J. Fluid Mech. (1985), *vol.* 151, *pp.* 55–79
Printed in Great Britain

The effect of variable stratification on linear doubly diffusive stability

By F. ZANGRANDO
The Solar Energy Research Institute, 1617 Cole Blvd, Golden, CO 80401
AND LEE A. BERTRAM
Sandia National Laboratories, Albuquerque, NM 87185

(Received 25 August 1983 and in revised form 8 May 1984)

The effect of variable stratification on the linear bifurcations of a doubly diffusive plane parallel layer is examined numerically by expanding in a Fourier series. Because the motivation is analysis of solar-pond stability, a Prandtl number of 7 and ratio of diffusivities of $\frac{1}{80}$ is used in the study, with (large) solute Rayleigh numbers Rs ranging from 10^4 to 10^{12}. Stratification of solute is cubic antisymmetric about midlayer; because temperature has a higher diffusivity, it is given a linear stratification. Exchange of stabilities results also solve the 'fingering' and thermal problems with cubic stratification. For the overstable case, the numerical results approach Walton's perturbation solution at large Rs, but differ significantly at smaller Rs ($< 10^8$). While both exchange of stabilities and overstable modes display an expected tendency to localize about the point of minimum solute gradient, the overstable modes behave in other, non-intuitive ways. Sublayers of reversed salinity gradient, if small enough, can be stable. Above $Rs = 10^{12}$ computations become prohibitively expensive as a continuous spectrum is approached. A simple sublayer scaling rule defines an infinite family of Rs and stratification parameters on which the localized eigensolution is nearly invariant.

1. Introduction

The aim of this investigation was to determine the effects of variable stratification on the doubly diffusive instability of large layers of brine heated and salted from below. We consider large values of the solute Rayleigh number given by

$$Rs = \frac{g\beta \Delta S h^3}{\nu\kappa},$$

where g is the acceleration due to gravity, β is the density coefficient of salting, ΔS is the salinity of the bottom of the layer minus the surface salinity, h is the depth of the layer, and ν and κ are respectively the kinematic viscosity and thermal diffusivity of the brine. The relevant Prandtl number $Pr = \nu/\kappa$ and ratio of diffusivities $\tau = D/\kappa$ are taken to have the nominal values of 7 and $\frac{1}{80}$ respectively, where D is the diffusivity of salt in water. Thus interest is focused on very large Rs and small τ.

The interest in this question arose in conjunction with the first author's dissertation research on the feasibility of salt-gradient solar ponds (Zangrando 1979) under field conditions. The pond, about two metres deep, consists of water with salt dissolved in it such that the bottom metre has a uniformly large salinity,

while the salinity of the topmost metre decreases with height. This top layer is thus made resistant to thermal convection because of the increase of density with depth caused by the salting, even when the layer is heated from below. Since water is a poor heat conductor, this metre layer serves to insulate the solar-heated bottom layer, even when the bottom temperature exceeds the surface temperature by 80 °C (i.e. the bottom layer has been brought to 109 °C). Under operating conditions, a mixed layer develops at the surface; its temperature remains within a few degrees of average ambient temperature, so there are in reality three layers. This arrangement serves as both a collector of solar energy and a storage medium, so long as the insulating layer (the layer with salt and temperature gradients) does not convect heat to the surface.

In the course of the research, suspended isothermal sublayers a few centimetres deep appeared in the interior of the insulating layer, clearly separated from the interfaces at the top and bottom of the layer. These isothermal sublayers were presumably the result of convective motions, and they displayed no oscillations in time, as would be expected from constant-gradients linear theory. Sometimes the sublayers would decay and disappear; sometimes they would evolve into a stairstep of further sublayers; and sometimes they would remain stationary. This behaviour suggests that the first sublayers appeared at a neutrally stable stratification of the gradient layer.

Theoretical prediction of neutral stability for infinite plane-parallel layers with constant gradients of temperature and salinity yields critical values of thermal Rayleigh number

$$Ra = \frac{g\alpha\,\Delta T\,h^3}{\nu\kappa},$$

where α is the coefficient of thermal expansion and ΔT is the bottom temperature minus the top temperature. For free–free boundary conditions, Stern (1960) and Weinberger (1962) obtained for exchange of stabilities (bifurcation to a steady solution) the critical value

$$Ra_{\mathrm{eos}} = \frac{Rs}{\tau} + Rb; \tag{1}$$

and, for overstable modes (Hopf bifurcation),

$$Ra_{\mathrm{os}} = \frac{Pr+\tau}{Pr+1}Rs + \frac{(1+\tau)(Pr+\tau)}{Pr}Rb, \tag{2}$$

where Rb is the Bénard eigenvalue for the same singly diffusive problem (Chandrasekhar 1961). 'Large Rs' will mean $Rs \gg Rb$ here, so the terms proportional to Rb in (1) and (2) will be negligible. For brine values of Pr and τ, Ra_{os} will be less than Ra_{eos}, as depicted in figure 1, and the critical mode should be oscillatory.

However, observed pond motions were apparently steady, and occurred at Ra as low as $\frac{1}{5}Ra_{\mathrm{os}}$ (Zangrando 1979). Our general aim here is to identify the source of this discrepancy.

One possible explanation is due to Joseph (1976), who gave the necessary and sufficient condition for the existence of growing finite-amplitude motions as

$$Ra \geqslant R_{\mathrm{f}} = \{(\tau Rs)^{\frac{1}{2}} + [(1-\tau^2)\,Rb]^{\frac{1}{2}}\}^2. \tag{3}$$

For large Rs, this gives $R_{\mathrm{f}} \sim \tau Rs$, a factor of 80 smaller than Ra_{os}. This drastic reduction of the critical Ra raises the question of how, if at all, the solutions growing from the linear bifurcation points (1) and (2) could evolve into a mode which reaches this energy limit; that is, what the form of the bifurcation diagram should be (figure 2).

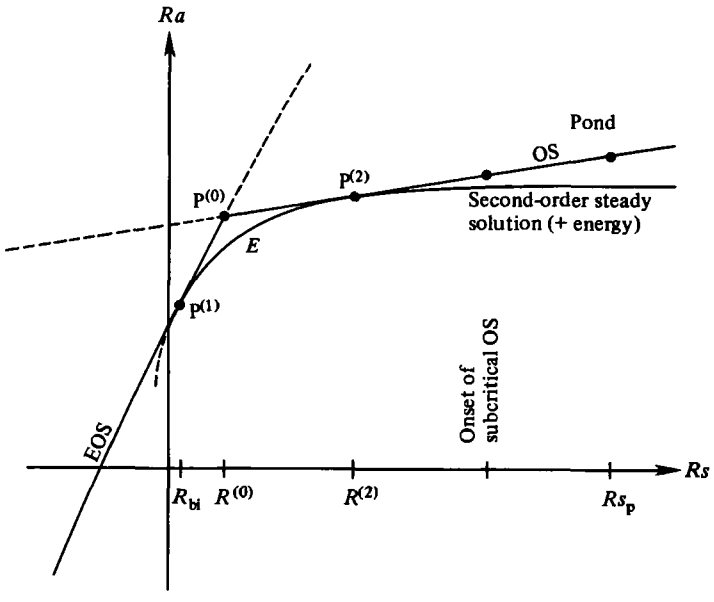

FIGURE 1. Stability loci (schematic). Lines EOS (exchange of stabilities, equation (1)) and OS (overstable, equation (2)) indicate linear neutral-stability loci. A typical pond observation parameter is indicated by Rs_p. Curve E is the minimum Ra locus which is necessary and sufficient for existence of second-order finite-amplitude steady solutions, and for global energy instability. $R_{bi} = \tau^3 Rb/(1-\tau)^2$ is the value at which steady bifurcation becomes subcritical and

$$R_{bi}\left(\frac{1+\tau}{\tau}\frac{Pr}{Pr+1}\right) = R^{(0)} = \frac{R^{(2)}}{\dfrac{1+\tau}{\tau}\dfrac{Pr}{Pr+1}}.$$

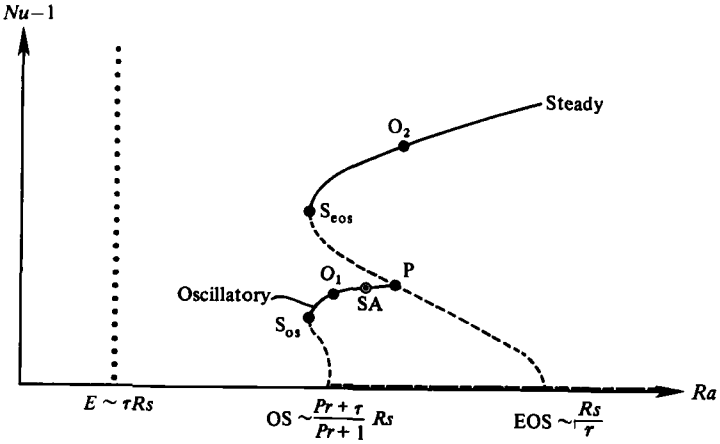

FIGURE 2. Schematic composite bifurcation diagram for doubly diffusive convection with given (fixed) large Rs, small τ. The vertical axis is Nusselt number minus one, so the conductive solution lies along the horizontal axis. Solid lines denote stable branches; broken lines denotes unstable branches. Bifurcations occur at OS and EOS with Ra given by (1) and (2), while the vertical dotted line indicates the energy stability limit (3). The finite-amplitude branches emanating from the bifurcation points are both subcritical for the (τ, Pr, Rs)-range of interest, because the overstable mode becomes subcritical between 8500 and 9000. The turning points (S_{os}, S_{eos}) on the subcritical branches are probably much nearer OS than E; both are to the left of the OS bifurcation. The oscillatory branch may contain a strange attractor (SA) and may end on the steady branch with infinite period at point P. Observed pond motions could be either finite-amplitude oscillatory motions (O_1), or steady motions (O_2), or even chaotic motions (SA), depending on disturbances applied after Ra exceeds the linear critical value at OS.

An early indication that this connection might occur at moderate amplitudes was given by Veronis (1965) by solving the second-order model problem (the doubly diffusive version of the Saltzman (1962)–Lorenz (1963) model of the Bénard problem). He found that the EOS branch was strongly subcritical, and turned around at $Ra = R_f$. He also clearly identified the mode of such a steady motion with the much more effective convective transport of salt as compared to heat when τ is small. However, in the purely thermal problem, subsequent elaboration of the Lorenz model indicated that the complications, including chaotic motions (strange attractors), appeared at larger and larger amplitudes for more accurate solutions (Curry 1977; Toomre, Gough & Spiegel 1977; Marcus 1981), raising questions as to how sensitive nonlinear solutions are to detail. Experimental observation of chaotic motion by Gollub & Benson (1980) suggests that this behaviour is realistic and should appear on the Bénard bifurcation diagram.

For the doubly diffusive problem, further developments by asymptotic analysis (Veronis 1968; Siegman & Rubenfeld 1975; Rubenfeld & Siegman 1977; Knobloch & Proctor 1981; Da Costa, Knobloch & Weiss 1981; Proctor 1981) and by numerical simulation (Huppert & Moore 1976) establish that both OS and EOS branches are subcritical with respect to Ra_{os}, but that the subcriticality and complication of the solutions, like that in the Bénard problem, decreases as the model is improved (subcriticality of OS mode at the turning point S_{os} in figure 2 declines to less than 10 % of λ^o in Huppert & Moore (1976), which has the case closest to our interest, i.e. $Pr = 10$, $\tau = 0.1$, $Rs = 1.4 \times 10^4$).

Thus, although the question is definitely still open, we assume that the onset of motion will occur near Ra_{os} since ponds have large $Rs \sim Rs_p$ of figure 1. Its form, as indicated in figure 2, could be either subcritical steady motion (point O_2), subcritical oscillatory (point O_1) or chaotic (SA). Support for this assumption comes from the absence of any experimental reports of extremely subcritical onsets of motion (Shirtcliffe 1968, 1970; Wright & Loehrke 1967; Zangrando 1979). Therefore we examine linear stability of a more pond-like basic state, with the stratifications of heat and salt being depth-dependent. These non-uniform gradients arise in ponds due to variability of thermophysical properties of the brine and due to construction and maintenance of the gradient layer, so they are almost always present.

