SIAM J APPL MATH.
Vol 29, No. 3, November 1975

A NONLINEAR MODEL FOR DOUBLE-DIFFUSIVE CONVECTION*

WILLIAM L. SIEGMANN AND LESTER A. RUBENFELD†

Abstract. We consider a simple fluid-loop model describing convection in a two-constituent fluid. The model permits explicit construction of linear stability and global stability boundaries in parameter space. A rigorous proof of global stability within the appropriate region is provided.

1. Introduction. Double-diffusive convection denotes convective motions in a fluid which possesses two constituents with different molecular diffusivities. A primary example is certain motions in the ocean and some lakes where the two properties are salt and heat, but other illustrations are known in metallurgy, astrophysics, and geology. Of particular interest is the situation when, in the absence of motion, constituent gradients are present which produce density gradients of opposite signs. An excellent account of known properties of double-diffusive convection in a Newtonian fluid is contained in a recent monograph by Turner [1].

In this paper we shall examine double-diffusive convection in a particularly simple fluid system. In our mathematical model, fluid is confined within a circular loop of very small cross-section, and the loop is embedded in a motionless medium. The model is enhanced with an enormously greater mathematical tractability, as compared with a fluid satisfying the Navier–Stokes equations, by two further assumptions. First, the fluid is required to move only one-dimensionally, in the direction around the loop. Second, it is permitted to exchange "heat" and "salt" with the surroundings at a rate proportional to the difference between the (instantaneous) salt and heat content of the medium and the loop. All these assumptions and others are described in detail in the following section.

Simple dynamical models have been used for some time to elucidate the properties of fluid convection. Keller [2] and Welander [3] used two-parameter, one-constituent fluid models to demonstrate the existence of periodic oscillations that are possibly analogous to experimentally observed oscillatory motions at moderate and high Rayleigh numbers (see next section for parameter definitions). Stommel and Rooth [4] developed a related but simpler model to demonstrate nonunique equilibrium states. Moore and Spiegel [5] and Baker, Moore and Spiegel [6] investigated a two-parameter, third order, nonlinear ordinary differential equation to model convection. By both analytical and numerical techniques, they were able to exhibit periodic solutions as well as a parameter region of highly aperiodic motions. Howard and Malkus [7] proposed a two-parameter circular loop model in order to obtain a closer resemblance to one-constituent convection governed by the Navier–Stokes equations. They were able to investigate the

* Received by the editors June 25, 1974, and in revised form October 11, 1974.

† Department of Mathematical Sciences, Rensselaer Polytechnic Institute, Troy, New York 12181.
This work was supported in part by the Army Research Office, Durham, under Grant DA 31-124 ARO-D-269.

properties of the resulting system of first order nonlinear equations in considerable detail over a wide range of parameter values; additional results will appear in a thesis [8]. Our model may be regarded as a two-constituent development of the Howard and Malkus equations, which evidently were derived and studied in another context earlier by Lorenz [9]. See also Malkus [17].

There are several aspects of double-diffusive convection which we are interested in clarifying by means of our model. For the Navier–Stokes equations, Joseph [10] has found a region in parameter space where the motionless or "conductive" solution is stable to perturbations of arbitrary initial amplitude. However, it is unclear whether his results are optimal. By this is meant that larger regions of stability may exist which his methods do not uncover, or that "stronger" decay of the perturbations may occur than he was able to demonstrate. Joseph noted these possibilities and, in particular, expressed concern because the boundary of his region was independent of one parameter, namely the Prandtl number. Sani [11] has used the Stuart–Watson method to describe the evolution of perturbations with infinitesimal initial amplitude for parameter conditions close to those for instability onset. His results predict steady and oscillatory instability growth and also sublinear instability in certain parameter regions. Nonetheless, a singularity in his amplitude expressions in a particularly interesting neighborhood of parameter space, where both oscillatory and steady instability can occur, indicates that the proper disturbance evolution was not provided. Another region, where higher order terms in the Stuart–Watson expansion evidently are required, was not fully explored. Finally, the development of convective motions further away from onset conditions has been very incompletely examined, because of numerical difficulties in treating the multiparameter problem governed by the Navier–Stokes equations. For example, recent calculations by Strauss [12], while directed toward interesting parameter regime, are based on fairly stringent assumptions for parameter values and for allowable spatial patterns of the motions. Properties of the convection very far from onset conditions are known only by extrapolations from more moderate parameter values.

In this paper we shall be concerned only with the stability of the conductive solution. In § 2 the model is formulated and the basic-parameter governing system of five ordinary differential equations is derived. The linear stability of the conductive solution is discussed in § 3, and the parameter region where steady convective solutions exist is derived in § 4. By a systematic procedure, we show in § 5 that the complement of this region is precisely the same as the region of global stability of the conductive solution. Other results on the nature of the stability are also given.

Another principal conclusion of this paper is that the parametric dependence of model solutions is strikingly similar to that known for solutions to the Navier–Stokes equations. The flow in the loop can be thought of as analogous to the flow of a single convective cell of a Newtonian fluid. There are of course inherent flaws in the analogy, among them that the convective length scale of the loop motion is prescribed rather than selected. Although the parametric evidence for such an identification is compelling, the ultimate test is the appropriateness of the predictions of the model. Additional properties of the model will be explored in future papers.

2. Formulation. We imagine a thin toroidal loop of central radius R_0 (see Fig. 1) filled with an incompressible Boussinesq fluid, i.e., one in which effects of density variation are retained only when coupled with gravity [13]. The state equation for the density of the fluid has the form

$$\rho = \rho_0(1 - \alpha(T - T_0) + \beta(S - S_0)), \qquad \alpha, \beta > 0, \tag{2.1}$$

where S and T are diffusing constituents referred to as salt concentration and temperature respectively; ρ_0, T_0 and S_0 are constant reference values; and α and β are positive constants.

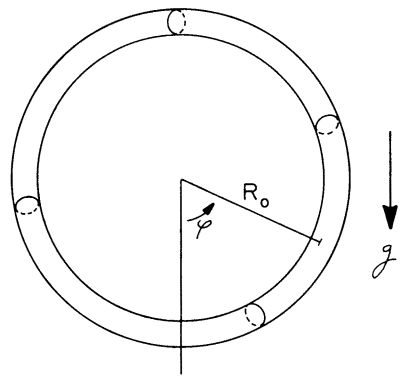

FIG. 1

Following Howard and Malkus [7] we assume that the fluid loop is embedded in a medium with which it can exchange salt and heat at a rate proportional to the difference between the salt concentration and temperature of the medium and loop. The essential feature of double-diffusive convection is that the proportionality constants are different.

