty$ for all ξ and η at some finite value of τ, as would be the case if A were initially independent of ξ and η. The singularity in the solution occurs at one point only, namely $\xi = \eta = 0$ for the initial condition (5.11). It should, of course, be remembered that the present theory cannot be extended right up to $|A| = \infty$, for the theory is only valid for $\epsilon \ll 1$ and A finite, but it nevertheless seems likely that the phenomenon reveals a substantial change in the character of the flow and may even presage the onset of turbulence.

6. Overstability

The linear theory of § 3 has shown that, when $q > 2$ and $\lambda < \lambda_E(q)$, marginally unstable disturbances are oscillatory or – as it is often said – convection arises first as overstability. Again an equation of the form (5.1) is obtained, but now the constants d_1 and k_1 are complex:

$$\text{if} \quad \lambda \leqslant 2/(1+q)\sqrt{3} \quad \begin{cases} d_1 = \dfrac{1}{2\sqrt{3}}\left[1-\dfrac{3\mathrm{i}}{\omega_c}\right], \\ k_1 = -\dfrac{1}{36(q-1)(q+1)^2}\left[q(3+5q)+\dfrac{3\mathrm{i}}{\omega_c}(16+15q-9q^2-6q^3)\right]; \end{cases}$$

if $\lambda_E(q) \geqslant \lambda \geqslant 2/(1+q)\sqrt{3}$

$$d_1 = \frac{m^3(1+q)}{[3m^2-2+q(m^2+2)]\sqrt{[2(m^2+1)]}}\left[1-\frac{\mathrm{i}(m^2+1)}{\omega_c q m^2}(m^2-2+2q)\right], \tag{6.1a}$$

$$\begin{aligned} k_1 = {} & \frac{m^2}{4(m^2+1)^2(q^2-1)[3m^2-2+q(m^2+2)]} \\ & \times [q\{2m^6-3m^4-3m^2-4+q(2m^6-m^4-9m^2-4)\}-\mathrm{i}(m^2+1) \\ & \times \{2m^8+m^6-m^4+10m^2+8+q(2m^8+m^6+3m^4+2m^2+4) \\ & +2q^2(m^2+1)(2m^4-5m^2-4)-q^3(m^2+1)(m^2+2)\}/\omega_c m^2]. \end{aligned} \tag{6.1b}$$

We see that, when the instability is in oblique rolls, the real part of k_1 is necessarily negative, A remains bounded for all t, and, as $t \to \infty$,

$$A \to \left(-\frac{d_{1r}}{k_{1r}}\right)^{\frac{1}{2}} \exp\left\{\mathrm{i}\left(d_{1i}-\frac{k_{1i}d_{1r}}{k_{1r}}\right)t+\text{constant}\right\}, \tag{6.2}$$

where the real and imaginary parts of k_1 have been written k_{1r} and k_{1i}, respectively, and similarly for d_1. According to (6.2) there is a sense in which exchange of stabilities occurs at $R = R_c$, but the new solution is unsteady.

In the case of the cross-modes it is, according to (6.1), not necessarily true that the real part of

k_1 is negative. Indeed, another branch of the curve $\lambda = \lambda_D(q)$ may be obtained by eliminating m between the first of (3.7 *b*) and

$$q = -\frac{(2m^6-3m^4-3m^2-4)}{(2m^6-m^4-9m^2-4)}. \tag{6.3}$$

This branch is also shown in figure 2. Below the curve $\lambda = \lambda_D(q)$ spectral evolution occurs, while above the instability is catastrophic.

These solutions may also be generalized to include initially centred infinitesimal disturbances. Assuming that the compatibility condition (2.10) is satisfied, the evolutionary process filters out, in a time $t = O(1)$, all but the most unstable oscillations of the linearized theory of § 2. When $t \gg 1$, but $\tau = \epsilon^2 t \ll 1$ (so that nonlinear terms have not had sufficient time to act significantly), the disturbance has become concentrated into a number of wave-packets. If (3.5 *a*) holds, there are four such packets, centred in the wave numbers $l = +l_c$ and $m = \pm m_c$ where, according to (3.5 *b*),

$$l_c = [2-\lambda(1+q)\sqrt{3}]^{\frac{1}{2}}, \quad m_c = [\lambda(1+q)\sqrt{3}]^{\frac{1}{2}}, \tag{6.4}$$

the corresponding frequency being $\omega_c = 3(q^2-2)^{\frac{1}{2}}/q$. In the wave-packet associated with $+l_c$ and $+m_c$, θ takes the form

$$\theta \sim \mathrm{Re}\left\{\frac{\delta_1}{t}\cos z\, e^{\Theta_1}[1+O(t^{-\frac{1}{2}})]\right\}, \tag{6.5}$$

for large times, where

$$\Theta_1 = i\omega_c t + d_1(R-R_c)\,t + il_c x + im_c y - \frac{b(x+u_x t)^2 - 2h(x+u_x t)\,(y+u_y t) + a(y+u_y t)^2}{4(ab-h^2)\,t}, \tag{6.6}$$

and δ_1 is determined by the initial disturbance. The dispersion relationship relating ω, l and m is given by

$$i(\omega-\omega_c) = d_1(R-R_c) + iu_x(l-l_c) + iu_y(m-m_c) \\ -a(l-l_c)^2 - 2h(l-l_c)\,(m-m_c) - b(m-m_c)^2 + \dots, \tag{6.7}$$

in the neighbourhood of $(l, m) = (l_c, m_c)$. The new terms appearing in (6.7) that were not present in (5.7) reflect the fact that variations in l and m, about their critical values for the overstable modes, alter the real part of ω as well as the imaginary part, and we have chosen R_c so that at the critical value of l and m the imaginary part of ω is zero, and a maximum. It follows that u_x and u_y are real. They are given by

$$u_x = 3(q^2-3)\,l_c/\omega_c q^2, \quad u_y = 3[m_c^2(q^2-3)+6(q^2-1)]/\omega_c q^2 m_c, \tag{6.8}$$

for $\lambda \leqslant 2/[(1+q)\sqrt{3}]$. It can readily be seen from (6.6) that the centre of the wave-packet (6.5) moves from the origin with the group velocity $(-u_x, -u_y, 0)$.

The fate of the three other wave-packets is similar, except for the change(s) in sign of l_c and/or m_c and hence, in virtue of (6.8), in direction of motion. At large times, each wave-packet is well separated from the others, and its interaction with them is formally negligible. There is therefore no need to restrict ourselves to the equivalent of $\delta_1\delta_2 = 0$, as was done in § 5.

If $\lambda_E \geqslant \lambda \geqslant 2/[(1+q)\sqrt{3}]$, the critical value of l is zero, while those of R_c, m_c and ω_c are given by (3.7); also

$$u_x = 0, \quad u_y = 2(m_c^2+1)\,[(3m_c^2+2)\,(q^2-1)-m_c^4]/\omega_c q^2 m_c^4. \tag{6.9}$$

Now there are only two wave-packets but they travel in opposite directions, parallel and antiparallel to the y-axis, and therefore their interaction at large times can again be neglected.