The effects of variable stratification can be treated as follows. Consider a layer with a quasisteady conductive solution $S_c(z)$, $T_c(z)$; let these functions be linear, except that $S_c(z)$ has a 'dent' in it somewhere in the layer's interior. Now a new lengthscale, the radius of curvature of the dent, has entered the problem. If this radius is very small, viscosity can prevent motion, even if the overall gradients are very close to neutral stability and the dent therefore has buoyant potential energy available locally. If the radius of curvature is large enough, motion will occur, but most of the layer, being stably stratified, will not participate except by secondary motions driven by the unstable region (weak penetrative convection). Thus the main motion will be confined to a sublayer of the scale of the dent in the salinity profile; i.e. localized vertically. This is precisely the behaviour of the pond described above. This vertical localization can be expected to be accompanied by horizontal localization of some scale if the sublayer is large enough. The principal aim of this paper is the determination of these lengthscales. The same question has been addressed by Walton (1982) by use of perturbation methods, employing our numerical results in §3.2 to guide the selection of the lengthscales that appear in the analysis of the overstable modes. No analogous problem seems to have been solved for the Bénard case, although Krishnamurti (1968*a*, *b*) considered a linearly varying temperature gradient.

The detailed goal is to determine how large the layer must be in order that the critical mode is effectively localized so that it has the eigenvalue given by Walton (1982). In addition, we seek to determine whether or not other modes are possible within the linear theory. These questions are answered by numerical solution of the eigenvalue problem posed in §2. The results are enumerated in §3 and discussed in §4. The conclusions of the analysis are summarized in §5.

2. Statement of the problem

Consider a static plane-parallel stratification of a solvent (water) with depth h, containing solute (salt) concentration $S^* = S_r + \Delta S S_c(z)$, and heated to $T^* = T_r + \Delta T T_c(z)$. Asterisks indicate dimensioned variables; S_r, T_r, ΔS, ΔT are specified reference values, and the dimensionless vertical coordinate $z = z^*/h$ is positive downward. The dimensionless equilibrium (conductive) stratification $S_c(z)$, $T_c(z)$ would in general depend on time $t = t^*/(h^2/\kappa)$ as well as z, but it is assumed that the evolution timescale of the stratification is sufficiently long compared to convective timescales that this dependence can be ignored. Then, if both S_c and T_c are nonlinear functions of z, the thermal diffusion timescale is controlling (since $\tau = D/\kappa$ is small here), while if $T_c(z)$ is linear, so that only S_c is evolving, the salt-diffusion timescale (a factor of $1/\tau$ longer) is controlling. Thus dimensionless timescales of the convection must be much less than $O(1)$ in the former case, $O(1/\tau)$ in the latter. Thermophysical properties are presumed constant through the layer depth h.

Let the dimensioned density ρ^* of the solution depend on temperature and salinity linearly:

$$\rho^* = \rho_r^*[1 - \alpha(T^* - T_r) + \beta(S^* - S_r)], \tag{4}$$

where

$$\alpha = -\frac{1}{\rho_r^*}\frac{\partial \rho^*}{\partial T^*}, \quad \beta = \frac{1}{\rho_r^*}\frac{\partial \rho^*}{\partial S^*}.$$

All variables subscripted r are taken at the same reference state. If in addition density changes are considered only in the body force term of the equations of motion (Boussinesq approximation), then standard manipulations (e.g. Turner 1973; Zangrando 1979) produce the following linearized equations for the evolution of small convective perturbations $\overline{w}$, T, S in vertical velocity, temperature and salinity respectively:

$$\left(\frac{\partial}{\partial t} - Pr\nabla^2\right)\nabla^2\overline{w} = -Pr\,Ra\,\nabla_2^2\overline{T} + Pr\,Rs\,\nabla_2^2\overline{S}, \tag{5a}$$

$$\left(\frac{\partial}{\partial t} - \nabla^2\right)\overline{T} = -f\overline{w}, \quad \left(\frac{\partial}{\partial t} - \tau\nabla^2\right)\overline{S} = -g\overline{w}. \tag{5b,c}$$

That is, the convective motions have $T^*(x,y,z,t) = T_r + \Delta T[T_c(z) + \overline{T}(x,y,z,t)]$, $S^* = S_r + \Delta S[S_c(z) + \overline{S}(x,y,z,t)]$, $w^* = (\kappa/h)\,\overline{w}(x,y,z,t)$, with the overbarred quantities governed by (5). The quantities f and g on the right-hand sides of (5b, c) are respectively the dimensionless temperature gradient $f(z) = \mathrm{d}T_c/\mathrm{d}z$ and salinity gradient $g(z) = \mathrm{d}S_c/\mathrm{d}z$ of the conductive solution, and it is their dependence on z that is novel here. The notation ∇_2^2 is the two-dimensional Laplacian, $\partial^2/\partial x^2 + \partial^2/\partial y^2$, while ∇^2 is the full Laplacian.

Because the convective layers are observed to appear in the interior of the gradient zone $0 \leqslant z \leqslant 1$, and because localization is expected to make the solutions insensitive to boundary conditions, free–free boundary conditions are specified:

$$\overline{w} = \frac{\partial^2\overline{w}}{\partial z^2} = \overline{T} = \overline{S} = 0 \quad \text{at } z = 0, 1. \tag{6}$$

With these boundary conditions, the system (5) and (6) can be solved by normal-mode analysis, wherein the behaviour of perturbations of horizontal wavenumber $a = (k_x^2 + k_y^2)^{\frac{1}{2}}$ is examined; i.e. one sets

$$[\overline{w}, \overline{T}, \overline{S}] = (w(z), T(z), S(z)] \exp\{\mathrm{i}(k_x x + k_y y) + pt\}.$$

In general, $p = p_r + \mathrm{i}p_i$ is a complex number, and the (complex) magnitudes $w(z), T(z), S(z)$ satisfy

$$[p - Pr(\mathrm{D}^2 - a^2)](\mathrm{D}^2 - a^2) w = a^2 Pr(Ra\, T - Rs\, S), \tag{7a}$$

$$[p - (\mathrm{D}^2 - a^2)] T = -fw, \quad [p - \tau(\mathrm{D}^2 - a^2)] S = -gw, \tag{7b,c}$$

as a consequence of (5). Here $\mathrm{D}^2 = \mathrm{d}^2/\mathrm{d}z^2$, and boundary conditions for (7) are given by (6) without overbars. The stability problem thus becomes the determination, as functions of (τ, Pr, Rs), of those values of Ra and p_i (say) at which $p_r = 0$ (neutral stability. The system (7) can be reduced to a single equation by elimination of T and S in favour of w, giving

$$(\mathrm{i}p_i - \tau\nabla^2)(\mathrm{i}p_i - \nabla^2)(\mathrm{i}p_i - Pr\nabla^2)\nabla^2 w = a^2 Pr[-(\mathrm{i}p_i - \tau\nabla^2)(Raf w) + (\mathrm{i}p_i - \nabla^2)(Rs\, gw)], \tag{8}$$

where ∇^2 now means $\mathrm{D}^2 - a^2$. Recasting (6) as boundary conditions in w alone yields

$$w = \mathrm{D}^2 w = \mathrm{D}^4 w = \mathrm{D}^6 w = 0 \quad \text{on } z = 0, 1. \tag{9}$$

We note explicitly that non-constant gradients prevent extending these conditions to all even-order derivatives, as could be done for constant gradients (Chandrasekhar 1961).

We treat both the overstable case ((8) and (9) with $p_i \neq 0$) and the exchange-of-stabilities case ($p = 0$) for which (8) reduces to the form

$$[-\epsilon(\mathrm{D}^2 - a^2)^3 + g] w = \lambda f w. \tag{10}$$

The reduced order results from assuming that $w(z)$ is real, and observing that $(\mathrm{D}^2 - a^2) F = 0$ with $F = 0$ on $z = 0, 1$ can have only the trivial solution. The eigenparameter in (10) is

$$\lambda = \tau\, Ra/Rs,$$

while the (known) parameter ϵ is defined by

$$\epsilon = \tau/a^2 Rs.$$

This parameter ϵ is very small, $O(10^{-15})$, for a metre of heated brine.

When $\epsilon \to 0$, (10) becomes singular. In fact, a continuous spectrum of λ may be generated in this way.

The problem (10) subject to the first three of (9) is positive-definite and self-adjoint if $f(x)$ and $g(x)$ are both forbidden to change signs in $z \in (0, 1)$. Therefore this restriction is imposed on the temperature and salinity gradients; i.e. neither temperature nor salinity is allowed to have an extremum in the gradient layer. Solution procedure then becomes very efficient and straightforward, as described in Zangrando & Bertram (1984).

In order to treat the overstability case, it seems advantageous to proceed by slightly rewriting (8) in the form

$$\epsilon^0(\mathrm{i}p_i - \nabla^2)(\mathrm{i}p_i - \tau\nabla^2)(\mathrm{i}p_i - Pr\nabla^2)\nabla^2 w - (\mathrm{i}p_i - \nabla^2) gw = -(\mathrm{i}p_i - \tau\nabla^2)\lambda^0 f w, \tag{11}$$

where the 'o' superscript indicates the overstable problem, with

$$\epsilon^{o} = \frac{1}{a^2 Pr\, Rs} \quad \text{and} \quad \lambda^{o} = \frac{Ra}{Rs}.$$

Boundary conditions are again (9). This problem is not symmetric, so general statements cannot be made about its spectrum *a priori*. However, the constant-gradient case ($f = g = 1$) has only a finite number of real eigenvalues with real p_i associated, so this same behaviour is anticipated here (and has been observed to be true for all cases computed.) Numerical solutions are obtained by expanding in Fourier series. Note that, if one sets $p_i = 0$, the exchange-of-stabilities problem (10) is recovered from (11).

At this point it is convenient to consider what forms of $f(z)$ and $g(z)$ are physically relevant to the solar-pond application. From inspection of preconvective experimental data, it was decided that a cubic posesses enough structure to model the experiments (Zangrando 1979), giving the simplest gradient with an interior minimum, so the forms

$$T_c(z) = a_0 + a_1 z + a_2 z^2 + a_3 z^3, \qquad S_c(z) = b_0 + b_1 z + b_2 z^2 + b_3 z^3$$

were adopted for the algorithm. The coefficients are not independent, but must satisfy the conditions $T_c(0) = S_c(0) = 0$ and $T_c(1) = S_c(1) = 1$. Further, since the thermal diffusivity of water is so much larger than the salt-diffusion coefficient, $T_c(z)$ is expected to be much smoother than $S_c(z)$.

For this reason, the stratification was simplified by making $g(z)$ symmetric about midlayer $z = \frac{1}{2}$ and defining $f(z) \equiv 1$. Thus, for the cases discussed below,

$$\left.\begin{aligned} T_c(z) &= z, \\ S_c(z) &= z + 4(1 - g_{\min})\, z(z - \tfrac{1}{2})\,(z - 1), \\ g(z) &= \frac{dS_c}{dz} = g_{\min} + 12(1 - g_{\min})\,(z - \tfrac{1}{2})^2 \end{aligned}\right\} \tag{12}$$

will be the standard stratification. Note that the variability of the salt gradient is characterized by the single parameter $g_{\min}$ here, and that centring the 'dent' in the gradient *minimizes* the effects of the boundary conditions on the solutions (see §3 below; this is in contrast with the cases studied by Shirtcliffe or Krishnamurti).