In addition we postulate that the fluid motion is retarded by a (viscous-type) force proportional to its velocity and take the normal of the plane of the loop to be perpendicular to the vertical direction of gravity (see Fig. 1). Our final assumptions are that the loop is so thin that salt concentration, temperature, and fluid velocity are constant over cross-sections; that the fluid velocity has only one component directed in the angular direction; and that the environment of the loop has constant vertical salt and heat gradients $\Delta S/R_0$ and $\Delta T/R_0$ respectively. These assumptions lead to

$$T = T(t, \phi), \quad S = S(t, \phi), \quad \mathbf{u} = R_0\omega(t, \phi)\mathbf{e}_\phi, \tag{2.2}$$

$$T_E = (\Delta T)\cos\phi + T_0, \qquad S_E = (\Delta S)\cos\phi + S_0, \tag{2.3}$$

where $\mathbf{u}$ is the fluid velocity in the loop, $\mathbf{e}_\phi$ is the unit vector in the ϕ-direction, and T_E and S_E are the surrounding temperature and salt concentration, respectively. We note that when ΔT and ΔS are both positive, the temperature and salt concentration are greatest at the bottom of the loop ($\phi = 0$), so that the environmental heating is destabilizing and the environmental salting is stabilizing. Thus, depending on the signs of ΔT and ΔS, there are possibilities of competing stabilization and destabilization mechanisms.

Using all of the previous assumptions we can arrive at a system of equations governing the fluid motion in the loop. It is possible to regard this system, given as follows, as the first approximation from an expansion of the fluid equations in terms of a small parameter measuring loop thickness:

$$R_0(\omega_t + \omega\omega_\phi) = -\frac{p_\phi}{R_0\rho_0} - g\frac{(\rho - \rho_0)}{\rho_0}\sin\phi - \nu R_0\omega, \tag{2.4}$$

$$\omega_\phi = 0, \tag{2.5}$$

$$\frac{DT}{Dt} \equiv T_t + \omega T_\phi = -K_T(T - T_E), \qquad K_T, K_S > 0, \tag{2.6}$$

$$S_t + \omega S_\phi = -K_S(S - S_E). \tag{2.7}$$

To these we must append (2.1) and (2.3). Equation (2.4) follows from the momentum balance using the assumption of a thin loop, with g denoting gravity and ν the dissipation coefficient. The pressure p is the deviation from the hydrostatic pressure p_0 which is in balance with the gravitational force acting on the average density ρ_0. Equation (2.5) follows from incompressibility, and finally (2.6) and (2.7) are the assumed forms of heat and salt balance in the loop with K_T and K_S positive exchange (or "diffusion") coefficients for heat and salt.

Using (2.5) in (2.4) and integrating over ϕ, we can eliminate the pressure term to get

$$\omega_t = -\frac{g}{2\pi R_0\rho_0}\int_0^{2\pi}(\rho - \rho_0)\sin\phi\, d\phi - \nu\omega. \tag{2.8}$$

Let us now define dimensionless variables and expand the quantities of interest in Fourier series as follows:

$$t = \tau/K_T, \tag{2.9}$$

$$\omega(t) = K_T c(\tau), \tag{2.10}$$

$$T - T_0 = (\Delta T)[a(\tau)\cos\phi + b(\tau)\sin\phi + \sum_{\substack{m=0\\(m\neq 1)}}^{\infty}(a_m(\tau)\cos m\phi + b_m(\tau)\sin m\phi)], \tag{2.11}$$

$$S - S_0 = (\Delta S)[d(\tau)\cos\phi + e(\tau)\sin\phi + \sum_{\substack{m=0\\(m\neq 1)}}^{\infty}(d_m(\tau)\cos m\phi + e_m(\tau)\sin m\phi)]. \tag{2.12}$$

Using (2.1), (2.3), and (2.9)–(2.12) in (2.6)–(2.8) we find:

$$a' + bc = -a + 1, \tag{2.13}$$

$$b' - ac = -b, \tag{2.14}$$

$$\sigma^{-1}c' = Rb - R_S e - c, \tag{2.15}$$

$$d' + ec = -\mu d + \mu, \tag{2.16}$$

$$e' - dc = -\mu e, \tag{2.17}$$

where we have defined

$$\begin{aligned} \sigma &= \nu/K_T, \\ R &= g\alpha(\Delta T)/2R_0\nu K_T, \\ R_S &= g\beta(\Delta S)/2R_0\nu K_T, \\ \mu &= K_S/K_T, \qquad 0 < \mu < 1. \end{aligned} \tag{2.18}$$

The parameter σ is a "Prandtl number" while R and R_S are "Rayleigh numbers". The ratio of diffusivities, μ, is a "Schmidt number" and is typically within the bounds indicated. The limiting case $\mu = 1$, when the diffusive and convective properties of $T$and S are identical, is readily reduced to the one-constituent model of Howard and Malkus [7]. When $\mu > 1$ the system may be converted to the $\mu < 1$ case by another nondimensionalization of c and t, new parameter definitions and identification of the heat fields with the salt fields. Let us also note that both R and R_S are defined with respect to a common scaling by K_T. This is consistent with Turner's [1] definitions but in conflict with Joseph's [10] definitions.

Equations (2.13)–(2.17), along with initial conditions, comprise a closed system for a, b, c, d, and e. The equations governing a_m, b_m, d_m, and e_m (see (2.11) and (2.12)) are just simple modifications of (2.13), (2.14), (2.16) and (2.17) without the constant (source) terms. It is trivial to show that all of these quantities decay in τ, and so we will not consider them further.

3. Linear stability of conductive solution. Equations (2.13)–(2.17) readily admit a steady motionless solution

$$a = 1, \quad b = 0, \quad c = 0, \quad d = 1, \quad e = 0. \tag{3.1}$$

We call (3.1) a "conductive" solution since the T field (a and b) and S field (d and e) in the loop merely acquire the environmental distributions T_E and S_E (see (2.3)), while the fluid in the loop is motionless ($c = 0$). Solutions of (2.13)–(2.17) which are "convective", i.e., which have a nonzero fluid velocity c, might be anticipated for ranges of the parameters R, R_S, μ, and σ for which the solution (3.1) becomes unstable. This section will be concerned with the question of the stability of the conductive solution to infinitesimal disturbances.