A parallel analysis to that of Davey *et al.* (1974) now shows that, since there is no mean flow to order ϵ^2, the nonlinear evolution of the wave-packet, defined by $+l_c$ and $+m_c$, is given by (5.10) except that the definition (5.9) must be replaced by

$$\xi = \epsilon(x+u_x t), \quad \eta = \epsilon(y+u_y t). \tag{6.10}$$

All the coefficients appearing in (5.10) are now complex. The properties of the solutions of (5.10) have been extensively, but not completely, discussed by Hocking & Stewartson (1971, 1972). Two broad classes of disturbances were studied, the quasi two-dimensional and the three-dimensional.

For the first of these, which we shall call 'skewed-plane solutions', it is supposed that A is a function of $\bar{\xi} = \beta_1\xi+\beta_2\eta$ and τ alone, where β_1 and β_2 are real constants, and it is required that $|A| \to 0$ as $\bar{\xi} \to \pm\infty$. The governing equation for A now reduces to

$$\frac{\partial A}{\partial \tau} - \bar{a}\frac{\partial^2 A}{\partial \bar{\xi}^2} = d_1 A + k_1|A|^2 A, \tag{6.11}$$

where $\bar{a} = a\beta_1^2 + b\beta_2^2 + 2h\beta_1\beta_2$. A full discussion of the fate of A as τ increases in the case when $k_{1r} \equiv \operatorname{Re} k_1 > 0$ has been given by Hocking & Stewartson (1972). They established that A becomes infinite in a centred burst at some finite time provided $\bar{a}$ and k satisfied certain conditions of which the most important are that one or both of the inequalities

$$3k_{1r}^2 - 4\bar{a}_i k_{1r} k_{1i}/\bar{a}_r - k_{1i}^2 > 0, \tag{6.12a}$$

$$(9+8\bar{a}_i^2/\bar{a}_r^2)\, k_{1r}^2 + 2\bar{a}_i k_{1r} k_{1i}/\bar{a}_r - k_{1i}^2 > 0, \tag{6.12b}$$

are obeyed. Except for a small domain, A remains finite for other values of $\bar{a}$ and k_1, and varies in a quasi-random way while spreading out in the $\bar{\xi}$ direction.

Hocking & Stewartson (1972) confined their discussion of the case $k_{1r} < 0$ to real values of $\bar{a}$. It appeared that $|A|$ then tends to a limit independent of $\bar{\xi}$ provided k_{1i}/k_{1r} is not too large. The special case of dissipationless problems, for which $a_r = b_r = h_r = d_{1r} = k_{1r} = 0$, has received extensive attention in connexion with a wide variety of situations arising in fluid mechanics and plasma physics. It is known (see, for example, Hasimoto & Ono 1972) that the uniform limit for the skewed-plane solutions is stable to long wave disturbances only if $\bar{a}_i k_{1i} > 0$. This result may be generalized. It may be shown that the conditions

$$\bar{a}_r k_{1r} + \bar{a}_i k_{1i} < 0, \quad k_{1r} < 0, \quad d_{1r} > 0, \tag{6.13a, b, c}$$

are sufficient to ensure the stability of the $\bar{\xi}$-independent solution, $|A|^2 = -d_1/k_{1r}$. Numerical computations by Karpman & Krushkal (1968) have shown that the dissipationless solution is unstable when $\bar{a}_i k_{1i} < 0$, and breaks up into solitons. There is still little evidence to decide how far (6.13) is sufficient to prevent the quasi-random behaviour reported by Hocking & Stewartson (1972) for cases in which $\bar{a}_i = 0$, $k_{1r} < 0$ and $|k_{1i}/k_{1r}|$ is large. Studies by Hayes (1973) and by Davey & Stewartson (1974) show that, if all skewed-plane solutions are stable, then so are the centred disturbances.

In the second class of solutions to which we referred above, A is required to vanish as $\xi^2+\eta^2 \to \infty$ for all values of ξ/η. Being truly centred, these are of the greatest relevance in our case. Unfortunately studies of this class have not been pursued to the same depth as for the skewed-plane solutions. If $k_{1r} > 0$, it has been shown that, provided $|k_{1i}/k_{1r}|$ is not too large, centred solutions will burst at one point at a finite value of τ; if $|k_{1i}/k_{1r}|$ is sufficiently large, the solutions that remain finite at all times can be computed. No information is available for the case $k_{1r} < 0$. We can anticipate that the results for the skewed-plane solutions may be paralleled by similar behaviour of the centred disturbances. Extensive numerical work will, however, be required before the boundaries of the domains of bursting can be located as precisely in parameter space.

In applying the results of the general theory to our present problem, it is clearly of importance to have values of a, b and h available when required. For any particular pair of values of q and λ, this is not difficult to achieve, one viable method being to solve (3.2) for $R = R_c$ and (l, m) in the neighbourhood of (l_c, m_c), and to obtain a, b and h by differencing. The derivation of formal algebraic expressions that are valid for all q and λ would be tedious, and has not been attempted. It is, however, relatively easy to find the leading terms in their expansions for large q. It is found that ia, ib and ih approach finite real values for all λ in the relevant range $0 < \lambda < \lambda_E$. Although the limiting values of a_i and b_i are positive, it may be shown that $a_i b_i < h_i^2$, if $\lambda q < \frac{6}{31}\sqrt{3}$. This implies that, for two values of β_1/β_2, the ratio $\bar{a}_r/\bar{a}_i$ is infinite.

The following inferences about the solutions of (5.10) are now seen to be in order. If

$$\lambda < 2/(1+q)\sqrt{3},$$

k_{1r} is negative for all q and λ, and as for the modal disturbances A remains finite for all τ. Given any fixed q and λ, it is not possible to be certain whether $|A|$ approaches a limit as $\tau \to \infty$, without determining the values of a, b and h and, in the event that (5.13) is obeyed, without carrying out a numerical integration of (5.10). At large values of q, however, we know from the results stated in the previous paragraph that a_i, b_i and h_i are such that $\bar{a}_i$ is finite and positive for all β_1/β_2 provided $(6\sqrt{3})/31q < \lambda < \lambda_E$. Hence (6.12) is violated when q is large, and we may conclude that there is a substantial range of values of q such that $|A|$ does not approach a limit as $\tau \to \infty$ neither in the skewed-plane nor in the centred cases. If $\lambda q < \frac{6}{13}\sqrt{3}$, there is a range of values of β_1/β_2 such that $\bar{a}_i < 0$, and this means that we cannot rule out the existence of a limit for $|A|$ as $\tau \to \infty$ for some skewed-plane solutions. Nevertheless $|A|$ for the centred disturbances still does not approach a limit as $\tau \to \infty$, since stability for centred disturbances requires all skewed-plane disturbances to be stable.

Suppose $2/(1+q)\sqrt{3} < \lambda < \lambda_E$, and let $q_D(\lambda)$ be the inverse of $\lambda_D(q)$. At the point (q_D, λ) in the (q, λ) plane, k_r is zero and therefore $|\delta_i| = |k_i/k_r| = \infty$. In general, both inequalities (6.12) are violated. Hence the solution remains finite at all times, but varies in a quasi-random way. If q is decreased with λ held fixed, either A behaves quasi-randomly or $|A|$ approaches a limit. It is plausible that the former occurs for larger q and the latter for smaller q. Since $k_i/k_r = O(q^{-\frac{1}{2}}) \to 0$ as $q \to \infty$, bursting must eventually occur if q is increased with λ held fixed. It is plausible that a transition point exists in $q_D < q < \infty$ which separates quasi-random from bursting solutions. These conclusions apply to skewed-plane disturbances, but the centred disturbances must behave similarly, as was demonstrated by the analytical work reported by Hocking & Stewartson (1971).

