J. Fluid Mech. (1969), *vol.* 38, *part* 1, *pp.* 203–224
Printed in Great Britain

Distant side-walls cause slow amplitude modulation of cellular convection

By LEE A. SEGEL
Rensselaer Polytechnic Institute, Troy, N.Y.

(Received 4 December 1968 and in revised form 24 March 1969)

The effect of vertical boundaries on convection in a shallow layer of fluid heated from below is considered. By means of a multiple-scale perturbation analysis, results for horizontally unbounded layers are modified so that they 'fit in' to a rectangular region. The critical Rayleigh number and critical wave-number are determined. Motion is predicted to have the form of finite 'rolls' whose axes are parallel to the shorter sides of the dish. Aspects of the non-linear development and stability of this motion are studied. The general question of convective pattern selection in a bounded layer is discussed in the light of available theoretical and experimental results.

1. Introduction

This paper deals with an aspect of the theory of convection in a bounded horizontal layer of fluid heated from below. Almost all previous theoretical investigations of this problem have assumed that the layer is unbounded. (When we speak of an *unbounded* or a *bounded* layer we shall be referring to whether or not the layer extends to infinity in horizontal directions.) A premise of such investigations has been that the inevitable presence of side-walls has little or no effect in a sufficiently shallow layer. What follows is a first attempt to demonstrate that the concept of slow amplitude modulation can be employed to show how shallow the layer need be if the side-walls are to have 'essentially' no effect and to estimate what their effect is if that effect is not too large.

Readers unfamiliar with developments in convection theory for unbounded layers might find it useful to consult one or more of the surveys by Brindley (1967), Görtler & Velte (1967), or Segel (1966). Directly relevant is the recent paper by Davis (1967), who considered rectangular dishes with length to depth ratios which varied from $\frac{1}{4}$ to 6. His linear stability analysis, using the Galerkin method, showed that convection should commence in the form of *finite rolls* parallel to the shorter side-wall of the dish. If co-ordinates are chosen so that the side-walls are parallel to the x- and y-axes then the two possible types of finite rolls are called *finite y rolls* and *finite x rolls*. The latter is a motion which is periodic (or nearly so) in the x-direction and has a zero velocity component in the y-direction ($v \equiv 0$). Finite y rolls have a corresponding definition. As Davis did, we emphasize the three-dimensional nature of finite rolls, which have no velocity component in the axial direction but depend on all three spatial co-ordinates.

This is to be contrasted with (ordinary) rolls in unbounded layers whose velocity does not vary in the axial direction.

This paper, like that of Davis, considers rectangular side-walls. It is based on the idea that only slow spatial changes need be introduced to 'fit' the solution for an unbounded layer into a shallow dish. The idea finds expression in a multiple-scale perturbation analysis which has close similarities to earlier work by Benney & Newell (1967) on slow spatial modulation of water waves (in an unbounded medium).

We shall show that in a shallow bounded layer the amplitude $A(\tau, X, Y)$ of x rolls satisfies

$$\left.\begin{aligned} &A_\tau = A + \delta_1 A_{XX} - 2i\delta_{12} A_{XYY} - \delta_2 A_{YYYY} - \beta A^2 \bar{A}, \\ &A = 0 \quad \text{at} \quad X = 0, \Lambda, \quad A = A_{YY} = 0 \quad \text{at} \quad Y = 0, \Sigma. \end{aligned}\right\} \tag{1.1}$$

Here τ, X and Y are stretched t (time), x and y variables and δ_1, δ_2, δ_{12}, β, Λ and Σ are constants. By neglecting non-linear terms in (1.1) we shall ascertain how the presence of vertical boundaries alters the results of linear stability theory for an unbounded layer. When variation with Y is negligible we shall determine the steady solution of (1.1) which is almost certainly the limiting flow (as $t \to \infty$) when the motionless layer becomes unstable. We shall conclude with some comparison with experiment and some remarks on the understanding which should emerge from a full exploitation of the ideas introduced here.

2. Derivation of the amplitude equation

In formulating dimensionless governing equations we shall use horizontal co-ordinates x and y and vertical co-ordinate z, corresponding velocity components (u, v, w), temperature θ, and time t. These dimensionless length, velocity, temperature and time variables will be referred to the scales d, κ/d, $\kappa\nu/\alpha_0 g d^3$ and d^2/κ respectively. Here d is the distance between two horizontal planes bounding a fluid with constant coefficients of thermal expansion, kinematic viscosity and thermal conductivity denoted by α_0, ν, and κ. g is the acceleration of gravity, which acts vertically downwards, in the direction of decreasing z. With this notation, the Boussinesq equations for finite x rolls ($v = 0$) can be manipulated in a familiar manner (Newell & Whitehead 1969) to give

$$u_x + w_z = 0, \tag{2.1}$$

$$(\Delta - \partial_t) T + \mathscr{R} w = N^{(1)}, \tag{2.2}$$

$$(\partial_t - \Delta)(\partial_t - \mathscr{P}\Delta)\Delta w - \mathscr{P}\mathscr{R}\Delta_1 w = (\partial_t - \Delta) N^{(2)} - \mathscr{P}\Delta_1 N^{(1)}, \tag{2.3}$$

where $$N^{(1)} \equiv uT_x + wT_z, \quad N^{(2)} \equiv (uu_x + wu_z)_{xz} - (uw_x + ww_z)_{xx}.$$

$$\Delta \equiv \Delta_1 + \partial_z^2, \quad \Delta_1 \equiv \partial_x^2 + \partial_y^2, \quad \mathscr{P} \equiv \nu/\kappa, \quad \mathscr{R} \equiv \alpha_0 (T_1 - T_2) g d^3 \nu^{-1} \kappa^{-1}. \tag{2.4}$$

The temperature has been decomposed into an unperturbed conductive mean part $-\mathscr{R}z$ and perturbation T. The top surface is kept at temperature T_2 and the bottom surface at $T_1, T_2 < T_1$, and both of these surfaces are considered 'free'

(normal velocity and tangential stress vanish), so that the boundary conditions on the horizontal planes are

$$T = w_{zz} = 0 \quad \text{on} \quad z = 0, 1. \tag{2.5a}$$

The vertical boundary conditions are

$$u = w = T = 0 \quad \text{on} \quad x = 0, \quad x = L, \quad y = 0, \quad y = S. \tag{2.5b}$$

When there are no vertical boundaries one can find the following solution to (2.1)–(2.4) and (2.5*a*) by one of the standard approaches to non-linear stability theory:

$$w = 2D(t) \sin \pi z \sin \pi\alpha x + O(D^3), \tag{2.6}$$

$$T = T_m D \sin \pi z \sin \pi\alpha x + O(D^2), \quad T_m = \tfrac{4}{3}\pi^{-2}\mathscr{R}, \tag{2.7}$$

$$u = 2\alpha^{-1} D \cos \pi z \cos \pi\alpha x + O(D^3), \quad v \equiv 0, \tag{2.8}$$

where

$$dD/dt = \sigma D - \gamma D^3 + O(D^5), \tag{2.9}$$

$$\sigma = \frac{2\pi^2\gamma(\alpha^2+1)\,[\mathscr{R} - \mathscr{R}_c(\alpha)]}{\mathscr{R}_c(\alpha)}, \quad \gamma = \frac{\mathscr{P}}{2(1+\mathscr{P})}, \quad T_m = \frac{2\pi^{-4}\mathscr{R}}{\pi^{-2}\sigma + \alpha^2 + 1}. \tag{2.10}$$

This solution represents a non-linear 'roll'. From linear theory

$$\mathscr{R}_c(\alpha) = \pi^4(\alpha^2+1)^3\alpha^{-2}, \quad \alpha_c^2 = \tfrac{1}{2}, \quad \mathscr{R}_c \equiv \min_\alpha \mathscr{R}_c(\alpha) \equiv \mathscr{R}_c(\alpha_c) = 27\pi^4/4. \tag{2.11a,b,c}$$

The solution to (2.9)

$$D = (\sigma/\gamma)^{\frac{1}{2}} e^{\sigma t}[C + e^{2\sigma t}]^{-\frac{1}{2}}, \quad C \text{ a constant}, \tag{2.12}$$

shows that when $\mathscr{R} > \mathscr{R}_c(\alpha)$ the amplitude $D(t)$ approaches a final equilibrium value

$$(\sigma/\gamma)^{\frac{1}{2}} = \pi(2\alpha^2+2)^{\frac{1}{2}}\epsilon, \quad \text{where} \quad \epsilon = [R - R_c(\alpha)]^{\frac{1}{2}}/[R_c(\alpha)]^{\frac{1}{2}}. \tag{2.13a,b}$$

The analysis is expected to be valid when $D(t)$ is sufficiently small. A natural small parameter to introduce is ϵ, so that $D(t)$ is $O(\epsilon)$.