Equations (2.13)–(2.17) can be written in terms of the deviation from the conductive solution (3.1) by defining a vector

$$\mathbf{x}^\top \equiv (x_1, x_2, x_3, x_4, x_5) = (a - 1, d - 1, b, c, e), \tag{3.2}$$

where $^\top$ denotes transpose. Then (2.13)–(2.17) become

$$\begin{aligned} &\text{(i)} \quad x_1' = -x_1 - x_3x_4, \\ &\text{(ii)} \quad x_2' = -\mu x_2 - x_5x_4, \\ &\text{(iii)} \quad x_3' = -x_3 + x_4 + x_1x_4, \\ &\text{(iv)} \quad x_4' = \sigma R x_3 - \sigma x_4 - \sigma R_S x_5, \\ &\text{(v)} \quad x_5' = -\mu x_5 + x_4 + x_2x_4. \end{aligned} \tag{3.3}$$

In vector form these equations are

$$\mathbf{x}' = A\mathbf{x} + x_4 N\mathbf{x},$$

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -\mu & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & \sigma R & -\sigma & -\sigma R_S \\ 0 & 0 & 0 & 1 & -\mu \end{bmatrix}, \qquad N = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}. \tag{3.4}$$

The conductive solution ($\mathbf{x} = \mathbf{0}$) is linearly stable if and only if the eigenvalues of the matrix A all have negative real parts. These eigenvalues λ satisfy the equation

$$(\lambda + 1)(\lambda + \mu)(\lambda^3 + \Delta\lambda^2 + \gamma\lambda + \Gamma) = 0, \tag{3.5}$$

where

$$\begin{aligned} \Delta &= \sigma + \mu + 1 > 0, \\ \gamma &= \sigma + \mu + \sigma\mu + \sigma(R_S - R), \\ \Gamma &= \sigma\mu + \sigma(R_S - \mu R). \end{aligned} \tag{3.6}$$

Since the roots of the two linear factors are $\lambda = -1, -\mu$, we need only check the conditions for which the cubic factor has roots with negative real parts. By the Hurwitz criteria [14] these roots have negative real parts for $\Gamma > 0$ and $\gamma\Delta > \Gamma$ (note that the third Hurwitz condition is implied by these). These two conditions yield a stability boundary in the (R, R_S)-plane as depicted in Fig. 2.

These curves are plotted for fixed values of the parameters μ and σ. The lines $\Gamma = 0$ and $\gamma\Delta = \Gamma$ are shown respectively as

$$\begin{aligned} L^{(1)}&: \quad R_S = \mu R - \mu, \\ L^{(2)}&: \quad R_S = \left(\frac{\sigma + 1}{\sigma + \mu}\right) R - \frac{(\sigma + 1)(\mu + 1)}{\sigma}, \end{aligned} \tag{3.7}$$

and the conductive solution is stable for points (R, R_S) in the region to the left of, and bounded by, the solid parts of both $L^{(1)}$ and $L^{(2)}$. We note that $L^{(1)}$ and $L^{(2)}$ intersect at the point

$$P^{(0)}: \quad (R^{(0)}, R_S^{(0)}) = \left(\frac{\sigma + \mu}{\sigma(1 - \mu)}, \frac{\mu^2(1 + \sigma)}{\sigma(1 - \mu)}\right). \tag{3.8}$$

On the lines $L^{(1)}$ and $L^{(2)}$ and at the point $P^{(0)}$ the solutions of (3.5) are given by

$$\begin{aligned} &\text{on } L^{(1)}: \quad \lambda = -1, -\mu, 0, \tfrac{1}{2}(-\Delta \pm \sqrt{\Delta^2 - 4\gamma}), \\ &\text{on } L^{(2)}: \quad \lambda = -1, -\mu, -\Delta, \pm i\sqrt{\Gamma/\Delta}, \\ &\text{at } P^{(0)}: \quad \lambda = -1, -\mu, 0, 0, -\Delta. \end{aligned} \tag{3.9}$$

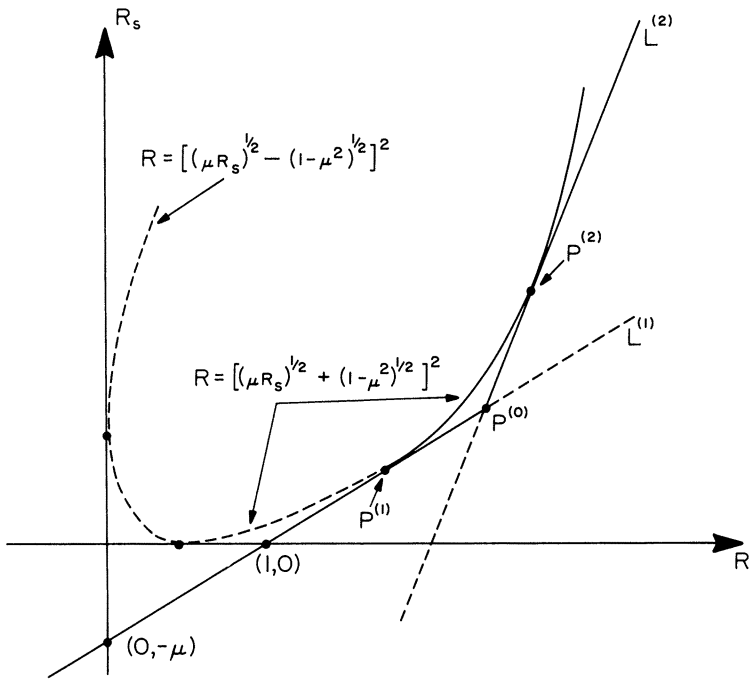

FIG. 2

Thus on $L^{(1)}$ there is one zero eigenvalue and four roots with negative real parts, and one of these roots coalesces with the zero root at $P^{(0)}$. On $L^{(2)}$ there are three negative and two imaginary roots, and the imaginary roots coalesce to two zero roots at $P^{(0)}$. Thus as one crosses $L^{(1)}$ from the stability region to the instability region, a solution of the linear system which was exponentially decaying becomes time independent on $L^{(1)}$ and then begins to grow exponentially just below $L^{(1)}$. If one passes across the line $L^{(2)}$, two solutions of the linear problem which are oscillatory and exponentially decreasing become purely oscillatory on $L^{(2)}$ and proceed to oscillate with exponentially growing amplitude just below $L^{(2)}$. In crossing from the region of stability to that of instability through the point $P^{(0)}$, two decaying solutions of the linear problem combine at $P^{(0)}$ so as to grow linearly in time, and then they begin to grow exponentially after passing through $P^{(0)}$. The dynamical aspects of the nonlinear problem in the neighborhood of this point are quite interesting and are discussed in [15].