We may summarize the fate of A for unstable disturbances in which the most rapidly increasing in amplitude are oscillatory as follows. Such disturbances are only relevant if $q > 2$ and $\lambda < \lambda_E$ and when q lies in some neighbourhood of $q = 2$, $|A|$ tends to a limit as $\tau \to \infty$ so that there is in some sense an exchange of stabilities although the new state is not *steady*. As q increases for fixed λ, A begins to take on a quasi-random behaviour while remaining finite, and at larger values of q its fate is to burst. The bursting can only be avoided as q increases if λ simultaneously $\to 0$ so that λq remains finite.

7. Discussion

Our objective in this paper has been to gain insight into the character of convection in a highly rotating, highly magnetic, system of a kind that abounds in the cosmos. We have been particularly interested in continuing the M.A.C.-wave theory of Braginsky (1964*c*, 1967). Although Braginsky (1967) was able to develop a very general formalism, that in principle recognized the approximately spherical structure of astrophysical applications, he investigated (Braginsky 1964*c*) in greatest depth a planar model which represented convection in the mid-latitudes of a spherical system by a horizontal layer in a uniform gravitational field; the predominantly zonal field of the spherical system was mimicked by a constant horizontal field, $B_0\hat{\mathbf{y}}$, and only the Coriolis forces arising from the component, $\Omega\hat{\mathbf{z}}$, of angular velocity perpendicular to the layer was recognized. The surface conditions necessary to complete the specification of the eigenvalue problem were considered to have at most a modest qualitative effect on the spectrum, and discussion was confined to the analytically simplest cases; for example, when viscosity was included, stress-free boundaries were postulated.

Further aspects of this specific model have been studied in this paper, for it provides an almost ideal medium by which to explore questions raised by Braginsky's work, one of which provided our motivation (see the introduction). These have led us to retain the effects of ohmic and thermal dissipation, and to study solutions whose character is decided by three dimensionless parameters R, λ and q defined in (2.5). The quantity λ is a rough measure of the relative importance of Coriolis and Lorentz forces, large values of λ corresponding to the rotationally dominant case and small values of λ to the magnetically dominant one. The quantity q measures the importance of thermal diffusion relative to ohmic diffusion, small values of q being relevant to a system such as the Earth's core which is resistively dominated, while large values of q are of interest in systems like the Sun and stars in which radiative diffusion is large. The parameter R has the nature of a Rayleigh number, and the linear stability of the system depends on whether R is greater or less than a certain critical value, $R_c(q, \lambda)$.

Braginsky placed his main emphasis on dissipationless systems in which therefore the quantity $R_B = R\lambda q$, independent of η and κ, played the part of the R of our theory. Since $\lambda \to 0$ as diffusion effects are ignored (i.e. as $\sigma \to \infty$), it might at first sight seem that Braginsky's results could be readily obtained as a special case of our own. That the limit is not completely straightforward may be appreciated by a simple observation: our results depend on the order in which the limits $\eta \to 0$ and $\kappa \to 0$ are taken, or more generally on the ratio $q = \kappa/\eta$ as either η or κ vanish. Such a situation is, of course, familiar in magnetohydrodynamics (see, for example, Stewartson 1960), but it is one that is obviously excluded from a theory that discards η and κ from the outset. Moreover, for fixed q, the marginal state of the dissipationless theory occurs at a particular value of R_B that is at an infinite value of R in the limit $\lambda \to 0$. The modes given by our theory, which involves finite values of R, must therefore (in the limit $\lambda \to 0$) differ from, and be subcritical with respect to, the modes of most interest in the diffusionless theory. This observation is related to the difficulty emphasized in § 1. The diffusionless theory always indicates that the most unstable modes are ones of infinite l; our theory in contrast predicts that l_c is finite in the marginal state.

The failure of the dissipationless theory to recognize q seems all the more serious when the dramatic dependence of our results on that parameter is recalled. For the geophysically interesting case of small q, linear theory predicts that the principle of the exchange of stabilities is valid. For the rotationally dominant case of large λ, the critical mode is one of convection in rolls, of small horizontal scale, transverse to the applied field. For small λ, the planform is, in general, the rectangular pattern implied by two families of rolls at the same oblique angle to $\boldsymbol{B}_0$. In the astrophysically interesting case of large q, the linear theory predicts that (unless λ exceeds 3.273...) overstability will occur at criticality. For the rotationally dominant case of large λq, the critical mode consists of travelling waves transverse to $B_0\hat{\boldsymbol{y}}$, and moving in the positive and negative y directions with a phase velocity of $2^{\frac{1}{2}}\pi\kappa/d$, their group velocities being 3 times greater. For small λq, the disturbance consists of four waves, one pair travelling in opposite senses along one direction inclined to $\boldsymbol{B}_0$ and the second pair along the other direction equally inclined to $\boldsymbol{B}_0$. Again the phase and group velocities are of order $\pi\kappa/d$.

A main conclusion of Braginsky's theory, that for slightly supercritical values of R_{B} the modes with $l \gg 1$ are the most unstable, strongly suggests that the controlling mechanism of the instability growth is dissipative in some sense and seriously limits the value of a dissipationless theory of its nonlinear development, quite apart from the mathematical difficulties inherent in such a study. By contrast the linear analysis of the dissipative system, which followed Eltayeb (1972), is immediately capable of extension to include nonlinear effects in marginally supercritical regimes. Excluding only the case of mixed oblique modes that can occur at small values of λ, we have shown that the amplitude A of the convective motions is governed by (5.10), an equation of a type that occurs with some frequency in non-linear stability theory. The constant k_1 is of particular significance even to the linear stability theory. For, when k_1 is negative as in the geophysically interesting case of small q, we may believe that the condition $R < R_{\mathrm{c}}$ is both necessary *and* sufficient for stability, and that (5.10) correctly gives the motions that will actually occur for small positive $R-R_{\mathrm{c}}$. This hope is strengthened by appeal to the analogous theory for classical Bénard convection or for Couette flow. The present problem is, however, unusual in that, as q increases from 0 to ∞, the system undergoes a gradual transition from this type of behaviour to another more reminiscent of instability in parallel flows. A transition point $q = q_{\mathrm{D}}(\lambda)$ is reached at which k_1 or Re (k_1) vanishes. For $q > q_{\mathrm{D}}$, the condition $R < R_{\mathrm{c}}$ is at most necessary for stability and it may be expected that the system is in fact unstable throughout a range $R_{\mathrm{m}} < R < R_{\mathrm{c}}$, provided the initial disturbance exceeds a certain critical amplitude that depends on R. It is not improbable that R_{m} is considerably smaller than R_{c} when q is large. There is at present time no general analytical method known by which such subcritical motions can be determined, and R_{m} located. The present problem is, however, remarkable in providing an exceptional case susceptible to proper analytic treatment. Few examples of this are known to us. If $\lambda_{\mathrm{D}}(q)$ denotes the inverse of $q_{\mathrm{D}}(\lambda)$, it is possible to show that, when $\lambda-\lambda_{\mathrm{D}}$ is of order $\sqrt{|R-R_{\mathrm{c}}|}$ the amplitude of weakly nonlinear disturbances is of order $|R-R_{\mathrm{c}}|^{\frac{1}{4}}$ and is governed by an equation of the form (A 18). The theory is developed, for the exchange of stability case only, in the Appendix.