Further, since $\sigma = O(\epsilon^2)$ it is apparent from (2.12) that the development of $D(t)$ is actually on a slow $O(\epsilon^2)$ time scale. It is thus useful to introduce properly scaled variables $A^{(i)}$ and τ by

$$\tau = 2(\alpha^2+1)\pi^2\epsilon^2\gamma t, \quad \epsilon A^{(i)}(\tau) = D(\tau/2(\alpha^2+1)\pi^2\epsilon^2\gamma). \tag{2.14a,b}$$

With this (2.9) becomes

$$\left.\begin{aligned} 2\gamma(\alpha^2+1)\pi^2\epsilon^3[dA^{(i)}/d\tau - A^{(i)} + \beta(A^{(i)})^3] = O(\epsilon^5), \\ \beta^{-1} \equiv 2(\alpha^2+1)\pi^2, \end{aligned}\right\} \tag{2.14c}$$

where

showing explicitly that

$$dA^{(i)}/d\tau = A^{(i)} - \beta(A^{(i)})^3$$

is the correct leading approximation to the amplitude equation satisfied by $A^{(i)}$.

The initial-value problem appropriate for the development of a roll requires terms varying on an $O(1)$ time scale in addition to the slowly varying terms considered above, but these terms all decay (Eckhaus 1965) and do not affect the amplitude equation (2.14*c*).

A slightly more general problem results if (2.6) is replaced by

$$w = 2\epsilon[A^{(r)}(\tau)\cos\pi\alpha x - A^{(i)}(\tau)\sin\pi\alpha x]\sin\pi z + \ldots,$$

or, in complex notation with $A \equiv A^{(r)} + iA^{(i)}$,

$$w = \epsilon[A\exp(i\pi\alpha x) + \bar{A}\exp(-i\pi\alpha x)]\sin\pi z + \ldots.$$

(Corresponding generalizations must be made in (2.7) and (2.8).) Writing

$$w = 2\epsilon[|A(\tau)|\cos(\pi\alpha x + \Theta(\tau))]\sin\pi z + \ldots, \quad \Theta \equiv \arg A,$$

explicitly brings out the fact that we are now allowing for the possibility of a phase shift. The amplitude equation, however, is

$$dA/d\tau = A - \beta A^2\bar{A}$$

or

$$d|A|/d\tau = |A| - \beta|A|^3, \quad d\Theta/d\tau = 0,$$

showing that no phase shift takes place (to lowest order). This is to be expected. In an unbounded layer phase is measured from an entirely arbitrary reference point. There can be no reason why rolls should begin at one undistinguished position and end up at another.

We wish to satisfy boundary conditions appropriate to vertical walls located at $x = 0$, $x = L$, $y = 0$, $y = S$ by assuming that the amplitude A varies slowly with the horizontal co-ordinates x and y as well as with time. It would not be surprising if there occurred a change of the initial position of the roll with respect to the vertical walls, so a phase shift is now *a priori* reasonable. *A posteriori*, a phase shift generally does occur. We thus assume

$$w = 2\epsilon[A^{(r)}(\tau, X, Y)\cos\pi\alpha x - A^{(i)}(\tau, X, Y)\sin\pi\alpha x]\sin\pi z + \epsilon^2 w_2 + \epsilon^3 w_3 + \ldots, \tag{2.15}$$

with corresponding assumptions for T and u. Here

$$X = \mu x, \quad Y = \eta y,$$

where μ and η are small parameters. Substituting (2.15) into (2.3) we find

$$2\mathscr{P}\pi^2\alpha^2\mathscr{R}_c[(-A_\tau + A)\epsilon^3 + \delta_1(\alpha)A_{XX}\epsilon\mu^2 - 2i\delta_{12}(\alpha)A_{XYY}\epsilon\mu\eta^2 - \delta_2(\alpha)A_{YYYY}\epsilon\eta^4] + \ldots = 0,$$

where ... denotes higher-order terms and

$$\left.\begin{aligned} \alpha^2\mathscr{R}_c(\alpha)\delta_1(\alpha) &= 3\pi^2(5\alpha^4 + 6\alpha^2 + 1) - \mathscr{R}\pi^{-2}, \\ \alpha^2\mathscr{R}_c(\alpha)\delta_2(\alpha) &= 3(\alpha^2 + 1), \\ \alpha^2\mathscr{R}_c(\alpha)\delta_{12}(\alpha) &= 6\alpha(\alpha^2 + 1)\pi. \end{aligned}\right\} \tag{2.16}$$

To retain the full effect of slow spatial variation we make all the above terms the same order of magnitude by choosing

$$\mu = \epsilon, \quad \eta = \epsilon^{\frac{1}{2}}.$$

It turns out that no new $O(\epsilon^3)$ non-linear terms proportional to $\sin\pi z\sin\pi\alpha x$ or $\sin\pi z\cos\pi\alpha x$ are introduced by the slow horizontal variation, so A satisfies

$$A_\tau = A + \delta_1 A_{XX} - 2i\delta_{12}A_{XYY} - \delta_2 A_{YYYY} - \beta A^2\bar{A}\ldots. \tag{2.17}$$

Conditions (2.5*b*) on the vertical boundaries $x = 0$, $x = L$, $y = 0$, $Y = S$ obviously require that

$$A = 0 \quad \text{on} \quad X = 0, \quad X = \Lambda, \quad Y = 0, \quad Y = \Sigma, \tag{2.18}$$

where

$$\Lambda = \epsilon L, \quad \Sigma = \epsilon^{\frac{1}{2}} S. \tag{2.19}$$

It turns out also to be necessary that

$$A_{YY} = 0 \quad \text{on} \quad Y = 0 \quad \text{and} \quad Y = \Sigma. \tag{2.20}$$

Details of the derivation of the amplitude equation (2.17) and the boundary conditions (2.18) and (2.20) will be found in the appendix.

3. Infinitesimal perturbations

To determine the initial behaviour of small disturbances we neglect the non-linear term in (2.17). It is convenient to re-introduce the unscaled time t by means of (2.14*a*). Normal mode solutions of the linearized version of (2.17) which satisfy the boundary conditions on $Y = 0$ and $Y = \Sigma$ can then be written

$$A = e^{\sigma t} \sin(m\pi Y/\Sigma)\, e^{ikX} a(X), \quad m \text{ an integer}, \tag{3.1}$$

where a satisfies a second-order equation. We choose

$$k = -\delta_{12}\delta_1^{-1} m^2 \pi^2 \Sigma^{-2},$$

so that the coefficient of a_X in this equation vanishes. The equation for a is then

$$a_{XX} + a[\delta_1^{-1} - (2\pi^2\gamma(\alpha^2+1)\epsilon^2\delta_1)^{-1}\sigma - \delta_1^{-2} F m^4 \pi^4 \Sigma^{-4}] = 0,$$

where

$$F(\alpha) \equiv \delta_1(\alpha)\,\delta_2(\alpha) - [\delta_{12}(\alpha)]^2. \tag{3.2}$$

Since

$$a = 0 \quad \text{at} \quad X = 0 \quad \text{and} \quad X = \Lambda,$$

we must have

$$a = \sin(n\pi X/\Lambda), \quad n \text{ an integer}.$$

For the 'most dangerous' mode ($m = n = 1$) we deduce, using (2.19), that

$$\sigma = 2\pi^2\gamma(\alpha^2+1)\epsilon^2[1 - \epsilon^{-2}(F'\delta_1' S^{-4} + \delta_1' L^{-2})], \tag{3.3}$$

where

$$\delta_1' = \pi^2\delta_1, \quad \delta_{12}' = \pi^3\delta_{12}, \quad \delta_2' = \pi^4\delta_2, \quad F' = \delta_1'\delta_2' - [\delta_{12}']^2. \tag{3.4}$$

We shall use the letter b as a subscript when a quantity has a value appropriate to the bounded layer. Thus, at marginal stability ($\sigma = 0$) we deduce from (3.3) and (2.13*b*) that the Rayleigh number has the value $\mathscr{R}_b(\alpha)$, where

$$\frac{\mathscr{R}_b(\alpha) - \mathscr{R}_c(\alpha)}{\mathscr{R}_c(\alpha)} = \frac{F'(\alpha)\,\delta_1'(\alpha)}{S^4} + \frac{\delta_1'(\alpha)}{L^2}. \tag{3.5}$$

When L and S are 'large', $\mathscr{R}_b(\alpha)$ has a minimum value when $\alpha = \alpha_b$, where

$$[\alpha_b]^2 = \tfrac{1}{2} - \tfrac{5}{3}L^{-2} - \tfrac{1}{2}S^{-4} + \ldots. \tag{3.6}$$

α_b is near $\alpha_c \equiv 2^{-\frac{1}{2}}$, the wave-number minimizing $\mathscr{R}_c(\alpha)$. Thus, to determine the leading terms in $\mathscr{R}_b$, the minimum value of $\mathscr{R}_b(\alpha)$, we substitute $\alpha = \alpha_c$ into the right side of (3.5) and into $\mathscr{R}_c(\alpha)$ on the left side. But

$$\delta_1'(\alpha_c) = \tfrac{8}{3}, \quad \delta_{12}'(\alpha_c) = \tfrac{4}{3}\sqrt{2}, \quad \delta_2'(\alpha_c) = \tfrac{4}{3}, \tag{3.7}$$

so

$$\delta_{12}'(\alpha_c) = \sqrt{[\delta_1'(\alpha_c)\,\delta_2'(\alpha_c)}, \quad F'(\alpha_c) = 0. \tag{3.8}$$

Thus

$$\mathscr{R}_b = \mathscr{R}_b[\alpha_b] = \mathscr{R}_c[1 + \delta_1'(\alpha_c) L^{-2} + \ldots]. \tag{3.9}$$

Further from formula (3.3) for the growth rate σ it follows that the maximum growth rate occurs at a wave-number α_{Mb} given by

$$\alpha_{Mb}^2 = \alpha_b^2 + \tfrac{1}{4}(\mathscr{R} - \mathscr{R}_b)/\mathscr{R}_c + \ldots. \tag{3.10}$$

Comparison with the corresponding wave-number α_M for an unbounded layer is facilitated if we write

$$\alpha_M^2 = \alpha_c^2 + \tfrac{1}{4}(\mathscr{R} - \mathscr{R}_c)/\mathscr{R}_c; \quad \alpha_{Mb}^2 = \alpha_b^2 + \tfrac{1}{4}[(\mathscr{R} - \mathscr{R}_b)/\mathscr{R}_b]\,[\mathscr{R}_b/\mathscr{R}_c];$$

so, for a given percentage increase of Rayleigh number above critical, the square of the wave-number giving maximum growth increases more in the bounded layer by the factor $\mathscr{R}_b/\mathscr{R}_c$. [Our analysis requires α^2 to be within $O(\epsilon^2)$ of a_c^2 (appendix). From (3.6), this means that L must be $O(\epsilon^{-1})$ and S must be $O(\epsilon^{-\frac{1}{2}})$ or, from (2.19), that Λ and Σ must be $O(1)$.]