The linear stability boundaries have interesting physical interpretations. From (2.18) we note that the signs of R and R_S are the same as those of ΔT and ΔS, defined in (2.3). Also, ΔT and ΔS both positive means that the environmental

temperature and salt concentration are increasing downward, producing a destabilizing temperature gradient and a stabilizing salt gradient. For ΔT and ΔS both negative the opposite effects are true. Thus in the second quadrant of Fig. 2 ($R < 0$ and $R_S > 0$) both constituents tend to produce stability, which is verified by the fact that this whole quadrant lies in the linear stability region. In the first quadrant of Fig. 2 ($R > 0$ and $R_S > 0$), we have the competition of unstable heating and stable salting. It is reasonable to expect that for large enough R_S ("very" stable salting), the fluid would experience a net stabilizing effect. This again is verified by the linear stability boundary in this quadrant. A similar argument can be made for the behavior in the third quadrant. A discussion of the physical mechanisms working to produce instability in a Newtonian fluid for the case of competing environmental gradients has been given by Turner [1]. His descriptions apply as well to our model.

In quadrant four of Fig. 2, where $R > 0$ and $R_S < 0$ and both constituents tend to produce instability, there is an additional feature. A region of linear stability exists, namely the small triangle in this quadrant. This region corresponds to "relatively small" values of R and R_S which from (2.18) means a "relatively large" value of the dissipation ν. Thus even though instability is produced by both heating and salting, these effects are dominated by the inherent damping in the system. This triangular region appears again in § 5, and is relevant when discussing regions of strong nonlinear stability of the basic conductive solution.

We may also ask how these linear dynamic stability results compare with a static stability analysis. For equilibrium between the fluid and its environment, $T = T_E$ and $S = S_E$. From (2.1), using (2.3) and (2.18) we find the equilibrium density given by

$$\rho_{\text{static}} = \rho_0\left[1 + \frac{2R_0\nu K_T}{g}(R_S - R)\cos\phi\right].$$

Static stability is interpreted as ρ_{static} decreasing upwards (light fluid over heavy) with static instability meaning ρ_{static} increasing upwards (heavy fluid over light). From Fig. 1 and the above result we have static stability for $R_S > R$ and static instability for $R_S < R$. If the curve $R_S = R$ were drawn in Fig. 2, it would intersect the linear stability lines $L^{(1)}$ and $L^{(2)}$ in the third and first quadrants respectively of the (R, R_S) plane since $\mu < 1$. Therefore, in these quadrants there are regions of static stability but dynamic instability, as well as static instability with dynamic stability. The steady convective motions which may arise above the line $R = R_S$ in the third quadrant are directly analogous to the "salt fingers" of a Newtonian fluid [1].

4. Steady solutions. In the sector of linear instability of the conductive solution of (3.3) (below the solid parts of both $L^{(1)}$ and $L^{(2)}$ in Fig. 2), perturbations must tend to some convective state of the system. The simplest possibilities are time independent solutions of (3.3) (or equivalently (2.13)–(2.17)), although whether these states are ever attained depends on their stability. These steady state solutions are easily shown to be of the form

$$\mathbf{x}_{\text{steady}} = \left(\frac{-p}{1+p}, \frac{-p}{\mu^2+p}, \frac{\pm p^{1/2}}{1+p}, \pm p^{1/2}, \frac{\pm \mu p^{1/2}}{\mu^2+p}\right), \tag{4.1}$$

where p is a nonnegative root of

$$p^{1/2}[p^2 + (\mu^2 + 1 + \mu R_S - R)p + (\mu^2 - \mu^2 R + \mu R_S)] = 0. \tag{4.2}$$

Aside from the root $p = 0$, which leads to the conductive state $\mathbf{x} = \mathbf{0}$, there are other possibilities if the quadratic factor in (4.2) possesses real positive roots. The roots of this term are given by

$$2p = -B \pm (B^2 - 4C)^{1/2}, \tag{4.3}$$

where

$$\begin{aligned} B &= \mu^2 + 1 + \mu R_S - R, \\ C &= \mu^2 - \mu^2 R + \mu R_S. \end{aligned} \tag{4.4}$$

The condition that there be real positive roots of the quadratic factor in (4.2) is just

$$-B \pm (B^2 - 4C)^{1/2} > 0.$$

An analysis of this condition reveals that there are:

no positive roots if

$$\begin{aligned} & B^2 - 4C < 0, \\ \text{or} \quad & C > 0, \quad B > 0 \quad \text{and} \quad B^2 - 4C > 0, \\ \text{or} \quad & C = 0 \quad \text{and} \quad B \geqq 0, \\ \text{or} \quad & B^2 - 4C = 0 \quad \text{and} \quad B \geqq 0, \end{aligned} \tag{4.5}$$

one positive root if

$$\begin{aligned} & C < 0, \\ \text{or} \quad & C = 0 \quad \text{and} \quad B < 0, \\ \text{or} \quad & B^2 - 4C = 0 \quad \text{and} \quad B < 0, \end{aligned} \tag{4.6}$$

two positive roots if

$$C > 0, \quad B < 0 \quad \text{and} \quad B^2 - 4C > 0. \tag{4.7}$$

The curve $B^2 - 4C = 0$ is a skewed parabola in the (R, R_S)-plane which is tangent to the R- and R_S-axes and also tangent to the stability boundaries $L^{(1)}$ and $L^{(2)}$ at the points (see Fig. 2) $P^{(1)}$ and $P^{(2)}$ given by

$$P^{(1)}: \quad (R^{(1)}, R_S^{(1)}) = \left(\frac{1}{1 - \mu^2}, \frac{\mu^3}{1 - \mu^2} \right), \tag{4.8}$$

$$P^{(2)}: \quad (R^{(2)}, R_S^{(2)}) = \left(\frac{(1 + \mu)(\sigma + \mu)^2}{\sigma^2(1 - \mu)}, \frac{\mu(\sigma + 1)^2(1 + \mu)}{\sigma^2(1 - \mu)} \right). \tag{4.9}$$

In fact we can factor $B^2 - 4C$ into

$$B^2 - 4C = [R - [(\mu R_S)^{1/2} + (1 - \mu^2)^{1/2}]^2][R - [(\mu R_S)^{1/2} - (1 - \mu^2)^{1/2}]^2]. \tag{4.10}$$

The two branches of the parabola are shown in Fig. 2 with the right branch, given by

$$R = [(\mu R_S)^{1/2} + (1 - \mu^2)^{1/2}]^2, \tag{4.11}$$

consisting of the solid curve (ending at $P^{(1)}$) and part of the dotted curve down to the R-axis. The left branch

$$R = [(\mu R_S)^{1/2} - (1 - \mu^2)^{1/2}]^2, \tag{4.12}$$

consists of the remaining portion of the dotted curve.