Besides allowing a single-parameter representation of the departure from the constant-gradient case, (12) also provide a stratification S_c, T_c that is antisymmetric about $z = \frac{1}{2}$. This gives solutions $w(z)$ which are either even functions of $z - \frac{1}{2}$ or odd ones; these can be separated and solved using a discrete problem of half the dimension of the full problem, a critical saving, since the parameter study was limited by computer time.

The solutions investigated will thus be dependent on the parameters $Pr, \tau, g_{\min}, Rs$. The strength of the dependence on the first two can be inferred from the analytical solutions to the constant-gradient problem; here we study only the effects of varying $g_{\min}$ and Rs, with the first being a shape parameter of the variable gradient and the second being viewed essentially as a scale parameter which varies as the fourth power of layer depth h when the average gradients over the layer are kept constant (1 °C/cm is typical). For heated brine, $Rs = 10^4$ corresponds to about $h = 1$ cm, and $Rs = 10^{12}$ to about $h = 100$ cm.

3. Results

The results summarized here are detailed in Zangrando & Bertram (1984), as are the numerical methods used. The procedure is to compute the eigenvalues λ, λ^0 for parameters in the ranges $0.05 \leqslant g_{\min} \leqslant 1.0$ and $10^4 \leqslant Rs \leqslant 10^{12}$, then vary a^2 until a minimum eigenvalue is achieved, and identify this mode with a 'c' (critical) subscript: $a^2 = a_c^2$ yields $Ra = Ra_c$.

Since the purpose of the parameter study was to quantify the anticipated contest between buoyant potential energy and viscous dissipation, it is useful to introduce the dimensionless measures:

$$\left.\begin{aligned} \left(\frac{d}{h}\right)_\lambda &= \text{fraction of layer depth in which } \frac{\lambda f(z)}{g(z)} \geqslant \lambda_r, \\ \left(\frac{d}{h}\right)_w &= \text{fraction of layer depth in which } |w(z)| \geqslant 0.1|w|_{\max}. \end{aligned}\right\} \tag{13}$$

The reference constant-gradient eigenvalue λ_r is defined for $Rs \to \infty$ and for the same total temperature and salinity changes across the layer as the non-constant-gradient layer under consideration. Thus it usually represents an upper bound on the non-constant-gradient eigenvalue, since the latter has $g(z) < 1$ in the active sublayer. For exchange of stabilities this will be given by

$$\lambda_r = \frac{\tau\, Ra_{\text{eos}}}{Rs} = 1;$$

and for overstable modes by

$$\lambda_r^0 = \frac{Ra_{\text{os}}}{Rs} = \frac{Pr+\tau}{Pr+1}.$$

For convenience in comparing eigenvalues obtained at different $g_{\min}$, a lower bound on eigenvalues is introduced by noting that no buoyant potential energy can be available in any sublayer until the weakest salinity gradient is small enough that the marginal-stability condition (3) is reached. That means, since $g(z) \geqslant g_{\min}$ everywhere, for exchange of stabilities the non-constant-gradient eigenvalue must exceed

$$\lambda_L = \lambda_r g_{\min} = g_{\min};$$

and for oscillatory bifurcation it must exceed

$$\lambda_L^0 = \frac{Pr+\tau}{Pr+1} g_{\min}.$$

The subscript 'L' indicates 'local' values, evaluated at the weakest salinity gradient of the layer. $(d/h)_w$ quantifies the size of the region in which the 'main' circulation (as opposed to viscously driven secondary motion) has significant magnitude. The eigenvalues plotted in the figures are $\Lambda = (\lambda - \lambda_L)/\lambda_L$ and $\Lambda^0 = (\lambda^0 - \lambda_L^0)/\lambda_L^0$. Attainment of the limit $\Lambda^0 \to 0$ with increasing Rs will be referred to as 'full localization'.

3.1. *Exchange of stabilities*

A particularly simple treatment of these results is possible for the case $f(z) \equiv 1$, when $g(z)$ is given by (12); then (10) becomes

$$\begin{aligned} -(\mathrm{D}^2 - a^2)^3 w &= a^2\left[\left(Ra - \frac{Rs}{\tau} g_{\min}\right) - 12(1-g_{\min})\frac{Rs}{\tau}(z-\tfrac{1}{2})^2\right] w \\ &= a^2[L - M(z-\tfrac{1}{2})^2]\, w, \end{aligned} \tag{10a}$$

where $M = 12(1-g_{\min})\,Rs/\tau$ is a specified constant independent of a^2, and

$$L = (\lambda - g_{\min})\,Rs/\tau$$

can be treated as the eigenvalue: $L = L(a^2, M)$. It follows immediately that we get the same L for every point on the locus $(1-g_{\min})\,Rs$ = constant for a given a^2, and therefore get the same $L_{\min}$ value at the same a_c^2. Note that $L \to 0$ for full localization.

The invariance of the eigenfunction $w(z)$ and critical eigenvalue $L_{\min}$ allows negative values of Ra and Rs (the 'fingering' regime) as well, so long as $g_{\min}$ is replaced by $g^*_{\max}$, which must be greater than one; then the transformation

$$Rs^*(1-g^*_{\max}) = Rs(1-g_{\min})$$

makes M = constant, and the critical Ra^* is given by

$$Ra^* = Ra + (Rs^* g^*_{\max} - Rs\, g_{\min})/\tau,$$

which is also negative when Rs^* is negative. Thus our numerical results are solutions for the heated- and salted-above case, as well as for the heated- and salted-below case for which they are derived.

The same argument can be applied to the purely thermal case with parabolic temperature gradient $f^*(z)$:

$$-(\mathrm{D}^2 - a^2)^3 w = a^2 Ra^* [f^*_{\mathrm{m}} + 12(1-f^*_{\mathrm{m}})(z-\tfrac{1}{2})^2]\, w. \tag{10b}$$

This problem has the same eigenfunction $w(z)$ as (10a) when its eigenvalue is

$$Ra^* = L/f^*_{\mathrm{m}}$$

and its stratification parameter is

$$f^*_{\mathrm{m}} = \frac{1}{1 - M/12L} = -\left(\frac{1-\lambda}{\lambda - g_{\min}}\right)^{-1} < 0.$$

Both f^*_{m} and Ra^* are negative because this problem must have an unstable central layer working against a strongly stable stratification near its boundaries, in order to be analogous to our problem. Because of this, $\lim f^*_{\mathrm{m}} \to 1$ is not possible, i.e. the Bénard problem is not a limit of (10b). Because both M and L must be varied to keep f^*_{m} constant, the critical a^2 for the thermal problem is found by seeking crossings of the f^*_{m} = constant rays and the numerically determined $L_{\min}[M, a_c^2(M)]$ locus in the (L, M)-plane. Thus we also solve the thermal problem when we solve (10a).

Returning to the doubly diffusive case (10a), we note that the solution of interest is the one with a_c^2, so it is natural to scale by the transformation $z_s = a_c(z-\frac{1}{2})$, resulting in a stretched version of the equation

$$-(\mathrm{D}_s^2 - 1)^3 w = (L_s - M_s z_s^2)\, w, \tag{10c}$$

where $\mathrm{D}_s^2 = \mathrm{d}^2/\mathrm{d}z_s^2$, $L_s = L/a_c^4$ and $M_s = M/a_c^6$, which is to be solved on the interval $-\frac{1}{2}a_c \leqslant z_s \leqslant \frac{1}{2}a_c$. All forms of (10), including (10c), are exact. As Rs is increased, the interval length of this version of the problem increases. When it is long enough, our hypothesis that the localized solution be insensitive to boundary data must be reflected in $w(z_s)$, L_s and M_s becoming asymptotically invariant. Consulting the numerical solutions, we find that this invariance holds for every case computed, with $M_s = 66.7 \pm 0.4$ and $L_s = 35.2 \pm 0.1$ for $\tau = \frac{1}{80}$. Translated into the original variables of the problem, with g_{m} implying either $g_{\min}$ or $g_{\max}$, these numbers give

$$a_c = \left[\frac{12(1-g_{\mathrm{m}})\,Rs}{\tau M_s}\right]^{\frac{1}{6}} = 1.560[(1-g_{\mathrm{m}})\,Rs]^{\frac{1}{6}} \tag{14a}$$

and $$\lambda_c = g_m + a_c^4 L_s \tau Rs^{-1} = g_m + 2.604(1-g_m)^{\frac{4}{3}} Rs^{-\frac{1}{3}}. \tag{14b}$$

Thus these simple considerations yield both the information that all the numerical solutions are fully localized, and the forms of the first correction terms as well. These are, of course, precisely the forms required to scale Walton's equation (18) into the form (10*c*). Also, since the smallest interval length over which (10*c*) was solved in our cases was 19.3, it is not very surprising that (10*c*) had become effectively asymptotic in form.

The eigenfunction, plotted as roll stream function, is shown in figure 3. Reducing the scale by a factor $100^{\frac{1}{6}} = 2.154$ at each 100-fold increase of Rs should leave the eigenfunction invariant; the graphically more convenient scale factor of 2 that is used in figure 3 clearly shows this is the case. Furthermore, the width/height aspect ratio of the central cells has the constant value 0.71–0.72 for all cases. The scale was chosen so that the maximum stream-function amplitude $|\Psi|_{\max} = 0.11$; this scale is chosen to facilitate comparison with Huppert & Moore's (1976) finite-amplitude solutions obtained for constant-gradient conditions. Figure 4 shows the density of the eigenvalues near λ_c, for $Rs = 10^{12}$. This density implies that interactions among modes are more easily excited as Rs increases.

The d/h measures of localization can be evaluated from (14). From the definition of $(d/h)_\lambda$ in (12) and from (13)

$$\left(\frac{d}{h}\right)_\lambda = 2\left[\frac{\lambda_c - g_m}{12(1-g_m)}\right]^{\frac{1}{2}} = 0.932[(1-g_m)\,Rs]^{-\frac{1}{6}}, \tag{14c}$$

which is a simple rearrangement of the result for λ_c. The numerical solutions uniformly show $(d/h)_w \approx 2.33(d/h)_\lambda$, so

$$(d/h)_w = 2.17[(1-g_m)\,Rs]^{-\frac{1}{6}} \tag{14d}$$

accurately represents our numerical results.

Carrying out the calculation for a_c^2 for the thermal problem (10*b*), using (14) to get $L_{\min}(M)$, results in

$$Ra_c^* = \frac{1410(f_m^*-1)^2}{f_m^{*3}}, \quad a_c^* = 2.52\left[\frac{f_m^*-1}{f_m^*}\right]^{\frac{1}{2}}. \tag{15}$$

In this expression, f_m^* is constrained to be small, $O(Rs^{-\frac{1}{3}})$, by equations (14*a*, *b*).

3.2. *Overstable modes*

For constant gradients the overstable modes may be obtained analytically. The result differs from the exchange-of-stabilities result in that there are only a finite number of neutrally stable modes, and these exist only for a finite range of wavenumbers $a_m^2 \leqslant a^2 \leqslant a_M^2$, where $a_{m,M}^2$ are functions of τ, Pr, Rs and eigenindex n. Because the non-constant-gradient problem (9) and (11) is not symmetric, no general statements can be made about its spectrum when $g_{\min} \neq 1$, but it does appear that it shares the properties of the spectrum of the $g_{\min} = 1$ case.

The numerical solutions for $g_{\min} < 1$ have in addition the following difficulties.

(1) To solve the problem (9) and (11), the frequency is iterated until $\mathrm{Im}\{\lambda^0\} = 0$ is satisfied at eigenfrequency $\hat{p}_i$. This iteration proves to be computationally difficult because of multivalued and crowded branches $\lambda_{(n)}^0(p_i)$ in the complex λ^0 plane.

(2) The multivalued and crowded λ^0 traces for a given a^2 are reflected in multivalued and crowded $\Lambda_{(n)}^0(a^2)$ curves. In fact, the eigenvalues and eigenfrequencies are triple-valued for some a^2, Rs values. This makes computational identification of critical modes difficult also.