We assert that we have considered the 'most dangerous modes', so that if L and S are the dimensionless lengths of the long and short sides of a shallow rectangular dish then (3.9) is the formula for the minimum critical Rayleigh number $\mathscr{R}_b$ in terms of the minimum critical Rayleigh number for an unbounded layer $\mathscr{R}_c$.

In support of this assertion we first point out that a lower critical Rayleigh number results from disturbances having the form of finite rolls parallel to the short side of the dish rather than from finite rolls parallel to the long side of the dish. Motion will ensue at the lowest possible value of $\mathscr{R}$. A glance at (3.9) shows that the fluid will 'prefer' to convect in the form of rolls parallel to the shorter side of the rectangular dish so that L, the length of the side perpendicular to the roll axes, is as large as possible. Although we have considered only modulated rolls parallel either to the short or to the long side of the rectangular box, we have dealt with the flows which intuition suggests are most likely to occur.

Expression (3.9) for the critical Rayleigh number has been obtained for a rectangular dish whose long and short sides are $O(\epsilon^{-1})$ and $O(\epsilon^{-\frac{1}{2}})$, respectively. If the dish has larger sides then $\delta_1 = \delta_{12} = \delta_2 = 0$ in (2.17) and, at lowest order, the dependence of A on X and Y is arbitrary except that the boundary conditions must be satisfied. For such 'very shallow' dishes the effect of vertical walls on the critical Rayleigh number is vanishingly small according to our asymptotic analysis.

To go to the other extreme, whatever the size of the dish there can be disturbances whose x and y variations are faster than $O(\epsilon^{-1})$ and $O(\epsilon^{-\frac{1}{2}})$. One's intuition suggests that such more rapidly varying modes will be less dangerous than the ones we have already considered. This is confirmed by our analysis. For example, if $A = A(\tau, \epsilon^{\frac{1}{2}}x, \epsilon^{\frac{1}{2}}y)$, to retain the time derivative term we must modify (2.14a), so that τ is proportional to ϵ, not ϵ^2. This yields

$$A_\tau - \delta_1 A_{XX} = 0. \tag{3.11}$$

Eigenfunctions vanishing at $X = 0$ and $X = \Lambda$ are proportional to

$$\exp(\sigma\tau)\sin(n\pi X/\Lambda) \quad \text{with} \quad \sigma = -\delta_1' n^2 \Delta^{-2}. \tag{3.12a, b}$$

The absence of the term proportional to A in (3.11) means that to a first approximation the modes of (3.12a) neither extract potential energy nor dissipate energy by viscosity; they are purely diffusive, so the decay shown in (3.12b) is to be expected.

It must be remembered that our entire analysis requires that ϵ be sufficiently small, i.e. that the Rayleigh number $\mathscr{R}$ be sufficiently close to $\mathscr{R}_c$, the minimum critical Rayleigh number for the unbounded layer. In particular, formula (3.9) for $\mathscr{R}_b$ can only be expected to be valid when the critical Rayleigh number is not altered too much by the presence of horizontal boundaries. To get some idea of what this involves, let us consider the arbitrary 'convergence criterion' $|\epsilon| \leqslant \frac{1}{2}$. From (3.9) this limit on $|\epsilon|$ requires that our analysis be restricted to dishes sufficiently shallow, so that

$$\delta_1' L^2 \leqslant \tfrac{1}{4}, \tag{3.13}$$

which means that $\mathscr{R}_b$ can at most be 25% above $\mathscr{R}_c$. Using (3.4), (3.13) is satisfied if $L \geqslant 3.3$. Further, S must be $O(\epsilon^{-\frac{1}{2}})$ and terms proportional to S^{-4} must be small. Taking $S > 2$ would seem to satisfy this requirement. These conditions are just rough estimates; they depend on the arbitrary requirement $|\epsilon| \leqslant \frac{1}{2}$ and on the values of δ_1' and δ_2' appropriate for free horizontal boundaries.

To summarize, we first define a *shallow* rectangular dish as one whose short side measures a little more than twice its depth and whose long side measures a little more than twice its short side. For shallow dishes, heated from below, convection should ensue in the form of spatially modulated finite rolls whose axes are parallel to the short side of the dish. The critical Rayleigh number for such a bounded domain should be well approximated by (3.9).

4. Steady finite amplitude solutions

If $A = A(\tau, X, Y)$ where $X = \epsilon x$, $Y = \epsilon y$ and τ is given in (2.14) then A must satisfy

$$A_\tau = A + \delta_1 A_{XX} - \beta A^3, \quad A(\tau, 0, Y) = A(\tau, \Lambda, Y) = 0. \tag{4.1a,b,c}$$

The boundary conditions

$$A = A_{YY} = 0 \quad \text{at} \quad Y = 0, \Sigma. \tag{4.2}$$

must still be satisfied, but we have taken Σ to be $O(\epsilon^{-1})$ rather than $O(\epsilon^{-\frac{1}{2}})$ so that we obtain (4.1a) rather than (2.17). Consideration of y variation can thus be deferred with a corresponding simplification which greatly facilitates analysis of non-linear effects. To be precise, we write

$$A(\tau, X, Y) = f(Y)\, B(\tau, X). \tag{4.3}$$

From (4.2), the boundary conditions for $f(Y)$ are $f(0) = f(\Sigma) = f''(0) = f''(\Sigma) = 0$. As is frequently the case in this type of calculation, the equation for $f(Y)$ only emerges when a 'resonance removal' condition is imposed upon consideration of higher-order terms in the original governing set of partial differential equations. We shall assume, as is physically reasonable, that the resulting problem for $f(Y)$ has a non-trivial solution.

From (4.3) and (4.1), B satisfies

$$B_\tau = B + \delta_1 B_{XX} - \beta B^3, \quad B(\tau, 0) = B(\tau, \Lambda) = 0. \tag{4.4a,b,c}$$

We look for steady solutions to (4.4). Equation (4.4a) has a steady solution in which B^2 has the constant value β^{-1}. This is the solution appropriate to an unbounded layer.

When boundary conditions (4.4b,c) are taken into account the only non-trivial steady solutions are

$$B(X) \equiv a \operatorname{sn} bX \quad \text{where} \quad a^2 = 2\beta^{-1}k^2(1+k^2)^{-1}, \quad b^{-2} = \delta_1(1+k^2). \tag{4.5a,b,c}$$

Here k is the modulus of the Jacobian elliptic function sn and is determined from (4.4c) by

$$b\Lambda = 2nK(k) \quad \text{or} \quad \tfrac{1}{2}\Lambda\delta_1^{-\frac{1}{2}} = n(1+k^2)^{\frac{1}{2}}K(k), \tag{4.6a,b}$$

where n is a positive integer.

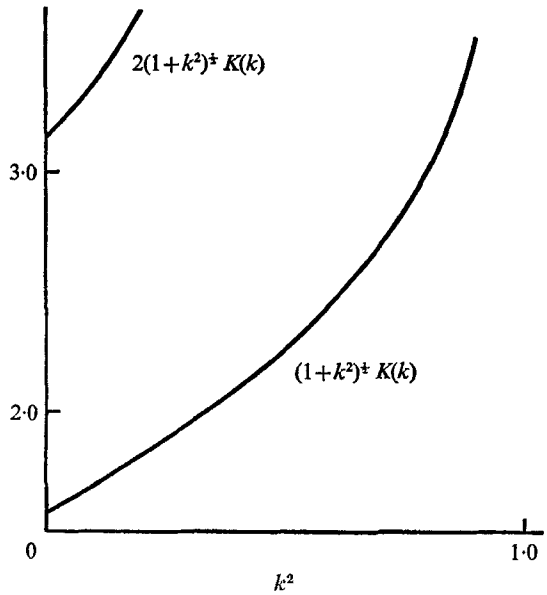

FIGURE 1. Plot of $n\,(1+k^2)^{\frac{1}{2}}K(k)$, where $4K(k)$ is the period of the elliptic function of modulus k, $n = 1, 2$.