The curve $C = 0$ is just the line $L^{(1)}$ in Fig. 2. The curve $B = 0$ is another line (not shown) which intersects the R_S-axis below the point of intersection of $L^{(1)}$ (since $\mu < 1$) and which intersects $L^{(1)}$ and the parabola at $P^{(1)}$.

It is now easy to show that the series of inequalities (4.5)–(4.7) yield the following situations in the (R, R_S) plane:

No convective steady states:

To the left of the solid part of the parabola and above $L^{(1)}$, and also that part of $L^{(1)}$ lying to the left of the point $P^{(1)}$.

One pair of convective steady states:

Below the line $L^{(1)}$, on that part of $L^{(1)}$ lying to the right of the point $P^{(1)}$, and on the solid part of the parabola.

Two pairs of convective steady states:

In the region between the solid portion of the parabola and the line $L^{(1)}$.

We note that there are additional steady states inside the region of linear stability. This is commonly referred to as the region of subcritical instability. The region of linear instability where there are additional steady states is the region of supercritical instability. We would expect that in the region of linear stability where there are no nontrivial steady states available, all solutions would decay to zero. This behavior is verified in the next section, thus demonstrating the absence of any other stable solutions, such as periodic motions.

For future reference let us note that the solid portion of the parabola can be represented by

$$R = [(\mu R_S)^{1/2} + (1 - \mu^2)^{1/2}]^2, \qquad R_S \geqq \frac{\mu^3}{1 - \mu^2}. \tag{4.13}$$

5. Finite-amplitude stability of conductive solution. We next investigate the stability of the conductive solution (3.1) to disturbances of arbitrary amplitude. The largest *possible* region of stability is above both $L^{(1)}$ and the solid part of the parabola in Fig. 2, where no steady convective solutions (4.1) exist. It is of course conceivable that other time-dependent solutions might occur there. We proceed to eliminate this possibility and thereby obtain comparable stability results as for a Newtonian fluid [10].

From (3.3) we easily derive the following four quadratic "energy" relations:

$$\frac{1}{2\sigma}(x_4^2)' = -x_4^2 + x_4(Rx_3 - R_S x_5), \tag{5.1}$$

$$\frac{1}{2}(x_1^2 + x_3^2)' = -(x_1^2 + x_3^2) + x_3x_4, \tag{5.2}$$

$$\frac{1}{2}(x_2^2 + x_5^2)' = -\mu(x_2^2 + x_5^2) + x_4x_5, \tag{5.3}$$

$$(x_1x_2 + x_3x_5)' = -(1 + \mu)(x_1x_2 + x_3x_5) + x_4(x_3 + x_5). \tag{5.4}$$

Equation (5.1) is derived by multiplying (3.3) (iv) by x_4, and (5.2) by multiplying (3.3) (i) by x_1, (3.3) (iii) by x_3 and adding. Equations (5.3) and (5.4) are derived in a similar manner.

If we now add to (5.1) arbitrary multiples of (5.2)–(5.4) and systematically complete the squares on both sides of the resultant equation we arrive at

$$\frac{1}{2}\frac{d\psi}{dt} = -P, \tag{5.5}$$

where

$$\psi = \alpha\left(x_1 + \frac{\gamma x_2}{\alpha}\right)^2 + \alpha\left(\frac{\beta}{\alpha} - \frac{\gamma^2}{\alpha^2}\right)x_2^2 + \alpha\left(x_3 + \frac{\gamma}{\alpha}x_5\right)^2 + \alpha\left(\frac{\beta}{\alpha} - \frac{\gamma^2}{\alpha^2}\right)x_5^2 + \frac{1}{\sigma}x_4^2, \tag{5.6}$$

$$P = \alpha\left(x_1 + \frac{\gamma(1 + \mu)}{\alpha}x_2\right)^2 + \alpha\left(x_3 - \frac{(R + \alpha + \gamma)}{2\alpha}x_4 + \frac{\gamma(1 + \mu)}{2\alpha}x_5\right)^2 + \left(1 - \frac{(R + \alpha + \gamma)^2}{4\alpha}\right)x_4^2 + \mu\alpha\left(\frac{\beta}{\alpha} - \frac{\gamma^2(1 + \mu)^2}{4\mu\alpha^2}\right)(x_2^2 + x_5^2) + \left(\frac{\gamma(1 + \mu)(R + \alpha + \gamma)}{2\alpha} - (\gamma + \beta - R_S)\right)x_4x_5. \tag{5.7}$$

We do not finish the completion of squares in P for reasons which will be clarified below. The quantities α, β and γ in (5.6) and (5.7) are the multiples of (5.2)–(5.4) referred to above.

We would now like to determine whether there exist α, β and γ such that the quadratic forms ψ and P in (5.6) and (5.7) are respectively positive definite and positive semidefinite. Clearly, the existence of such α, β and γ will depend on the parameters R, R_S, μ and σ. In fact, we would like to find the maximum value of R, for fixed R_S, for which we are assured of finding an α, β and γ which yield the required properties for ψ and P.

From (5.6) and (5.7) it is clearly *sufficient* for ψ to be positive definite and P positive semidefinite, that the following system of inequalities on α, β and γ hold:

$$\gamma^2 = \frac{4\mu\alpha\beta}{(1 + \mu)^2}, \tag{5.8}$$

$$\gamma + \beta - R_S = \frac{\gamma(1 + \mu)(R + \alpha + \gamma)}{2\alpha}, \tag{5.9}$$

$$\alpha > \frac{(R + \alpha + \gamma)^2}{4}. \tag{5.10}$$

It is possible to specify other conditions on α, β and γ to ensure that the desired conditions on P and ψ are satisfied; such conditions could be found by finishing the completion of the squares in (5.7) for example. However, it was found that the expected stability results could be obtained without considering these additional conditions.

We can now prove the following theorem.

