J. Fluid Mech. (1980), *vol.* 101, *part* 4, *pp.* 759–782
Printed in Great Britain

Nonlinear Rayleigh–Bénard convection between poorly conducting boundaries

By C. J. CHAPMAN
School of Mathematics, University of Bristol
AND M. R. E. PROCTOR
Department of Applied Mathematics and Theoretical Physics, University of Cambridge

(Received 3 December 1979 and in revised form 14 May 1980)

The convective instability of a layer of fluid heated from below is studied on the assumption that the flux of heat through the boundaries is unaffected by the motion in the layer. It is shown that when the heat flux is above the critical value for the onset of convection, motion takes place on a horizontal scale much greater than the layer depth. Following Childress & Spiegel (1980) the disparity of scales is exploited in an expansion scheme that results in a nonlinear evolution equation for the leading-order temperature perturbation. This equation which does not depend on the vertical co-ordinate, is solved analytically where possible and numerically where necessary; most attention is concentrated on solutions representing two-dimensional rolls. It is found that for any given heat flux a continuum of steady solutions is possible for all wave numbers smaller than a given cut off. Stability analysis reveals, however, that each mode is unstable to one of longer wavelength than itself, so that any long box will eventually contain a single roll, even though the most rapidly growing mode on linear theory has much shorter wavelength.

1. Introduction

The convective instability of a layer of Boussinesq fluid heated uniformly from below has long been recognized as a problem of crucial importance in many fields of fluid mechanics, and any discussion of energy transport in the earth's atmosphere, its oceans, mantle and core, or the outer layers of the sun, must take account of this instability. In attempts to understand the phenomenon theoretically, the model usually adopted is that of a fluid layer of infinite horizontal extent, whose upper and lower boundaries are held at fixed temperatures. This problem was first addressed by Rayleigh (1916) and has been extensively studied since. (See Chandrasekhar 1961 for a full account of the linearized problem.) In recent years many authors have tried to understand more about the nonlinear aspects of the instability as a prelude to an understanding of thermal turbulence. Malkus & Veronis (1958) have found the relation between temperature difference and amplitude of the instability in the weakly nonlinear regime, and Schlüter, Lortz & Busse (1965) have used these weakly nonlinear solutions to determine which of the several planforms permitted by the linear problem is in fact realized. All analytic work on nonlinear effects has perforce to be accomplished by means of a perturbation expansion in the amplitude, since the fully

0022–1120/80/4636–2810 $02.00

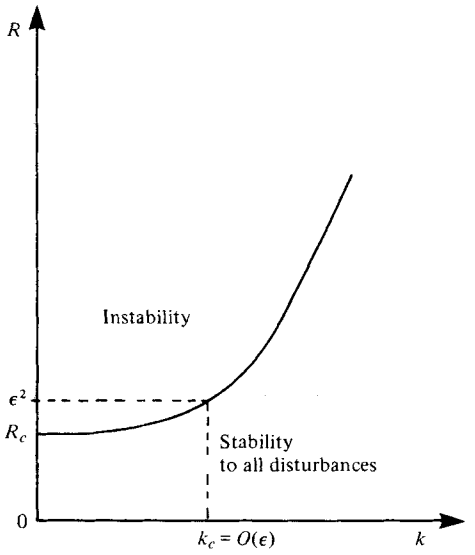

FIGURE 1. The boundary for the transition from stability to instability in a periodic box of length $2\pi/k$ as a function of k. For $R = R_c + \epsilon^2 (\epsilon^2 \ll 1)$ instability is possible only for $k < k_c = O(\epsilon)$.

nonlinear problem usually defies analytical solution. Other studies have investigated the stability of steady solutions at all amplitudes using numerical techniques. (For a recent review, see Busse 1978.)

It is clear, though, that the assumption of fixed boundary temperature is sometimes quite inappropriate – for example, when convection occurs in a relatively thin layer of highly conducting fluid embedded in a very poorly conducting region. Horizontal temperature variations in the Earth's mantle due to deep-seated convection indicate that in this case also a fixed temperature boundary condition is inappropriate.† One might consider a sequence of conditions of the 'Newton's Law of Cooling' type. As a first step in this path, bearing in mind that the fixed temperature problem is well understood, it seems natural to investigate the problem of convective instability when the boundaries of the layer are much poorer conductors of heat than the fluid itself. The linear stability problem was first investigated by Jeffreys (1926) in the limit for which the boundaries are almost perfect insulators. His analysis is incorrect, however, owing to an error in the velocity boundary conditions. More recently Sparrow, Goldstein & Jonsson (1964) and Hurle, Jakeman & Pike (1967) made comprehensive studies, the former for temperature boundary conditions of a 'Newton's Law of Cooling' type, and the latter for a layer enclosed between poorly conducting solid slabs. Related work by Nield (1967) and Jakeman (1968) should also be mentioned. These authors noted that as the ratio of conductivities of fluid to boundary becomes very large, the critical horizontal wavenumber for the onset of convection approaches zero, in sharp contrast to the fixed temperature case. In the so-called 'perfect problem' (both boundaries perfectly insulating so that the heat flux through the layer is unchanged by the motion), the critical wavenumber *is* zero. It is then clear that if the

† The relevance of the present study to mantle convection is discussed by Chapman, Childress & Proctor (1980) who also summarize the results presented here.

heat flux exceeds the critical value (for steady motion at zero horizontal wavenumber) only by a small amount, then motion can take place only on a large horizontal scale (figure 1). This was noted first for the related problem of convection due to swimming micro-organisms by Childress, Levandowsky & Spiegel (1975): more recently Childress & Spiegel (1980) have exploited this disparity of scales to develop an expansion scheme for their equations in powers of the (small) horizontal wavenumber, and suggest the applicability of their method to other types of convection problem. (Such an expansion was foreshadowed by Nield (1975), who solved the linear problem in this way, but made no attempt to include the nonlinear terms). Although Childress & Spiegel's method yields only subcritical (unstable) solutions for the problem they treat, in the present problem the technique can be used to obtain stable finite-amplitude solutions.†

In this paper we carry out an expansion procedure of the type suggested by Childress & Spiegel for the 'perfect problem' (see above). The more involved 'imperfect problem' has yet to be solved, although the first steps are given in Chapman (1980).

We suppose that the layer has large horizontal extent and seek time-dependent solutions with a long horizontal length scale when the imposed heat flux through the layer is just larger than that needed for infinitesimal motion to occur. Most of the study is concerned with two-dimensional (roll) planforms, though we do indicate the procedure for hexagon-type modes as well. The three-dimensional problem has been addressed in a different way by Busse & Riahi (1980). They consider weakly nonlinear solutions of the problem solved by Sparrow *et al.* when the boundaries are almost perfect insulators. By using analysis similar to that of Schlüter *et al.* (1965) they show that the preferred horizontal planform takes the form of rectangular cells when the heat flux is very close to its critical value for the onset of convection. The methods of the present paper can be used to extend greatly the parameter range in which the preferred planform can be found, but work on this is at an early stage. We must emphasize, in the light of many comments from those who have read a preliminary report of this work (Chapman 1978), that the expansion is not a 'small amplitude' expansion of the Malkus & Veronis (1958) type; for example, the first non-trivial differential equation to appear is in fact fully nonlinear. (Of course, for certain wavenumbers the amplitude of the motion is small and in this case the solution can also be found by an amplitude expansion.) Because of the approximation, it is easy to deal with any combination of boundary conditions on the velocity at the top and bottom of the layer of fluid. We treat the obvious cases of fixed–fixed, free–free and free (top)–fixed (bottom). There is no important difference between the first two, but the last problem, involving as it does an asymmetry between up and down motions (even though the fluid is taken to be Boussinesq), introduces some new effects. In particular, the Prandtl number $\sigma = \nu/\kappa$ where ν is the kinematic viscosity and κ the thermal diffusivity, enters the analysis only in this case.

† The techniques employed resemble those of Newell & Whitehead (1969) who studied packets of closely related modes in the fixed temperature problem, and obtained envelope equations for slow modulation that are similar to ours; the details are different, however, as they are perturbing about a finite wavenumber, and we perturb about zero wavenumber.

2. The expansion scheme

We consider a layer of Boussinesq fluid of depth $2d$ lying between the planes $z = \pm d$; (x, y, z) are Cartesian co-ordinates. Gravity $\mathbf{g} = -g\hat{\mathbf{z}}$ is perpendicular to the boundaries. The fluid has velocity $\mathbf{U}$, and pressure p. Its kinematic viscosity is ν and and thermal diffusivity κ. The temperature gradient at the top and bottom boundaries is β, and is assumed to be unaffected by the motion (the top and bottom boundaries being poor conductors of heat). If the temperature of the fluid is $T = T_0 - \beta z + \theta$ where T_0 is some reference temperature (at which the density of the fluid is ρ_0), and α is the coefficient of expansion, then the non-dimensional equations of motion and heat conduction are

$$\sigma^{-1}\left(\frac{\partial \mathbf{U}}{\partial t} + \mathbf{U}.\nabla\mathbf{U}\right) + \nabla p = R\theta\hat{\mathbf{z}} + \nabla^2\mathbf{U}, \tag{2.1}$$

$$\frac{\partial \theta}{\partial t} + \mathbf{U}.\nabla\theta = \mathbf{U}.\hat{\mathbf{z}} + \nabla^2\theta, \tag{2.2}$$

$$\nabla.\mathbf{U} = 0, \tag{2.3}$$

where $|\mathbf{U}|$ is scaled with κ/d, p with $\rho_0\nu\kappa/d^2$ time t with d^2/κ, lengths with d, and θ with βd. The dimensionless parameters are the Prandtl number

$$\sigma = \nu/\kappa, \tag{2.4}$$

and the Rayleigh number

$$R = g\alpha\beta d^4/\kappa\nu. \tag{2.5}$$

It will be noted that the boundaries of the layer are now at $z = \pm 1$. The boundary conditions on $\mathbf{U}$ and θ at $z = \pm 1$ are

$$\partial\theta/\partial z = 0 \tag{2.6}$$

$$\left.\begin{aligned} &\text{and either } \mathbf{U}.\hat{\mathbf{z}} = \frac{\partial}{\partial z}(\mathbf{U}.\hat{\mathbf{z}}) = 0, \quad z = 0, 1, \quad \text{case } A, \text{ two stress-free boundaries} \\ &\text{or} \qquad \mathbf{U} = 0; \quad z = 0, 1, \quad \text{case } B, \text{ two rigid boundaries} \\ &\text{or} \qquad \mathbf{U} = 0, \quad z = 0; \quad \mathbf{U}.\hat{\mathbf{z}} = \frac{\partial}{\partial z}(\mathbf{U}.\hat{z}) = 0, \quad z = 1 \quad \text{case } C. \end{aligned}\right\} \tag{2.7}$$

There is clearly a solution to these equations with $\mathbf{U} = 0$ and horizontal isotherms ($\theta = 0$). As with fixed temperature boundaries, this static state is unstable if R is sufficiently large. It can be shown that in the present case, the critical value of R for the onset of convection is a monotonically decreasing function of wavelength, so that the lowest critical value occurs for infinitely long horizontal scales. Figure 1 gives a sketch of the stability boundary as a function of $k \equiv 2\pi/L$, where L is the wavelength of the disturbance. Above the curve shown, some disturbances can grow. Below the curve, all disturbances decay, even if they are finite amplitude, provided L is now defined as the dimension of a 'periodic box'. This very powerful result, exactly analogous to one for Bénard convection between perfectly conducting boundaries, can be proved by 'energy integral' methods of the type pioneered for fluid mechanical problems by Backus (1958) and Serrin (1959); a full description of modern applica-

tions can be found in the monograph of Joseph (1976). The proof in the case at hand is given in appendix B.