Recall that sn has period $4K$ and has zeros at $2nK$. As k traverses $[0, 1)$, $K(k)$ increases monotonically from its minimum value of $\frac{1}{2}\pi$, rising relatively slowly over most of the k range but rapidly approaching infinity as k nears 1. Hence $(1+k^2)^{\frac{1}{2}}K(k)$ has the appearance depicted in figure 1.

Because the minimum distance between zeros of sn is π (when $k = 0$), (4.6) has no solution if $\Lambda\delta_1^{-\frac{1}{2}} < \pi$. The minimum length Λ_m permitting satisfaction of (4.4c) thus satisfies

$$\Lambda_m^2 = \pi^2\delta_1 \quad \text{or} \quad \Lambda_m^2 = \delta_1'.$$

From (3.3) with $\Sigma = 0$, infinitesimal disturbances governed by the linearized version of (4.4) will decay if $\Lambda^2 < \delta_1'$. We thus have the satisfactory situation that only $A \equiv 0$ is a possible steady solution (of small amplitude) to the non-linear problem (4.4) when the layer is so narrow that all infinitesimal disturbances are damped.

For $2\pi < \Lambda\delta_1^{-\frac{1}{2}} < 3\pi$ there are two solutions to (4.6), corresponding to $n = 1$ and $n = 2$ (figure 1). The general situation is best summed up by regarding (4.4) as a non-linear eigenvalue problem for Λ. In figure 2 we give graphs allowing determination of the maximum amplitude a as a function of Λ, for each permitted value of n, by plotting $\frac{1}{2}\beta a^2$ *vs.* $\frac{1}{2}\Lambda\delta_1^{-\frac{1}{2}}$. The graphs are easily constructed by considering them to be parameterized by k. For each n, given k, we can determine $\frac{1}{2}\Lambda\delta_1^{-\frac{1}{2}}$ by (4.6). Further, $\frac{1}{2}\beta a^2 = k^2(1+k^2)^{-1}$ by (4.5). The complete elliptic integral $K(k)$ is tabulated by Milne-Thompson (1950, pp. 106–9) among others.

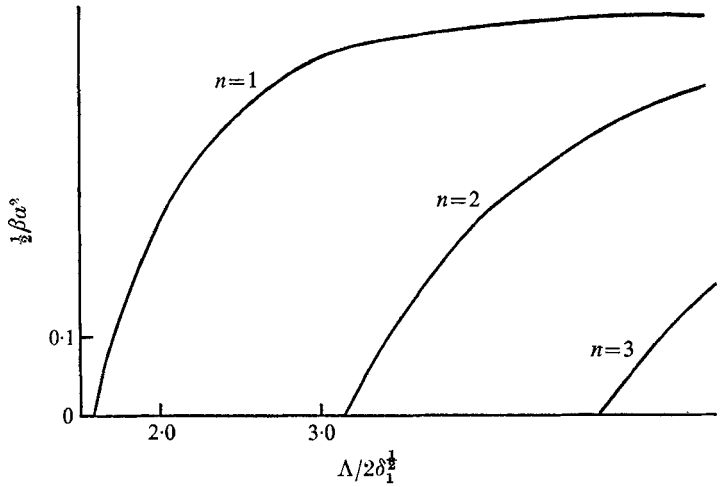

FIGURE 2. Relation of maximum amplitude a to dimensionless box length Λ, as obtained from a non-linear eigenvalue problem. β and δ_1 are constants. More solutions become possible as Λ increases. Bifurcation occurs at the eigenvalues of the linearized problem. It is conjectured that the only stable solution is the one corresponding to $n = 1$.

Note that a new solution appears (bifurcation) each time Λ exceeds an integral multiple of $\pi\delta_1^{\frac{1}{2}}$. As is familiar in non-linear eigenvalue problems, the bifurcation points occur at eigenvalues of the linearized problem, which in this case is

$$\delta_1 d^2B/dX^2 + B = 0, \quad B(0) = B(\Lambda) = 0.$$

Amplitudes $B(X)$ satisfying (4.6) for $n > 1$ vanish at one or more vertical planes in the interior of the dish as well as on its vertical bounding walls. We call the corresponding solutions *imaginary wall solutions*. That such solutions exist is evidently due to the fact that if the dish is wide enough then the amplitude $B(X)$ can slowly change so that it vanishes on fictitious vertical walls which the fluid 'imagines' are equally spaced between the actual boundaries of the dish. It is natural to conjecture that the imaginary wall solutions are unstable to suitable disturbances. We shall speak of the solution associated with $n = 1$ as the *primary solution*. This solution should be stable.

From figure 1 it is clear that as the dish size increases beyond its minimum value Λ_m the corresponding value of k rapidly approaches unity: for example, when $\Lambda > 2\Lambda_m$ then $k^2 > 0{\cdot}95$. But, when $k^2 \approx 1$,

$$B(X) \approx a \tanh bX \quad \text{for} \quad 0 \leqslant X \leqslant \tfrac{1}{2}\Lambda \tag{4.7}$$

(Milne-Thomson 1950, p. 21). Symmetry about the midpoint gives corresponding behaviour for $\frac{1}{2}\Lambda \leqslant X \leqslant \Lambda$. Using (4.7), it is found that $B(X)$ increases from zero to within about 10 % of its maximum value of a in a boundary layer of thickness b^{-1}, where $b^{-1} \approx \sqrt{2}\,\delta_1^{\frac{1}{2}}$. Outside boundary layers of this thickness on each side of the dish, the amplitude is well approximated by the constant value a, where, from (4.5b) with $k^2 \approx 1$, $a^2 \approx \beta^{-1}$. As expected this is the same value taken on by the constant solution appropriate to the unbounded domain. (Compare the second sentence under (4.4).)

The boundary-layer thickness $\sqrt{2}\,\delta_1^{\frac{1}{2}}$ is nearly the same as half the minimum width ($\frac{1}{2}\Lambda_m = \frac{1}{2}\pi\delta_1^{\frac{1}{2}}$). We deduce that, regardless of how wide the layer is, as long as it is wide enough so that slow amplitude modulation is possible, wall effects are almost entirely confined to a layer of thickness approximately given by $X = \sqrt{2}\,\delta_1^{\frac{1}{2}}$ or $x = \sqrt{2}\,\delta_1^{\frac{1}{2}}\epsilon^{-1}$. For free–free boundary conditions, $\delta_1^{\frac{1}{2}} \approx \frac{1}{2}$ and $\alpha_c^2 = \frac{1}{2}$ so the dimensionless boundary-layer thickness becomes $\frac{1}{2}(\epsilon\alpha_c)^{-1}$. As an example, when $\mathscr{R} = 1{\cdot}07\mathscr{R}_c$, $\epsilon \approx \frac{1}{4}$ and this dimensionless boundary-layer thickness becomes $2/\alpha_c$, the critical wavelength. In assessing these numbers we must keep in mind the inevitable arbitrariness in defining boundary-layer thickness and the fact that we have used values of α_c and δ_1 for free–free boundary conditions which differ somewhat from the values appropriate to more realistic boundary conditions.

We thus conclude with the rough statement that a representative result of our calculations is that when the Rayleigh number is 5 or 10 % above its critical value for an unbounded medium then the effects of lateral walls should be confined to a boundary layer whose thickness is about a wavelength. That is, the distortion should disappear within the width of two rolls (or one hexagon).

5. Stability of finite amplitude solutions

To see whether solutions of (4.4) which are initially 'near' $B(X)$ approach $B(X)$ as $\tau \to \infty$ we write

$$B(\tau, X) = B(X) + C(X, \tau).$$

We substitute into (4.4a) and linearize. Writing

$$C(X, \tau) = \exp(\sigma\tau)\, D(\xi), \quad \text{where} \quad \xi = bX, \tag{5.1}$$

we obtain

$$d^2D/d\xi^2 + D[H - m(m+1)\,k^2 \operatorname{sn}^2 \xi] = 0, \quad D(0) = D(b\Lambda) = 0, \tag{5.2a,b,c}$$

where

$$m = 2, \quad H = (1+k^2)(1-\sigma). \tag{5.3a,b}$$

Equation (5.2a) is the Lamé equation. Most of the literature on the Lamé equation concerns 'polynomial' solutions which exist when m is an integer for suitable corresponding values of H (see Arscott 1964). In examining the stability

of the solution $B(X) = a\,\mathrm{sn}\,bX$ of (4.5a) we are faced with the different task of solving (5.2a) subject to the homogeneous boundary conditions (5.2b, c). It is noteworthy that the stability problem depends on the modulus k associated with $B(X)$, not directly on the constants β and δ_1 of the governing equation (4.4a). The problem also depends on the integer n of (4.6), where $n-1$ is the number of zeros which $B(X)$ has in the interior of its domain of definition $[0, \Lambda]$. This is brought out more clearly if we use (4.6a) to write the boundary conditions (5.2b, c) as

$$D(0) = D(2nK) = 0. \tag{5.4}$$

From (5.3b)

$$\sigma = 1 - H(1+k^2)^{-1}, \tag{5.5}$$

so we are assured of stability if $\sigma < 0$ when H is the lowest eigenvalue of (5.2). This eigenvalue should be associated with an eigenfunction with no interior zeros.