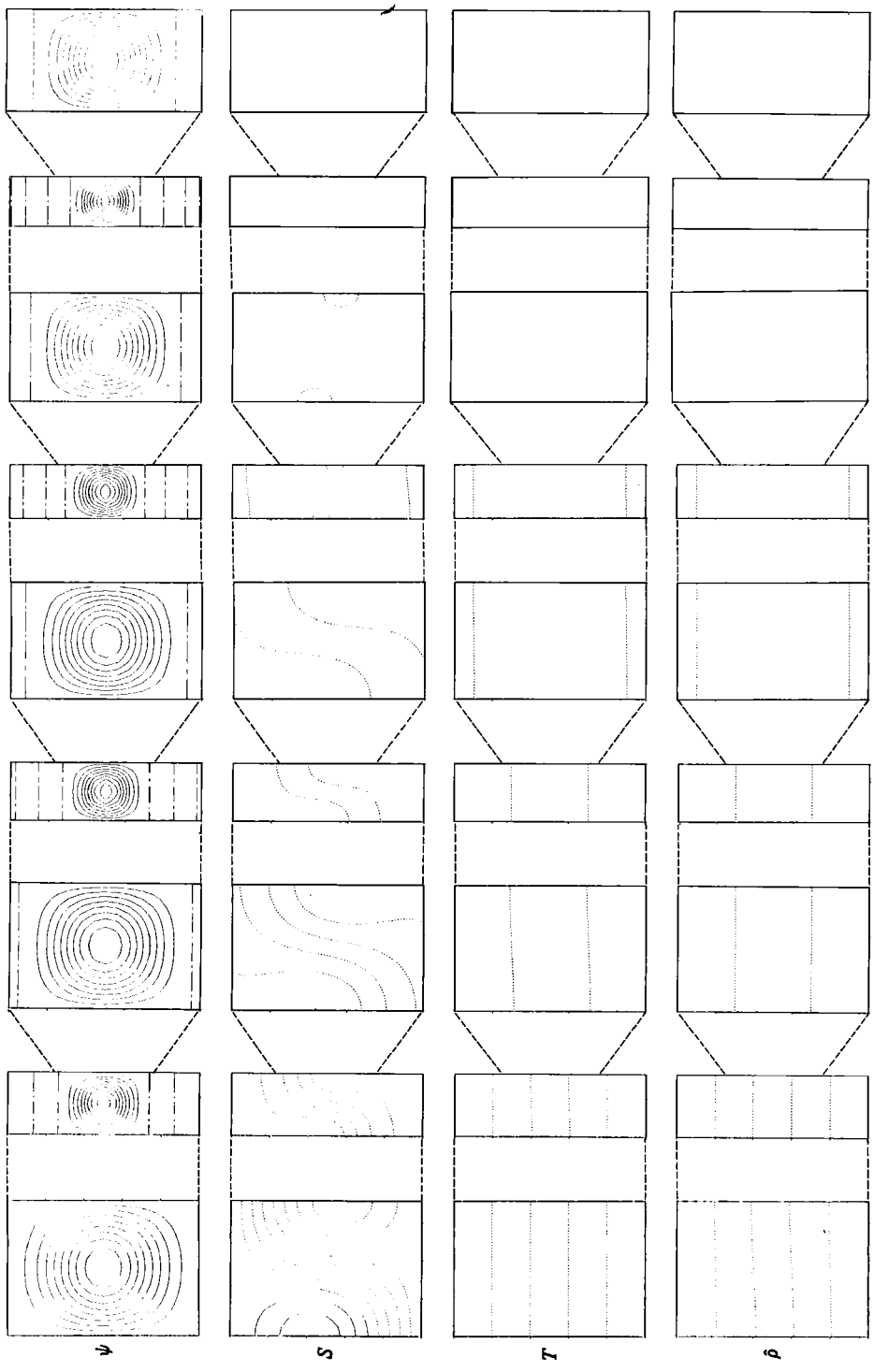

FIGURE 3. Exchange-of-stabilities eigenfunction localization with increasing size (Rs); $Pr = 7$; $\tau = \frac{1}{60}$; $g_{\min} = 0.95$. Broken-line contours have positive values; solid contours have negative values; chain-dashed contours are zero values. First row is streamlines with $\Delta\psi' = \frac{1}{60}$; remaining rows are respectively dimensionless salinity, dimensionless temperature and dimensionless density $\hat{\rho}$, each with spacing 0.2. Magnifications by 2× are indicated by diverging lines between frames. (*a*) $Rs = 10^4$, $Ra_c = 7.732\,53 \times 10^5$, $a_c^2 = 19.3$; (*b*) 10^6, $7.628\,25 \times 10^7$, 90; (*c*) 10^8, $7.606\,09 \times 10^9$, 415; (*d*) 10^{10}, $7.601\,311 \times 10^{11}$, 1927; (*e*) 10^{12}, $7.600\,28 \times 10^{13}$, 8924.

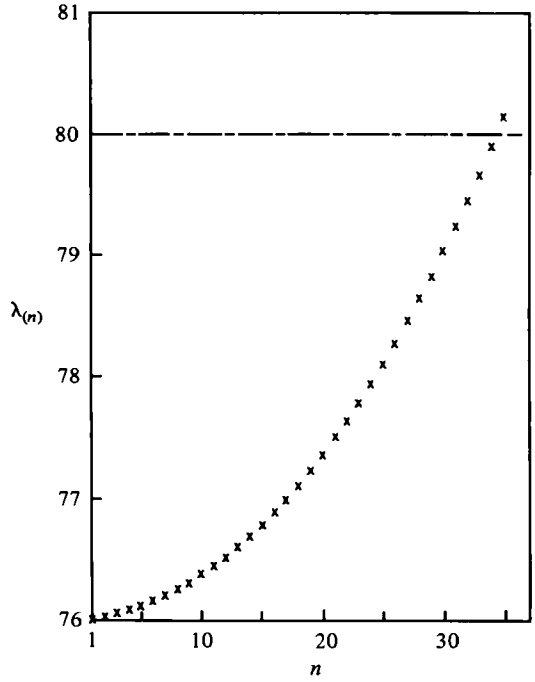

FIGURE 4. First 35 eigenvalues for exchange-of-stabilities modes at $Rs = 10^{12}$, $g_{\min} = 0.95$, $a^2 = 8924$. These 35 fall between lowest eigenvalue of non-constant-gradient problem and eigenvalue of constant-gradient solution with $f = g = 1$. Approach to continuous spectrum is apparent.

(3) Because of the difficult behaviour of the $\Lambda^0_{(n)}(a^2)$ curves, and the similar behaviour of eigenfrequency curves $\hat{p}_i(a^2)$, no simple method of assigning eigenindex to a particular numerical solution exists. A very tedious and expensive procedure was used here: index was assigned by ordering branches in ascending order of eigenvalue at low a^2 where the eigenbranches are widely separated. Then the continuity of the branch, inferred from the fact that the coefficients of (11) are entire functions of the parameters, is used to identify the index appropriate for a particular solution at a large a^2 value.

Each of these difficulties is illustrated in Zangrando & Bertram (1984), which describes the basic numerical procedure used to obtain individual eigensolutions; i.e. the set of values $\{\lambda^0_{(n)}\}$ for a given parameter set $(\tau, Pr, Rs, g_{\min}, a^2)$. The technique by which the critical mode is extracted from these solutions is also described there; that is, the process of obtaining

$$\lambda^0_c \equiv \min_{a^2,\, n} \{\lambda^0_{(n)}(a^2)\}.$$

The resulting critical wavenumbers are displayed in figure 5.

When the layer is small, as in the case $Rs = 10^4$, viscous forces clearly should favour motions with minimal spatial structure, and indeed the critical mode is very similar to the constant-gradient $n = 1$ mode with $\sin \pi z$ vertical structure and $a_c^2 = \frac{1}{2}\pi^2$. The critical wavenumber at this Rs-value departs from $\frac{1}{2}\pi^2$ only for very large nonuniformity of salt stratification, with $g_{\min} < 0.5$.

On the other hand, when $Rs = 10^6$, the layer is sufficiently deep that more spatial structure appears in the eigenfunctions when $g_{\min}$ is decreased from unity. At first, for $1.0 > g_{\min} > 0.823$, the whole layer motion with $w(z) \sim \sin \pi z$ appears, but at a

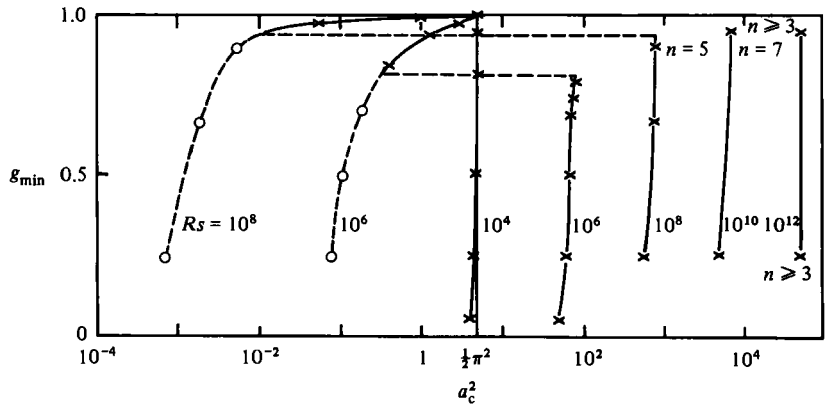

FIGURE 5. Critical horizontal wavenumber a_c^2 of overstable mode, as a function of minimum salinity gradient g_{min}. The × denotes critical mode; i.e. absolute minimum over a^2. The ○ denotes relative minimum on $\Lambda^o_{(1)}(a^2)$ curve at low a^2 values; i.e. at $a^2 < \frac{1}{2}\pi^2$. Eigenindex is noted if $n \neq 1$.

significantly reduced wavenumber. Then at lower g_{min} the critical mode has a very different vertical structure and much larger a_c^2 value, and this a_c^2 is only slightly reduced by further reductions in g_{min}. All these critical solutions, at both large and small a^2 values, belong to the $n = 1$ eigenbranch. Loops in $\lambda^o_{(1)}(a^2)$ sometimes appear between the relative minima.

The behaviour of the streamlines for one example from each of the high- and low-a^2 branches is shown in figure 6. Because these are linear oscillatory modes, it suffices to display a quarter-period of the motion. Therefore the streamfunction is given at time intervals of $\frac{1}{16}$ period, starting from $t = 0$. Note that the high-a^2 solution, in addition to displaying some localization of the main circulation, has weak but noticeable motions at $t = 0$. Essentially, these are a measure of the streamfunction phase, as a function of depth z. This phase difference, which results from the variable stratification, becomes more pronounced with larger a^2 in general.

Increasing Rs to 10^8 enhances the shift in critical wavenumber between low-a^2 and high-a^2 critical solutions, as shown in figure 5. The jump between branches occurs at $g_{min} \approx 0.94$, much nearer constant-gradient conditions than the jump in the $Rs = 10^6$ branches. In addition, the nature of the high-a^2 eigenfunction has changed – rather than being a member of the $n = 1$ branch, it comes from the $n = 5$ branch at $g_{min} = 0.90$, as determined by tracing eigenvalue curves $\lambda^o_{(n)}(a^2)$.

The low-a^2 relative minima of $\lambda^o(a^2)$ were located only for $g_{min} = 0.25$ and 0.95 for the very large Rs-values of 10^{10} and 10^{12}, because of computational expense. Since these are not the critical modes for these Rs-values, no motivation exists for tracing out the low-a^2 branches in detail.

Returning to figure 5, it can be seen that the high-a^2 critical mode at $Rs = 10^{10}$, $g_{min} = 0.95$ has $n = 7$, while the $g_{min} = 0.25$ mode has $n = 1$. At $Rs = 10^{12}$ the effort of tracing out a sufficient number of branches in enough detail to determine eigenindices of the critical modes was excessive, so calculation was done only to establish that at least one crossover of branches occurs. Thus, as indicated in the figure, all that can be said for the two cases of g_{min} at $Rs = 10^{12}$ is that $n \geqslant 3$.