Inspection of figure 1 shows that the only disturbances that can grow have a horizontal wavelength that is long compared with the depth of the layer. Indeed since $R \sim A_1 + A_2/L^2$ (where A_1, A_2 are constants) for $L \to \infty$, if $R = R_c + O(\epsilon^2)$ where $\epsilon \ll 1$ and R_c is the critical value for infinite L, then $L \geqslant O(\epsilon^{-1})$ for growing solutions to be possible. This suggests that in order to investigate the latter we should set

$$\left.\begin{aligned} &R = R_c + \mu^2\epsilon^2, \quad \mu = O(1), \\ &\partial/\partial x = \epsilon\partial/\partial X, \quad \text{and similarly for } \partial/\partial y. \end{aligned}\right\} \tag{2.8}$$

We also need to find appropriate scales for temperature perturbation, velocity and time. It will emerge from the analysis that the correct scaling is

$$\frac{\partial}{\partial t} = \epsilon^4 \frac{\partial}{\partial \tau}, \tag{2.9}$$

the power of ϵ that appears is plausible since (i) R is only just greater than R_c so growth rates are slow (this gives a factor ϵ^2) and (ii) horizontal scales are long so information takes a time $O(\epsilon^{-2})$ to diffuse across the system. One of the consequences of linearized theory is that $|\mathbf{U}|/|\theta| = O(\epsilon)$ (the rate of vorticity generation is small due to the long horizontal scales). However, we have no guide as to the correct amplitude to take for θ. In what follows, we suppose that $\theta = O(1)$. The analysis then leads to a non-linear equation for θ which has periodic solutions. If we had set θ smaller, then nonlinear terms would not have appeared at leading order. Although the scaling actually adopted is hard to see from the equations in advance, it turns out to be the only appropriate one when $\mu = O(1)$.

Since ϵ is a small parameter, it is natural to attempt a solution by expansion in powers of ϵ. Such an expansion is in general asymptotic and has no formal radius of convergence. In this respect it differs from the analysis made by Malkus & Veronis (1958) for perfectly conducting boundaries, where a formal zone of convergence exists. It seems likely that, as with many asymptotic expansions, the first terms do in fact provide an accurate description if ϵ is sufficiently small, but the proof of such a proposition seems impossible without detailed computation.

The bulk of the analysis we undertake will be for two-dimensional rolls. Although Busse & Riahi (1980) have shown that square cells may be the preferred planform in an infinite layer when the boundaries are slightly conducting, there are configurations for which only rolls are possible, although they are somewhat artificial. The main conclusions of the theory presented here appear to apply to general three-dimensional planforms, as discussed in the next section. The investigation of Busse & Riahi, based as it is on an expansion about a single horizontal wavenumber, is valid in a parameter range that is much more restrictive than that studied here (in particular, $\mu \ll 1$ is necessary for their expansions to be valid), and there is no guarantee that their results remain valid for the full range of μ. Since the two-dimensional analysis is so much simpler, it seems natural to use it as a starting point for understanding the general behaviour. We therefore suppose that $U_y = 0$ and that all variables are independent of y. Then $\mathbf{U}$ can be written in terms of a stream function $\psi(X, z)$;

$$\mathbf{U} = \left(\frac{\partial\psi}{\partial z}, 0, -\epsilon\frac{\partial\psi}{\partial X}\right). \tag{2.10}$$

If we now adopt the scaling of (2.8), (2.9) and in addition set $\psi = \epsilon\phi$, then (2.2) and the curl of (2.1) may be written, without approximation, as

$$\frac{1}{\sigma}\left\{\epsilon^6\frac{\partial}{\partial\tau}\left(\frac{\partial^2\phi}{\partial X^2}\right)+\epsilon^4\frac{\partial}{\partial\tau}(D^2\phi)+\epsilon^4\frac{\partial(\phi,\partial^2\phi/\partial X^2)}{\partial(X,z)}+\epsilon^2\frac{\partial(\phi,D^2\phi)}{\partial(X,z)}\right\}$$
$$=(R_c+\mu^2\epsilon^2)\frac{\partial\theta}{\partial X}+\epsilon^4\frac{\partial^4\phi}{\partial X^4}+2\epsilon^2\frac{\partial^2}{\partial X^2}(D^2\phi)+D^4\phi, \quad (2.11)$$

and

$$\epsilon^4\frac{\partial\theta}{\partial\tau}+\epsilon^2\frac{\partial(\phi,\theta)}{\partial(X,z)}=\epsilon^2\frac{\partial\phi}{\partial X}+\epsilon^2\frac{\partial^2\theta}{\partial X^2}+D^2\theta, \quad (2.12)$$

where $D\theta \equiv \partial\theta/\partial z$, etc. and $\partial(\ldots,\ldots)/\partial(X,z)$ is the usual Jacobian.

Since odd powers of ϵ do not appear in the expansion, we may expand θ and ϕ as power series in ϵ^2; thus we write

$$\left.\begin{aligned}\theta &= \theta_0(X,z,\tau)+\epsilon^2\theta_2(X,z,\tau)+\ldots,\\ \phi &= \phi_0(X,z,\tau)+\epsilon^2\phi_2(X,z,\tau)+\ldots,\end{aligned}\right\} \quad (2.13)$$

if we substitute these expressions into (2.11), (2.12) and equate powers of ϵ^2, the resulting sequence of problems can be solved successively. We undertake their solution in the following section: the aim is to find the dependence of θ_0 and ϕ_0 on X, z and τ, and this will require discussion of the equations for θ_2, θ_4 and ϕ_2.

3. Derivation of the governing equation

We now seek solutions to (2.11), (2.12) in powers of ϵ^2. We shall give solutions valid for all three sets of velocity boundary conditions discussed above. The differences arise only in the details of certain polynomials that appear in the analysis, and these are relegated to appendix A.

At $O(1)$ equation (2.12) yields

$$D^2\theta_0 = 0 \quad (3.1)$$

with $D\theta_0 = 0$, $z = \pm 1$. Thus

$$\theta_0 = f(X,\tau) \quad (3.2)$$

and the purpose of the subsequent analysis is to determine an equation for f.

Equation (2.11) gives

$$0 = R_c\frac{\partial\theta_0}{\partial X}+D^4\phi_0, \quad (3.3)$$

which is solved by

$$\phi_0 = R_cP(z)f', \quad (3.4)$$

where $f' \equiv \partial f/\partial X$, $D^4P(z) = -1$ and $P(z)$ satisfies appropriate boundary conditions.

At $O(\epsilon^2)$ we obtain

$$\frac{\partial(\phi_0,\theta_0)}{\partial(X,z)}=\frac{\partial\phi_0}{\partial X}+\frac{\partial^2\theta_0}{\partial X^2}+D^2\theta_2 \quad (3.5)$$

or

$$D^2\theta_2 = -R_cDPf'^2-f'''(R_cP+1). \quad (3.6)$$

This equation is an inhomogeneous boundary-value problem, and solutions only exist when the right-hand side satisfies a solvability condition. In the case at hand, we have, from the boundary condition on θ_2,

$$\int_{-1}^{1} D^2\theta_2 dz = [D\theta_2]_{-1}^{1} = 0 \tag{3.7}$$

and so, since $P(1) = P(-1) = 0$,

$$f''' \int_{-1}^{1} (R_c P(z) + 1)\, dz = 0. \tag{3.8}$$

Equation (3.8) determines R_c. As we shall see from appendix A, $R_c = 15/2$, 45 and 20 in cases A, B and C, respectively. With the appropriate value of R_c, (3.6) can be solved to give

$$\theta_2 = f_2(X,\tau) + f'^2 W(z) + f'' Q(z) \tag{3.9}$$

where

$$D^2 Q = -(R_c P + 1), \quad D^2 W = -R_c DP, \tag{3.10}$$

and f_2 is unknown. Then at $O(\epsilon^2)$, (2.11) yields

$$\frac{1}{\sigma}\frac{\partial(\phi_0, D^2\phi_0)}{\partial(X,z)} = R_c \frac{\partial\theta_2}{\partial X} + \mu^2 \frac{\partial\theta_0}{\partial X} + 2\frac{\partial^2}{\partial X^2}(D^2\phi_0) + D^4\phi_2. \tag{3.11}$$

If (3.4) and (3.9) are used to substitute for ϕ_0, θ_2 we can solve (3.11) to obtain

$$\phi_2 = R_c f_2' P + f''' U + \mu^2 f' P + f'' f' S, \tag{3.12}$$

where

$$\left.\begin{aligned} D^4 U &= -R_c Q - 2R_c D^2 P, \\ D^4 S &= -2R_c W + \sigma^{-1} R_c^2 \{P D^3 P - DP D^2 P\}. \end{aligned}\right\} \tag{3.13}$$

At $O(\epsilon^4)$ we have

$$\frac{\partial\theta_0}{\partial\tau} + \frac{\partial(\phi_2,\theta_0)}{\partial(X,z)} + \frac{\partial(\phi_0,\theta_2)}{\partial(X,z)} = \frac{\partial\phi_2}{\partial X} + \frac{\partial^2\theta_2}{\partial X^2} + D^2\theta_4 \tag{3.14}$$

and the solvability condition for this equation yields

$$\frac{\partial f}{\partial\tau} = -A\mu^2 f'' - Bf'''' + C(f'^3)' - D(f'f'')' \tag{3.15}$$

where

$$\left.\begin{aligned} A &= -\tfrac{1}{2}\int_{-1}^{1} P dz = 1/R_c, \quad B = -\tfrac{1}{2}\int_{-1}^{1} (U+Q)\, dz, \\ C &= \tfrac{1}{2}R_c^2 \int_{-1}^{1} P^2 dz, \quad D = \tfrac{1}{2}R_c \int_{-1}^{1} PDQ dz - \tfrac{1}{2}\int_{-1}^{1} (S+2W)\, dz. \end{aligned}\right\} \tag{3.16}$$

These non-negative constants depend on the boundary conditions. It is immediately clear from the equations that if both boundaries are fixed or both are free, then P and Q are even functions of z and S is odd. Thus $D = 0$ in both these cases. If we group these two problems as the 'symmetric case', then we see that in this case (3.15) is independent of σ, since the latter enters only through $S(z)$. The unknown function $f_2(X,\tau)$ does not appear in (3.15) and is determined at higher order in the expansion scheme. Table 1 shows the values of A, B, C, D in the different cases, and the method for determining them is shown in appendix A.