When $n = 1$ it happens (Arscott 1964, p. 205) that such an eigenfunction is the Lamé polynomial $\mathrm{sn}\,\xi\,\mathrm{dn}\,\xi$ with associated eigenvalue $H = 1+4k^2$. The corresponding growth rate

$$\sigma = 1-(1+4k^2)(1+k^2)^{-1} \tag{5.6}$$

being negative, the primary solution [(4.5) with $n = 1$] represents a stable steady solution to (4.4).

The further apart the vertical bounding walls, the closer k approaches unity and the closer the value of σ in (5.6) approaches $-\frac{3}{2}$. By contrast, consider an infinitesimal perturbation to the equilibrium solution of the amplitude equation

$$dA/d\tau = A - \beta A^3$$

appropriate to an unbounded layer. Here one finds $\sigma = -2$. In this instance, there is a quantitative difference between the behaviour of an unbounded layer and the behaviour of a layer whose boundaries are infinitely far apart. The qualitative behaviour is the same: both finite amplitude solutions are stable.

When $n = 2, 3, \ldots$ (imaginary wall solutions) the lowest eigenvalue will be associated with an eigenfunction whose zeros are $2Kn$ apart. This increase in separation of the zeros should be accompanied by a decrease in the eigenvalue, which is the tendency required if the imaginary wall solutions are to be unstable. In contrast to the case $n = 1$ there seems nothing in the literature which will give us the answer at once, so we have left for future work confirmation of the conjecture that imaginary wall solutions are unstable.

6. Summary and discussion

We have shown that a multiple scale approach leads to a relatively simple analytic treatment of side-wall effects if the walls are not too close together. For finite x rolls the problem was reduced to an analysis of the amplitude equation and boundary conditions of (1.1). Neglecting non-linear terms we derived a formula relating the critical Rayleigh number in an unbounded layer ($\mathscr{R}_c$) with the corresponding quantity for a bounded layer ($\mathscr{R}_b$),

$$\mathscr{R}_b/\mathscr{R}_c = 1 + \delta_1' L^{-2} + \ldots. \tag{3.9}$$

It is noteworthy that the (longer) walls perpendicular to the roll axes have much less effect than the (shorter) walls parallel to the roll axes. This is evidenced in two ways. First, the effect of the shorter walls can be ascertained by a perturbation analysis when they are separated by an $O(\epsilon^{-1})$ amount while the effect of the longer walls is a perturbation of the same order when they are only $O(\epsilon^{-\frac{1}{2}})$ apart. Secondly, (3.9) shows that to lowest order $\mathscr{R}_b$ is unaffected by the distance S separating the longer walls. The paper of Davis (1967) shows the same predominance of 'width effect' over 'length effect' and discusses the physical reason for the differing response of a finite roll to having its length and width varied. We can sum up by saying that the rolls emerge parallel to the shorter side so that the former effect is minimized.

The values of δ_1', δ_{12}' and δ_2' at $\alpha = \alpha_c$ for 'free' horizontal boundaries are given in (3.7). An extension to the case of rigid horizontal boundaries will allow comparison to be made with the graphs given by Davis (1967). For fixed S the graph of $\mathscr{R}_b/\mathscr{R}_c$ *vs.* L should at first decrease rather rapidly as L increases and then should decrease more slowly to its asymptotic value of unity. Davis' results give the first portion of this graph and the beginning of the second portion. The second portion, where $\mathscr{R}_b \approx \mathscr{R}_c$, should correspond to (3.9). When modified for rigid horizontal boundaries our results should merge smoothly into those of Davis. As it is, our results have the same tendency as one would expect from Davis' work. In the relevant special case, they have the qualitative behaviour predicted by Sani (1964).

There do not seem to be any experimental estimates of the effect of vertical boundaries on the critical Rayleigh number for fluid in a shallow dish. Experiments are typically performed in dishes whose horizontal dimensions are about ten times their depth. For such dishes, the predicted effect of the vertical walls on the critical Rayleigh number should be about 1 %. This is within usual experimental error and so is undetectable. To check predictions concerning the effects of vertical boundaries one would have to perform a series of experiments in dishes wherein L and S can vary and are not too large.

Formula (3.9) assumes that motion ensues in the form of rolls parallel to the short side of the rectangular dish. We have shown that such a motion is preferred to rolls parallel to the long side of the dish. Koschmieder (1966) observed rolls parallel to the short side of the dish and Davis' (1967) linear theory also predicts them. Graphs of radial velocity distribution given by Koschmieder (1967, figure 4) show the type of modulated sinusoidal behaviour on which our theory is based.

We have found that for sufficiently shallow rectangular dishes the effect of the vertical wall is confined to a boundary layer. Photographs of layers with a free upper surface like those of Koschmieder (1967) and Cowley & Rosensweig (1967) show the presence of such a boundary layer by depicting the striking way hexagons fit into a circular dish; the vertical boundary seems to distort only the outer ring of hexagons.

Many of the essentials of a theoretical explanation of Koschmieder's (1967) experiments were given in the non-linear analysis of Scanlon & Segel (1967) which treated an unbounded semi-infinite fluid. The instability was assumed to

be driven by surface tension variation with temperature. Viewed broadly, the analysis is essentially the same as earlier ones for gravity-driven instabilities. Cowley & Rosensweig's experiments concern destabilization by a normal magnetic field of the flat interface between a magnetizable and a non-magnetic fluid. A theoretical analysis for a horizontally unbounded layer along the lines of that by Scanlon & Segel should be applicable, so for the experiments of both Koschmieder (1967) and Cowley & Rosensweig (1967) the effect of vertical boundaries could be treated by the methods of this paper.

Two different types of rolls have been observed in circular dishes. With rigid horizontal boundaries, Koschmieder (1966) finds 'round rolls' whose boundaries are concentric circles while Chen & Whitehead (1968) depict 'straightish rolls' on which the effect of the circular boundary seems limited to curving the outermost roll or two. (Uneven heating may be an important cause of the latter observations.)

This selection of recent experimental results leads one to ask what convective pattern should be observed in a dish of given shape. For $\mathscr{R}$ near $\mathscr{R}_c$ existing theory gives the appropriate predictions for an unbounded layer. These predictions are presumably valid for a sufficiently shallow dish. But how shallow is 'sufficiently shallow'? What effect do different-shaped vertical boundaries have on pattern selection? Such questions can be answered, when the vertical boundaries are not too close together, by the methods introduced here. Although this is a subject for future work we shall set forth some conjectures as to the results; to give experimenters something to aim for, and to show sceptics how the results for the unbounded layer might form part of the desired theory for finite dishes.

To enhance simplicity of exposition our conjectures will be given for a layer which is bounded by circular vertical walls and which is subject to the changing mean temperature studied by Krishnamurti (1968). Three parameters will enter: (i) the *aspect ratio* ρ is the ratio of the depth of the layer to the radius of the circular dish; (ii) *the vertical asymmetry measure* η is Krishnamurti's dimensionless rate of change of mean temperature. In other contexts, a non-zero value of η is associated with viscosity variation (Palm 1960), variation of any of the other fluid properties (Busse 1967*b*), variation from planar of the free surface position (Davis & Segel 1968), and variation of the surface tension with temperature (Scanlon & Segel 1967); (iii) the *amplitude measure* (for the unbounded layer) ϵ is defined by

$$\epsilon = \operatorname{sgn}(\mathscr{R}-\mathscr{R}_c)\sqrt{|\mathscr{R}-\mathscr{R}_c|}/\sqrt{\mathscr{R}_c}.$$

This definition coincides with the definition of ϵ when $\mathscr{R} > \mathscr{R}_c$ which was given in (2.13*b*). Remember that $\mathscr{R}_c$ is the critical Rayleigh number for the unbounded layer, so that motion will not ensue in a bounded layer until ϵ attains a sufficiently large positive value.

Using clues gleaned from available results of theory and experiment we shall try to designate the various forms of convection possible at each point of (ρ, η, ϵ)-space. We shall consider *modulated hexagons*, *modulated rolls* and *wall modes*. By the first two of these forms of convection we mean the hexagonal and roll patterns valid for the horizontally unbounded layer subject to slow spatial variation of

the type considered in previous sections. By wall modes we mean any pattern which cannot be obtained by spatial modulation of a stable steady flow appropriate to an unbounded layer. For a given geometry there may be several different types of wall modes but we shall not distinguish among them. The round (toroidal) rolls observed by Koschmieder (1966) in circular dishes are an example of wall modes. Koschmieder's (1967) pictures are an example of modulated hexagons.

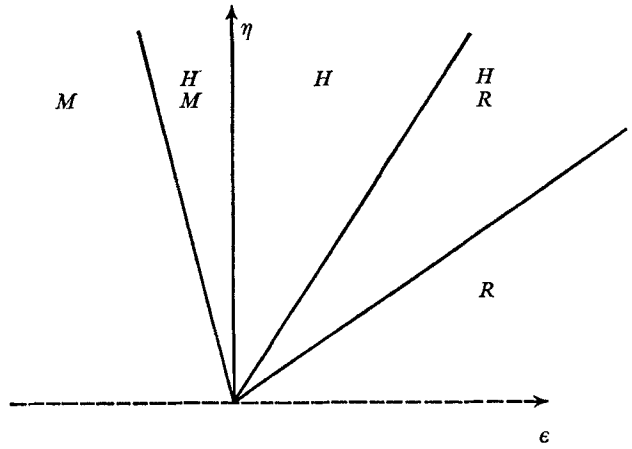

FIGURE 3. For legend see facing page.