THEOREM. *For a given μ, σ and R and R_S in the linear stability region, suppose that there exist values of α, β and γ which make ψ in* (5.6) *positive definite and P in* (5.7) *positive semidefinite. Then the initial value problem for* (3.4) *possesses a unique solution, bounded in $0 \leqq \tau < \infty$ and tending to zero as $\tau \to \infty$ (the zero solution of* (3.4) *is globally asymptotically stable for these values of R, R_S, μ, and σ).*

Proof. Since ψ is assumed positive definite and P positive semidefinite, from (5.5) all solutions $\mathbf{x}$, where they exist must be bounded. Therefore, the right side of (3.4) is continuous, has bounded continuous partial derivatives, and is thus Lipschitzian. Therefore, by Theorem 2.1, the discussion on page 18, and Theorem 3.1 of [16, Chap.1], the initial value problem for (3.4) possesses a unique, bounded solution for $0 \leqq \tau < \infty$.

Now, by Theorem 1.3 of [16, Chap. 10], all solutions of (3.4) tend to the hypersurface defined by $P = 0$. With our choice of conditions (5.8)–(5.10), this hypersurface is given by the relations

$$x_4 = 0,$$

$$x_3 - \frac{(R + \alpha + \gamma)}{2\alpha} x_4 + \frac{\gamma(1 + \mu)}{2\alpha} x_5 = 0,$$

$$x_1 + \frac{\gamma(1 + \mu)}{\alpha} x_2 = 0.$$

It is now an easy matter to show that, on this hypersurface, all solutions of (3.4) decay to the origin as $\tau \to \infty$. Q.E.D.

We will now show that for fixed μ and for R and R_S taking values inside of the region of linear stability where there are no additional steady states we can find a set α, β and γ satisfying inequalities (5.8)–(5.10). Note σ does not appear in (5.8)–(5.10).

We first observe from (5.8)–(5.10) that α and β must be positive, so we can define

$$\alpha = y^2, \qquad \beta = z^2. \tag{5.11}$$

We now solve (5.8) for γ in terms of y and z as

$$\gamma = -\frac{2\mu^{1/2} yz}{(1 + \mu)}. \tag{5.12}$$

We have arbitrarily chosen the negative sign since we only want to find some solution to (5.8)–(5.10).

We now replace the inequality (5.10) by an equality using (5.11), (5.12) and defining a new variable w by

$$w = \frac{R + y^2 - (2\mu^{1/2} yz)/(1 + \mu)}{2y}. \tag{5.13}$$

Then (5.10) requires that

$$w^2 < 1. \tag{5.14}$$

Next, we use the remaining equation (5.9), together with (5.11)–(5.13) to solve for R and y in terms of z and w to get

$$y = \frac{(1+\mu)}{2\mu^{1/2}}\left(z - \frac{R_S}{z} + 2\mu^{1/2}w\right), \tag{5.15}$$

$$R(z;w) = (1-\mu^2)w^2 + \frac{(1+\mu^2)}{2\mu}R_S + \mu^{1/2}(1-\mu)wz$$

$$- \frac{(1-\mu)^2}{4\mu}z^2 + \frac{\mu^{1/2}(1+\mu)R_S w}{z} - \frac{(1+\mu)^2 R_S^2}{4\mu z^2}, \tag{5.16}$$

$$0 < w < 1, \tag{5.17}$$

$$0 < z. \tag{5.18}$$

Note that we arbitrarily restrict w and z to be positive. Again, we are only looking for one solution.

We now ask, for a given R_S, what is the maximum value which R can take so that (5.16) is satisfied for some z, with w fixed. We then vary w between 0 and 1 to get a sequence of maximum values for R and choose the supremum of these. Then, for any value of R less than this we will have found a z satisfying (5.16), a w satisfying (5.17) and can then define y by (5.15). This in turn will give α, β and γ from (5.11) and (5.12).

From (5.16) we find that

$$\frac{\partial R}{\partial z} = -\frac{(1-\mu)^2}{2\mu z^3}\left[z^2 - \frac{(1+\mu)}{(1-\mu)}R_S\right]\left[z^2 - \frac{2w\mu^{3/2}z}{(1-\mu)} + \frac{(1+\mu)}{(1-\mu)}R_S\right], \tag{5.19}$$

$$\frac{\partial^2 R}{\partial z^2} = -\frac{(1-\mu)^2}{2\mu} + \frac{2\mu^{1/2}(1+\mu)R_S w}{z^3} - \frac{3(1+\mu)^2 R_S^2}{2\mu z^4}. \tag{5.20}$$

Note that we were able to factor the quartic $\partial R/\partial z$.

We now set $\partial R/\partial z = 0$ to find the maximum values of R for each $w < 1$ and fixed R_S. We consider below the cases when one or both of the factors of (5.19) vanish.

Case A. $R_S \geqq \mu^3/(1-\mu^2)$. In this case, the second factor in (5.19) has no real zeros for $w < 1$ and the only positive root of $\partial R/\partial z = 0$ is

$$z_1 = \left[\frac{(1+\mu)R_S}{(1-\mu)}\right]^{1/2}. \tag{5.21}$$

Using (5.16) and (5.20), we can show that

$$R(z_1;w) = ((1-\mu^2)^{1/2}w + (\mu R_S)^{1/2})^2,$$

$$\frac{\partial^2 R}{\partial z^2}(z_1;w) = \frac{2\mu^{1/2}(1-\mu)^{3/2}}{(1+\mu)^{1/2}}\left[\frac{w}{R_S^{1/2}} - \frac{(1-\mu^2)^{1/2}}{\mu^{3/2}}\right]. \tag{5.22}$$

Clearly, for $0 < w < 1$, $(\partial^2 R/\partial z^2)(z_1; w) < 0$ in this range of R_S. Thus $z = z_1$ yields a maximum of R, and the plot of R vs. z for fixed w is given in Fig. 3. We note that the supremum of R taken over $z > 0$ and $0 < w < 1$, is given by

$$R_{\text{sup}} = ((\mu R_S)^{1/2} + (1 - \mu^2)^{1/2})^2, \qquad R_S \geqq \frac{\mu^3}{1 - \mu^2}. \tag{5.23}$$

This is just the solid part of the parabola in Fig. 2 which separates the regions where there are convective steady states from where there are none (see, e.g., (4.13)).