Case	A	B	C
R_c	15/2	45	20
A	2/15	1/45	1/20
B	1091/1386	34/231	232/693
C	155/126	10/7	760/567
D	0	0	$\frac{1}{3}+5/21\sigma$

TABLE 1

4. Steady periodic solutions of equation (3.15)

4.1. *The symmetric case*

We shall now suppose that $f(X,\tau)$, the leading order temperature perturbation, is independent of time, and thus obtain the possible forms of steady finite-amplitude convection. Which of these forms is stable must be decided by consideration of the full time-dependent equations.

First, we reduce (3.15) to canonical form by writing

$$\xi = \sqrt{\frac{A}{B}}X, \quad T = \frac{A^2}{B}\tau, \quad F = \sqrt{\frac{C}{B}}f, \tag{4.1}$$

whence (3.15) becomes

$$F_T = -\mu^2 F_{\xi\xi} - F_{\xi\xi\xi\xi} + (F_\xi^3)_\xi - \alpha(F_\xi F_{\xi\xi})_\xi, \tag{4.2}$$

where $\alpha = D/(BC)^{\frac{1}{2}}$. We now set $F_T = 0$ and (for the moment) $\alpha = 0$ also. One integration with respect to ξ then yields

$$0 = G_{\xi\xi} + \mu^2 G - G^3 + \beta; \quad G \equiv F_\xi, \tag{4.3}$$

where β is a constant. Recall that $G(\xi)$ is proportional to the horizontal velocity at any value of z. If we now seek solutions in a periodic box of length $\epsilon^{-1}(B/A)^{\frac{1}{4}}L$, where $L = 2\pi/k$, say is $O(1)$, then clearly not only must G be periodic, but it must have zero mean (otherwise F would not be periodic). If we multiply (4.3) by G_ξ and integrate we obtain

$$0 = \tfrac{1}{2}G_\xi^2 + \tfrac{1}{2}\mu^2 G^2 - \tfrac{1}{4}G^4 + \beta G + \gamma, \tag{4.4}$$

where γ is another constant. This is the equation for the 'position' G of a 'particle' in a potential well

$$W(G) = \tfrac{1}{2}\mu^2 G^2 - \tfrac{1}{4}G^4 + \beta G \tag{4.5}$$

as a function of 'time', ξ.

Figure 2 shows configurations of the well for $\beta > 0$, $\beta = 0$, $\beta < 0$. For a periodic solution the 'particle' must oscillate in the well. It is clear that if $\beta \neq 0$ the 'particle' oscillates about a non-zero value of G and it clearly spends more 'time' on one side of the origin than on the other; G then has a non-zero mean, and F is not periodic. Therefore $\beta = 0$ for periodic solutions.

The particle analogy shows that there are infinitely many periodic solutions of different periods corresponding to different total 'energies'. For the 'particle' to stay in the well, we need $|G| < \mu$, and the solutions of very large period clearly have $|G| \simeq \mu$ for arbitrarily long times.

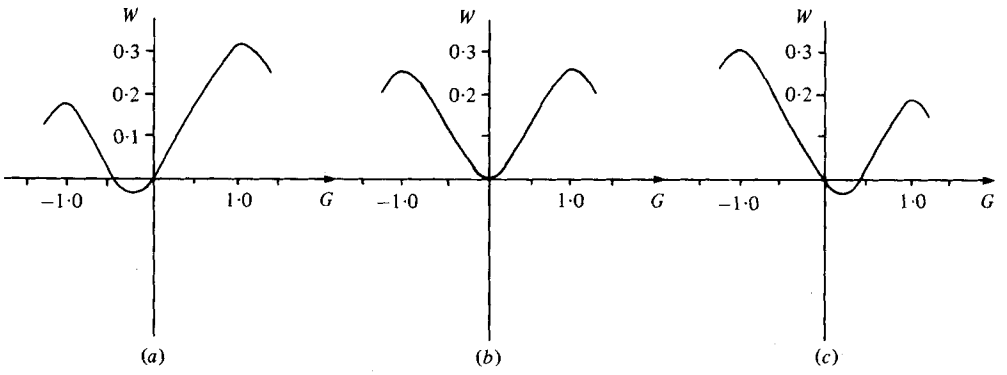

FIGURE 2. Sketches of the form of the potential ($W(G) = \frac{1}{2}G^2 - \frac{1}{4}G^4 + \beta G$, for (*a*) $\beta > 0$, (*b*) $\beta = 0$, (*c*) $\beta < 0$. Only in the second case can G have zero mean.

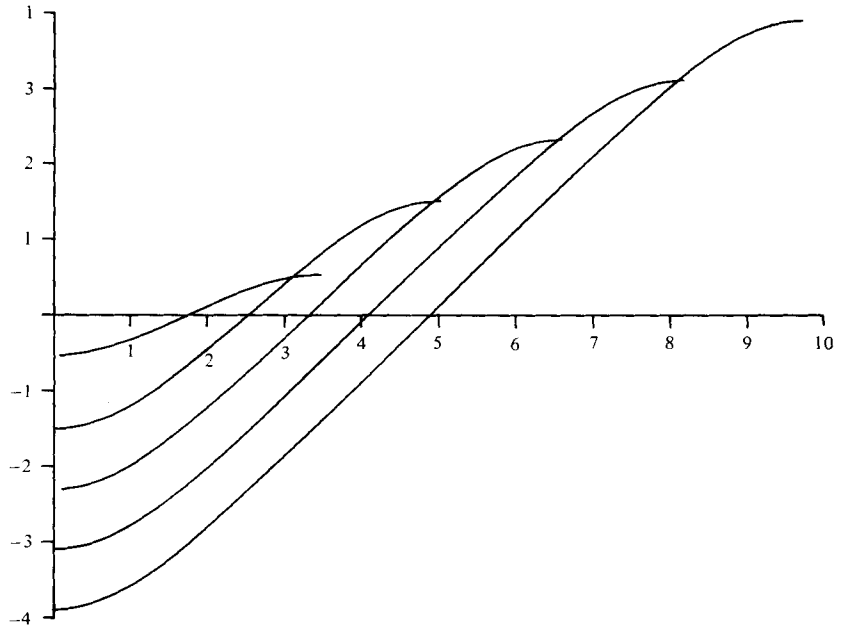

FIGURE 3. The temperature pertubation $F(X)$ in the steady state for various widths of box (solutions are impossible for widths $< \pi$) and $\alpha = 0$.

The boundary conditions at the edges of the periodic box, corresponding to zero horizontal motion, heat flux and tangential stress, can be written

$$F_\xi = F_{\xi\xi\xi} = 0 \quad \text{at} \quad \xi = 0, L \tag{4.6}$$

(the first of these conditions ensures satisfaction of both the first two constraints) and we adopt these henceforward. Now the region $0 \lesssim \xi \lesssim L$ contains two cells, between which the velocity reverses sign. Since the average of G is zero, G must vanish in the interval, at $\xi = \frac{1}{2}L$ by symmetry, and since $\beta = 0$, $G_{\xi\xi} = 0$ there also. Thus (4.6) may be applied at $\xi = 0, \frac{1}{2}L$, between which values there is just one cell.

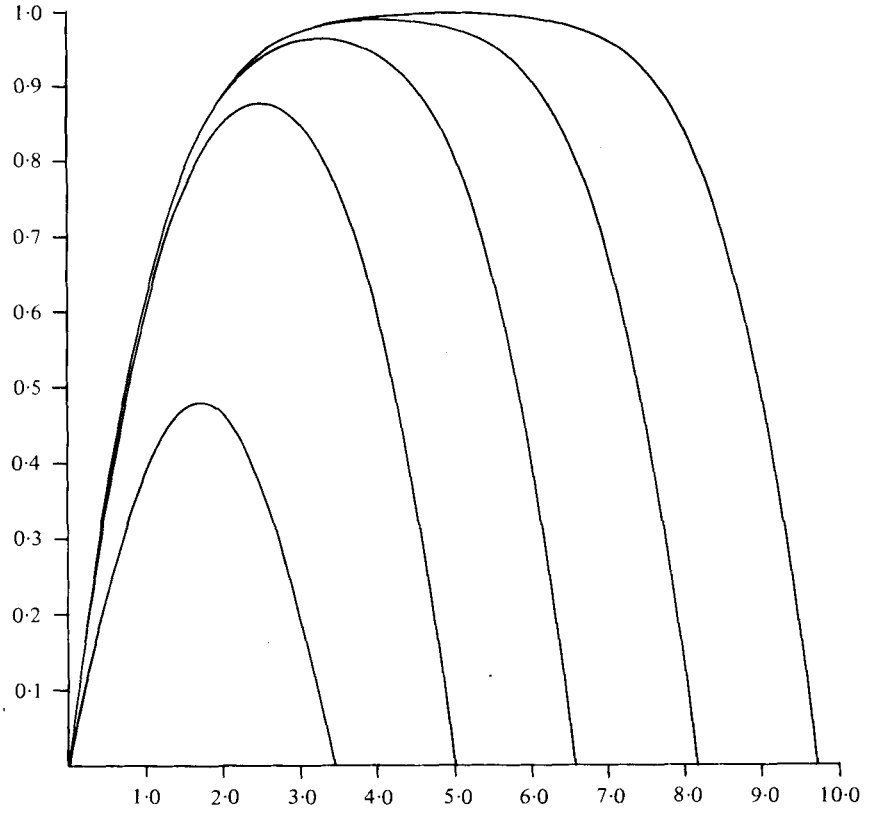

FIGURE 4. The function G, proportional to the horizontal velocity at any given value of z, in the same cases as for figure 3.