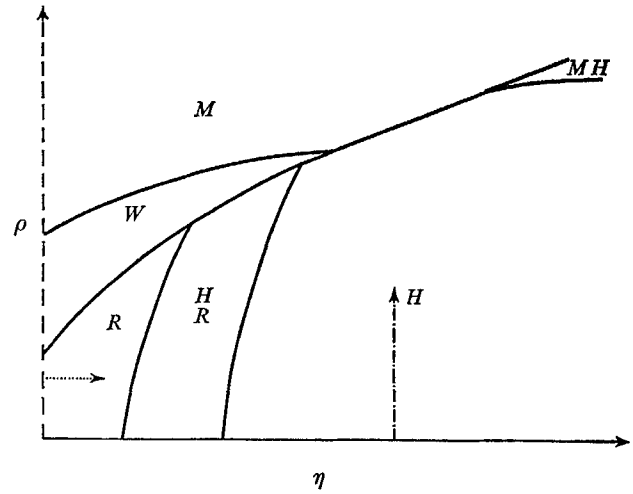

FIGURE 4. For legend see facing page.

Results for $\rho = 0$ (unbounded layer), have been determined by Krishnamurti (1968). These perturbation results, displayed in figure 3, form the foundation of our speculations. They are valid for sufficiently small values of η and ϵ, but it is reasonable to expect that their qualitative nature does not change markedly even when the perturbation methods lose their convergence or asymptotic nature. There has been some theoretical confirmation of this (Busse 1967*a*).

Nevertheless, one must always bear in mind that, the further a point in (ρ, η, ϵ)-space is from the origin, the more likely it is that our predictions are in error.

Typical sections of (ρ, η, ϵ)-space at $\epsilon = \epsilon_0$, $\rho = \rho_0$, $\eta = \eta_0$ ($\epsilon_0, \rho_0, \eta_0$ constants) are given in figures 4–6. These figures were constructed on the basis of the

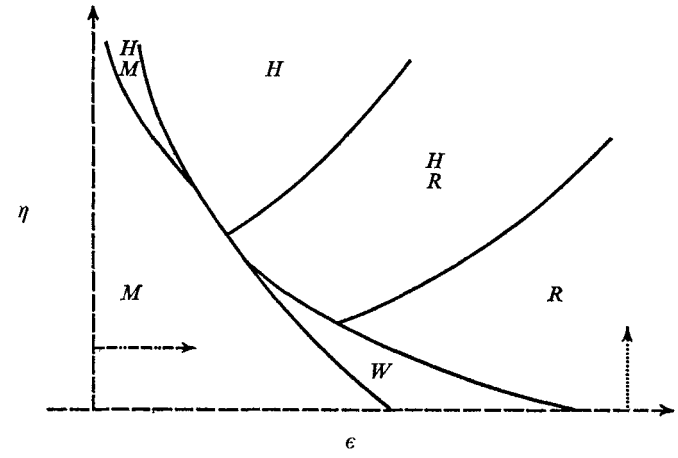

FIGURE 5

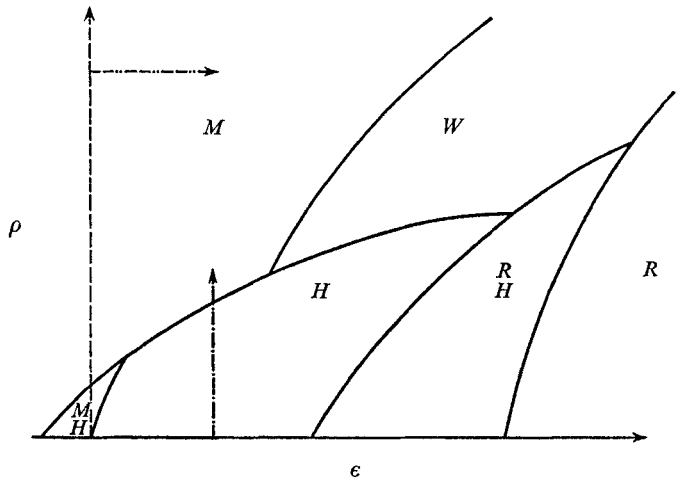

FIGURE 6

FIGURES 3–6. Various sections of (ρ, η, ϵ)-space with conjectured stable states. ρ is depth-to-radius ratio for circular dish, η is measure of vertical asymmetry, ϵ is magnitude of convection (see text). M, motionless layer; H, hexagonal pattern; R, roll pattern; w, wall modes (pattern determined by wall shape). FIGURE 3. $\rho = 0$ (unbounded layer), see Krishnamurti (1968). FIGURE 4. ϵ fixed and positive. FIGURE 5. ρ fixed and not too small. FIGURE 6. η fixed.

following facts and conjectures. (i) When $\rho = 0$ we must recover the known results for the unbounded layer. (ii) When ρ is sufficiently small we still get the same results as for the infinite layer. (This is not inevitable; there could be a non-uniform limit as $\rho \to 0$.) (iii) For fixed ϵ and η, as ρ increases the layer becomes more and more confined so one eventually must attain a motionless state. (iv) When $\eta = 0$ modulated hexagons will not appear (as is the case for the unbounded layer)

but for fixed ϵ and ρ, as η increases, hexagons are increasingly likely, due to second-order equilateral triangle interactions (Segel 1965). (v) For fixed ρ and η, as ϵ increases modulated rolls become more likely (as is the case again for the unbounded layer).

In figure 4 the possible states suggested for very large and very small ρ follow from (iii) and (i). As ρ decreases the motionless state should give way to wall modes unless η is very large, in which case (iv) may mean that the motionless state is succeeded by hexagons.

The dotted line segments in figures 4 and 5 are both projections of the same segment of the line $\rho = \rho_0$, $\epsilon = \epsilon_0$. Figures 4 and 5 should be (and are) such that the same sequence of states is met as η increases on the two projections. (The arrows point in the direction of increasing η.) In figure 4, a line $\rho = \rho_1$, parallel to the dotted line shown, would be the projection of a line like the dotted line in figure 5, but on a plane $\rho = \rho_1$ different from the plane of figure 5. It is a reasonable assumption, however, that the general appearance of the various sections only changes slowly as the sectioning plane is moved parallel to itself. If so, then, as η increases along various lines parallel to the dotted line of figure 4, roughly the same sequence of states should be encountered as η increases along corresponding lines in figure 5. To see the tendency of corresponding lines, note that both when ρ is relatively small and when ϵ is relatively large we have a situation of rather vigorous convection with a relatively large amount of fluid buoyancy presumably dominating in selecting a flow pattern which is not much affected by the walls. For this reason, the dotted line in figure 4 along which ρ has a relatively small value corresponds to the dotted line in figure 5 along which ϵ has a relatively large value. As the former line moves up the latter line should move to the left.

In like manner, roughly the same sequence of states will probably be met in traversing the dot-dashed lines and the double-dot-dashed lines. The way the dot-dashed lines correspond emerges from the fact that a relatively large tendency towards hexagons and small tendency towards rolls is associated both with η large and with ϵ small. The positions of corresponding double-dot-dashed lines is as is shown in figures 5 and 6 because with both η small and ρ large there is a relatively small amount of vertical asymmetry and a relatively large amount of horizontal modulation.

Using the above reasoning as a final tool in constructing figures 4–6, we have required that the same sequence of states is met if one moves in the direction of the arrows for any corresponding pair of lines. Exceptions are allowed in transition regions. Here is a typical feature which emerges. As the dotted line moves in the direction of decreasing ϵ in figure 5, the first three sequences of states which appear are R, H or R, H; W, R, H or R, H; W, H or R, H. In order that the same sequence of states appears as the dotted line in figure 4 moves in the direction of increasing ρ, the lower boundary of the W region must be concave downward. Such features are details and are subject to future modification. It seems worth while to present some detail, however, if only to show what a complicated sequence of states is likely to be obtained as a given parameter is varied while the others are held constant.

(i) When $\rho = 0$ there are values of ϵ and η such that either rolls or hexagons are obtained, depending on initial conditions. For other values of ϵ, η the fluid may either be motionless or convecting in a hexagonal pattern. Non-uniquenesses of this kind will doubtless occur for $\rho > 0$. In figure 4, for example, there are probably regions where one can obtain either wall modes or hexagons. So as not to complicate the diagrams further, we have not indicated regions where the presence or absence of wall modes might depend on initial conditions. (ii) Predictions for negative η can be obtained from figures 4 and 5 by reflexion in the ρ- and ϵ-axes respectively. (iii) If ρ is fixed at a sufficiently small value we expect a situation like that of figure 3. In figure 5, ρ is taken large enough so that wall modes are possible.

In concluding we recall some of the topics which have been left for future work. It remains to modify the results to take into account more realistic horizontal boundary conditions and different dish shapes. This paper deals only with modified rolls but the work should be extended to deal with modified hexagons and with more general combinations of rolls. The stability of the imaginary wall solutions must be determined. If these solutions are unstable as conjectured, this would be another instance of the preferred mode being one of maximum amplitude (see figure 2).