Another interesting aspect of the instabilities for $q > q_{\mathrm{D}}$ is that of bursting, that is a strong tendency of an initial perturbation to refocus into a concentrated disturbance somewhat resembling the turbulent spots observed in shear flows. It would serve no useful purpose here to describe again the diverse behaviours possible at different q and λ, according to § 6. Instead we offer a few remarks about their possible astrophysical significance, taking the Sun as principal example and first summarizing an idealized theory of the solar cycle.

The existence of a quasi-regularly reversing solar field with a period as short as 22 years strongly suggests that the field itself is created by a dynamo process driven by turbulence in the convection zone. Unlike the case of the Earth, the magnetic energy density in the zone seems to be small compared with the kinetic. The field exerts comparatively little reaction on the flow, and shows a greater degree of randomness. For example, it is not uncommon for flux of both signs (in and out) to emerge from a polar cap, the net signed flux being perhaps only 10 % of the net unsigned flux (see Stenflo 1972; Howard 1972). As in other instances of turbulence, theory has concentrated on attempts to understand the behaviour of the average field. (Strictly this average is taken over an ensemble of identical Suns, but in effect it is the average over a number of repetitions of the solar cycle.) Steenbeck *et al.* (1966) initiated the study of a new subject, mean field electrodynamic, with which to describe the behaviour of the averaged magnetic field. Steenbeck & Krause (1966) used this theory to construct a number of models of the mean solar field some of which reproduced the principal features of the solar cycle with remarkable fidelity (see also Roberts 1972*a*).

Broadly, the mean field models of Steenbeck & Krause (1966) can be pictured as a combination of two processes. First, a comparatively weak poloidal field, symmetric with respect to the solar rotation axis, is sheared by mean differential zonal flows prevailing in the convection zone and forms a toroidal field of increasing strength. This is the analogue of the way by which toroidal field is thought to be generated in the Earth (Elsasser 1947), and which was mentioned in § 1 in connexion with the $\alpha\omega$-dynamos; it appears that Professor T. G. Cowling was aware of its significance in solar dynamo theory at about the same time. In the second process of the Steenbeck–Krause dynamo, the fields induced by the turbulence from the toroidal field interact with the turbulence itself to create a mean poloidal field of opposite polarity. The two processes are repeated with all fields oppositely directed to complete the cycle.

The second of the processes just described, that of creating poloidal field from toroidal, requires the emergence of new flux from the solar surface, and it is interesting to note that emerging flux regions (e.f.rs) are frequently seen on the solar surface (see Zirin 1972; Frazier 1972; Vorpahl 1973) Parker (1955*a*) was the first to suggest that 'magnetic buoyancy' might be responsible. He pointed out that the high radiative conductivity prevailing in the Sun would tend to equalize the temperatures inside and outside any isolated tube of magnetic flux that happened to form in the convection zone. Since, however, mechanical equilibrium ensures equality of total pressure (that is, the sum of gas pressure and magnetic pressure), the gas density inside the flux tube would, according to the gas law, be less than that of the surroundings. The flux tube would therefore experience a buoyancy force upwards, and be brought to the solar surface, with an emergence of its flux. The argument can similarly explain how a loop of toroidal flux can break through the solar surface as an e.f.r. Sunspots occur, on this picture, when for some reason the flux contained in the rising tube is highly concentrated. Although Parker's theory is the one that has won greatest acceptance, the present work prompts us to raise two points.

The first concerns the physical process itself. Even though the theory of magnetic buoyancy does recognize thermal diffusion – essentially κ is regarded as infinite – the mode of maximum instability is found to be of zero wavelength perpendicular to the field, i.e. $l_c = \infty$ (cf. Gilman 1970). The analogy with the results for the dissipationless M.A.C.-waves is impressive and suggests that a totally satisfactory theory of magnetic buoyancy, and particularly of its nonlinear aspects, will require other diffusive effects to be restored. In our opinion, particular interest attaches to the finite amplitude theory, for to our knowledge no adequate account has so far been given of the process through which a perturbation, which was presumably at first dispersed over large areas of the solar surface, becomes concentrated into sunspots perhaps in a manner reminiscent of the appearance of turbulent spots in turbulent shear flows. We may conjecture that the enormous size of q near stellar surfaces will favour overstability and bursting irrespective of whether the buoyancy mechanism is magnetic or thermal, and that this provides the basic reason for the highly localized form of sunspots. It has sometimes been suggested that a parallel phenomenon occurs in the Earth and that particularly the larger fluctuations in the length of the day are created by the magnetic stress of 'core spots' erupting through the base of the mantle (see, for example, Roberts 1972*b*). This view seems less attractive after the analysis of this paper, bearing in mind the probability that q is tiny in the core.

Our second observation is that astrophysical circumstances may exist in which the vertical gradients of toroidal field are insufficiently large for significant magnetic buoyancy, and in which the diffusive M.A.C.-wave instabilities of the present paper provide the instabilities necessary to regenerate the poloidal field. We may note particularly that, like the magnetic buoyancy process, the system becomes increasingly unstable as the toroidal field strengthens. [Equations (3.5) and (3.7) show that for fixed q, the Rayleigh number qR appropriate for large q, decreases monotonically with λq from $2\lambda q$ at large λ, to $3\sqrt{3}$ for small λ.) Thus, like magnetic buoyancy, the process possesses the desirable characteristic of leaving the layer stable until the toroidal field has grown to an amplitude sufficient to recreate an appreciable poloidal field.

While we believe that the present model is of theoretical interest, combining as it does so many diverse features of other stability problems, it is clear that, like the M.A.C.-wave theory from which it arose, it provokes a number of further questions. In particular, the significance of the Taylor condition (2.10), the relevance of the model to processes of field generation and of sunspot formalism requires further attention: we hope to consider these in future studies.

We wish to thank Emeritus Professor T. G. Cowling, F.R.S., for helpful correspondence about the solar dynamo (§ 7). Our thanks are also due to an anonymous referee for his constructive criticisms.