Equation (4.3) can be solved in terms of elliptic functions. If the parameter $m (0 < m < 1)$ is defined by

$$\left.\begin{aligned} \mu L &= 4(1+m)^{\frac{1}{2}} K(m) \\ K(m) &= \int_0^{\frac{1}{2}\pi} \frac{d\theta}{(1-m\sin^2\theta)^{\frac{1}{2}}} \end{aligned}\right\} \quad (4.7)$$

where

is the complete elliptic integral of the first kind, then

$$G(\xi) = \mu \left(\frac{2m}{1+m}\right)^{\frac{1}{2}} \mathrm{sn}\left(\frac{\mu\xi}{(1+m)^{\frac{1}{2}}} \middle| m\right). \quad (4.8)$$

Thus we see that $\mu L \geqslant 2\pi$ for a solution to be possible. When μL is close to 2π, m is small and so $G(\xi)$ is sinusoidal and of small amplitude. For large μL, on the other hand, $m \simeq 1$ and $G \sim \mu \tanh(\mu\xi/\sqrt{2})$. It will be seen that, although G is formally $O(1)$ in the scaling we have adopted, $G \to 0$ as $\mu \to 0 (R \to R_c)$ at fixed L.

Furthermore, there is no qualitative change in the solutions for different values of μ: indeed, μ can be eliminated from the solution by setting $\hat{\xi} = \mu\xi$, $\hat{G} = \mu G$. Since in any case we are interested in an experimental situation in which R is fixed and the

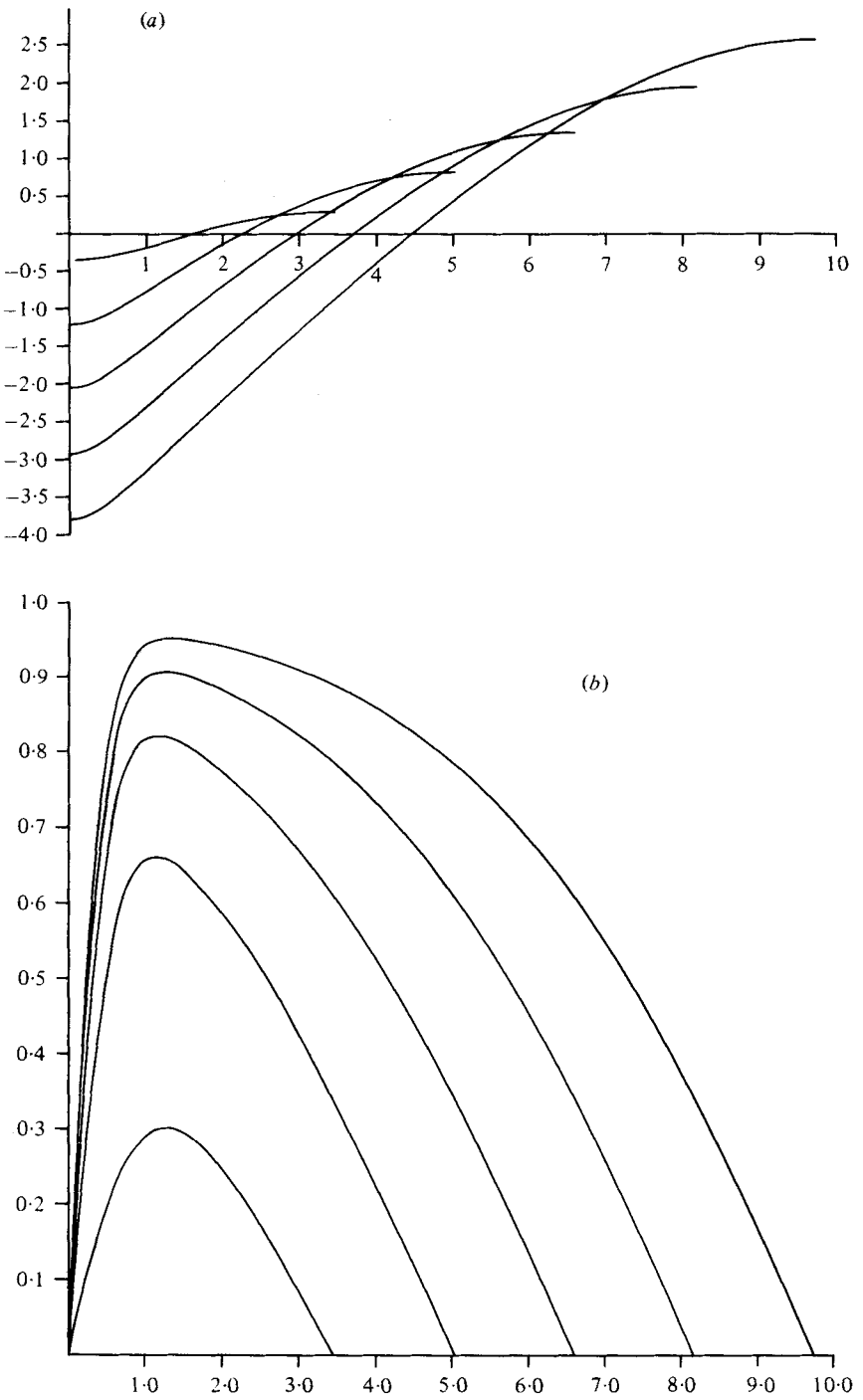

FIGURE 5. The functions (*a*) F and (*b*) G for various box widths (cf. figures 3 and 4) and $\alpha = 4$.

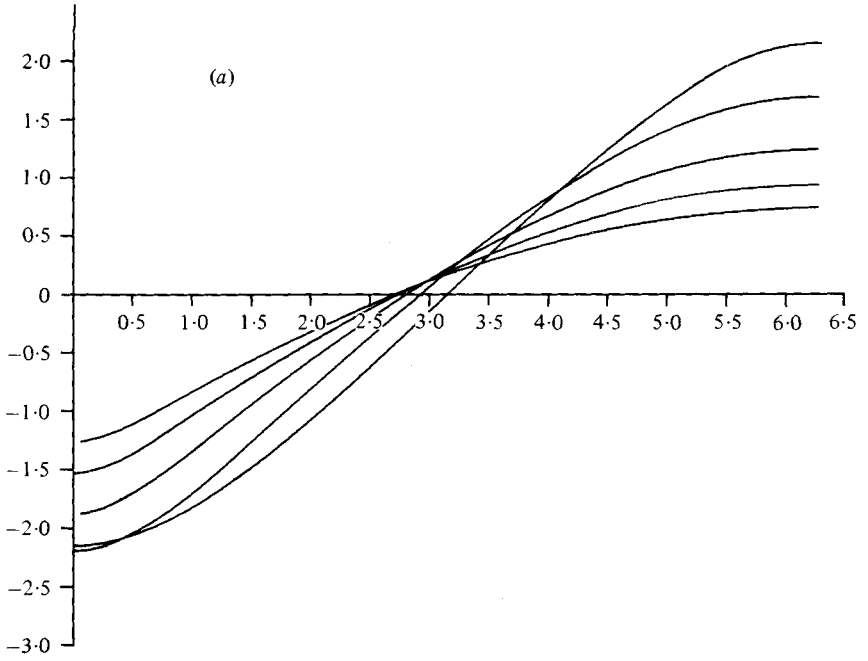

FIGURE 6(a). For legend see facing page.

disturbance is left to find its own length scale ,we now and henceforth set $\mu = 1$, so that ϵ is *defined* as $(R - R_c)^{\frac{1}{2}}$. (We could, as an alternative, fix L and regard μ as a parameter. Then there is a bifurcation at $\mu = 2\pi/L$, and the maximum amplitude of G increases monotonically with μ.)

For $\mu = 1$, (4.8) can be integrated once to yield

$$F(\xi) = \sqrt{2}\ln\left\{\left[\mathrm{dn}\left(\frac{\xi}{(1+m)^{\frac{1}{2}}}\middle| m\right) - m^{\frac{1}{2}}\,\mathrm{cn}\left(\frac{\xi}{(1+m)^{\frac{1}{2}}}\middle| m\right)\right]\Big/(1-m)^{\frac{1}{2}}\right\} \tag{4.9}$$

Figures 3 and 4 show F and G for different values of L. It should be emphasized again that these solutions are *all* possible in an infinite layer. If however, the convection is confined to a periodic 'box' of length L_0 the number of solutions depends on L_0. For $2\pi \leqslant L_0 \leqslant 4\pi$ only one solution is possible, for $4\pi \leqslant L_0 \leqslant 6\pi$ two solutions are possible, and so on.