Finding a solution for an initial-value problem appropriate to rolls, and to combinations thereof, has been left for the future. A goal of such investigations should be to determine how long it takes for a new convective pattern to be established after a change in the conditions of the problem. In experiments on the related problem of flow between rotating cylinders Snyder (1969) has found this time to be approximately $L^2/6\nu$, where ν is the kinematic viscosity and L is the length of the cylinders. The present work strongly suggests that such a diffusion time is involved because of the diffusive nature of the new terms which appear in the amplitude equation.

As is shown by the complicated nature of the conjectures in figures 4–6 a good deal of work will be required before one can predict with confidence the form which convection will take in a bounded layer at various values of ρ, ϵ and η. Such predictions will be easier at small ρ when, as we have seen, the flow has a boundary-layer character. The limit $\rho \to 0$ is not uniform in all respects (see the paragraph under (5.6)) so one cannot be sure of the results even for very small ρ.

There is little doubt that the effect of side boundaries is responsible for the discrepancy between the round rolls observed by Koschmieder (1966) in a fairly shallow round dish and the straight rolls or hexagons predicted by the theory of unbounded layers. In another type of disagreement between theory and experiment Koschmieder (1966) finds that the observed wave-number decreases as $\mathscr{R}$ increases but this cannot happen according to Schlüter, Lortz & Busse's (1965) theory for the unbounded layer. It is less obvious that the discrepancy here is associated with the presence or absence of vertical boundaries; there are (inevitably) many simplifications in the theoretical model. Yet wave-number selection is certainly influenced by vertical boundaries: (3.9) shows how the critical wave-number is principally changed by the 'width effect' ($L \neq \infty$). For

an unbounded region, the wave-number at the onset of motion minimizes the ratio of potential energy release to dissipation *per cell*, but to explain (3.9) this ratio will have to be computed for the total number of cells in the dish.

The calculation, in § 3, of maximum growth rate for large L and S shows no profound difference between the unbounded and bounded cases, although there is a slight tendency in the right direction. In any case, wave-number selection is a non-linear problem. It will be interesting to see whether a combination of this analysis and that of Schlüter, Lortz & Busse (1965) can resolve the discrepancy between theoretical and experimental predictions regarding the change of preferred wave-number with increasing $\mathscr{R}$. The 'shape assumption' work of Davis (1968) is encouraging.

The idea that slow spatial modulation should prove useful in the theory of convection arose in a conversation with D. Benney. While this paper was in its later stages of preparation the author benefited from seeing a version of work by Newell & Whitehead (1969) which begins with essentially the same amplitude equation (2.17) but then goes on to a valuable investigation of 'sideband' instabilities in unbounded layers. Earlier blunders connected with the incorporation of Y variation were corrected thanks to seeing this work and a private communication from Newell.

Much of this paper was written in the summer of 1968 when the author enjoyed the hospitality of the Imperial College Department of Mathematics. The work was largely supported by the Office of Naval Research (Mechanics Branch), and also by the Army Research Office (Durham). It was completed while the author was on leave from (but partially supported by) Rensselaer Polytechnic Institute, at the Biomathematics Division, Graduate School of Medical Sciences, Cornell University, and Sloan-Kettering Institute, New York, N.Y.

Appendix. Details of solution

We look for a solution to (2.1)–(2.5) of the form

$$w = 2\epsilon[\mathrm{A}^{(r)}\cos\pi\alpha x - A^{(i)}\sin\pi\alpha x]\sin\pi z + \epsilon^2 w^{(2)}(\tau, X, Y, x, y) + \epsilon^3 w^{(3)}(\tau, X, Y, x, y) + \ldots, \quad \text{(A 1)}$$

$$T = T_m\epsilon[A^{(r)}\cos\pi\alpha x - A^{(i)}\sin\pi\alpha x]\sin\pi z + \epsilon^2 T^{(2)}(\tau, X, Y, x, y) + \epsilon^3 T^{(3)}(\tau, X, Y, x, y) + \ldots, \quad \text{(A 2)}$$

$$u = -2\alpha^{-1}\epsilon[A^{(r)}\sin\pi\alpha x + A^{(i)}\cos\pi\alpha x]\cos\pi z + \epsilon^2 u^{(2)}(\tau, X, Y, x, y) + \epsilon^3 u^{((3)}(\tau, X, Y, x, y) + \ldots, \quad \text{(A 3)}$$

where, to satisfy (2.5*b*),

$$A = 0 \quad \text{on} \quad x = 0, L \quad \text{and} \quad y = 0, S.$$

It would be best to specify the problem further by imposing initial conditions at $t = 0$:

$$w = 2\epsilon[A^{(r)}(0, X, Y)\cos\pi\alpha x - A^{(i)}(0, X, Y)\sin\pi\alpha x]\sin\pi z,$$

$$T = T_m\epsilon[A^{(r)}(0, X, Y)\cos\pi\alpha x - A^{(i)}(0, X, Y)\sin\pi\alpha x]\sin\pi z,$$

$$u = -2\alpha^{-1}\epsilon[A^{(r)}(0, X, Y)\sin\pi\alpha x - A^{(i)}(0, X, Y)\cos\pi\alpha x]\cos\pi z,$$

where $A(0, X, Y)$ is given. But a preliminary exploration indicates that a number of subtleties then arise, connected with the multiple-scale nature of the analysis. We do not choose to pursue these subtleties at this time. Rather, we proceed under the assumption that the particular solution we shall exhibit differs from the solution to the initial-value problem only in details of the velocity and temperature fields, but that the fundamental amplitude equation (2.17) is unaffected. The above-mentioned preliminary exploration indicates that this assumption is correct, as does the fact that it holds when there is no spatial modulation of the amplitudes (Eckhaus 1965).

To lowest order all equations and boundary conditions are satisfied by (A 3) for arbitrary $A \equiv A^{(r)} + iA^{(i)}$ vanishing on the vertical boundaries. Non-uniqueness already enters at this first stage. The continuity equation (2.1) remains satisfied at order ϵ if we add to u any x independent function satisfying the boundary conditions. It may be that this function is uniquely determined by higher-order conditions or it may be that we have an instance of the non-uniqueness in multiple-scale expansions pointed out by Erdelyi (1968). We find no contradiction in taking this function to be identically zero.

At $O(\epsilon^2)$ the forcing terms on the right side of (2.3) add to zero (a simplifying peculiarity associated with free–free horizontal boundary conditions). No $O(\epsilon^2)$ forcing terms arise from the linear left side of (2.3) (see below), so

$$\Delta^3 w_2 - \Delta_1 w_2 = 0,$$

and we can take $w_2 \equiv 0$. We substitute (A 1), (A 3) into (2.1) and find at $O(\epsilon^2)$

$$u_x^{(2)} = 2\alpha^{-1}[A_X^{(i)} \cos \pi\alpha x + A_X^{(r)} \sin \pi\alpha x] \cos \pi z,$$

so
$$u^{(2)} = 2\pi^{-1}\alpha^{-2} \cos \pi z[A_X^{(i)} \sin \pi\alpha x - A_X^{(r)} \cos \pi\alpha x + h(\tau, X, Y, y)].$$

Because u must vanish on the boundary [since $A_X = 0$ on $y = 0, S$],

$$\text{at} \quad y = 0, \Sigma, \quad h = 0; \quad \text{at} \quad x = 0, \quad h = A_X^{(r)}(\tau, 0, Y);$$

$$\text{at} \quad x = \Lambda, \quad h = A_X^{(r)}(\tau, \Lambda, y) \cos \pi\alpha L - A_X^{(i)}(\tau, \Lambda, y) \sin \pi\alpha L;$$

h is otherwise arbitrary.

Substituting into (2.2), at $O(\epsilon^2)$

$$\Delta T^{(2)} = T_m \sin \pi z Q(\tau, X, Y, x) + \pi T_m |A|^2 \sin 2\pi z,$$

$$Q \equiv (2\pi\alpha A_X^{(i)} - A_{YY}^{(r)}) \cos \pi\alpha x + (2\pi\alpha A_X^{(r)} + A_{YY}^{(i)}) \sin \pi\alpha x,$$

so
$$T^{(2)}/T_m = -\pi^{-2}(\alpha^2 + 1)^{-1} \sin \pi z[Q - g(\tau, X, Y, x, y] - \tfrac{1}{4}\pi^{-1}|A|^2 \sin 2\pi z.$$

g is the solution of the homogeneous equation

$$\Delta_1 g - \pi^2 g = 0, \tag{A 4a}$$

and the boundary conditions

$$\text{on} \quad x = 0, \quad g = 2\pi\alpha A_X^{(i)}(0, Y); \tag{A 4b}$$

$$\text{on} \quad x = L, \quad g = 2\pi\alpha[A_X^{(i)}(\Lambda, Y) \cos \pi\alpha L + A_X^{(r)}(\Lambda, Y) \sin \pi\alpha L]; \tag{A 4c}$$

$$\text{on} \quad y = 0, \quad g = -A_{YY}^{(r)}(X, 0) \cos \pi\alpha x + A_{YY}^{(i)}(X, 0) \sin \pi\alpha x; \tag{A 4d}$$

$$\text{on} \quad y = S, \quad g = -A_{YY}^{(r)}(X, \Sigma) \cos \pi\alpha x + A_{YY}^{(i)}(X, \Sigma) \sin \pi\alpha x; \tag{A 4e}$$

which follow from the requirement that $T^{(2)}$ vanish on the boundaries. Since X and Y are to be considered constants in (A 4*a*), the problem of (A 4) seems to be standard. Using separation of variables, we expect to be able to solve it by superposing the solutions of four sub-problems in which g is required to vanish on three of the boundaries and one of the non-homogeneous boundary conditions (A 4*b*)–(A 4*e*) is applied on the appropriate fourth boundary. For example, if g is required to vanish on $x = 0$, $X = L$, and $Y = 0$, we would satisfy (A 4*e*) by appropriate choice of the constants C_n in the following formal solution of (A 4*a*):

$$g = \sum_{n=1}^{\infty} C_n \sin[n\pi L^{-1}x] \sinh[\pi(1+n^2L^{-2})^{\frac{1}{2}}y].$$