References

Braginsky, S. I. 1964*a* *Soviet Phys. JETP* **20**, 726.
Braginsky, S. I. 1964*b* *Geomag. Aeron.* **4**, 572.
Braginsky, S. I. 1964*c* *Geomag. Aeron.* **4**, 698.
Braginsky, S. I. 1967 *Geomag. Aeron.* **7**, 851.
Busse, F. 1973 *J. Fluid Mech.* **57**, 529.
Chandrasekhar, S. 1961 *Hydrodynamic and hydromagnetic stability*. Oxford University Press.
Childress, S. 1967 *Courant Inst. Reports* AFOSR-67-0124 and 0976.
Childress, S. & Soward, A. M. 1972 *Phys. Rev. Lettrs.* **29**, 837.
Cowling, T. G. 1933 *Mon. Not. R. Astr. Soc.* **94**, 39.
Cowling, T. G. 1946. *Mon. Not. R. Astr. Soc.* **106**, 218.
Davey, A., Hocking, L. M. & Stewartson, K. 1974 *J. Fluid Mech.* **63**, 529.
Davey, A. & Stewartson, K. 1974 *Proc. R. Soc. Lond.* A **338**, 101.
Elsasser, W. M. 1947 *Phys. Rev.* **72**, 821.
Eltayeb, I. A. 1972 *Proc. R. Soc. Lond.* A **326**, 229.
Frazier, E. 1972 *Solar Phys.* **26**, 130.
Gilman, P. A. 1970 *Astropys. J.* **162**, 1019.
Gubbins, D. 1973 *Geophys. J. R.A.S.* **33**, 57.
Hasimoto, H. & Ono, H. 1972 *J. Phys. Soc. Japan* **33**, 805.
Hayes, W. D. 1973 *Proc. R. Soc. Lond.* A **332**, 199.
Hocking, L. M. & Stewartson, K. 1971 *Mathematika* **18**, 219.
Hocking, L. M. & Stewartson, K. 1972 *Proc. R. Soc. Lond.* A **326**, 289.
Hocking, L. M., Stewartson, K. & Stuart, J. T. 1972 *J. Fluid Mech.* **51**, 705.
Howard, R. 1972 *Solar Phys.* **25**, 5.
Karpman, V. I. & Krushkal, E. M. 1968 *Soviet Phys. JETP* **61**, 118.
Moore, D. R. & Weiss, N. O. 1973 *J. Fluid Mech.* **61**, 553.
Parker, E. N. 1955*a* *Astrophys. J.* **121**, 491.
Parker, E. N. 1955*b* *Astrophys. J.* **122**, 293.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Roberts, G. O. 1970 *Phil. Trans. R. Soc. Lond.* A **266**, 535.
Roberts, P. H. 1972*a* *Phil. Trans. R. Soc. Lond.* A **272**, 663.
Roberts, P. H. 1972*b* *J. Geomag. Geoelec.* **24**, 231.
Roberts, P. H. & Soward, A. M. 1972 *Ann. Rev. Fl. Mech.* **4**, 117.
Soward, A. M. 1972 *Phil. Trans. R. Soc. Lond.* A **272**, 431.
Soward, A. M. 1974 *Phil. Trans. R. Soc. Lond.* A (to appear).
Steenbeck, M. & Krause, F. 1966 *Z. Naturf.* **21** A 1285.
Steenbeck, M., Krause, F. & Rädler, K.-H. 1966 *Z. Naturf.* **21** A 369.
Stenflo, J. 1972 *Solar Phys.* **23**, 307.
Stewartson, K. 1960 *J. Fluid Mech.* **8**, 82.
Stewartson, K. & Stuart, J. T. 1971 *J. Fluid Mech.* **48**, 529.
Stuart, J. T. 1958 *J. Fluid Mech.* **4**, 1.
Taylor, J. B. 1963 *Proc. R. Soc. Lond.* A **274**, 274.
Vorpahl, J. A. 1973 *Solar Phys.* **28**, 115.
Zakharov, V. E. & Shabat, A. B. 1972 *Soviet Phys. JETP* **61**, 118.
Zirin, H. 1972 *Solar Phys.* **26**, 130.

Appendix. Weakly nonlinear solutions near the critical curve $\lambda = \lambda_D(q)$ for values of q for which the principle of the exchange of stabilities holds

For simplicity we at first suppose that $\lambda = \lambda_D(q)$ so that, by definition, k_1 is zero, and the terms proportional to $|A|^2 A$ in (5.1) and (5.10) are absent. Finite amplitude effects now occur at the level $|A|^4 A$, and it is convenient to adopt a different definition of ϵ, replacing (4.2), (4.3) and (5.9) by

$$\epsilon^4 = |R-R_c|, \quad \tau = \epsilon^4 t, \quad \xi = \epsilon^2 x, \quad \eta = \epsilon^2 y, \tag{A 1}$$

which, however, in terms of $R-R_c$ leaves the large length and time scales unaltered. It is necessary to continue expansions such as (4.4) to order ϵ^5, and to apply the consistency condition (4.10) at that level. In the case of θ, the terms required in the expansion are

$$\begin{aligned}\theta = {} & \epsilon(\theta_{11}E_1+\theta_{11}^*E_1^{-1})\cos z+\epsilon^2(\theta_{02}\sin 2z+\theta_{22}E_1^2+\theta_{22}^*E_1^{-2})\\ & +\epsilon^3[(\theta_{13}E_1+\theta_{13}^*E_1^{-1})\cos z+(\tilde\theta_{13}E_1+\tilde\theta_{13}^*E_1^{-1})\cos 3z+(\theta_{33}E_1^3+\theta_{33}E_1^{-3})\cos z]\\ & +\epsilon^4(\theta_{04}\sin 2z+\theta_{24}E_1^2+\theta_{24}^*E_1^{-2})+\epsilon^5(\theta_{15}E_1+\theta_{15}^*E_1^{-1})\cos z.\end{aligned} \tag{A 2}$$

In order to display the explicit dependence of the coefficients on z, we have here adopted a notation that differs in a small but obvious way from (4.4) and (4.5). From now onwards we revert to the old notation. Some of the terms shown in (A 2), namely θ_{22}, $\tilde\theta_{13}$ and θ_{24}, happen to be zero, but are given since corresponding terms for $\boldsymbol{V}$ and $\boldsymbol{B}$ do not vanish. Terms which make no contribution to the consistency condition are not given. For example, it is clear that terms proportional to $\epsilon^4E_1^4$ and $\epsilon^4E_1^{-4}$ exist, but (whether zero or not) they cannot make a contribution to θ_{15} etc. It is less evident, but in fact true, that terms proportional to $\epsilon^4E_1^2\sin 2z$ and $\epsilon^4E_1^{-2}\sin 2z$ though present play no part in the consistency condition: the same is true of all similar terms in the expansions of the other variables. In future, similar irrelevant terms will be omitted from the working without comment.

The first and second order solutions have already been presented in (3.8), (4.6) and (4.7) above. We now proceed to third order. We have

$$(k^2+1)\,\theta_{13}+V_{13z} = \tfrac{1}{2}(k^2+1)^{-1}\,|A|^2A\cos z, \tag{A 3a}$$

$$(k^2+1)\,B_{13z}-\mathrm{i}mV_{13z} = -\tfrac{1}{2}\mathrm{i}(k^2+1)^{-1}\,q^2m|A|^2A\cos z, \tag{A 3b}$$

$$(k^2+1)j_{13z}-im\omega_{13z} = \tfrac{1}{2}im^{-1}(k^2+1)^{-2}\lambda q^2(k^4+8k^2+3)\,|A|^2 A\sin z, \tag{A 3c}$$

$$imj_{13z}+\lambda\,\partial V_{13z}/\partial z = -\tfrac{1}{2}(k^2+1)^{-2}\lambda q^2(3k^2+1)\,|A|^2 A\sin z, \tag{A 3d}$$