4.2. *The asymmetric case* ($D \neq 0$)

Again we can integrate (4.2) once to obtain (with $\mu = 1$)

$$0 = G_{\xi\xi} + G - G^3 + \alpha G G_\xi + \beta. \tag{4.10}$$

We may set $\beta = 0$ if we assume the same boundary conditions on G as in the symmetric case. This choice gives G point symmetry about $\xi = \frac{1}{2}L$, which makes physical sense since one would not expect a cell rotating anticlockwise to look essentially different from one rotating clockwise. Furthermore, construction of the phase diagram for (4.10) reveals that the mean of G is zero only if $\beta = 0$. The solution to (4.10) can be reduced to quadratures (we are indebted to a referee for pointing this out), but it

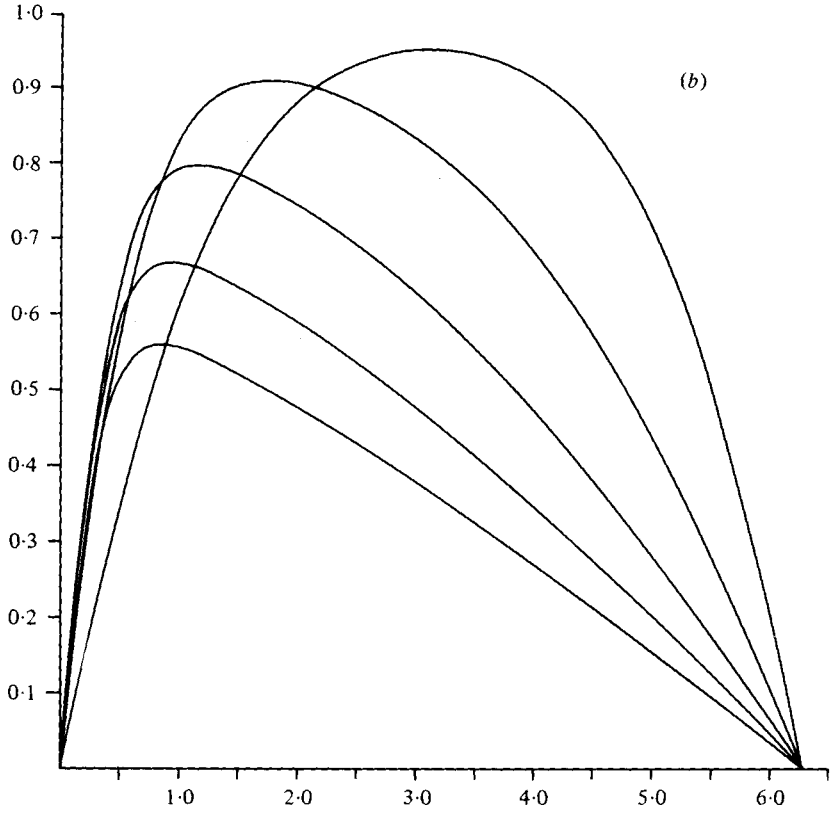

FIGURE 6. The functions (*a*) F and (*b*) G for a box of width 2π for $\alpha = 0$ (2) 8.

turns out that a numerical solution using a space discretization is the most efficient way of obtaining the form of G. If the region $0 \leqslant \xi \leqslant \frac{1}{2}L$ is divided into 40 mesh intervals, solutions can be obtained by an accelerated Newton–Raphson type procedure in less than 1*s* of CPU time on the Cambridge IBM 370/165. Figures 5 and 6 show solutions for F and G with various values of α and box length. It will be seen that as α increases the maximum of G moves to the left when $G > 0$ and that the region with $F \leqslant 0$ is narrower than that with $F \geqslant 0$. This is presumably because the slippery upper boundary allows the horizontal temperature gradients to drive more rapid motions, thus compressing the cooling fluid into a relatively narrow cold (descending) plume.

It has been noted elsewhere (Chapman *et al.* 1980) that (4.10) also arises, with $\alpha > 0$, when convection is due to any source of internal heating additional to or instead of heating from below. As pointed out in that paper, the internally heated case is relevant to the problem of convection in the earth's upper mantle. It seems probable that the term in α in (4.10) will parametrize a wide range of effects involving asymmetry between the top and bottom of the layer.

4.3. *The Prandtl-number dependence*

We have shown that the solution (4.8) does not contain the Prandtl number, even when the latter is $O(1)$; in the non-symmetric case, $\alpha = 0{\cdot}4976 + 0{\cdot}3554/\sigma$, and thus α for water ($\sigma = 6{\cdot}8$) differs significantly from that for mercury ($\sigma = 0{\cdot}025$). We do not expect the theory to remain accurate for very small σ, since if $\sigma = O(\epsilon)$ the Reynolds number of the evolved flow is $O(1)$. It is thus potentially subject to shear flow instability, which will presumably destroy the assumption of long horizontal length scales. Thus $\sigma \gg \epsilon$ seems to be a necessary condition for the analysis to be valid. It is also likely that the condition is sufficient, though this is hard to prove.

4.4. *Three-dimensional solutions at infinite Prandtl number*

The analysis of this paper would not be complete without a brief discussion of solutions that depend on y as well as on x. Clearly the appropriate length scale in the y direction will be $O(\epsilon^{-1})$ times the layer depth. If we suppose that $\sigma \to \infty$ (a circumstance approximated closely in the Earth's upper mantle), then the resulting velocity field has no vertical vorticity, so that it is purely poloidal. It may thus be represented in terms of a single scalar function $\tilde{P}(x, y, z)$, where

$$\mathbf{U} = \nabla \wedge \nabla \wedge (\tilde{P}\hat{z}). \tag{4.11}$$

(If $\partial \mathbf{U}/\partial y = U_y = 0$, $\tilde{P}$ is related to the ϕ of the last section by $\phi = -\partial \tilde{P}/\partial X$.) The analysis can then be carried out in terms of $\tilde{P}$ and θ. In the symmetric case, this yields an equation for the dominant temperature perturbation $F(\xi, \eta, T)$ (where $\eta = (A/B)^{\frac{1}{2}} \epsilon y$, etc. and A, B, etc. are all as in § 3), in the canonical form

$$F_T = -\nabla_H^2 F - \nabla_H^4 F + \nabla_H . (\nabla_H F |\nabla_H F|^2) \tag{4.12}$$

where $\nabla_H = (\partial/\partial\xi, \partial/\partial y)$.

Equation (4.12) reduces to equation (4.2) if $\partial F/\partial\eta = 0$, $\alpha = 0$ and $\mu = 1$. No simple first integral of the equations appears to be available even when $F_T = 0$, but it is clear that steady solutions are possible with hexagonal and rectangular planforms in addition to the roll solutions already found. Busse & Riahi (1980) show that in the case of imperfectly insulating boundaries square cells are preferred when the amplitude of the perturbation is very small. In the present case it seems unlikely that roll solutions will be preferred in an infinite layer, since the system selects the longest horizontal scale available to it, as shown below. However, there are circumstances when only roll solutions are possible and a somewhat wider range in which rolls are stable. We amplify and justify these remarks in the next section.

5. Time-dependent solutions and stability

5.1. *Linearized theory*

We have shown in the last section that two-dimensional disturbances can grow provided their period is sufficiently long. When these modes have small amplitude their growth rate can be obtained by linearizing equation (4.2) (for $\mu = 1$), so that F satisfies

$$F_T = -F_{\xi\xi} - F_{\xi\xi\xi\xi}, \tag{5.1}$$

and solutions exist in the form

$$F(\xi, T) = e^{ik\xi + sT}, \tag{5.2}$$

where

$$s = k^2 - k^4. \tag{5.3}$$

Thus when the disturbance is weak, the most rapidly growing mode has $k = 1/\sqrt{2}$. Linear theory then predicts that cells of width $\xi = \pi\sqrt{2}$ will be present at large times. However the linear theory is misleading, as we shall see. Investigation of the nonlinear equation (4.2) reveals that all modes in a periodic box are unstable to ones of greater wavelength. Thus eventually there will be only one roll in the box. The time taken for this 'eating up' of smaller cells by larger ones is in general larger than the time taken for the $k = 1/\sqrt{2}$ solution to be established. If the top and bottom boundaries are not perfectly insulating, wavelengths cannot become arbitrarily long, as previously mentioned (see Hurle *et al.* 1967); however, in that case the most stable wavelength is in general much larger than that corresponding to maximum growth rate, so that the perfect problem represents an appropriate limit.

We first of all derive an exact (if (4.2) is assumed) evolution equation in the case $\alpha = 0$ which shows that a steady solution cannot lose stability to a mode with a shorter wavelength. This does not tell us whether a mode of any given wavelength is stable to *small* disturbances, even of longer wavelength, and a variational problem is then solved numerically to show that in fact all solutions are locally unstable to modes longer than themselves. As a final test, numerical solutions of (4.2) were found by marching forward in time from an arbitrary initial state. The use of the computer allowed the case $\alpha \neq 0$ to be studied: in all runs the eventual steady state was one with as few nodes as possible.

5.2. *The evolution equation for* $\alpha = 0$

We now consider (4.2) with $\alpha = 0$ in a periodic box, at the boundaries of which $F_\xi = F_{\xi\xi\xi} = 0$. Let angle brackets denote a space average over this box. Then, following a suggestion of S. Childress, consider the functional

$$V[F, T] = \langle \tfrac{1}{4}F_\xi^4 + \tfrac{1}{2}F_{\xi\xi}^2 - \tfrac{1}{2}F_\xi^2 \rangle. \tag{5.4}$$

(Note that V is the mean 'Lagrangian' in the particle analogy.)

From the equation for $F(\xi, T)$ we have

$$\frac{dV}{dT} = \langle F_\xi^3 F_{\xi T} + F_{\xi\xi} F_{\xi\xi T} - F_\xi F_{\xi T} \rangle \tag{5.5}$$

$$\left.\begin{aligned} &= \langle F_T \{F_{\xi\xi} + F_{\xi\xi\xi\xi} - (F_\xi^3)_\xi\} \rangle \\ &= -\langle F_T^2 \rangle \end{aligned}\right\} \tag{5.6}$$

after integrating by parts and using (4.2). Thus V always decreases while the system evolves to a steady state. In such a state, $dV/dT = 0$ and (from (4.2) again)

$$0 = \langle F_\xi^2 - F_{\xi\xi}^2 - F_\xi^4 \rangle \tag{5.7}$$

so that if $dV/dT = 0$ then

$$V = -\tfrac{1}{4}\langle F_\xi^4 \rangle = -\tfrac{1}{4}\langle G^4 \rangle. \tag{5.8}$$

From (4.8) it now follows that V in a steady state is a decreasing function of wavelength. We can also see that the Euler–Lagrange equation for a stationary value of V (at fixed T) subject to the boundary conditions on F is

$$-(F_\xi^3)_\xi + F_{\xi\xi\xi\xi} + F_{\xi\xi} = 0, \tag{5.9}$$

which is just the condition that F is a steady solution of (4.2). Thus for such a solution V is stationary with respect to small perturbations. If V is a local maximum, the solution is unstable, since V can never increase; if V is a (local) minimum the solution is (locally) stable. Thus if the box is sufficiently long to contain several different steady solutions of different periods, there is no possible perturbation to an initial periodic state of given wavelength that can result in a final steady state of shorter wavelength (if it did, V would have to increase at some point).

Thus the only possibilities left are (i) that all possible modes in the box are unstable to those of *larger* wavelength, if such exist, and (ii) that some modes are locally stable to all small disturbances. To find out more we must consider the second variation of V.