In order to display the dependence of the overstable critical values on g_{min}, the scaled critical value

$$\frac{\lambda_c^o}{\lambda_r^o} = \frac{Ra_c}{Ra_{os}} = \frac{\Delta T_{max}}{\Delta T_{max,\, g_{min}=1}}$$

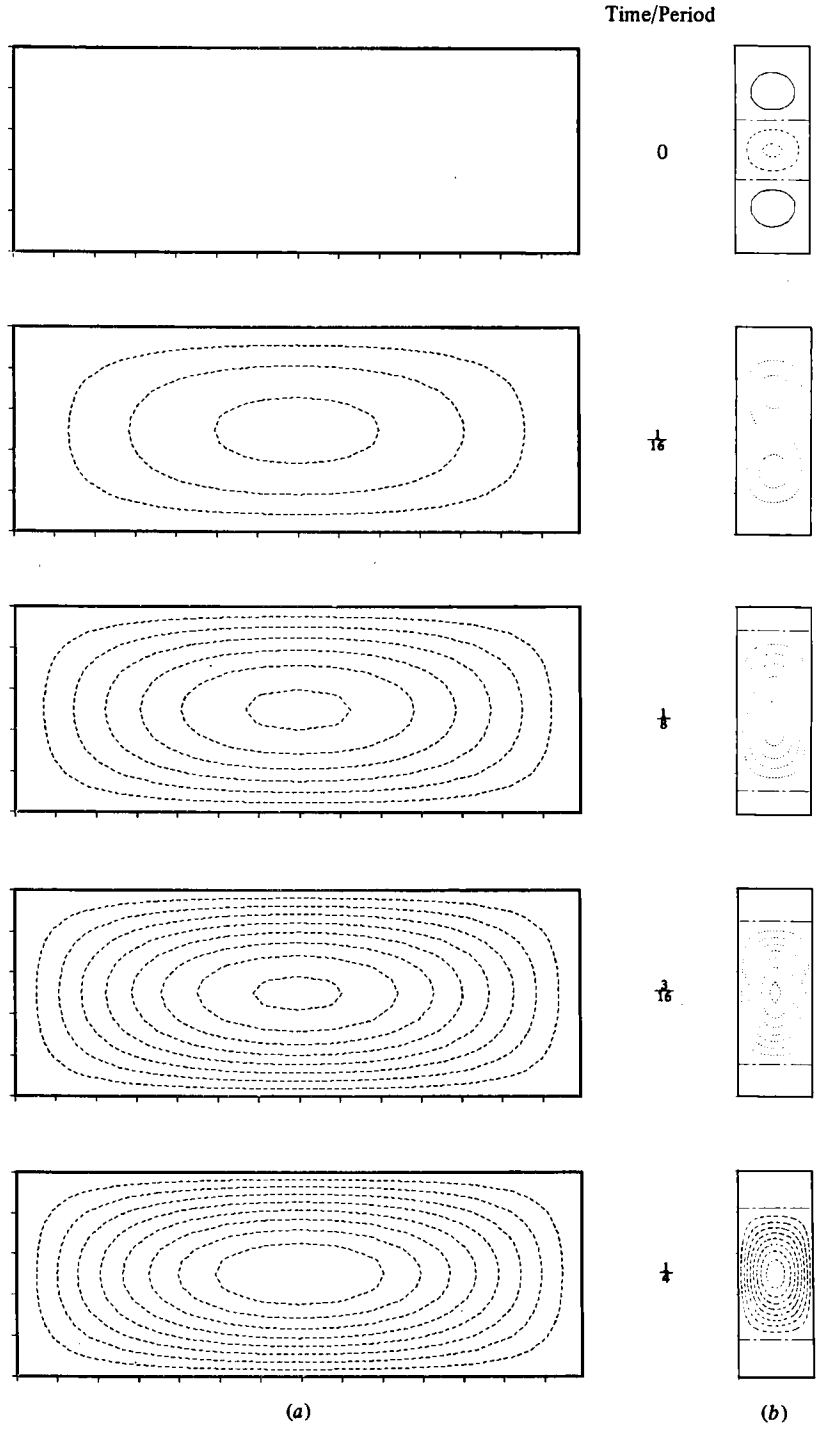

FIGURE 6. Stream functions of critical overstable modes for $Rs = 10^6$. (*a*) $g_{\min} = 0.95$, $a^2 = a_c^2 = 1.273$, $\hat{p}_1 = 308.26$; i.e. a low-a^2 (whole layer) solution. (*b*) $g_{\min} = 0.75$, $a^2 = a_c^2 = 76.0$; i.e. a high-a^2 (localized) solution. Chain-dashed line is $\psi = 0$ streamline. The enclosing rectangle has $z = 0$ at the top, $z = 1$ at the bottom, $x = 0$ at the left, and $x = \pi/a$ at the right. Time increases from top to bottom in timesteps of $\frac{1}{16}$ period.

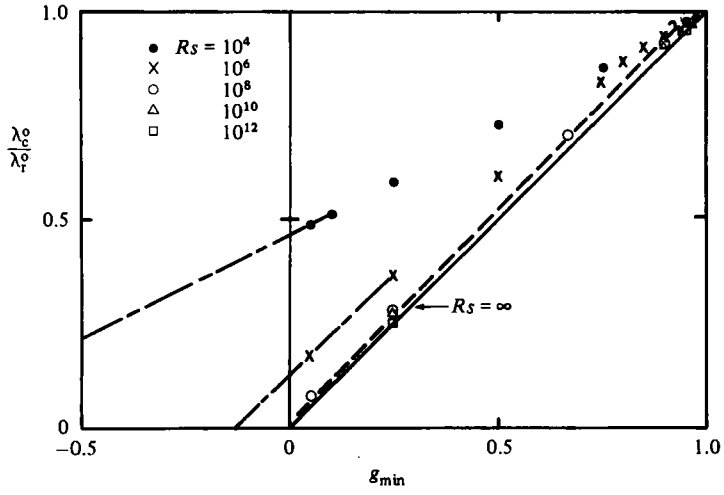

FIGURE 7. Dependence of overstable critical Ra_c on g_{min}. The straight solid line represents the infinite-Rs limit. The broken line is Walton's solution for $Rs = 10^8$ (equation (21)). This solution is valid up to $g_{min} = 0.94$. The chain-dashed lines extrapolating $Rs = 10^4$ and 10^6 indicate that *negative* g_{min} is stable for small-enough layer sizes.

is used (see figure 7). This choice makes explicit how much the thermal insulating capacity of a pond gradient layer is decreased by the redistribution of the salt. It also allows full display of the numerical points.

As anticipated, the calculated eigenvalues fall between the constant-gradient values $\lambda_r^0(R_{ac}/R_{os} = 1.0)$ and the local value $\lambda_L^0(Ra_c/Ra_{os} = g_{min})$, for the range of parameters covered. As expected, large Rs causes the eigenvalue to approach the value λ_L^0, which is plotted as the solid line in the figure. At smaller Rs it is notable that extrapolations of the curves to zero eigenvalue requires negative g_{min}. This would imply that small reversed-salt-gradient zones can be tolerated, apparently because they are small enough that viscous stabilization is possible. At every Rs, the intercept of the eigenvalue curve with $g_{min} = 0$ appears to be positive, so these reversed-gradient solutions are possible for every finite Rs.

The dependence of these solutions on Rs is made more explicit in figure 8. The plots indicate parameter values for the relative minima at both high and low a^2 (figure 8*a*); thus the critical mode is the relative minimum with the smaller A^0 value. For $g_{min} = 0.95$ this means that the critical mode is the low-a^2 branch for $Rs \lesssim 8 \times 10^8$, which is the Rs-value at which the eigenvalues of the two branches are equal in figure 8(*b*). Above this value, the high-a^2 branch is critical.

The eigenfrequency $\hat{p}_i$ is plotted in figure 8(*c*). For future reference, the constant-gradient ($g_{min} = 1$) value, which is given by

$$\hat{p}_i^2 = (1-\tau)\frac{Pr}{Pr+1}\frac{Rs}{3} - \left[\frac{3\pi^2}{2}\right]^2 \tau^2, \tag{16}$$

is plotted as the chain-dashed line. The low-a^2 branches for $g_{min} = 0.25$ and 0.95 have constant $\hat{p}_i$, and therefore appear as horizontal lines in the figure. With all these branches tending to converge at the left of the figure, the question as to how they are to be connected naturally arises. The connections indicated in the figure will be discussed below in §4. The localization tendency of the eigenfunction is apparent in figure 8*d*.

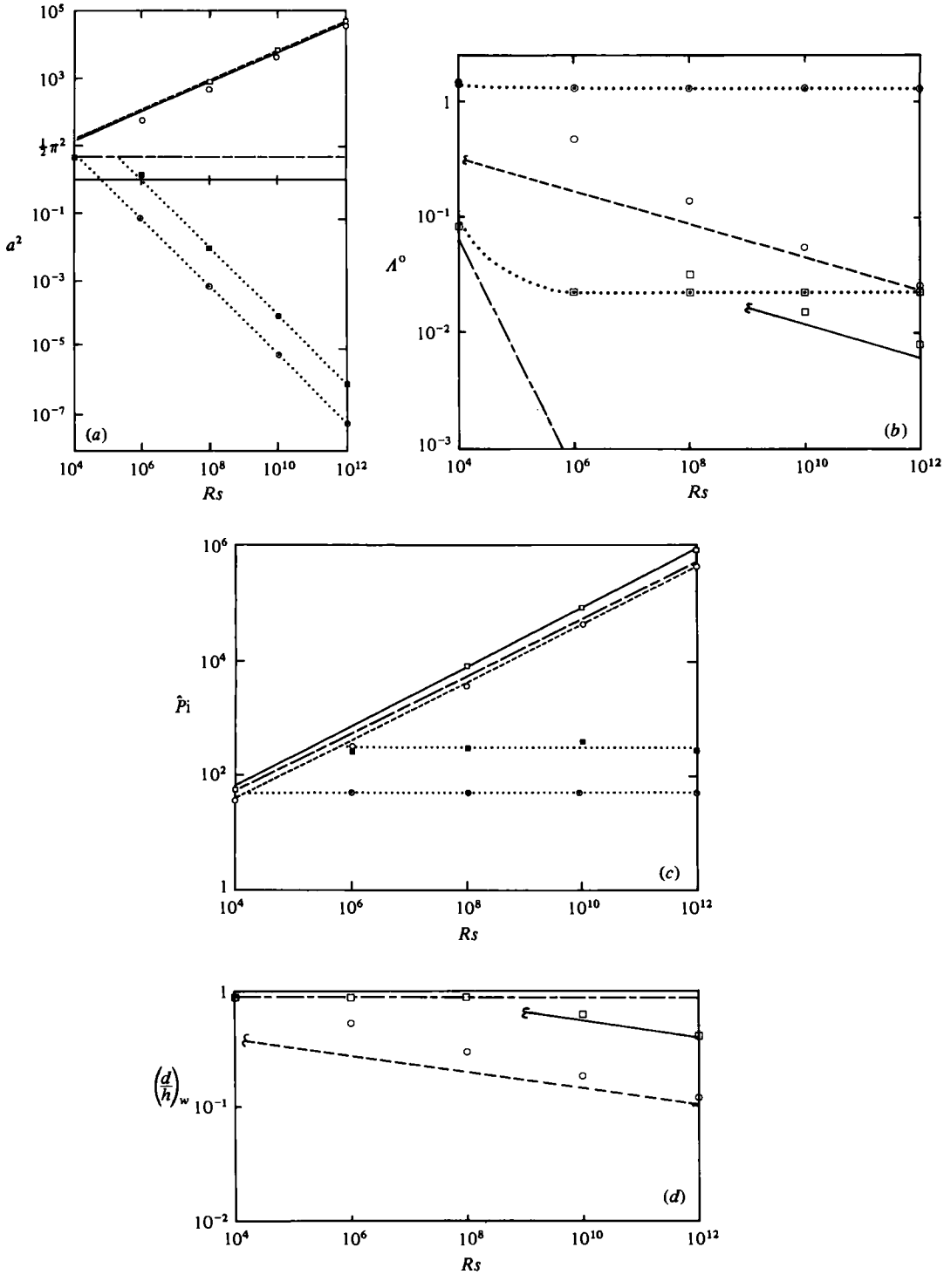

FIGURE 8. For caption see facing page.