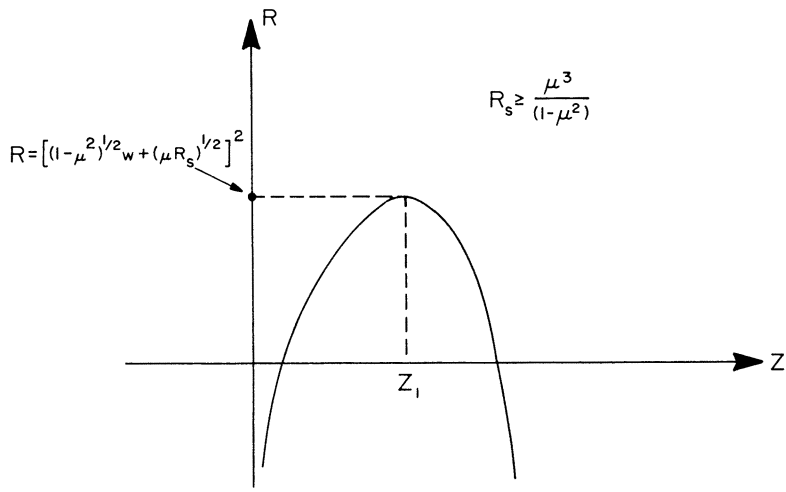

FIG. 3

On the basis of Fig. 3, we can now give the following prescription for finding a y, z and w (and thus α, β, γ) so that (5.15)–(5.18) are satisfied:

(i) For $R = R_0 \leqq 0$, choose for w any $0 < w < 1$ and for z either of the two intersection points of the curve in Fig. 3 with $R = R_0$.

(ii) For $R = R_0$ with $0 < R_0 < ((\mu R_S)^{1/2} + (1 - \mu^2)^{1/2})^2$, choose any $w > 0$ and

$$\frac{R_0^{1/2} - (\mu R_S)^{1/2}}{(1 - \mu^2)^{1/2}} < w < 1$$

and either of the two intersection points for z.

Case B. $0 < R_S < \mu^3/(1 - \mu^2)$. If we choose a w such that $0 < w < ((1 - \mu^2)^{1/2}/\mu^{3/2})R_S^{1/2}$, then the second factor in (5.19) will again have no roots and the result will be as in Case A. That is, the largest R, for given R_S, will be

$$R_{\text{sup}} = ((\mu R_S)^{1/2} + (1 - \mu^2)^{1/2})^2, \qquad 0 < R_S < \frac{\mu^3}{1 - \mu^2}.$$

This is the dotted portion of the right branch of the parabola in Fig. 2. In this range of R_S, this curve does not separate the region of steady states from that without steady states. Therefore, we must look for values of w in the range

$$\frac{(1 - \mu^2)^{1/2} R_S^{1/2}}{\mu^{3/2}} < w < 1.$$

The second factor of (5.19) will now have two roots:

$$z_{\pm} = \frac{\mu^{3/2}}{(1-\mu)}\left[w \pm \left(w^2 - \frac{(1-\mu^2)R_S}{\mu^3}\right)^{1/2}\right]. \tag{5.24}$$

We note that, in the range of w and R_S considered here, both of $z_{\pm}$ are positive. In addition, the first factor of (5.19) still has the positive root z_1 given by (5.21). The situation is now shown in Fig. 4. In this case we have that

$$R_{\text{sup}} = 1 + \frac{1}{\mu}R_S, \qquad 0 < R_S < \frac{\mu^3}{1-\mu^2}, \tag{5.25}$$

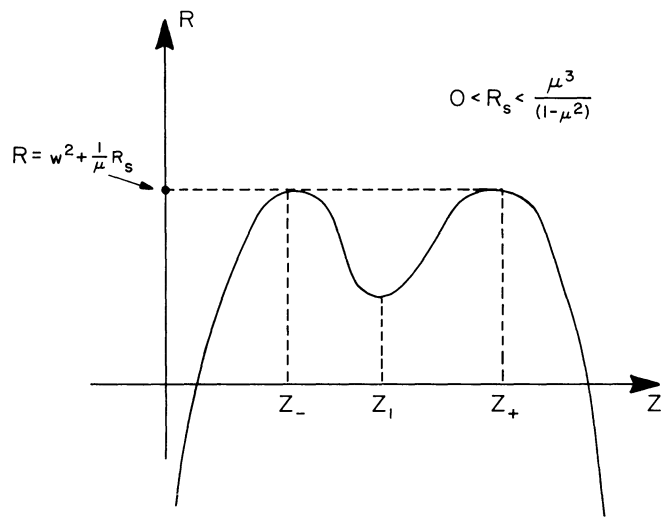

FIG. 4

which is just the desired curve $L^{(1)}$. Again we have the following prescription for finding values of y, z and w satisfying (5.15)–(5.18):

(i) For $R = R_0 \leqq 0$, choose any w satisfying $(1-\mu^2)^{1/2}R_S^{1/2}/\mu^{3/2} < w < 1$, and choose for z either of the two intersection points from Fig. 4.

(ii) For $R = R_0$ with $0 < R_0 < 1 + R_S/\mu$, choose any w satisfying $0 < w < 1$ and

$$w^2 > \max\left\{R - \frac{1}{\mu}R_S, \frac{(1-\mu^2)R_S}{\mu^3}\right\},$$

and for z, choose any of the possible intersection points from Fig. 4.

Case C. $R_S \leqq 0$. In this case only the second factor of (5.19) has real roots, and the only positive root of (5.19) is at $z = z_+$ (see, e.g., (5.24)). The graph of R vs. z is shown in Fig. 5 (the case $R_S = 0$ is different than the case $R_S < 0$ as indicated by (5.16)). As before we now have

$$R_{\text{sup}} = 1 + \frac{1}{\mu}R_S, \qquad R_S \leqq 0, \tag{5.26}$$

and this is again the desired curve $L^{(1)}$. The quantities z and w are chosen as in Case B.

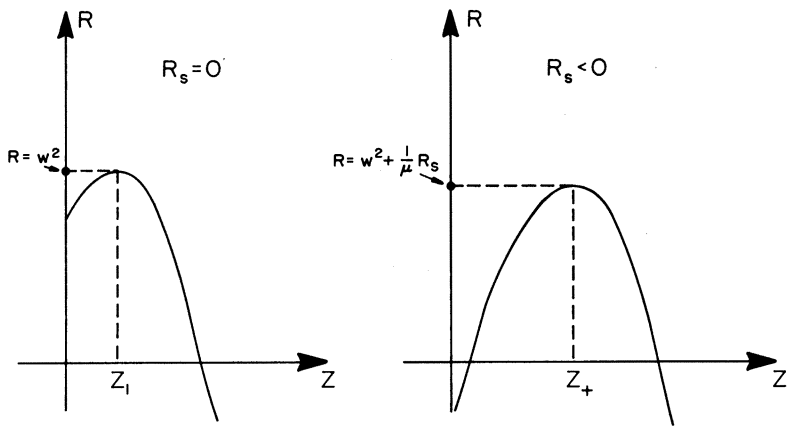

FIG. 5

To complete the discussion let us try to seek the regions in the (R, R_S)-plane where the solution to (3.3) exhibits "strong stability" (see [10]). That is, for what values of R and R_S can we find a positive definite energy functional which decays exponentially from its initial values?