5.3. *The stability criterion*

Let $F_0(\xi)$ be a steady solution of (4.2) with $\alpha = 0$ and let $G_0 = F_{0\xi}$. Then consider the value of V for $F = F_0 + \zeta(\xi)$ at fixed T, where $|\zeta|$ is so small that all terms of $O(|\zeta|^3)$ or smaller may be ignored; we have

$$V[F_0+\zeta] = V[F_0] + \langle \zeta_\xi (F_{0\xi}^3 - F_{0\xi\xi\xi} - F_{0\xi}) \rangle + \langle \tfrac{3}{2} F_{0\xi}^2 \zeta_\xi^2 + \tfrac{1}{2}\zeta_{\xi\xi}^2 - \tfrac{1}{2}\zeta_\xi^2 \rangle + O(|\zeta|^3). \tag{5.10}$$

The second term on the right-hand side vanishes by virtue of (4.2) and so the second variation of V is

$$\begin{aligned} \delta^2 V &= V[F_0+\zeta] - V[F_0] \\ &= \tfrac{1}{2}\langle (3G_0^2 - 1)\,\zeta_\xi^2 + \zeta_{\xi\xi}^2 \rangle \end{aligned} \tag{5.11}$$

and we are interested in whether $\delta^2 V$ is positive for all ζ that satisfy the conditions on F at the ends of the box. ζ, like F, is only determined up to a constant, so let us write $\chi = \zeta_\xi$ and make the problem homogeneous by seeking

$$M(\chi) = \min_\chi \frac{\langle (3G_0^2 - 1)\,\chi^2 + \chi_\xi^2 \rangle}{\langle \chi^2 \rangle} \tag{5.12}$$

subject to the conditions $\chi = \chi_{\xi\xi} = 0$ at the ends of the box. If $M < 0$ there is some disturbance χ which will make $\delta^2 V$ negative. Then the solution F_0 is unstable since V cannot increase. If $M > 0$ then F_0 is locally stable. (In fact exactly the same variational problem can be derived from direct consideration of the linear perturbation equations satisfied by ζ.)

The Euler–Lagrange equation for (5.12) is

$$\chi_{\xi\xi} + (1 + M - 3G_0^2)\chi = 0, \tag{5.13}$$

and it can be seen that the four boundary conditions given after (5.12) can be satisfied provided only that $\chi = 0$ at the ends.

We first note that if the box is just long enough ($2\pi + \delta$ where $\delta \ll 1$) to contain either one cell or two cells, then M is certainly negative. For in this case $3G_0^2 \ll 1$ for the two-cell mode [$G_0 = (8\delta/3)^{\frac{1}{2}} \sin\xi + 0(\delta^{\frac{3}{2}})$] and so there is a solution $\chi \simeq \sin\frac{1}{2}\xi$ when $M \simeq -0{\cdot}75$. We have not been able to solve (5.13) analytically when G_0 is finite, but

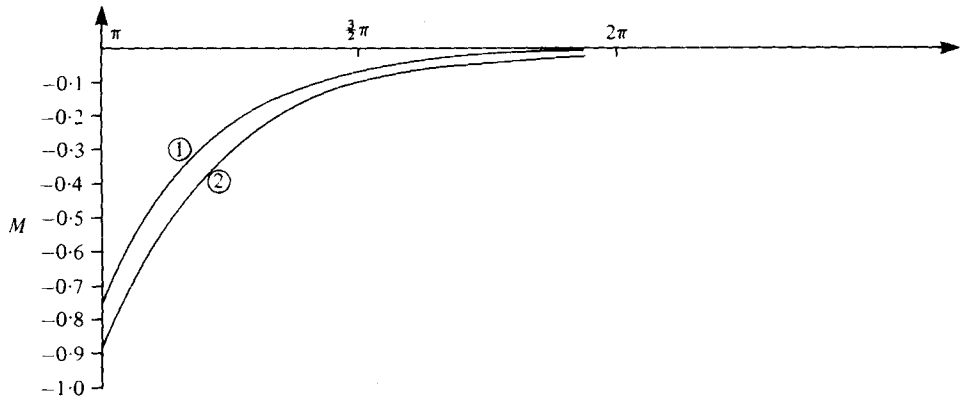

FIGURE 7. The stability parameter M plotted against the cell width of the lower wavelength mode when the perturbation has (1) twice and (2) three times its wave length. It will be seen that $M < 0$, implying instability. The wavelengths for which the perturbation grows as fast as the basic solution on linear theory are $\frac{1}{2}\pi\sqrt{5}$ and $\frac{1}{3}\pi\sqrt{10}$ respectively.

M was found for various lengths of box by discretizing equation (5.13) using finite differences. If, for some integer $n > 0$

$$\Delta\xi = L/n, \quad \xi_i = i\Delta\xi, \text{etc.} \tag{5.14}$$

where L is the length of the box, then, correct to $O(\Delta\xi^2)$, (5.13) becomes

$$A_{ij}\chi_j + M\chi_i + B_{ij}\chi_j = 0 \quad (0 \leqslant i \leqslant n)$$

where

$$\left.\begin{aligned} B_{ii} &= 1 - 3G_{0i}^2, \quad B_{ij} = 0, \quad i \neq j, \\ A_{ii} &= -2/\Delta\xi^2, \quad A_{i+1,j} = A_{i-1,j} = 1/\Delta\xi^2, \\ & \qquad A_{ij} = 0, \quad |i-j| > 1. \end{aligned}\right\} \tag{5.15}$$

Thus the problem is reduced to that of finding the (algebraically) smallest eigenvalue of a tridiagonal matrix, which can be accomplished very rapidly (even for large n), for example by the method of bisections. Computations were carried out in boxes twice as long and three times as long as the cell under study, for various box lengths. The results are plotted in figure 7. It will be seen that M is always negative in both cases, though it quickly approaches zero as the period becomes larger. For very long periods $|M|$ is so small that the method is no longer accurate. It seems clear, however, that M is a smooth function of the period in each case, and so the transition from stability to instability (if it occurs) should take place via a 'neutral' mode with $M = 0$.

There is in fact never such a mode, for any box length: if $M = 0$, then the resulting equation

$$\chi_{\xi\xi} + (1 - 3G_0^2)\chi = 0 \tag{5.16}$$

has the solution $G_{0\xi}$. This is periodic with the correct period but does not satisfy the boundary conditions at $\xi = 0, L$. It is then a consequence of Floquet's theorem that the other independent solution of (5.16) can not have period L (Minorsky 1974). Thus an $M = 0$ mode does not exist, and so there seems no way in which M can pass through zero.

Thus we have shown that all steady solutions are unstable to modes of longer wave-

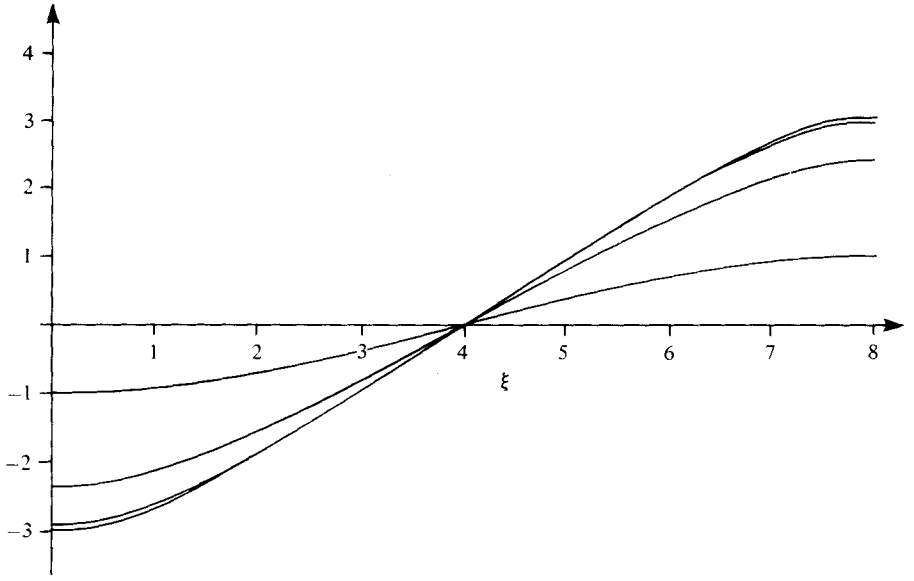

FIGURE 8. Time development of the function $F(\xi,T)$ in a box of length 8, starting from $F = -\cos(\pi\xi/8)$, with $\alpha = 0$; F is plotted at $T = 0$, 10, 20, 30 units.

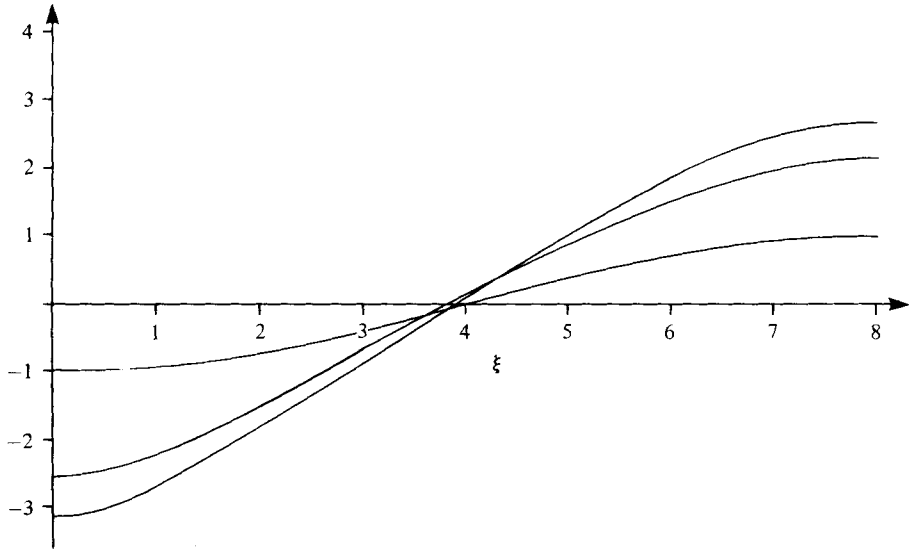

FIGURE 9. As figure 8, but with $\alpha = 1$. Convergence occurred after only 20 units in this case.

length, if the box is sufficiently long. This would not be surprising if the linear growth rate of the longer mode was the greater, but the phenomenon persists well beyond box lengths greater than $\pi\sqrt{5}$, which is the point at which cells of width L have slower growth rates than those of period $L/2$. Thus the linear theory is misleading in this context.