The stream function derived from the eigenfunction is plotted for $g_{\min} = 0.95$ in figure 9. Again, rolls are assumed, and $|\psi|_{\max} = 0.11$ is assigned, as was done in the exchange of stabilities case in figure 3. The wavenumbers of these solutions are those identified as critical in figures 5 and 8(*a*).

If the critical eigenfunctions for $g_{\min} = 0.25$ are plotted in the same way, only high-a^2 modes with $n = 1$ are observed (see figures 5 and 10). Their structure develops from zero-phase streamfunctions similar to the one in figure 9(*a*), to complex structures that differ from those in figure 9(*d*) only in that they are somewhat more localized, and have somewhat different phase than the latter. Unlike the exchange-of-stabilities modes, the central cells for both $g_{\min}$ values display a central-cell aspect ratio (width/height) that decreases strongly with Rs.

4. Discussion

The critical modes for exchange-of-stabilities case proved to be quite conventional, with $n = 1$ in all cases, and a single relative minimum on the $\lambda^0_{(1)}(a^2)$ curves. The progressive localization with increasing Rs or decreasing $g_{\min}$ is made explicit by the stretching defined in (10*c*).

For the overstable critical modes, on the other hand, the numerical solutions display considerable complication. Both high- and low-a^2 branches appear, and the high-a^2 modes undergo mode switching to higher eigenindex as Rs or $g_{\min}$ is increased. Nevertheless, considerable order can be inferred by inspecting the relations of these solutions in the $(Rs, g_{\min})$-plane (see figure 11).

First, we argue that even eigenfunctions $w(z)$ (i.e. symmetric about the midplane $z = \frac{1}{2}$) are always critical, because an odd eigenfunction will require more driving power to sustain the dissipation occurring in the additional velocity gradient at midlayer, but will be receiving less driving power because the displacement is zero where buoyant forces are maximum at midlayer. Next, we attribute the appearance of low-a^2 critical modes to two causes. When the layer is shallow, higher velocity gradients of localized $w(z)$ solutions as compared with the smooth $w(z) \sim \sin \pi z$ solutions mean that the high-a^2 solutions are harder to sustain. When the layer is deep and $g_{\min}$ is near unity, the critical mode must be nearly $\sin \pi z$, since this is known to be the constant-gradient solution. Therefore the hatched region in figure 11 has the correct general shape. Note that this shape suggests that the high-a^2 branch is isolated from the constant-gradient solutions in the sense that the jump to low-a^2 depicted in figure 5 is necessary before $g_{\min} = 1$ can be approached. Therefore the

FIGURE 8. Dependence of critical-mode parameters on Rs for overstable modes. (*a*) Wavenumbers of relative minima on high- and low-a^2 branches. The dotted lines are curve fits to low-a^2 solutions as $a^2 Rs = k$, with $k = 10^6$ and 0.7×10^5 respectively for $g_{\min} = 0.95$ and 0.25. The chain-dashed lines are constant-gradient-case analytical solutions for $g_{\min} = 1$. (*b*) Eigenvalue curves. The dotted lines are curve fit (25) with $A^0 = 0.023$ and 1.313 respectively. Note absence of high-a^2 numerical solution at $Rs \leqslant 10^6$. (*c*) Eigenfrequency curves. Low-a^2 curve fits are $\hat{p}_1 = \text{const} = 290.3$ and 51.466 respectively. (*d*) Localization of eigenfunctions, based on sublayer fraction in which $|w| \geqslant 0.1|w|_{\max}$. Sublayer fraction with buoyant potential energy available, $(d/h)_\lambda$, is given by $(d/h)_\lambda \approx 0.42(d/h)_w$.

	High-a^2	Low-a^2	Walton's solution
$g_{\min} = 0.95$	□	⊡	——
$g_{\min} = 0.25$	○	⊙	- - - - -

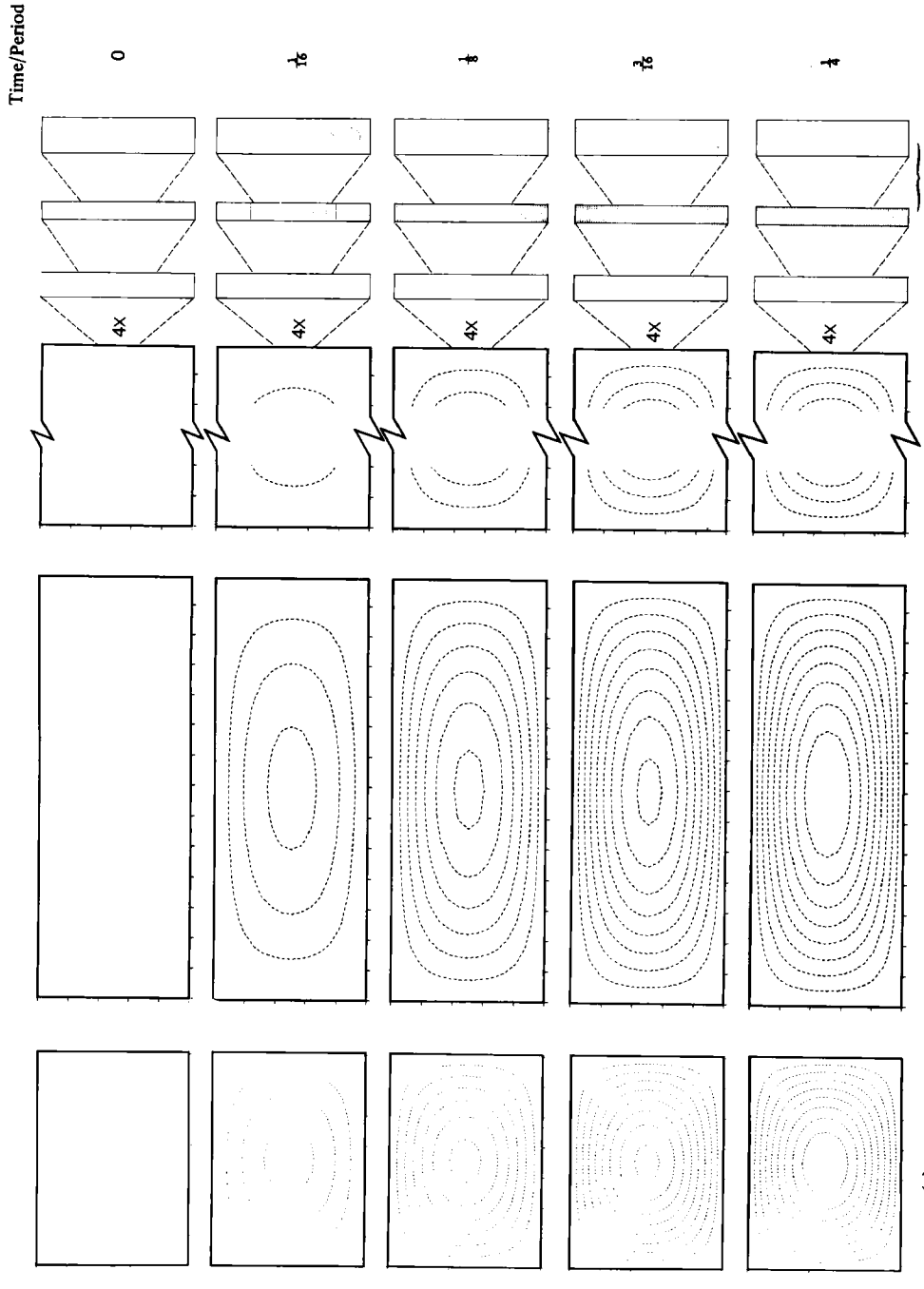

FIGURE 9. Critical overstable modes for $g_{\min} = 0.95$ with increasing Rs. Stream function is scaled to the same maximum as exchange-of-stabilities case in figure 3. (*a*) $Rs = 10^4$, $a_c^2 = 4.935$, $\hat{p}_1 = 52.822$, $n = 1$; (*b*) 10^6, 1.273, 308.3, 1; (*c*) 10^8, 0.012, 317.7, 1; actual cell has width $\pi/a_c = 28.68$ and is not plotted; (*d*) 10^{10}, 5900, 88223, 7, scale is increased by 4 ×; (*e*) 10^{12}, 46360, 894780, $n \geqslant 3$, scale is expanded twice, by 2 ×. Time increases from top row to bottom in time steps of $\frac{1}{16}$ period.

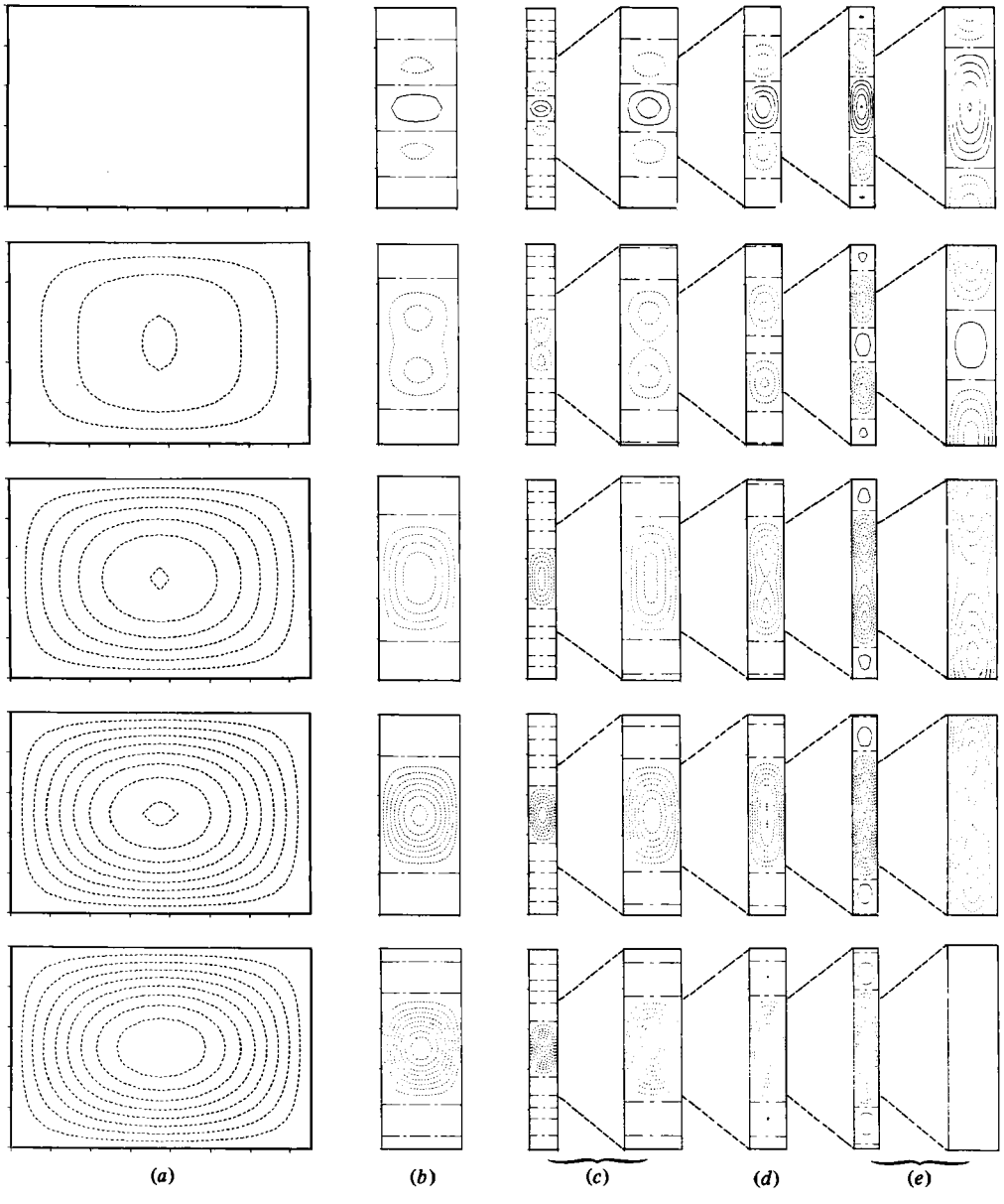

FIGURE 10. Critical modes for $g_{\min} = 0.25$ with increasing Rs. Streamfunction scale is as before. (*a*) $Rs = 10^4$, $a_c^2 = 4.36$, $\hat{p}_1 = 36.37417$, $n = 1$; (*b*) 10^6, 61.075, 342.13, 1; (*c*) 10^8, 478, 3738.3, 1; (*d*) 10^{10}, 4750, 41848, 1; (*e*) 10^{12}, 40421, 440970, $n \geqslant 3$. Broken lines indicate scale expansions by $2\times$. Time increases from top row to bottom in time steps of $\frac{1}{16}$ period.

low-a^2 branch has been shown in each case in figure 8 as connecting to the constant-gradient solutions, and the high-a^2 branches have been shown as isolated.