But this approach is inconsistent because

$$\sin[n\pi L^{-1}x] \equiv \sin[n\pi X]$$

is a function of X, not of x. A little thought shows that the only way out of the difficulty is to require

$$A_{YY} = 0 \quad \text{on} \quad Y = 0 \quad \text{and} \quad Y = \Sigma,$$

and to satisfy the conditions of (A 4) by taking

$$g = G_1(X, Y)\sinh\pi x + G_2(X, Y)\cosh\pi x,$$

where

$$G_i(X, 0) = G_i(X, \Sigma) = 0 \quad (i = 1, 2);$$

$$G_2(0, Y) = 2\pi\alpha A_X^{(i)}(0, Y);$$

$$G_1(\Lambda, Y)\sinh\pi\Lambda + G_2(\Lambda, Y)\cosh\pi\Lambda = 2\pi\alpha[A_X^{(i)}(\Lambda, Y)\cos\pi\alpha L + A_X^{(r)}(\Lambda, Y)\sin\pi\alpha L].$$

As is characteristic of multiple-scale methods, only the boundary conditions for the functions G_i are prescribed at this stage. The governing equations will emerge when higher-order terms are considered.

We now substitute into (2.3) and obtain (2.17) from the existence condition which requires that there be no forcing terms proportional to the eigenfunctions $\sin\pi\alpha x \sin\pi z$ and $\cos\pi\alpha \sin\pi z$ of the lowest-order (self-adjoint) operator $\Delta^3 - \Delta_1$. Note that in $u^{(2)}$ and $T^{(2)}$ all new terms associated with the slow spatial variation (which are proportional to X and Y derivatives of A) have a vertical variation given by $\sin\pi z$. They therefore cannot give a term proportional to $\sin\pi z$ in $O(\epsilon^3)$ non-linear terms like $w^{(1)}T_x^{(2)}$ and $w^{(1)}T_z^{(2)}$. This is why the coefficient of $A^2\bar{A}$ in (2.17) is the same as the coefficient which appears when no spatial variation of the amplitude is present.

In determining $O(\epsilon^3)$ terms in (2.3) we make use of the fact that

$$(\Delta^3 - \mathscr{R}\partial_x^2)(A^{(i)}\sin\pi z \sin\pi\alpha x) = \sin\pi z[P(\partial_x)A^{(i)}\sin\pi\alpha x],$$

where $P(\partial_x)$ is an even-order polynomial operator with constant coefficients

$$P(\xi) \equiv (\xi^2 - \pi^2)^3 - \mathscr{R}\xi^2.$$

It is easy to show that if $Q(\xi) \equiv P(i\xi)$ then

$$\begin{aligned} P(\partial_x) A^{(i)} \sin \pi\alpha x &= [P(\partial_x) \sin \pi\alpha x] A^{(i)} + [P'(\partial_x) \sin \pi\alpha x] \partial_x A^{(i)} \\ &\quad + \tfrac{1}{2}[P''(\partial_x) \sin \pi\alpha x] \partial_x^2 A^{(i)} + \dots \\ &= Q(\pi\alpha) A^{(i)} \sin \pi\alpha x - \mu Q'(\pi\alpha) A_X^{(i)} \cos \pi\alpha x \\ &\quad - \tfrac{1}{2}\mu^2 Q''(\pi\alpha) A_{XX}^{(i)} \sin \pi\alpha x + O(\mu^3). \end{aligned} \tag{A 5}$$

But the neutral curve is given by $Q(\pi\alpha) = 0$ and the critical wave-number satisfies $Q'(\pi\alpha_c) = 0$. Thus the term in (A 5) proportional to $\cos \pi\alpha x$ is zero when $\alpha = \alpha_c$. Actually, since

$$Q'(\pi\alpha) = (\alpha - \alpha_c) Q''(\pi\alpha_c) + O(\alpha - \alpha_c)^2,$$

the $\cos \pi\alpha x$ term is negligible if $\alpha - \alpha_c = O(\epsilon^2)$, and our analysis is valid for any fixed wave-number α in this range. The coefficient of A_{YY} in (2.17) is zero for essentially the same reason (just discussed) that the coefficient of A_X is zero.

Insulating boundaries at $x = 0, L$ can easily be handled. To obtain a solution satisfying the appropriate boundary condition $T_x = 0$ one merely needs to modify the boundary conditions in (A 4); equation (2.17) is unchanged. Insulating boundaries at $y = 0, S$ would require a non-trivial modification of our work since the normal derivative T_y introduces a term proportional to $\epsilon^{\frac{3}{2}}$ which is inconsistent with our present analysis.

REFERENCES

ARSCOTT, F. 1964 *Periodic Differential Equations*. New York: Pergamon.

BENNEY, D. & NEWELL, A. 1967 The propagation of nonlinear wave envelopes. *J. Math. Phys.* **46**, 133–139.

BRINDLEY, J. 1967 Thermal convection in horizontal fluid layers. *J. Inst. Math. Appl.* **3**, 313–343.

BUSSE, F. 1967*a* On the stability of two-dimensional convection in a layer heated from below. *J. Math. Phys.* **46**, 140–150.

BUSSE, F. 1967*b* The stability of finite amplitude cellular convection and its relation to an extremum principle. *J. Fluid Mech.* **30**, 625–649.

CHEN, M. & WHITEHEAD, J. 1968 Evolution of two-dimensional periodic Rayleigh convection cells of arbitrary wave-numbers. *J. Fluid Mech.* **31**, 1–16.

COWLEY, M. & ROSENSWEIG, R. 1967 The interfacial stability of a ferromagnetic fluid. *J. Fluid Mech.* **30**, 671–688.

DAVIS, S. 1967 Convection in a box: linear theory. *J. Fluid Mech.* **30**, 465–478.

DAVIS, S. 1968 Convection in a box: on the dependence of preferred wave-number upon the Rayleigh number at finite amplitude. *J. Fluid Mech.* **32**, 619–624.

DAVIS, S. & SEGEL, L. 1968 Effects of surface curvature and property variation on cellular convection. *Phys. Fluids* **11**, 470–476.

ECKHAUS, W. 1965 *Studies in Nonlinear Stability Theory*. New York: Springer.

ERDELYI, A. 1968 Two-variable expansions for singular perturbations. *J. Inst. Math. Appl.* **4**, 113–119.

GÖRTLER, H. & VELTE, W. 1967 Recent mathematical treatments of laminar flow and transition problems. *Phys. Fluids Suppl.* **10**, S3–10.

KOSCHMIEDER, L. 1966 On convection on a uniformly heated plane. *Beit. z. Phys. Atmos.* **39**, 1–11.

KOSCHMIEDER, L. 1967 On convection under an air surface. *J. Fluid Mech.* **30**, 9–15.

KRISHNAMURTI, R. 1968 Finite amplitude convection with changing mean temperature. *J. Fluid Mech.* **33**, 445–463.

MILNE-THOMPSON, L. M. 1950 *Jacobian Elliptic Function Tables*. New York: Dover.

NEWELL, A. & WHITEHEAD, J. 1969 Finite bandwidth, finite amplitude convection. *J. Fluid Mech.* (To be published.)

PALM, E. 1960 On the tendency towards hexagonal cells in steady convection. *J. Fluid Mech.* **8**, 183–192.

SANI, R. 1964 Note on flow instability in heated ducts. *ZAMP* **15**, 381–387.

SCANLON, J. & SEGEL, L. 1967 Finite amplitude cellular convection induced by surface tension. *J. Fluid Mech.* **30**, 149–162.

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

SEGEL, L. 1965 The nonlinear interaction of a finite number of disturbances to a layer of fluid heated from below. *J. Fluid Mech.* **21**, 359–84.

SEGEL, L. 1966 Nonlinear hydrodynamic stability theory and its applications to thermal convection and curved flows. *Non-equilibrium Thermodynamics, Variational Techniques, and Stability* (R. Donnelly, R. Herman, and I. Prigogine, eds.). University of Chicago Press, 165–197.

SNYDER, H. 1969 Wave-number selection at finite amplitude in rotating Couette flow. To appear in *J. Fluid Mech.*

Note added in proof. The instability of imaginary wall solutions has been confirmed in a contribution by the author to the Proceedings of the 1969 IUTAM Symposium on Instability of Continuous Systems to be published by Springer.