$$im(k^2+1)\,B_{13z}+\lambda\,\partial\omega_{13z}/\partial z-\lambda k^2 R_c\theta_{13} = -\tfrac{1}{2}(k^2+1)^{-2}q^2m^2(k^2-3)\,|A|^2 A\cos z, \tag{A 3e}$$

where $\boldsymbol{\omega}$ = curl $\boldsymbol{V}$ and $\boldsymbol{j}$ = curl $\boldsymbol{B}$. These equations are soluble since the consistency condition is obeyed by supposition, and for the same reason they possess an infinity of solutions all of which lead to the same final equation (A 18) governing A. Without loss of generality we may take $V_{13z} = 0$ and obtain

$$\boldsymbol{V}_{13} = \frac{i\lambda q^2(k^4-2k^2-1)\,|A|^2 A}{m^2k^2(k^2+1)^2}(m,\ -l,\ 0)\sin z, \tag{A 4a}$$

$$\boldsymbol{B}_{13} = -\frac{q^2m|A|^2 A}{2k^2(k^2+1)^2}\left[\left\{l+\frac{\lambda(3k^2+1)}{m}\right\}\sin z,\ \left\{m-\frac{\lambda l(3k^2+1)}{m^2}\right\}\sin z,\ ik^2\cos z\right], \tag{A 4b}$$

$$\theta_{13} = \frac{|A|^2 A}{2(k^2+1)^2}\cos z. \tag{A 4c}$$

Turning to the terms distinguished by the tilde, we obtain

$$(k^2+9)\,\tilde{\theta}_{13}+\tilde{V}_{13z} = \tfrac{1}{2}(k^2+1)^{-1}\,|A|^2 A\cos 3z, \tag{A 5a}$$

$$(k^2+9)\,\tilde{B}_{13z}-im\tilde{V}_{13z} = -\tfrac{1}{2}i(k^2+1)^{-1}\,q^2m|A|^2 A\cos 3z, \tag{A 5b}$$

$$(k^2+9)\tilde{j}_{13z}-im\tilde{\omega}_{13z} = \tfrac{3}{2}m^{-1}i\lambda q^2|A|^2 A\sin 3z, \tag{A 5c}$$

$$im\tilde{j}_{13z}+\lambda\,\partial\tilde{V}_{13z}/\partial z = -\tfrac{1}{2}(k^2+1)^{-1}\lambda q^2|A|^2 A\sin 3z, \tag{A 5d}$$

$$im/(k^2+9)\,\tilde{B}_{13z}+\lambda\,\partial\tilde{\omega}_{13z}/\partial z-\lambda k^2R_c\tilde{\theta}_{13} = -\tfrac{1}{2}(k^2+1)^{-2}q^2m^2(k^2-3)\,|A|^2 A\cos 3z. \tag{A 5e}$$

Writing, for brevity

$$\tilde{D} = \tfrac{1}{8}[m^4+\lambda^2(k^2+7)\,(k^2+13)]^{-1}, \tag{A 6}$$

we find that the solution of (A 5) is

$$\tilde{\boldsymbol{V}}_{13} = \left[\frac{il}{k^2}\frac{\partial\tilde{V}_{13z}}{\partial z}+\frac{im}{k^2}\tilde{\omega}_{13z},\ \frac{im}{k^2}\frac{\partial\tilde{V}_{13z}}{\partial z}-\frac{il}{k^2}\tilde{\omega}_{13z},\ \tilde{V}_{13z}\right], \tag{A 7a}$$

$$\tilde{\boldsymbol{B}}_{13} = \left[\frac{il}{k^2}\frac{\partial\tilde{B}_{13z}}{\partial z}+\frac{im}{k^2}\tilde{j}_{13z},\ \frac{im}{k^2}\frac{\partial\tilde{B}_{13z}}{\partial z}-\frac{il}{k^2}\tilde{j}_{13z},\ \tilde{B}_{13z}\right], \tag{A 7b}$$

$$\tilde{V}_{13z} = \{-\tfrac{1}{2}[m^4+\lambda^2(k^2+1)]+q^2(k^2+9)\,(k^2+1)^{-2}[m^4(k^2-1)-3\lambda^2(k^2+1)\,(k^2-3)]\}$$
$$\times\tilde{D}|A|^2 A\cos 3z, \tag{A 7c}$$

$$\tilde{\omega}_{13z} = \{\tfrac{3}{2}(k^2+9)\,[m^4+\lambda^2(k^2+1)]-q^2(k^2+1)^{-2}[m^4(3k^6+59k^4+173k^2-267)$$
$$-\lambda^2(k^2+1)^3\,(k^2-3)]\}\lambda m^{-2}\tilde{D}|A|^2 A\sin 3z, \tag{A 7d}$$

$$\tilde{B}_{13z} = \{-\tfrac{1}{2}[m^4+\lambda^2(k^2+1)]+q^2(k^2+1)^{-2}[m^4(k^4+4k^2-13)-\lambda^2(k^2+1)$$
$$\times(7k^4+98k^2+283)]\}im(k^2+9)^{-1}\tilde{D}|A|^2 A\cos 3z, \tag{A 7e}$$

$$\tilde{j}_{13z} = \{\tfrac{3}{2}[m^4+\lambda^2(k^2+1)]-q^2(k^2+1)^{-2}[m^4(3k^4+20k^2-31)-\lambda^2(k^2+1)^2\,(13k^2+121)]\}$$
$$\times i\lambda m^{-1}\tilde{D}|A|^2 A\sin 3z, \tag{A 7f}$$

$$\tilde{\theta}_{13z} = \{\tfrac{1}{2}[m^4+9\lambda^2(k^2+9)]-q^2(k^2+1)^{-1}[m^4(k^2-1)-3\lambda^2(k^2+1)\,(k^2-3)]\}$$
$$\times(k^2+1)^{-1}\tilde{D}|A|^2 A\cos 3z. \tag{A 7g}$$

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

The final set of equations posed at the third level are

$$(9k^2+1)\,\theta_{33}+V_{33z} = 0, \tag{A 8a}$$

$$(9k^2+1)\,B_{33z}-3imV_{33z} = 0, \tag{A 8b}$$

$$(9k^2+1)\,j_{33z}-3im\,\omega_{33z} = 3im^{-1}(k^2+1)^{-2}\,\lambda q^2(k^2-1)\,A^3\sin z, \tag{A 8c}$$

$$3imj_{33z}+\lambda\,\partial V_{33z}/\partial z = 3(k^2+1)^{-2}\,\lambda q^2A^3\sin z, \tag{A 8d}$$

$$3im(9k^2+1)\,B_{33z}+\lambda\,\partial\omega_{33z}/\partial z-9\lambda k^2R_c\,\theta_{33} = 0. \tag{A 8e}$$