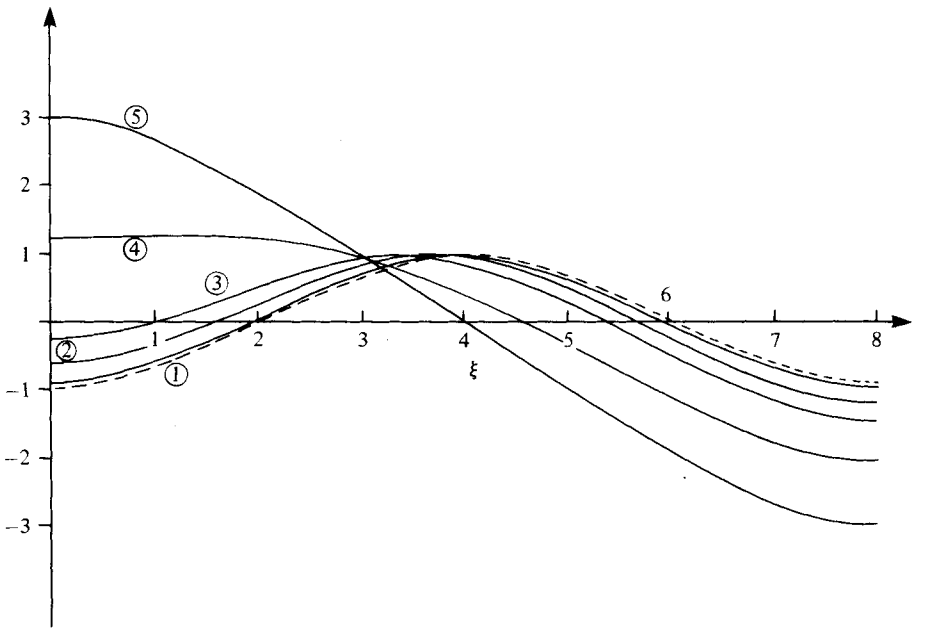

FIGURE 10. Showing the breakdown of a two-cell mode and its replacement by a one-cell mode in a box of length 8. The dotted line shows the original steady state before perturbation. Graphs are shown on $T = 0, 20, 40, 60, 80$.

5.4. *Numerical solution*

We decided to test the predictions of the last section in an entirely independent way by solving (4.2) as an initial value problem. Such a scheme also allows consideration of the case $\alpha \neq 0$, which defies analytical solution. An explicit leap-frog scheme was used for the time derivative, and the space derivatives were obtained using a scheme of DuFort-Frankel type (see, for example, Moore & Weiss 1973). Naive application of the method led to a rapidly growing instability on the scale of the mesh interval, but this was removed with a slight adjustment to the time derivative. The final steady states of F and G obtained by the theory were in excellent agreement with the steady solutions found earlier by a different method, and the growth rates in the linear regime correspond well with (5.3). Forty mesh points were used in most runs; this gave errors less than 1 % (though double-precision arithmetic had to be used).

We first investigated the approach to the steady state by starting from a small value of F and marching forwards in time. Figures 8 and 9 show a typical development for $\alpha = 0$ and $\alpha \neq 0$ respectively. No anomalous behaviour was found. The numerical scheme could be 'fooled', by appropriate symmetry in the initial conditions, into reaching a steady state with half the period of the usual steady solution. Having obtained this 'fooled' steady state we perturbed it slightly with a larger period mode and allowed the system to evolve in time. The box length chosen was 8; this is greater than $\pi\sqrt{5}$ so the two-cell mode has faster growth rate (see discussion in §5.3). Nonetheless for both $\alpha = 0$ and $\alpha \neq 0$, the perturbation grew from a very small level and

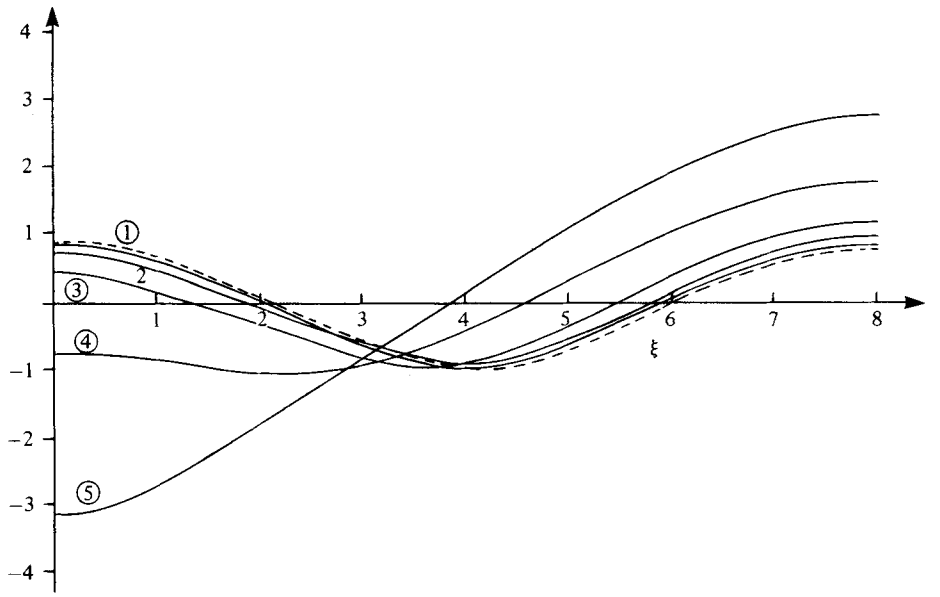

FIGURE 11. As for figure 10, but with $\alpha = 1$.

the eventual outcome was a long wavelength mode filling the box. The time development in these two cases is shown in figures 10 and 11. Other runs were performed for $\alpha \neq 0$; in no case did we find that the long wavelength perturbations decayed.

5.5. *The three-dimensional problem with* $\alpha = 0$ *and* $\sigma \to \infty$

The evolution equation approach can also be applied here. Consider

$$V_3[F,T] = \langle \tfrac{1}{4}|\nabla_H F|^4 + \tfrac{1}{2}(\nabla_H^2 F)^2 - \tfrac{1}{2}|\nabla_H F|^2 \rangle \tag{5.17}$$

where the angle brackets now denote an average over a periodic box in (ξ, η) space (or possibly the whole space if a hexagonal tesselation is being considered). Then, just as before

$$\frac{dV_3}{dT} = -\langle F_T^2 \rangle \tag{5.18}$$

and in a steady state

$$V_3 = -\tfrac{1}{4}\langle |\nabla_H F|^4 \rangle \tag{5.19}$$

just as for the two-dimensional case. Thus the globally stable mode in any periodic box is the one with the smallest value of V_3. It is immediately obvious that for many boxes two-dimensional rolls are globally stable. Consider for example a box $0 \leqslant \xi \leqslant L$, $0 \leqslant \eta \leqslant L$. Then we know that a two-dimensional roll solution can exist in the box if $L \geqslant \pi$. However, the steady linear problem for square cells has the form

$$0 = (\nabla_H^4 + \nabla_H^2) F, \tag{5.20}$$

and this has a solution of the form $F = \cos a\xi \cos b\eta$ if $a^2 + b^2 = 1$. Thus a square cell solution can exist only for $a = b = 1/\sqrt{2}$ and so $L \geqslant \pi\sqrt{2}$ for the cell to fit in the box.

So for $\pi \leqslant L \leqslant \pi\sqrt{2}$ a roll is the only solution: for a wide range of values of $L \leqslant \pi\sqrt{2}$ the roll solution will have much larger amplitude and thus a smaller value of V_3. This is the basis of the remark made in the last section that rolls may be stable in the case $\alpha = 0$, but a full proof will have to await the development of a suitable numerical scheme.

6. Conclusion

In previous sections we have shown that when the Rayleigh number is close to the critical value R_c for the existence of stationary disturbances with very small wavenumbers, convection can take place only on long horizontal length scales. Using the expansion technique of Childress & Spiegel (1980), we have been able to exploit this fact to obtain a description of the convection that does not involve the vertical displacement except as a passive parameter. This simplification allows an analytic solution to be obtained in certain circumstances even when the disturbance is fully nonlinear. It should once again be emphasized that the expansion is *not* of the Landau or 'modified-perturbation' type based on a single wavenumber; the expansion parameter is the horizontal wavenumber (related to the magnitude of $R - R_c$) but fully nonlinear solutions are permitted at leading order. The nonlinear analysis, backed up by numerical computation, reveals that the solution will always eventually exhibit the largest wavelength available to it, in contradiction to the predictions of linear theory alone.

When the boundaries are not perfect insulators, the wavelength cannot be arbitrarily large. Preliminary work on the 'imperfect problem' indicates, however, that in this case the most stable wavelength is much greater than that of the most rigidly growing linear mode. The question of three-dimensional perturbations has only been considered in passing: it is clearly desirable to have a numerical scheme that would treat the general three-dimensional geometry as an initial-value problem. The boundary conditions at the edges of the box used in this work are somewhat abstract, and are really no more than periodicity conditions. More realistic boundary conditions would be $F_\xi = F_{\xi\xi} = 0$ corresponding to solid, perfectly insulating boundaries. There is a perfectly good steady solution of (4.2) with these conditions corresponding to a single cell in the box, but there are no non-trivial solutions corresponding to two or more cells. This is because at leading order the equation is of fourth rather than sixth order in space and so the full set of boundary conditions will in general overdetermine the problem. In a long box containing many cells there is presumably a relatively thin boundary region where horizontal length scales are $O(1)$ and the extra derivatives provide a means of matching the boundary conditions. In the interior the solutions will resemble closely those found already for the periodic conditions. It seems likely that for the viscous edge conditions the most stable mode will be one whose wavelength is of the order of the width of the box, though it may not be the largest one possible. This aspect certainly deserves further study.

The analysis presented in this paper has shown the enormous influence of the thermal boundary condition on the form and aspect ratio of the developing and developed convection cells. The two-scale method, applied here for the first time to nonlinear thermal convection, can be used also in the cases where the boundaries are imperfectly insulating, and makes possible for the first time the detailed treatment of

fully nonlinear three-dimensional cellular convection with more general boundary conditions. The possibilities are manifold; the present paper represents only a first step.