From (5.1)–(5.3), for $R, R_S > 0$, the positive definite functional ϕ defined by

$$\phi \equiv R(x_1^2 + x_3^2) + \frac{1}{\sigma}x_4^2 + R_S(x_2^2 + x_5^2), \tag{5.27}$$

satisfies

$$\begin{aligned} \tfrac{1}{2}\phi' &= -R(x_1^2 + x_3^2) - x_4^2 - \mu R_S(x_2^2 + x_5^2) + 2Rx_3x_4 \\ &= -(1 - R^{1/2})[R(x_1^2 + x_3^2) + x_4^2] - \mu R_S(x_2^2 + x_5^2) \\ &\quad -R^{3/2}x_1^2 - R^{1/2}(R^{1/2}x_3 - x_4)^2. \end{aligned} \tag{5.28}$$

For $0 < R < 1$, it is easily shown from (5.28) that

$$\tfrac{1}{2}\phi' \leqq \begin{cases} -\mu(1 - R^{1/2})\phi, & \sigma \geqq 1, \\ -(1 - R^{1/2})\sigma\mu\phi, & \sigma < 1. \end{cases} \tag{5.29}$$

With slight modifications in the functional ϕ we can show that a similar result to (5.29) holds for $R < 0$ and $R_S > -\mu$. Thus we have exhibited strong stability in the regions

$$0 < R < 1, \qquad R_S > 0$$

and

$$R < 0, \qquad R_S > -\mu.$$

These regions include the second quadrant in the (R, R_S)-plane, where, since both salt and temperature are stabilizing, we expect there to be strong stability. It also includes that region in the first quadrant lying directly above the triangular region in Fig. 2. These values of R (i.e., $0 < R < 1$) correspond to large dissipation (see discussion in § 3). Thus this region of strong stability in the first quadrant, where temperature produces instability while salt produces stability, is physically

accounted for by the large dissipation ($0 < R < 1$) dominating the tendency of the temperature field to produce instability. A similar argument applies in the third quadrant where $R_S > -\mu$, which lies to the left of the triangular region.

6. Summary and future work. We have considered a simplified model of a two-constituent fluid which exhibits most phenomena currently known to be present in the more realistic model of a Newtonian fluid [1], [10]. Our model permits explicit determination not only of linear stability curves but also of the boundary in parameter space which delineates the region of existence of nontrivial steady solutions. We have rigorously shown that the expected region of global stability for the conductive (trivial) solution is precisely given by the analogue of the energy methods of Joseph [10]. Specifically, we have shown that the subregion of linear stability, in which there are no nontrivial steady solutions, is a region of global stability where all solutions decay.

Several important questions remain. We are presently providing, through asymptotic methods, a formal resolution of how, in parameter regions near the linear stability boundaries, solutions to the initial value problem (2.13)–(2.17) with finite but small initial conditions progress in time. We have also explored the phenomenon of dynamic bifurcation near the point $P^{(0)}$ of Fig. 2, where two linear eigenvalues coalesce, and have derived time evolution equations. These asymptotic results will appear in another paper [15].

We are also studying the stability of the steady convective states in § 4 and integrating (2.13)–(2.17) numerically in order to understand further transitions to periodic and nonperiodic unsteady motions. In addition, we are determining the asymptotic nature of the solutions of these equations for large Rayleigh numbers R and R_S and small values of the Schmidt number μ. Finally we hope to apply the techniques developed for this model to other more complicated and realistic ones.

Acknowledgment. The authors would like to thank Professors Clement McCalla and Harry McLaughlin, of the Mathematical Sciences faculty at Rensselaer Polytechnic, and Professor Louis Howard of M.I.T., for several helpful discussions relating to the work presented here.

REFERENCES

[1] J. S. TURNER, *Buoyancy Effects in Fluids*, Cambridge Univ. Press, London, 1973.
[2] J. B. KELLER, *Periodic oscillations in a model of thermal convection*, J. Fluid Mech., 26 (1966), pp. 599–606.
[3] P. WELANDER, *On the oscillatory instability of a differentially heated fluid loop*, Ibid., 29 (1967), pp. 17–30.
[4] H. STOMMEL AND C. ROOTH, *On the interaction of gravitational and dynamic forces in simple circulatory models*, Deep-Sea Res., 15 (1968), pp. 165–170.
[5] D. W. MOORE AND E. A. SPIEGEL, *A thermally excited non-linear oscillator*, Astrophys. J., 143 (1966), pp. 871–887.
[6] N. H. BAKER, D. W. MOORE AND E. A. SPIEGEL, *Aperiodic behavior of a non-linear oscillator*, Quart. J. Mech. Appl. Math., 24 (1971), pp. 391–422.
[7] L. N. HOWARD AND W. V. R. MALKUS, 1972, unpublished.
[8] K. ROBBINS, Ph.D. thesis, Dept. of Math., Mass. Inst. of Tech., submitted.
[9] E. N. LORENZ, *Deterministic nonperiodic flow*, J. Atmospheric Sci., 20 (1963), pp. 130–141.
[10] D. D. JOSEPH, *Global Stability of Fluid Motions*, Springer Tracts in Natural Philosophy, to appear.

[11] R. L. SANI, *On finite amplitude roll cell disturbances in a fluid layer subjected to heat and mass transfer*, Amer. Inst. Chem. Eng. J., 11 (1965), pp. 971–980.
[12] J. M. STRAUSS, *Finite amplitude doubly diffusive convection*, J. Fluid Mech., 56 (1972), pp. 353–374.
[13] E. A. SPIEGEL AND G. VERONIS, *On the Boussinesq approximation for a compressible fluid*, Astrophys. J., 131 (1960), pp. 442–447.
[14] E. A. GUILLEMIN, *The Mathematics of Circuit Analysis*, John Wiley, New York, 1949.
[15] L. A. RUBENFELD AND W. L. SIEGMANN, *Nonlinear dynamic theory for a double-diffusive convection model*, to be submitted.
[16] J. K. HALE, *Ordinary Differential Equations*, Interscience, New York, 1969.
[17] W. V. R. MALKUS, *Non-periodic convection at high and low Prandtl number*, Mem. Soc. Roy. Sci. Liege, 4 (1972), pp. 125–128.