Within the unhatched zone of figure 11 where high-a^2 solutions are critical, the new feature to be explained is the occurrence of mode switching. As argued in §1, vertically localized solutions should prefer to have increasing horizontal localization as well, which means these solutions automatically occur at high a^2 values. In this zone, at $Rs \leqslant 10^6$ and small $g_{\min}$ values, an increase in the eigenindex results in a

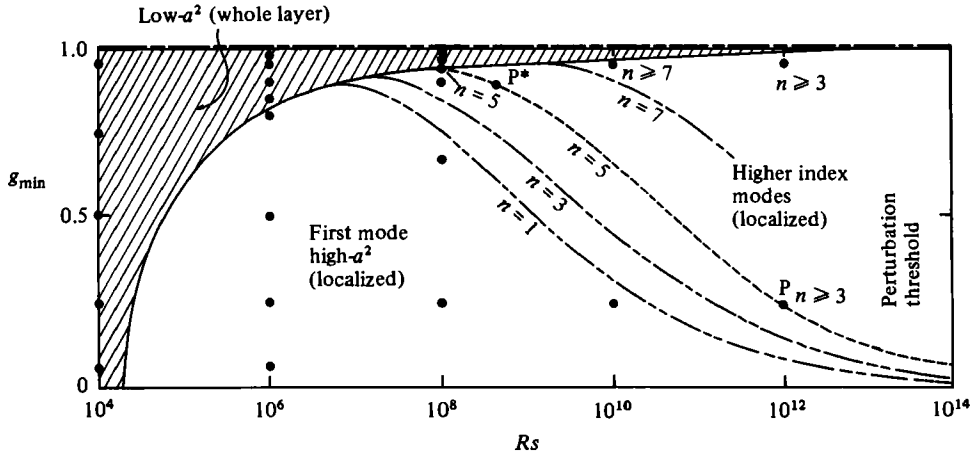

FIGURE 11. Critical-mode characterization for overstability at $\tau = \frac{1}{80}$, $Pr = 7$. Dots denote computation points: $n = 1$ etc. refers to eigenindex of critical mode. Chain-dashed line is constant-gradients case, double chain-dashed line is separatrix between zones with different critical eigenindex, broken line is sublayer scaling rule (eq. 20) for initial point P. Solutions at P and P* are compared in figure 12.

clear increase of spatial structure in the main cell. This may be attributed to the branches being well separated at all a^2, presumably because the more elaborate structures imply greater dissipation. As Rs increases, though, this viscous separation of branches must decrease. Simultaneously, increased localization (more vertical structure) becomes possible for solutions with larger a^2 (more horizontal structure), but these are conjectured to be high-index solutions because of their fine-scale vertical structure. This tendency, combined with decreased viscous separation, implies that mode switching should occur. However, the numerical critical and near-critical modes are so complicated that no clear correlation to eigenindex is discernible. Identification of the mode switches with this contest between localization and dissipation also suggests the qualitative shape of the separatrices between indices.

The separatrix shape can in fact be made quantitative by hypothesizing that the circulation in the main cell is insensitive to the detailed boundary conditions, so long as they are applied outside this cell. To see this, consider a sublayer of depth h^*, centred at midlayer,

$$\tfrac{1}{2}-\tfrac{1}{2}h^* \leqslant z \leqslant \tfrac{1}{2}+\tfrac{1}{2}h^*,$$

The same solution should be obtained by solving the problem for this sublayer as for the whole layer, if the 'conductive' salinity and temperature profiles are appropriately rescaled. The rescaled salinity difference would be given by

$$\Delta S^* = \Delta S[S_c(\tfrac{1}{2}+\tfrac{1}{2}h^*)-S_c(\tfrac{1}{2}-\tfrac{1}{2}h^*)] = \Delta S\,[h^*+4(1-g_{\min})\,h^*(h^{*2}-1)],$$

so that the solute Rayleigh number is

$$Rs^* = Rs\,h^{*4}[1+(1-g_{\min})\,(h^{*2}-1)], \tag{17}$$

and its corresponding minimum dimensionless gradient becomes

$$g^*_{\min} = g^*(z^* = \tfrac{1}{2}) = \frac{\Delta S\, g(\frac{1}{2})}{\Delta S^*}\frac{\mathrm{d}z}{\mathrm{d}z^*} = \frac{g_{\min}}{1+(1-g_{\min})\,(z^{*2}-1)}, \tag{18}$$

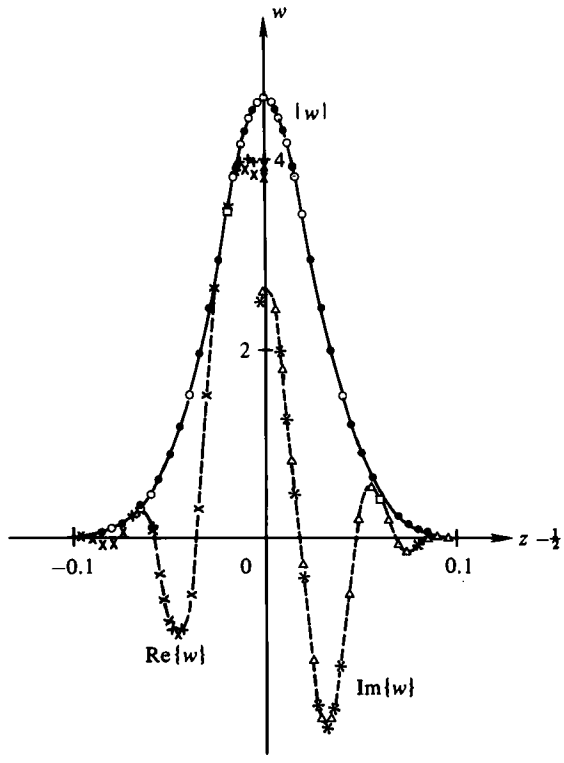

FIGURE 12. Eigenfunction (vertical velocity) in sublayer for points P and P* of figure 11, related by sublayer-scaling rule (20). Only minor differences are visible at $z = \frac{1}{2}$.

where
$$z^* = \frac{1}{2} + \frac{z - \frac{1}{2}}{h^*}.$$

Solving the neutral-stability problem with $(Rs^*, g^*_{\min})$ should then give the eigenvalue

$$\lambda^{0*} = \frac{\lambda^0 g^*_{\min}}{g_{\min}}. \tag{19}$$

As a check, such a calculation has been carried out for the points P ($Rs = 10^{12}$, $g_{\min} = 0.25$) and P* ($Rs^* = 4.48 \times 10^8$, $g^*_{\min} = 0.8929$, $h^* = 0.20$) in figure 11. The agreement between eigenfunctions is shown in figure 12, and the eigenvalues satisfy (19) to four significant figures. This is taken as numerical confirmation of the hypothesis that sublayer solutions are insensitive to boundary conditions applied outside the sublayer.

Now, (17) and (18) are parametric forms of a locus $Rs^*(g^*_{\min})$ in the plane of figure 11. Elimination of h^* results in the explicit 'sublayer-scaling rule'

$$\frac{Rs^*}{Rs} = \left(\frac{g_{\min}}{g^*_{\min}}\right)^3 \left(\frac{1 - g^*_{\min}}{1 - g_{\min}}\right)^2. \tag{20}$$

Along locus (20) the eigensolutions are essentially invariant in the sublayer zone, and too small to be significant outside it. If eigenindex corresponds to structure within the sublayer, then mode-switch separatrices should also be of form (20). Therefore the separatrices of figure 11 are given the shape of this locus, and placed on the $(Rs, g_{\min})$-plane to be consistent with the indices of the numerical solutions. All have

the same shape because this is a universal curve in $g_{\min}$ which merely shifts its position for different choices of the initial point $(Rs, g_{\min})$.

Clearly the locus (20) cannot be extended indefinitely in the direction of decreasing Rs in figure 11. In particular, when h^* corresponds to the central cell size $(d/h)_w$ of the original problem, the boundary conditions will be influencing the sublayer solution, so that (17)–(19) no longer hold. If the jump to a low-a^2 critical solution hasn't already occurred for some other reason, it must occur by this h^* value. As a working hypothesis, we suppose that each curve (20) is terminated at the left when $h^* = k(d/h)_w$, where k is a constant only slightly greater than unity. The solid line in the figure has $k = 1.2$, based on the values for P and P*. This turns out to be consistent with the intersections of the high-a^2 and low-a^2 eigenvalue branches in figure 8(*b*), increasing confidence in it.

The sublayer-scaling rule (20) also applies to the exchange of stabilities solutions in the 'fingering' regime; and it specifies precisely how a numerically difficult problem with large Rs may be reduced to a more accessible problem with smaller Rs^*. A reasonably chosen sublayer with $h^* \gtrsim 1.2(d/h)_w$ centred on $g_{\min}$ can then be used to capture the critical mode of the original layer.

With the numerical results ordered as in figure 11, it is now convenient to discuss their relation to Walton's (1982) perturbation solutions for the high-a^2 modes. We start by collecting the relevant results from Walton in the following forms:

$$\lambda_c^0 = \lambda_L^0 + 0.6514 g_{\min}^{\frac{4}{7}} (1-g_{\min})^{\frac{2}{7}} Rs^{-\frac{1}{7}} + \dots, \tag{21}$$

with $\lambda_L^0 = [(Pr+\tau)/(Pr+1)]g_{\min}$ as before, and

$$a_c = 0.7243(1-g_{\min})^{\frac{1}{14}} g_{\min}^{\frac{1}{7}} Rs^{\frac{3}{14}} + \dots, \tag{22}$$

$$p_c = 0.92955(g_{\min} Rs)^{\frac{1}{2}} + \dots, \tag{23}$$

$$(d/h)_w = 1.1417 g_{\min}^{\frac{2}{7}} (1-g_{\min})^{-\frac{5}{14}} Rs^{-\frac{1}{14}} + \dots. \tag{24}$$

In these expressions, (Pr, τ) have been given the values appropriate to brine, $g_{\min}$ is written in place of Walton's G_0, and the '0' subscript of Walton has been replaced by 'L' (i.e. 'fully localized') on the leading terms of the expansions (21)–(24). The value $(d/h)_w$ is obtained by determining the lengthscale required to make $|w(d^*)| = 0.1|w(0)|$ in Walton's equation (40) for the Gaussian envelope to the eigenfunctions. The quantitative success of these expansions is apparent in figure 8.