This work was begun at the 1978 Woods Hole Geophysical Fluid Dynamics Summer Program. Thanks are due to the Director, W.V.R. Malkus, and to the National Science Foundation, who made our visits possible. It was completed while one of us (C.J.C.) was in receipt of an S.R.C. Studentship. We acknowledge a great debt to S. Childress and E.A. Spiegel, who provided many of the ideas in the paper, and communicated the results of their work on bio-convection prior to publication, and we also thank L.N. Howard, N.O. Weiss and students at the Summer Programme for helpful discussion.

Appendix A. Evaluations of polynomials and integrals occurring in §§ 2 and 3

We first note that all the integrals in (3.16) (not just A and C) can be expressed in terms of the two polynomials $P(z)$ and $DQ(z)$ which eliminates the need to evaluate $U(z)$, $W(z)$, $S(z)$. If we write $\int_{-1}^{1} U dz = \int U$, etc. then the equation $D^4P = -1$, together with the boundary conditions on P and U, give

$$\int U = -\int D^4P.U = -\int PD^4U \tag{A 1}$$

$$= \int P(R_cQ + 2R_cD^2P) \quad \text{(from (3.13))} \tag{A 2}$$

$$= -\int (1+D^2Q)\,Q - 2R_c\int (DP)^2 \quad \text{(from (3.10))} \tag{A 3}$$

so

$$B = -\tfrac{1}{2}\int (U+Q) = R_c\int (DP)^2 - \tfrac{1}{2}\int (DQ)^2. \tag{A 4}$$

Similarly

$$\int S = -\int PD^4S$$

$$= 2R_c\int WP - \sigma^{-1}R_c^2\int (P^2D^3P - PDPD^2P) \quad \text{(from (3.13))}$$

$$= -2\int (1+D^2Q)\,W - \tfrac{3}{2}\sigma^{-1}R_c^2\int (DP)^3 \quad \text{(from (3.10))}$$

$$= -2\int W - 2\int R_cPDQ - \tfrac{3}{2}\sigma^{-1}R_c^2\int (DP)^3 \quad \text{(from (3.10))} \tag{A 5}$$

so

$$D = \tfrac{1}{2}R_c\int PDQ - \tfrac{1}{2}\int (S+2W) = \tfrac{3}{2}R_c\int PDQ + \tfrac{3}{4}\sigma^{-1}R_c\int (DP)^3. \tag{A 6}$$

All that remains is to give P and DQ in the three different cases, using $D^4P = -1$ and equation (3.10):

(i) Free–free: $P = D^2P = DQ = 0$, $z = \pm 1$,

$$P(z) = -\tfrac{1}{24}(4 - 6z^2 + z5), \tag{A 7}$$

$$R_c = -2/\int P = 1\tfrac{5}{2}, \tag{A8}$$

$$DQ = \tfrac{1}{16}(9z - 10z^3 + z^5). \tag{A 9}$$

(ii) Solid–solid: $P = DP = DQ = 0,\ z = \pm 1,$

$$P(z) = -\tfrac{1}{24}(z^4 - 2z^2 + 1), \tag{A 10}$$

$$R_c = 45, \tag{A 11}$$

$$DQ = \tfrac{1}{8}(7z - 10z^3 + 3z^5). \tag{A 12}$$

(iii) Solid–free: $P = DQ = 0,\ z = \pm 1;\ DP = 0,\ z = -1;\ D^2P = 0,\ z = +1,$

$$P(z) = -\tfrac{1}{24}(z^4 - z^3 - 3z^2 + z + 2), \tag{A 13}$$

$$R_c = 20, \tag{A 14}$$

$$DQ = \tfrac{1}{24}(-5 + 16z + 10z^2 - 20z^3 - 5z^4 + 4z^5). \tag{A 15}$$

A, B, C, D, are now obtainable by elementary integration.

Appendix B. Energy-stability theory for an arbitrary disturbance

We take the scalar product of (2.1) with $\mathbf{U}$, the product of (2.2) with θ, use (2.3) and take an average over the whole layer, assuming the disturbance is bounded so the averages exist. This leads to the integral relations

$$\frac{1}{\sigma}\frac{dE_U}{dt} = R\langle\theta\mathbf{U}.\hat{\mathbf{z}}\rangle - \langle|\nabla\mathbf{U}|^2\rangle \tag{B 1}$$

and

$$\frac{dE_\theta}{dt} = \langle\theta\mathbf{U}.\hat{\mathbf{z}}\rangle - \langle|\nabla\theta|^2\rangle, \tag{B 2}$$

where $E_U = \frac{1}{2}\langle|\mathbf{U}|^2\rangle$, $E_\theta = \frac{1}{2}\langle|\theta|^2\rangle$ and brackets denote averages. If we introduce the parameter γ^2 we can easily derive the inequality

$$\frac{d}{dt}(\sigma^{-1}E_U + \gamma^2 E_\theta) \leqslant (<|\nabla\mathbf{U}|^2\rangle + \gamma^2\langle|\nabla\theta|^2\rangle)\left[\frac{R+\gamma^2}{2H\gamma} - 1\right] \tag{B 3}$$

where

$$H^2 = \min_{\mathbf{U},\theta}\frac{\langle|\nabla\theta|^2\rangle\langle|\nabla\mathbf{U}|^2\rangle}{\langle\theta\mathbf{U}.\hat{\mathbf{z}}\rangle^2}$$

among all fields θ, $\mathbf{U}$ satisfying the boundary conditions and $\nabla.\mathbf{U} = 0$. (We have used the fact that $\langle|\nabla\mathbf{U}|^2\rangle + \gamma^2\langle|\nabla\theta|\rangle \geqslant 2\gamma\langle|\nabla\mathbf{U}|^2\rangle^{\frac{1}{2}}\langle|\nabla\theta|^2\rangle^{\frac{1}{2}}$.)

Now $(R+\gamma^2)/2H\gamma^{-1}$ has a minimum value $(R-H^2)/2H^2$, attained when $\gamma = H$. Thus if $R < H^2$ the right-hand side of (B 3) is negative for $\gamma = H$; since $\langle|\nabla\mathbf{U}|^2\rangle + \gamma^2\langle|\nabla\theta|^2\rangle/(\sigma^{-1}E_U + \gamma^2 E_\theta)$ is bounded below when the boundary conditions on $\mathbf{U}, \theta$ are taken into account, this implies exponential decay of $\mathbf{U}, \theta$. But the Euler-Lagrange equations for the variational problem determining H are precisely the linearized forms of (2.1)–(2.3) with $\partial/\partial t = 0$ and H^2 replacing R. Thus for any 'periodic box' of length L, H^2 is equal to the critical Rayleigh number corresponding to that value of L. This shows that the short-wave cut off that holds when R is very close to R_c applies to nonlinear as well as linear disturbances.

REFERENCES

ABRAMOWITZ, M. & STEGUN, I. A. 1964 *Handbook of Mathematical Functions*. Washington: National Bureau of Standards.

BACKUS, G. E. 1958 A class of self-sustaining dissipative spherical dynamos. *Ann. Phys.* **4**, 372–447.

BUSSE, F. H. 1978 Non-linear properties of thermal convection. *Rep. Prog. Phys.* **41**, 1929–1937.

BUSSE, F. H. & RIAHI, N. 1980 Nonlinear convection in a layer with nearly insulating boundaries *J. Fluid Mech.* **96**, 243–256.

CHANDRASEKHAR, S. 1961 *Hydrodynamic and Hydromagnetic Stability*, Clarendon.

CHAPMAN, C. J. 1978 Bénard convection with constant heat flux boundaries. *Woods Hole Summer Fellowship Lectures*, vol. 2, 1–17.

CHAPMAN, C. J. 1980 Bénard convection between insulating boundaries. Ph.D. thesis, Bristol University.

CHAPMAN, C. J., CHILDRESS, S. & PROCTOR, M. R. E. 1980 Long wavelength thermal convection between non-conducting boundaries. *Earth & Planet. Sci. Letters* (in press).

CHILDRESS, S., LEVANDOWSKY, M. & SPIEGEL, E. A. 1975 Pattern formation in a suspension of swimming micro-organisms: equations and stability theory. *J. Fluid Mech.* **63**, 591–613.

CHILDRESS, S. & SPIEGEL, E. A. 1980 Pattern formation in a suspension of swimming micro-organisms: nonlinear aspects. (To be published.)

HURLE, D. T. J., JAKEMAN, E. & PIKE, E. R. 1967 On the solution of the Bénard problem with boundaries of finite conductivity. *Proc. Roy. Soc.* A **296**, 469–475.

JAKEMAN, E. 1968 Convective instability in fluids of high thermal diffusivity. *Phys. Fluids* **11**, 10–14.

JEFFREYS, H. 1926 The stability of a layer heated from below. *Phil. Mag.* **2**, 833–844.

JOSEPH, D. D. 1976 *Stability of Fluid Motions* (2 vols), Springer.

MALKUS, W. V. R. & VERONIS, G. 1958 Finite-amplitude cellular convection. *J. Fluid Mech.* **4**, 225–260.

MINORSKY, N. 1974 *Nonlinear Oscillations*, Krieger.

MOORE, D. R. & WEISS, N. O. 1973 Two-dimensional Rayleigh–Bénard convection. *J. Fluid Mech.* **58**, 289–312.

NEWELL, A. C. & WHITEHEAD, J. A. 1969 Finite bandwidth finite-amplitude convection. *J. Fluid Mech.* **38**, 279–303.

NIELD, D. A. 1967 The thermohaline Rayleigh–Jeffreys problem. *J. Fluid Mech.* **29**, 545–553.

NIELD, D. A. 1975 The onset of transient convective instability. *J. Fluid Mech.* **71**, 441–454.

RAYLEIGH, LORD 1916 On convective currents in a horizontal layer of fluid when the higher temperature is on the under side. *Phil. Mag.* **32** (6), 529–546.

SCHLÜTER, A., LORTZ, D. & BUSSE, F. 1965 On the stability of steady finite-amplitude convection. *J. Fluid Mech.* **23**, 129–144.

SERRIN, J. 1959 On the stability of viscous fluid motions. *Arch. Rat. Mech. Anal.* **3**, 1–13.

SPARROW, E. M., GOLDSTEIN, R. J. & JONSSON, V. H. 1964 Thermal instability in a horizontal fluid laver: effect of boundary conditions and nonlinear temperature profile. *J. Fluid Mech.* **18**, 513–528.