Writing for brevity
$$D = \tfrac{3}{4}(k^2+1)^{-2}[81k^2m^4-2\lambda^2(9k^2+5)]^{-1}, \tag{A 9}$$

we find that the solution to (A 8) is

$$\boldsymbol{V}_{33} = \left[\frac{il}{3k^2}\frac{\partial V_{33z}}{\partial z}+\frac{im}{3k^2}\omega_{33z},\ \frac{im}{3k^2}\frac{\partial V_{33z}}{\partial z}-\frac{il}{3k^2}\omega_{33z},\ V_{33z}\right], \tag{A 10a}$$

$$\boldsymbol{B}_{33} = \left[\frac{il}{3k^2}\frac{\partial B_{33z}}{\partial z}+\frac{im}{3k^2}j_{33z},\ \frac{im}{3k^2}\frac{\partial B_{33z}}{\partial z}-\frac{il}{3k^2}j_{33z},\ B_{33z}\right], \tag{A 10b}$$

$$V_{33z} = -q^2(6k^2-1)\,(9k^2+1)\,DA^3\cos z, \tag{A 10c}$$

$$\omega_{33z} = 9m^{-2}\lambda q^2(6k^2-1)\,[\lambda^2(k^2+1)^2-8k^2m^4]\,DA^3\sin z, \tag{A 10d}$$

$$B_{33z} = -3i\lambda^2q^2m(6k^2-1)\,DA^3\cos z, \tag{A 10e}$$

$$j_{33z} = im^{-1}\lambda q^2[-108k^2m^4+\lambda^2(18k^4+23k^2+13)]\,DA^3\sin z, \tag{A 10f}$$

$$\theta_{33} = \lambda^2q^2(6k^2-1)\,DA^3\cos z. \tag{A 10g}$$

We next consider the fourth order equations. It happens (see above) to be necessary to solve only the equations for θ_{04}, etc., and the terms independent of z in θ_{24}, etc. We obtain

$$\boldsymbol{V}_{04} = (\alpha l+\beta m,\ \alpha m-\beta l,\ 0)\cos 2z, \tag{A 11a}$$

$$\boldsymbol{B}_{04} = (\gamma l+\delta m,\ \gamma m-\delta l,\ 0)\cos 2z, \tag{A 11b}$$

$$\theta_{04} = \phi\sin 2z, \tag{A 11c}$$

where

$$\alpha = \frac{2iq^3\,|A|^4}{(k^2+1)^3}-\frac{2qA^*}{k^2m}\tilde{B}_{13z}+\frac{2qmA^*}{\lambda k^2(k^2+1)}\tilde{j}_{13z}, \tag{A 12a}$$

$$\beta = \frac{8qmA^*}{\lambda k^2(k^2+1)}\tilde{B}_{13z}, \tag{A 12b}$$

$$\gamma = \frac{q^3m|A|^4}{2k^2(k^2+1)^2}+\frac{iqA^*}{k^2}\tilde{B}_{13z}-\frac{qmA^*}{k^2(k^2+1)}\tilde{V}_{13z}, \tag{A 12c}$$

$$\delta = \frac{q^3\lambda(3k^2+1)\,|A|^4}{2m(k^2+1)^3}-\frac{iq\lambda(k^2+1)\,A^*}{2m^2k^2}\tilde{B}_{13z}-\frac{iqA^*}{2k^2}\tilde{j}_{13z}+\frac{q\lambda A^*}{2mk^2}\tilde{V}_{13z}+\frac{qmA^*}{k^2(k^2+1)}\tilde{\omega}_{13z}, \tag{A 12d}$$

$$\phi = \frac{|A|^4}{4(k^2+1)^2} + \frac{A^*}{2}\tilde{\theta}_{13} - \frac{A^*}{2(k^2+1)}\tilde{V}_{13z}. \tag{A 12e}$$

In (A 12), and in (A 14) below, we understand that the functions of z are extracted from $\tilde{B}_{13z}$ etc. before substitution.

The second relevant set of equations yield

$$\boldsymbol{V}_{24} = \psi(m, -l, 0), \quad \boldsymbol{B}_{24} = \chi(m, -l, 0), \quad \theta_{24} = 0, \tag{A 13}$$

where

$$\psi = -\frac{\mathrm{i}q^3\lambda(4k^4-k^2-1)\,|A|^2A^2}{2m^2k^2(k^2+1)^3} - \frac{\mathrm{i}q\lambda A^*}{6m^2k^2}V_{33z} - \frac{\mathrm{i}qA^*}{6k^2(k^2+1)}\omega_{33z} + \frac{q\lambda(k^2-1)\,A^*}{6m^3k^2}B_{33z} - \frac{q(3k^2+1)\,A^*}{6mk^2(k^2+1)}j_{33z}, \tag{A 14a}$$

$$\chi = \frac{q^3\lambda(2k^2+1)\,|A|^2A^2}{2mk^2(k^2+1)^3} + \frac{\mathrm{i}q\lambda A^*}{6m^2k^2}B_{33z} - \frac{\mathrm{i}qA^*}{6k^2(k^2+1)}j_{33z}. \tag{A 14b}$$

We are now in a position to proceed to the fifth order. The set of equations governing θ_{51} etc. are of the same form as (4.8) but, to the Q_1 to Q_5 given by (4.9) we must add Q_1' to Q_5', where

$$Q_1' = \frac{qmk^2A}{2(k^2+1)}\left[\frac{\lambda(k^2+1)}{m^2}\gamma + 2\delta\right]\frac{2qmk^2A^*}{(k^2+1)}-\chi - \frac{\mathrm{i}q^2m|A|^2}{2(k^2+1)}\left[\frac{2\lambda(k^2+1)}{m^2}\tilde{B}_{13z} - \tilde{j}_{13z}\right] + \frac{q^4\lambda(5k^2+1)|A|^4A}{4(k^2+1)^3}, \tag{A 15a}$$

$$Q_2' = \frac{qmk^2(k^2-3)\,A}{2(k^2+1)}\gamma + \frac{\mathrm{i}q^2m(k^2+5)\,|A|^2}{2(k^2+1)}\tilde{B}_{13z} + \frac{q^4m^2(k^2-3)\,|A|^4A}{4(k^2+1)^3}, \tag{A 15b}$$

$$Q_3' = \frac{\mathrm{i}k^2A}{2(k^2+1)}\alpha - A\phi + \frac{|A|^2}{2(k^2+1)}\tilde{V}_{13z}, \tag{A 15c}$$

$$Q_4' = -\frac{\mathrm{i}qk^2mA}{(2k^2+1)}\alpha + \frac{qk^2A}{2}\gamma - \frac{q^2m|A|^2}{2(k^2+1)}\tilde{V}_{13z}, \tag{A 15d}$$

$$Q_5' = -\frac{\mathrm{i}qmk^2A}{(k^2+1)}\left[\frac{\lambda(k^2+1)}{m^2}\alpha - 2\beta\right] + \frac{qk^2A}{2}\left[\frac{\lambda(k^2+1)}{m^2}\gamma - 2\delta\right] - 2qk^2A^*\chi - \frac{2\mathrm{i}qmA^*}{(k^2+1)}\psi - \frac{q^2m|A|^2}{2(k^2+1)}\left[\frac{2\lambda(k^2+1)}{m^2}\tilde{V}_{13z} + \tilde{\omega}_{13z}\right] + \frac{\lambda q^4(k^4-4k^2-1)}{2m(k^2+1)^3}|A|^4A. \tag{A 15e}$$

We now relax our initial restriction that k_1 is zero. Denoting by $q_{\mathrm{D}}(\lambda)$ the value of q for which k_1 is zero, we suppose that $q-q_{\mathrm{D}}$ is small, or more precisely that

$$q - q_{\mathrm{D}} = \xi\,|R-R_{\mathrm{c}}|, \tag{A 16}$$

so that

$$|k_1| = \xi_1(\lambda)\,|R-R_{\mathrm{c}}| \quad \text{where} \quad \xi_1 = \xi(\partial k_1/\partial q)_{q_{\mathrm{D}}}. \tag{A 17}$$