Qualitatively, it is clear that, even though the fully localized eigenvalue is the same as for a constant-gradient layer with Rs replaced by $g_{\min} Rs$, the a_c and p_c values are not derivable in this manner. It follows that the constant-gradient solution cannot be recovered by taking the limit $g_{\min} \to 1$. In short, the perturbation solution, like the numerical solution, indicates that the high-a^2 branch is isolated from the $g_{\min} = 1.0$ case. The reason for this is readily traced back in Walton's derivation to the scaling of the vertical coordinate $\bar{z}$ as compared with the horizontal lengthscaling $(\bar{a})^{-1}$; that is, the choice of $\delta = \epsilon^3$, so that vertical diffusion is far weaker than horizontal diffusion. In the constant-gradient solution the two are comparable.

A secondary issue raised by the isolation of the high-a^2 solutions from the constant-gradient solutions is their actual region of validity. The perturbation scheme took no explicit notice of eigenindex, so no mode switching is taken into account. However, the quantitative agreement between (21)–(24) and the numerical results in figure 8 extends to the $n \neq 1$ cases; in fact is even better there. Thus validity is apparently guaranteed everywhere Rs exceeds 10^{12} in figure 11, as indicated by the 'perturbation threshold' marking.

Overall, the numerical and perturbation solutions are complementary in the sense that the numerical parameters cover the regions inaccessible to the perturbation analysis, and *vice versa*. They are mutually supporting in that they agree very well in the zone of overlap around $Rs \sim 10^{12}$. Both show that localization occurs for the high-a^2 neutral overstable modes, and that the secondary flows outside the turning points of (11) are exponentially weak.

Having gained confidence from this dual approach that the linear-stability problem is correctly solved, and that we now know its solutions' properties, we return to the questions of its relevance to the observations. Wright & Loehrke (1976) attempted measurements of the constant-gradient free–free case in a small tank, but the actual top and bottom boundary conditions were uncertain. Results generally indicated an oscillatory onset of motion, but data records were too short to give a reliable period measurement, or to determine what finite-amplitude state was being reached as the system evolved. Horizontal scale a_c was not examined. Shirtcliffe (1967, 1969) used a variable stratification, with parabolic $g(z)$ having its maximum at the midplane. When motion began at the boundaries, he attempted to estimate the depth of the layer in motion, and to infer equivalent constant-gradient Ra, Rs values from it. This procedure was discussed by Wright & Loehrke; we note also that our results above indicate that the Rs values of 10^6–10^7 attained would be neither fully localized nor the low-a^2 type (which is insensitive to stratification details) even if the boundary conditions had been controlled. Nevertheless, this system also showed oscillations at the onset of motion. As already mentioned, the solar pond observations (Zangrando 1979) indicated that the motions in that case (very large Rs, variable stratification, and large but finite horizontal extent of about 15 m) evolved to steady amplitude. However, starting transients were not observed, and spatial structure of the circulating layers was not determined. The eigenvalues, though, agree with the fully localized value (21) within the experimental precision of $\pm 10\,\%$.

Experimental evidence, then, supports an overstable onset of motion in the neighbourhood of the localized overstable bifurcation, and, under pond field conditions, an evolution toward a finite-amplitude steady circulation. Data does not provide any reliable spatial or temporal detail on the critical modes.

The lack of experimental detail leaves some uncertainty about the effects of finite amplitude. The numerical simulations of Huppert & Moore (1976) can be used to evaluate whether a subcriticality of 10 % or less, as suggested by solar pond data, is to be expected. Although they could not reach the extreme parameters of interest here, $(Pr, \tau, Rs) = (10, 0.1, 1.4 \times 10^4)$ was reached. Not only did the subcriticality observed prove to be less than 10 %, but also the structure of their solutions was qualitatively different from any of the linear eigensolutions for stratified or constant-gradient conditions, in that their values of $|S|/|\psi|$ and $|T|/|\psi|$ were $O(1)$ both in exchange of stabilities and overstable solutions, while linear solutions have quite different ratios.

5. Conclusions

The numerical solution of the linear stability problem for variable stratification in a doubly diffusive system has resulted in the conclusions that the overstable modes provide the critical eigenvalue, and that there are two eigenmode types with quite different forms.

The first type is a localized, high-a^2 mode that is accurately described by Walton's (1982) solutions (21)–(24). The family of localized overstable solutions is invariant

by the sublayer-scaling rule (20), i.e. the family is characterized by one free parameter instead of independent $(Rs, g_{\min})$.

The second occurs at small Rs ($\leqslant 10^4$) or near $g_{\min} = 1$ (constant gradients) for all Rs; it is a low-a^2 mode that is accurately described as having a nonlocalized eigenfunction $w(z) = \sin \pi z + O(10^{-6})$ and an eigenvalue

$$\lambda^0 = \frac{Pr+\tau}{Pr+1} g_{\min} + \frac{(1+\tau)(Pr+\tau)}{Pr} \frac{Rb}{Rs} + l(g_{\min}), \tag{25}$$

where $l(g_{\min})$ is independent of Rs. Like the eigenvalue, the critical eigenfrequency is independent of Rs beyond 10^6, while critical wavenumber decreases like $Rs^{-\frac{1}{2}}$.

Exchange-of-stabilities solutions scale as described in (10a), with numerical values given by (14) for both the diffusive and fingering cases, and are also invariant by the sublayer-scaling rule (20). The analogous, purely thermal problem is solved by (15).

Experience in solar ponds indicates that the observed motions occur near the overstable eigenvalue, with a vertical scale compatible with the overstable solutions, but the motion appears to be steady. Evaluation of stability under field conditions is best carried out by use of $g_{\min}$ in (20)–(25), with (20) being used to map the physically interesting problem onto easy-to-solve numerical or perturbation parameters in figure 11.

The relevance of Joseph's energy limit (3) is apparently limited to quite large perturbations, not the modest amplitudes of Huppert & Moore (1976), nor the sometimes substantial, uncontrolled perturbations acting on the solar pond (Zangrando 1979).

Finally, our solutions offer a resolution to the observation/theory impasse described by Huppert & Turner (1981) in their review. Since any specific boundary conditions prove to be difficult for either theory or experiment to handle, the indifference of the solutions given here to the boundary conditions should prove to be a great advantage experimentally. To match precisely the case analysed, an experiment need only provide the same temperature and salinity stratification in the sublayer. Such a salinity profile can be stably set up using isothermal fluid, and will be distorted by diffusion only very slowly. Then temperature gradients can be established by changing the boundary temperatures, thereby creating the most nearly linear gradient at the interior sublayer. Such control may allow one to measure the hysteresis of figure 2 with the same level of precision already achieved in the doubly diffusive Soret driven problem (e.g. Caldwell 1974, 1976).

The authors wish to acknowledge suggestions made by C. Bretherton for the development of §3.1, as well as discussions held with H. C. Bryant, H. E. Huppert, E. Knobloch, R. E. Kelly, and L. Romero.

This work was supported by the US Department of Energy under contracts EG-77-C-01-4042 (SERI) and DE-AC04-76DP00789 (Sandia Labs).

REFERENCES

Bretherton, C. S. 1981 Double diffusion in a long box. *Woods Hole Geophysical Fluid Dynamics Course Lectures*, WHOI-81-102.

Busse, F. H. 1967 Non-stationary finite amplitude convection. *J. Fluid Mech.* **28**, 223–239.

Caldwell, D. R. 1974 Experimental studies on the onset of thermohaline convection. *J. Fluid Mech.* **64**, 347–367.

Caldwell, D. R. 1976 Thermosolutal convection in a solution with large negative Soret coefficient. *J. Fluid Mech.* **74**, 129–142.

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

CURRY, J. H. 1978 A generalized Lorenz system. *Commun. Math. Phys.* **60**, 193–204.

DA COSTA, L. N., KNOBLOCH, E. & WEISS, N. O. 1981 Oscillations in double-diffusive convection. *J. Fluid Mech.* **109**, 25–43.

GOLLUB, J. P. & BENSON, S. V. 1980 Time dependent instabilities and the transition to turbulent convection. *J. Fluid Mech.* **100**, 449–470.

HUPPERT, H. E. & LINDEN, P. F. 1979 On heating a stable salinity gradient from below. *J. Fluid Mech.* **95**, 431–464.

HUPPERT, H. E. & MOORE, D. R. 1976 Nonlinear double diffusive convection. *J. Fluid Mech.* **78**, 821–854.

HUPPERT, H. E. & TURNER, J. S. 1981 Double-diffusive convection. *J. Fluid Mech.* **106**, 229–329.

JOSEPH, D. D. 1976 *Stability of Fluid Motions, II*, p. 43. Springer.

KNOBLOCH, E. & PROCTOR, M. R. E. 1981 Nonlinear periodic convection in double-diffusive systems. *J. Fluid Mech.* **108**, 291–316.

KRISHNAMURTI, R. 1968*a* Finite amplitude convection with changing mean temperature. Part 1. Theory. *J. Fluid Mech.* **33**, 445–455.

KRISHNAMURTI, R. 1968*b* Finite amplitude convection with changing mean temperature. Part 2. An experimental test of the theory. *J. Fluid Mech.* **33**, 457–463.

LINDEN, P. F. & WEBER, J. E. 1977 The formation of layers in a double-diffusive system with a sloping boundary. *J. Fluid Mech.* **81**, 757–773.

LORENZ, E. N. 1963 Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.

MARCUS, P. S. 1981 Effects of truncation in modal representations of thermal convection. *J. Fluid Mech.* **103**, 241–255.

PROCTOR, M. R. E. 1981 Steady subcritical thermohaline convection. *J. Fluid Mech.* **105**, 507–521.

RUBENFELD, L. A. & SIEGMANN, W. L. 1977 Nonlinear dynamic theory for a double-diffusive convection model. *SIAM J. Appl. Maths* **32**, 871–894.

SALTZMAN, B. 1962 Finite amplitude free convection as an initial value problem. *J. Atmos. Sci.* **19**, 329–341.

SANI, R. L. 1965 On finite amplitude roll cell disturbances in a fluid layer subjected to heat and mass transfer. *AIChE J.* **11**, 971–980.

SHIRTCLIFFE, T. G. L. 1967 Thermosolutal convection: observation of an overstable mode. *Nature* **213**, 489–490.

SHIRTCLIFFE, T. G. L. 1969 An experimental investigation of thermosolutal convection at marginal stability. *J. Fluid Mech.* **35**, 788–688.

SIEGMANN, W. L. & RUBENFELD, L. A. 1975 A nonlinear model for a double-diffusive convection. *SIAM J. Appl. Maths* **29**, 540–557.

TOOMRE, J., GOUGH, D. O. & SPIEGEL, E. A. 1977 Numerical solutions of single-mode convection equations. *J. Fluid Mech.* **79**, 1–31.

TURNER, J. S. 1973 *Buoyancy Effects in Fluids*. Cambridge University Press.

VERONIS, G. 1965 On finite amplitude instability in thermohaline convection. *J. Mar. Res.* **23**, 1–17.

VERONIS, G. 1968 Effect of a stabilizing gradient of solute on thermal convection. *J. Fluid Mech.* **34**, 315–336.

WALIN, G. 1964 Note on the stability of water stratified by both salt and heat. *Tellus* **16**, 389–393.

WALTON, I. C. 1982 Double-diffusive convection with large variable gradients. *J. Fluid Mech.* **125**, 123–135.

WRIGHT, J. H. & LOEHRKE, R. I. 1976 The onset of thermohaline convection in a linearly-stratified horizontal layer. *Trans. ASME* C: *J. Heat Transfer* **98**, 558–563.

ZANGRANDO, F. 1979 Observation and analysis of a full-scale experimental salt gradient solar pond. Ph.D. dissertation, University of New Mexico.

ZANGRANDO, F. & BERTRAM, L. A. 1984 Doubly-diffusive linear stability with nonconstant gradients. *SERI* TR-252-2052.