On substitution of (4.9), amended by (A 15), into (4.10), we now obtain in place of (5.1) the equation

$$\frac{\mathrm{d}A}{\mathrm{d}\tau} = \pm d_1A \pm \xi_1|A|^2A + \eta_1|A|^4A, \tag{A 18}$$

where the first $\pm$ takes the sign of $R - R_c$ and the second the sign of $q - q_D$ (i.e. the sign of k_1), d_1 and ξ_1 are positive and

$$
\begin{aligned}
&[\lambda^2(q+1)(k^2+1) - (q-1)m^4]\,\eta_1 \\
&\quad = -q^4\lambda^2[18k^2m^4(27k^2+7) - \lambda^2(108k^6+164k^4+79k^2+15)] \\
&\qquad \div \{2(k^2+1)^4[81k^2m^4 - 2\lambda^2(9k^2+5)]\} + [\![[m^4+\lambda^2(k^2+1)] \\
&\qquad \times \{(k^2+9)[(3k^2+11)m^4+\lambda^2(19k^4+254k^2+811)] + 8q(k^2+7)[m^4+\lambda^2(k^2+1)] \\
&\qquad - 2q^2(k^2+1)^{-1}[m^4(6k^6+k65k^4+176k^2-139) - \lambda^2(k^2+1)(9k^6+127k^4+203k^2-1451)] \\
&\qquad - 16q^3(k^2+1)^{-2}[m^4(k^6+11k^4+17k^2-73) - \lambda^2(k^2+7)(7k^6+103k^4+337k^2+49)]\} \\
&\qquad + 4q^4(k^2+1)^{-2}[m^8(3k^8+31k^6+77k^4-31k^2-272) \\
&\qquad - \lambda^2m^2(21k^{10}+480k^8+3506k^6+9224k^4+6065k^2-4976) \\
&\qquad - 8\lambda^4(3k^{12}+74k^{10}+546k^8+672k^6-4413k^4-2962k^2-656)]]\!] \\
&\qquad \div \{32(k^2+1)^2(k^2+9)[m^4+\lambda^2(k^2+7)(k^2+13)]\},
\end{aligned}
\tag{A 19}
$$

an expression in which q must be set equal to q_D.

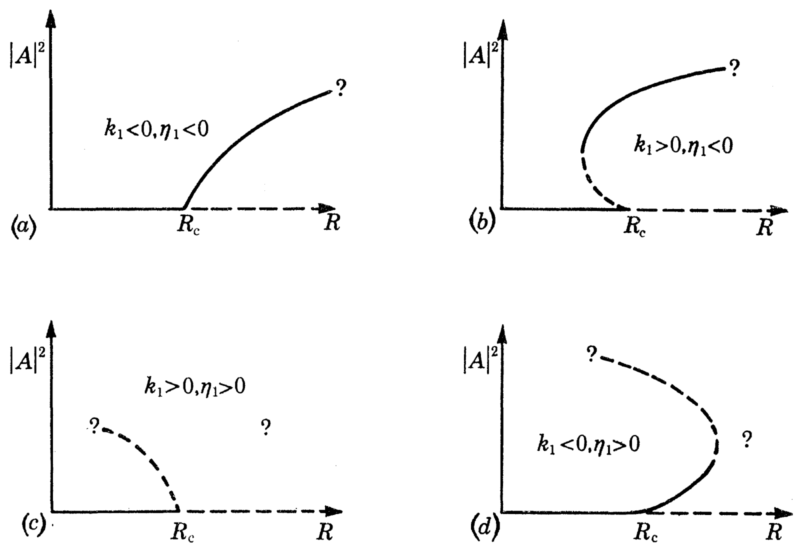

FIGURE 3. Properties of solutions of equation (A 18) in different cases. Dotted curves indicate linearly unstable solutions. Question marks suggest regions in which the theory fails to give valid solutions because $|A|^2$ is too great.

It is worth reviewing the properties of solutions to (A 19). One admissible steady-state solution is $A = 0$, and this is linearly unstable if $R > R_c$, and stable if $R < R_c$. If $\eta_1 < 0$, $R < R_c$ and $l_1 < 0$, the *only* admissible steady solution, bearing in mind that $|A|^2$ cannot be negative, is $A = 0$ and it is stable: similarly if $\eta_1 > 0$, $R > R_c$ and $k_1 > 0$, only $A = 0$ is possible, but it is unstable. If $\eta_1 < 0$ and $R > R_c$ there is, in addition to the unstable solution $A = 0$, a stable solution. The same is true if $\eta_1 > 0$ and $R < R_c$, but in this case the new solution is unstable, and $A = 0$ is stable. Finally, if $\eta_1 < 0$, $R < R_c$ and $k_1 > 0$, there are in all three solutions, but (apart from $A = 0$) only the solution

$$|A|^2 = [\xi_1 \pm \surd(\xi_1^2 - 4d_1|\eta_1|)]/2|\eta_1| \tag{A 20}$$

of the larger amplitude is stable: if $\eta_1 > 0$, $R > R_c$ and $k_1 < 0$, only that of smaller amplitude is stable. The possible steady states are sketched in figure 3, the unstable states being shown dashed.

It is found that, in the case of oblique rolls ($\lambda < 2/\sqrt{3}$) the constant k_1 is negative and

$$\xi_1 = \sqrt{2}\xi/3,$$

$$k_1 = -\frac{89}{384} + \frac{(195471 - 96606\sqrt{2})}{36432} \approx -1.8471 < 0. \qquad \text{(A 21)}$$

If, in contrast, we consider the transverse rolls in the limit of large m, we obtain

$$\xi_1 = 2(\sqrt{2} - 1)\,\xi,$$

$$k_1 = \frac{\sqrt{2} - 1}{8\lambda^2\sqrt{2}} \approx 0.03661\lambda^{-2} > 0. \qquad \text{(A 22)}$$

Thus each of the situations depicted in figure 3 pertain for some λ.

Note added in proof, 16 *August* 1974.

Professor R. Hide has kindly pointed out to us that the statement made in the first complete sentence of p. 291 rests on (1.2) which is not a uniformly valid approximation to the dispersion relation when $l \to \infty$ since inertial effects, neglected in deriving (1.2), may then become significant. The interplay between inertial and diffusive effects in modifying (1.2) is quite subtle and beyond the scope of this paper but our standpoint is consistent provided we interpret the neglect of inertial terms and the retention of diffusive effects to mean, in the notation of (2.5), that $\delta \to 0$ while R, λ, q remain finite. Braginsky's result (1.2) then follows on letting $R \to \infty$ and $\lambda \to 0$, holding q finite. The necessity for considering finite values of λ and R arises because ω^2/δ^2 is real finite and negative when (1.3) is satisfied but $\to -\infty$ as $l/n \to \infty$.

It is observed that although this argument provided the motivation for our present study, it transpired that a new resistive mode was of greater significance than the waves of the type governed by (1.2) when modified by diffusion. In other words, stability is controlled by a separate branch of the dispersion relation.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES