THE DEVELOPMENT OF LAYERS IN DIRECT DOUBLE-DIFFUSIVE CONVECTION

by

Mark David Handel

A.B.Hon. Physics
The College, The University of Chicago
(1979)

Submitted to the Department of
Earth, Atmospheric, and Planetary Sciences
in Partial Fulfillment of the Requirements of the Degree of

Master of Science

at the

Massachusetts Institute of Technology

February 1984

Signature of author: _____
Mark David Handel
January 20, 1984

Certified by: _____
Erik L. Mollo-Christensen
Thesis Supervisor

Accepted by: _____
Theodore R. Madden
Chairman, Departmental Committee on Graduate Students

The development of layers in direct double-diffusive convection

by

Mark David Handel

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences on January 20, 1984, in partial fulfillment of the requirements for the Degree of Master of Science in Physical Oceanography.

ABSTRACT

Direct double-diffusive convection (salt fingering) is known under many circumstances to set its own vertical scale smaller than that imposed by the physical boundaries. In the oceans, a domain essentially unbounded in the vertical, the vertical scale of the salt finger cells is determined by the local dynamics, rather than the location of the boundaries. A new mechanism is proposed for the formation of layers in direct double-diffusive convection. Convective modes of higher vertical wave number than the fastest growing fundamental, called harmonic instabilities, create regions of reduced, or even inverted, density gradient alternating with regions of increased density gradient. The regions of reduced density gradient become turbulent because of additional convective instabilities. The resulting configuration is alternating layers of laminar cells and turbulent convection. Qualitatively this agrees well with observations. Calculations of the effects of growing linear modes on the mean fields are presented to illustrate this mechanism. Some preliminary results from numerical simulations of very supercritical direct double-diffusion are presented and discussed.

Thesis Supervisor: Erik L. Mollo-Christensen, Sc.D.

Title: Professor of Oceanography

to J. G. Ferretti

for the wine, the path, and the warnings

"Hank, if we open the faucet it will run forever."

Arnold B. Arons to Henry Stommel

Acknowledgments

The support of the Fannie and John Hertz Foundation during the preparation of this thesis is gratefully acknowledged. Erik Mollo-Christensen, my advisor, has been great at keeping science fun. Steve Piacsek of the Navy Ocean Research and Development Activity provided some of the computer codes used and many hours of his time helping me begin the simulations. Diana Spiegel has helped me fight with more than one computer. Ed Spiegel, Melvin Stern, and Willem Malkus asked searching questions which helped the progress of this work. My thanks to Steve Meacham for occasionally filling our office with jasmine. Ray Schmitt introduced me to the subject of salt fingers. Computer time for the simulations has been provided by the National Aeronautics and Space Administration (NASA) through its High Speed Computer Facility at Goddard Space Flight Center. Some of the results have been checked using MACSYMA and other facilities of the Mathlab Group of the Laboratory for Computer Science at MIT which was supported in part by NASA under grant NSG 1323, by the Office of Naval Research under grant N00014-77-C-0641, by the U. S. Department of Energy under grant ET-78-C-02-4687 and by the U. S. Air Force under grant F49620-79-C-020. The figures from Schmitt (1979a) are reproduced with the permission of Pergamon Press. The figures from Linden (1978) are reproduced with the permission of the American Geophysical Union. My thanks to the crew on the 8th floor at 545 Tech Square for their assistance with the Maximum Confusion machine. Susan Deming and I now have each survived the other's degree crunch — gracefully. The continued moral support of Mark Holzbach and my parents, Irma and Mort Handel, has gotten me through the moments of frustration. My thanks to all of them.

Table of Contents

Abstract 2
Acknowledgments 4
List of Tables and Figures 6
List of Symbols 11
1. Introduction 15
2. History and Applications 25
3. Basic Equations and Linear Theory 33
4. Nonlinear Theory 63
5. Layer Related Theory and Observations 67
6. Harmonic Instabilities 83
7. Numerical Simulations 120
8. Discussion and Collected Thoughts 137
9. Summary and Conclusions 150
References 154

List of Tables and Figures

Fig. 1. Illustration of double-diffusive instability: p. 17.
a) direct case
b) overstable case.

Fig. 2. Growth rate as a function of horizontal and vertical wave numbers for direct double-diffusive convection heat-salt case, density ratio = 2.0. p. 39.

Fig. 3. Growth rate as a function of horizontal and vertical wave numbers for heat-salt case, expanded scale: p. 40.
a) density ratio = 1.5
b) density ratio = 1.2
c) density ratio = 1.1
d) density ratio = 1.0.

Fig. 4. Growth rate as a function of horizontal and vertical wave numbers for direct double-diffusive convection salt-sugar case, density ratio = 2.0. p. 42.

Fig. 5. Growth rate as a function of horizontal and vertical wave numbers for salt-sugar case, expanded scale: p. 43.
a) density ratio = 1.5
b) density ratio = 1.1
c) density ratio = 1.0
d) density ratio = 0.95.

Fig. 6. Maximum growth rate as a function of density ratio (from Schmitt 1979a): p. 47.
a) heat-salt case
b) salt-sugar case.

Fig. 7. Wave number of fastest growing mode and marginal wave number as a function of density ratio (from Schmitt 1979a): p. 48.
a) heat-salt case
b) salt-sugar case.

Fig. 8. Derivative of growth rate with respect to square of vertical wave number: p. 50.
a) heat-salt case, density ratio = 1.2
b) heat-salt case, density ratio = 1.2, detail.

Fig. 9. Derivative of growth rate with respect to square of vertical wave number: p. 51.
a) heat-salt case, density ratio = 1.1
b) salt-sugar case, density ratio = 1.1.

Fig. 10. Growth rate as a function of horizontal and vertical wave numbers for one component convection: p. 54.
a) salt case
b) thermal case.

Fig. 11. Buoyancy flux ratio as a function of horizontal and vertical wave numbers: p. 58.
a) heat-salt case, density ratio = 1.2
b) salt-sugar case, density ratio = 1.2.

Fig. 12. Flux ratio for fastest growing mode and flux ratio for neutral mode as a function of density ratio (from Schmitt 1979a): p. 60.
a) heat-salt case
b) salt-sugar case.

Fig. 13. There is no fig. 13.

Fig. 14. Time lapse photographs of salt-sugar fingers without layering (from Linden 1978): p. 75.
a) $t = 1h01$
b) $t = 7h33$.

Fig. 15. Time lapse photographs of salt-sugar fingers with layering (from Linden 1978): p. 77.
a) $t = 0h23$
b) $t = 0h42$
c) $t = 1h05$
d) $t = 1h22$
e) $t = 2h04$.

Fig. 16. Time lapse photgraphs of salt-sugar fingers with layering (from Linden 1978): p. 81.
a) $t = 4h08$
b) $t = 5h01$
c) $t = 21h43$.

Fig. 17. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for heat-salt case, density ratio = 1.5: p. 90.
a) $t = 5$
b) $t = 10$
c) $t = 20$
d) $t = 25$.

Fig. 18. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for heat-salt case, density ratio = 1.1: p. 92.
a) $t = 10$
b) $t = 15$
c) $t = 20$
d) $t = 25$.

Fig. 19. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar case, density ratio = 1.5: p. 96.
a) $t = 40$
b) $t = 60$
c) $t = 80$
d) $t = 100$.

Fig. 20. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar fingers, density ratio = 1.1: p. 98.
a) $t = 10$
b) $t = 20$
c) $t = 30$
d) $t = 40$
e) $t = 50$
f) $t = 60$.

Fig. 21. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar fingers, density ratio = 1.05: p. 101.
a) $t = 10$
b) $t = 20$
c) $t = 30$
d) $t = 40$.

Fig. 22. Flux divergence ratio as a function of horizontal and vertical wave numbers: p. 104.
a) heat-salt case, density ratio = 1.2
b) salt-sugar case, density ratio = 1.2.

Fig. 23. Effect on mean potential energy density by growing modes as a function of horizontal and vertical wave numbers for salt-sugar fingers, density ratio = 1.5: p. 105.
a) $t = 20$
b) $t = 30$
c) $t = 40$
d) $t = 50$.

Fig. 24. Effect on mean potential energy density by growing modes as a function of horizontal and vertical wave numbers for salt-sugar fingers, density ratio = 1.1: p. 107.
a) $t = 10$
b) $t = 20$
c) $t = 30$
d) $t = 40$.

Fig. 25. Vertical mode with greatest effect on the mean density ratio as a function of time p. 111.

Fig. 26. Development over time of mean salinity field in numerical simulation *A*. p. 124.

Fig. 27. Development over time of mean density field in numerical simulation *A*. p. 125.

Fig. 28. Development over time of salt Reynolds fluxes in numerical simulation *A*. p. 126.

Fig. 29. Development over time of mean salinity field in numerical simulation *B.* p. 128.

Fig. 30. Development over time of mean density field in numerical simulation *B.* p. 129.

Fig. 31. Development over time of salt Reynolds fluxes in numerical simulation *B.* p. 130.

Fig. 32. Development over time of mean salinity field in numerical simulation *C.* p. 133.

Fig. 33. Developemt over time of mean density field in numerical simulation *C.* p. 134.

Fig. 34. Development over time of salt Reynolds fluxes in numerical simulation *C.* p. 135.

List of Symbols

The variables are labeled for the heat-salt system. For other systems the faster diffusing component corresponds to heat and the slower diffusing one to salt. When examining one component convection only the component labeled S is present.

No notational distinction is made between dimensional and nondimensional growth rates, wave numbers, wavelengths, and other quantities; context makes the choice unambiguous. Numerical quantities are dimensionless unless units are specified.

a = Stern's parameter in eq. (27)

A_o = initial linear temperature gradient when constant

$A(z)$ = $\frac{\partial \overline{\theta_\bullet}}{\partial z}$ = mean temperature gradient, may be time dependent

B_o = initial linear salinity gradient when constant

$B(z)$ = $\frac{\partial \overline{S_\bullet}}{\partial z}$ = mean salinity gradient

C = area of horizontal integration or averaging

F = $\beta F_S - \alpha F_T$ = total buoyancy flux

F_S = $\overline{wS} - K_S \frac{\partial \overline{S_\bullet}}{\partial z}$ = total salinity flux

F_T = $\overline{w\theta} - K_T \frac{\partial \overline{\theta_\bullet}}{\partial z}$ = total heat flux

$\mathcal{F}_S$ = $\overline{wS}$ = salt Reynolds flux

$\mathcal{F}_T$ = $\overline{w\theta}$ = heat Reynolds flux

g = gravitational acceleration magnitude

h = thickness of a finger layer

H = height of experimental box, or thickness of turbulent region

$J(\psi, \varsigma)$ = $\frac{\partial \psi}{\partial x}\frac{\partial \varsigma}{\partial z} - \frac{\partial \varsigma}{\partial x}\frac{\partial \psi}{\partial z}$ = Jacobian determinant; ψ and ς are arbitrary functions

k = horizontal wave number magnitude, $k^2 = k_1^2 + k_2^2$

k_f = fastest growing wave number

k_1, k_2 = x and y wave numbers, respectively

K_S = salt, or slower diffusing component, diffusivity

K_T = thermal conductivity, or faster diffusing component diffusivity

l = horizontal wavelength

l_f = horizontal wavelength of fastest growing modes

L = length scale = $[g\beta B/(K_T\nu)]^{-1/4}$ for double-diffusion
= $[g\beta B/(K_S\nu)]^{-1/4}$ for one component convection

m = vertical (z) wave number

m_1 = vertical wave number of fundamental for system of given height

m_l = vertical wave number for $\Gamma_{\max}$, or vertical wave number for layering

N = $[-(g/\rho)\frac{\partial\rho}{\partial z}]^{1/2}$ = Brunt-Väisälä, or buoyancy, frequency

Nu_T = $F_T/(K_T A)$ = heat Nusselt number

Nu_S = $F_S/(K_S B)$ = salt Nusselt number

p = pressure

q = ratio of nonlinear to linear terms, see eq. (38)

Q = $\partial_z F_T/\partial_z F_S$ = flux divergence ratio

R_c = critical Rayleigh number for convection
for convection between free slip plates = $\frac{27}{4}\pi^4$

Re = $|v|l/\nu$ = Reynolds number

Ri = $[-(g/\rho)(\frac{\partial\rho}{\partial z})]/(\frac{\partial u}{\partial z})^2$ = Richardson number

R_S = $(gH^4\beta B)/(\nu K_S)$, or $(gH^3\beta\,\Delta S)/(\nu K_S)$ = salt Rayleigh number

R_T = $(gH^4\alpha A)/(\nu K_T)$, or $(gH^3\alpha\,\Delta T)/(\nu K_T)$ = thermal Rayleigh number

R_* = $R_S - R_T$ = double-diffusive Rayleigh number

S = $S_* - B_o z$ = salinity change from base initial state

S_* = salinity measured from reference, S_o

S = $g\beta BL$ = salinity scale

t = time

T = time scale = $(g\beta BK_T/\nu)^{-1/2}$ for double-diffusion
= $(g\beta BK_S/\nu)^{-1/2}$ for one component convection

U = potential energy of density field

$\mathcal{U}$ = potential energy density of density field

$\mathbf{v}$ = (u, v, w) = velocity vector in (x, y, z) directions

$\hat{w}$ = vertical velocity perturbation amplitude

x, y, z = spatial coordinates; with carets, *e.g.*: $\hat{z}$, for unit vectors

α = $-\frac{1}{\rho_o}\frac{\partial \rho}{\partial \theta_*}\Big|_{\substack{\theta_*=0 \\ S_*=0}}$ = coefficient of thermal expansion, *N.B.* sign convention is maintained even for components other than heat so $\alpha < 0$ is allowed

β = $\frac{1}{\rho_o}\frac{\partial \rho}{\partial S_*}\Big|_{\substack{\theta_*=0 \\ S_*=0}}$ = density coefficient for salt

γ = $(k^2 + m^2)^{1/2}$ = total wave number

$\delta\mathcal{U}$ = change in potential energy density

ϵ = $\Lambda(z) - 1$ = measure of density stratification

ϵ_o = $\Lambda_o - 1$

ς = $\nabla^2\psi$ = 2-dimensional vorticity function of (x, z)

η = $K_T/\nu = \sigma^{-1}$ = reciprocal of Prandtl number

θ_* = temperature measured from reference, θ_o

θ = $\theta_* - A_o z$ = temperature difference from initial base state

Θ = $\Lambda_o S = g\alpha AL$ = temperature scale

λ = complex growth rate

Λ_o = $(\alpha A_o)/(\beta B_o)$, or $(\alpha\,\Delta T)/(\beta\,\Delta S)$ = density ratio

$\Lambda(z)$ = $[\alpha A(z)]/[\beta B(z)]$ = vertically local, horizontally averaged, density ratio

μ = $K_S/\nu = \sigma_S^{-1}$ = reciprocal Schmidt number

ν = kinematic viscosity

σ = $\sigma_T = \nu/K_T$ = Prandtl number, or Schmidt number if using a solute instead of heat

σ_S = ν/K_S = salinity Schmidt number

ρ = density

ρ_o = $\rho(\theta_o, S_o)$ = reference density

τ = K_S/K_T = diffusivity ratio, or Lewis number

ϕ = $\rho_o^{-1} p - \overline{w^2}$

χ = $(\alpha F_T)/(\beta F_S)$, or approximately $(\alpha \mathcal{F}_T)/(\beta \mathcal{F}_S)$ = flux ratio

ψ = 2-dimensional stream function of (x, z)

$\overline{\psi}$ = overbar indicates horizonal average; ψ is an arbitrary function

$\frac{d}{dt}$ = $\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla$ = $\frac{\partial}{\partial t} + J(\psi,)$ = advective time derivative

ΔS = salinity difference across a layer or between boundaries

ΔT = temperature difference across a layer of between boundaries

∇_2^2 = $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}$ = two dimensional Laplacian in x and z

1. Introduction

The possibility of buoyantly driven convective motion in a fluid with a negative average density gradient has been known for some time, though this is usually referred to as a stable stratification. There are several ways to create such situations, of which direct double-diffusive convection is one. Direct double-diffusive, or differential-diffusive, convection can occur if there is a destabilizing component and a more rapidly diffusing stabilizing component, even if the stabilizing component contributes more to the overall density field. The modes are referred to as direct because the initial perturbation grows monotonically and exponentially. If the component gradients are reversed but the density gradient is kept negative, the convection begins with a growing oscillation. For the oceanic case, where the destabilizing component is salt and the stabilizing one heat, the phenomenon is known as salt fingers. Historically, this was the first case of double-diffusion that was examined theoretically (Stern 1960).

The direct cellular modes have some resemblance to low Rayleigh number single component convection, such as in the Bénard, or Rayleigh-Jeffreys problem, as it is variously known. However in double-diffusive convection, diffusion is needed throughout the fluid to drive the instability even at large Prandtl number, not just at the boundaries where vertical motions are prevented. This is clearly essential in run-down experiments, but the balance of boundary and interior effects in a real parallel plate experiment might be different. The visually striking features of salt fingers are tall narrow cells and distinct layering. Layers of laminar cells alternate with regions resembling high Rayleigh number convection. The process of layer formation, the instabilities that lead to it, and the resulting vertical structures

are the subjects of this thesis.

This introduction gives a qualitative description of the basic instability and lays out the problem to be attacked. Chapter 2 gives a brief history of the study of direct double-diffusive convection. It also examines where the phenomenon is important and notes variations also being studied. The basic equations and linear theory, both previously published and new, are presented in chapter 3. Previous nonlinear work is reviewed in chapter 4, while previous work on layering is presented in chapter 5. Chapter 6 presents the details of the harmonic instability theory of layering being introduced, including determinations of flux divergences, their effects on the mean fields, and the resulting change in potential energy, using the growing linear modes. Preliminary numerical simulations of salt fingers are discussed in chapter 7. Chapter 8 is devoted to a discussion and several loosely coupled conjectures. Chapter 9 contains a summary and conclusions along with further problems to persue. Many symbols and named parameters are not defined in the text, but all are included in the list of symbols.

Qualitatively, as initially realized by Stern (1960), the instability may be described as follows. Start with a body of fluid having both temperature and salinity increasing upwards. If only the temperature stratification was present in the given sense the fluid would be stable, while if only the salinity stratification was present the fluid would be unstable to overturning convection. Note that the thermal conductivity is nearly one hundred times greater than the salt diffusivity. Initially the density effect of the temperature is taken to be greater than that from the salt, so the net density stratification is decreasing upwards and hence usually referred to as stable. A parcel of fluid displaced upwards, by any small random cause, gains heat much faster than salt, becomes positively buoyant, and continues

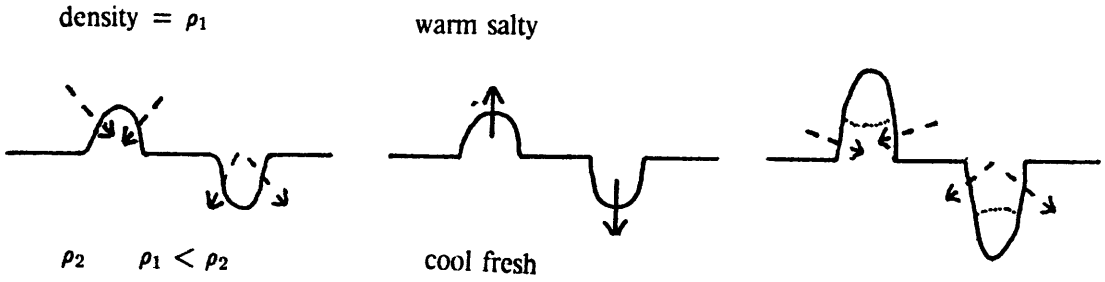

Fig. 1a.

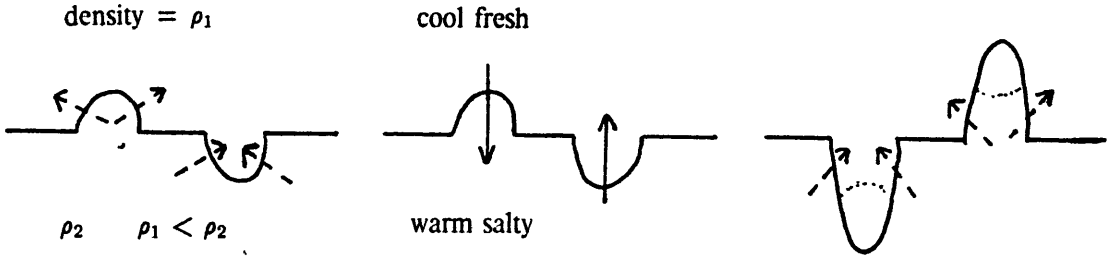

Fig 1b.

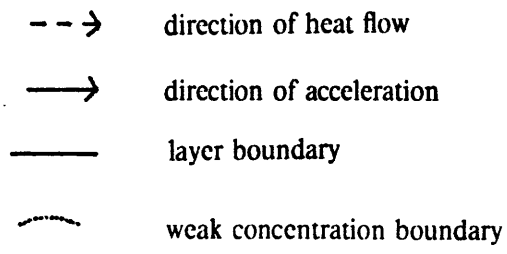

Fig. 1. Illustration of double-diffusive instabilities:
a) direct case — An initial upwards perturbation is cool fresh water surrounded on almost all sides by warm salty water. Heat, which diffuses faster than salt, enters the perturbation which soon becomes relatively warm while remaining fresh. It is now positively buoyant and continues upwards. The water filling its place is subjected to the same process. A downwards perturbation looses heat to its cooler surroundings , becomes cool while remaining salty, and is accelerated down because of its negative relative buoyancy.

b) overstable or oscillatory case — An initial upwards perturbation is cooled by its surroundings but keeps its salt and so becomes negatively buoyant. When it drops back to its original position it is still accelerated downwards giving a growing oscillation. Conversely, a downward perturbation gains heat but not salt and from this added positive buoyancy is still accelerating upwards when it passes its original position.

upwards. This differential diffusion process can occur faster than the overall density stratification can send the parcel back to its initial position by internal oscillations. Complimentarily, a parcel displaced downwards loses heat before losing salt and continues to descend. An illustration of this process is in fig. 1a. The thinner the intrusions headed in opposite directions are, the more rapid is the differential diffusion process. However, for a set of many thin intrusions shears are also greater, increasing viscous dissipation, or friction. This balance leads to a minimum unstable wavelength. On the other hand, very large intrusions are not viscously damped but are returned toward their initial position by the overall stable stratification, before differential diffusion can create local regions of relative positive buoyancy. These opposing processes lead to some optimal size intrusion for growth.

It is not neccessary for the vertical density gradient to invert anywhere, even locally, for this process to occur, though it might. As is particularly clear from the vorticity equation (11), it is the pattern of the lateral density gradients combined with the continuity requirement that drives the vertical motions. The difference in the diffusivities need not be so extreme as the factor of nearly one hundred between heat and salt. The salt-sugar case has a diffusivity ratio of three and fairly vigorous convection is still observed, though double-diffusive convection on the whole is still considerably tamer than convection with a positive density gradient. Double-diffusion has even been observed with a diffusivity ratio of less than 1.3 using potassium chloride and sodium chloride (Stern and Turner 1969) (however the ratio may be understated since only the relative diffusivities of the cations is relevant with the anions identical).

Typically salt finger widths in the laboratory are slightly less than a centimeter. For the weaker mean gradients in the ocean fingers several centimeters wide are

observed. Salt-sugar fingers created in the lab or kitchen have wavelengths of a couple of millimeters or less. These scales depend on several fluid and solute parameters which are the viscosity of water $\sim 0.01\, cm^2/sec$, thermal conductivity $\sim 1.3 \cdot 10^{-3}\, cm^2/sec$, salt (sodium chloride) diffusivity $\sim 15 \cdot 10^{-6}\, cm^2/sec$, and sugar (sucrose) diffusivity $\sim 5 \cdot 10^{-6}\, cm^2/sec$. Density gradients in the lab are usually of order parts per thousand in centimeters while in the ocean much weaker gradients of parts per million over centimeters to meters are the norm.

Double-diffusion also allows another interesting instability which begins as a growing oscillation, first presented by Turner and Stommel (1964) and independently by Walin (1964). In this case salt, the more slowly diffusing component, is stabilizing, while the temperature gradient is destabilizing. Now a parcel displaced upwards loses heat more rapidly than salt and gains density. It is strongly accelerated back downwards and when it reaches its initial position it is still feeling a downwards acceleration because of its increased density. This situation is sometimes referred to as overstability and is illustrated in fig 1b. Sharply defined layers with internal fronts also occur in this type of double-diffusion (Turner 1965, Huppert and Linden 1979). However very different processes seem to cause the layers though they may be maintained in similar manners. The oscillatory case will not be considered further except tangentially. As a further example of a convection with a hydrostatically "stable" mean density gradient, Wunsch (1970) showed that a sloping insulated boundary forces fluid motions because the isopycnals must be perpendicular to the insulated boundary. Convection is a far more complicated and varied set of phenomena than motions driven by a simple density inversion.

Though the work cited above describes the nature of the initial salt finger instability, it gives few clues as what to expect for a finite amplitude steady, or

quasi-steady, state. It is known that the cells caused by thermohaline instability often do not extend to the physical top and bottom boundaries of the unstable fluid but rather form discrete layers of laminar cells separated by turbulent regions (Stern and Turner 1969). Linden's (1978) work clearly demonstrates the evolution from linear gradients, to tall cells, and then to layers of cells. This layering is considered a signature of salt fingers in the ocean. The photographs in Turner 1967, Shirtcliffe and Turner 1970, and Linden 1978 (reproduced here in figs. 14 – 16), are intriguing examples of the phenomena under discussion. Williams (1975) provides an excellent description of oceanic cases of vertical structure caused by salt fingers. Vertical scales for salt fingers in the lab range from centimeters to tens of centimeters. Oceanic layers range from tens of centimeters to several meters. Sugar-salt fingers have heights from several millimeters to several centimeters in typical experiments. Currently the only theory on layering is that of Stern (1969a) on the collective instability which was refined by Holyer (1981), whereby they show that salt fingers can drive internal waves and then suggest that the resulting shear breaks up the cells.

This thesis presents a new mechanism for layering, or lateral front development, in direct double-diffusive convection. Convective modes of higher vertical wave number than the fastest growing fundamental, called primary harmonic instabilities, create regions of reduced, or even inverted, density gradient alternating with regions of increased density gradient. This is the result of the difference in the vertical flux divergences of heat and salt from any given mode. The regions of reduced density gradient, because of further small scale convective instabilities, called secondary harmonic instabilities, become turbulent. The resulting configuration is alternating layers of laminar cells and turbulent convection.

These two phases of laminar cells and turbulent convection are each seen separately in the Bénard problem with one component. With just one component convection roughly goes, with increasing Rayleigh number, from diffusive transport and no motion, to laminar convection rolls with width and height comparable, to square convection cells (often called cross rolls in the convection literature), to time dependent oscillating convections, to high Rayleigh number or turbulent convection. (For a more complete description see Busse 1978a.) This ignores the dependence on the Prandtl number and applies particlularly at large Prandtl numbers, as with water. In the ordered laminar motions the gradients are concentrated slightly at the boundaries and at moderate Rayleigh number can sometimes even lead to the inversion of the interior gradient so that diffusion balances regions with excessive Reynolds fluxes (fluxes from the correlation of velocity and temperature disturbances). In turbulent convection at high Prandtl number there are thin boundary layers and almost no interior gradient. Transport is by plumes or blobs of fluid passing through the interior that have broken from the boundary (a model of this is presented by Howard 1964). At low Prandtl number diffusion remains important throughout the fluid and the interior gradients are more convoluted (see the numerical simulations by Grötzbach 1982).

Beyond the issue of understanding layering in some specific situation, such as the ocean, the question is of interest in a relatively abstract sense for the pure study of fluid motion and the partial differential equations associated with that study. Sharply defined layers are frequently the result of secondary instabilities as well as from saturation at large amplitude of an initially unstable mode. A secondary instability is a perturbation that grows because of conditions created by the finite amplitude effects of a primary disturbance different from the initial conditions. The

primary disturbance can be either growing, as is relevant here, or decaying. (For a fascinating example of a secondary instability on a decaying mode see Orszag and Patera 1983 on a theory of low Reynolds number instability in wall bounded flows.) Stern's collective instability is clearly a secondary instability. The status of the theory of harmonic instability presented here is more vague for reasons that will be discussed below. There is a great difficulty in distinguishing *a priori* which unstable initial conditions will lead to a simple saturation of a growing mode in the manner of a Landau equation, which will lead to complicated time dependent amplitude behaviour in the manner of some coupled amplitude equations or in the manner of chaotic Lorenz type systems, and which will allow new instabilities to grow on the backs of other instabilities. Thankfully, the laboratory frequently provides an efficient computer for analysing the behaviour of certain sets of partial differential equations.

An unusual property of the direct double-diffusive system is the manner in which length scales are determined. It appears that if the experimental cell is sufficiently large, the length scales in all three spatial dimensions of the instability phenomena become independent of the boundary geometry. The boundary conditions are still controlling, but possibly only the solute fluxes are important. This is greatly different than in single component convection. Though internal fluid conditions frequently set length scales in two spatial dimensions, it is difficult to find other systems where the internal dynamics determine all three. If one considers what kinds of processes can break symmetries in fluid problems this is not so surprising. Then again we may eventually find out that salt finger layers are long lived transients with the final state a single cell stretching between the top and bottom boundaries.

It is important to note that layering in double-diffusive convection is not just the appearance of standing cells with a vertical wave number greater than the smallest allowed for the fundamental, though such a system would be layered. For example, the double-diffusive problem experimentally examined by Calman (1977), using heat and angular momentum, has a primary instability with the fastest growing mode having large vertical wave number and smallest allowed radial wave number, which leads to a set of vertically stacked cells but not sharp fronts. The convection treated here leads to the development of internal horizontal fronts of great lateral extent. These fronts appear to be quite sharp and divide regions with qualitatively different behaviour.

In both the structure of the physics and the mathematics, many similarities can be found between one component and double-diffusive convection. However, there are some important qualitative differences. Most noticeably, direct double-diffusion leads to tall, narrow convection cells that form sharply defined layers of broad horizontal extent, separated by turbulent regions. This propensity for layering is not usually found in one component convection, though the two different phases within the double-diffusive layers are each seen separately in different parameter ranges of one component convection. Also, unless otherwise constrained the cellular regime with only one convecting component has comparable vertical and lateral scales.

Comparisons between one and two component convection will be made here more extensively than has been done in the past. Several features of layering will be speculated about because the experiments are as yet undone. On what time scale are layers a steady state phenomenon if not indefinitely? If one performed a double-diffusive experiment akin to the Bénard experiment would eventually

one high Rayleigh number type cell result? Do density or concentration gradients ever invert in the layers? How much of the apparent differences between one and two component convection are simply the result of the difference in time scales, because of the relative magnitude of the buoyant forces, and not structural distinction in the system equations? Some of the differences may provide clues as to why double-diffusive convection demonstrates complex layered vertical structure.

2. History and Applications

The first known work on double-diffusion is that of Jevons in 1857. He was attempting to model the cause of "cirrous" clouds and performed an experiment using sugar and heat creating fingers, which he also correctly explained. After this insightful piece of work diffusive convection was promptly forgotten for one hundred years. The estimated importance of double-diffusive processes has grown immensely in the past twenty-five years. What started in the present round of research as an "oceanographic curiosity," has led to a complex set of phenomena with myriad applications. The curious idea was of a passive pipe that could pump up colder, fresher, water in the ocean (Stommel *et al.* 1956)* by allowing diffusion of heat but not salt, in an area where the gradients of temperature and salinity were both positive. Stern (1960) then showed that the pipe was unneccessary and that free convection in a fluid interior was possible, based on the inherent difference of the diffusivities.

The first observation, in the modern work, that salt fingers are in fact realized was an experiment by Stommel and Faller (Stern 1960). Turner (1967) provided the next experiments and was the first to demonstrate the layering effect and clearly document the large height to width aspect ratio. The cells tend to be square and meet at right angles (Shirtcliffe and Turner 1970) while layers have a large, level, horizontal coherence. The first modern applications were on oceanic mixing of heat and salt (Stern 1967). Oceanic applications will be discussed in more detail below. Though it is apparently not responsible for cirrus clouds, Jevons might be

* A description of this discovery can be found in Arons 1981, p. xvi. Stommel, Louis Howard, and Dave Nergaard eventually built an oceanic salt fountain of uncertain success with a long piece of tubing (see Stong 1971).

pleased to know that double-diffusion may be responsible for mammas on anvil clouds (E. Mollo-Christensen, personal communication; see photographs in Scorer 1972). Heat-moisture double-diffusion may also exist in the atmosphere other than in clouds (Merceret 1977). Application has been made in stellar dynamics (*e.g.*: Ulrich 1972). Double-diffusive convection causes problems in the solidification of alloys (Copley *et al.* 1970) and in separation processes (Mason 1976), where one might like to eliminate it. Schmitt (1983) gives some behaviors of the growth rate, wave number, and other dependent properties, covering Prandtl numbers and diffusivity ratios appropriate for such systems.

Complications could be compounded by having an insulated sloping boundary (Linden and Weber 1977) or by adding a third component [as Griffiths (1979) did for the overstable case]. Other variants include non-Fickian conservation equations of the components, such as with chemical reactions. Systems with density and angular momentum differentially diffusing (McIntyre 1970a, 1970b) are possible and the primary instability exhibits coherent layers (Calman 1977), though as was noted above, these are not fronts. This case may be important in the dynamics of Gulf Stream eddies and rings (Lambert 1974). A differential-diffusive convection is possible with both gradients stably stratified when the diffusion equations are coupled linearly (McDougall 1983) or nonlinearly (Schaefer 1975). To the two component case we started with, the effects of shear (Linden 1974), rotation (Schmitt and Lambert 1981, Pearlstein 1981), or both (Worthem *et al.* 1982) may be added. Including these makes singular changes to the system of equations by increasing the order of the dispersion relation and adds new sets of modes, some of which have been studied in other contexts and for which diffusion is an uninteresting perturbation. These complications, however, may greatly affect the

essentially diffusion driven modes in some parameter ranges.

As Munk (1966) noted, in most of the world ocean the observed turbulence levels are insufficent to provide mixing rates as large as are observed if heat and salt are passive tracers. (Though in the area he was primarily looking, the mid-Pacific, salt fingers would not easily come to his rescue.) The first work to show that salt fingers were strong enough, despite their small scale, to have a real effect on ocean dynamics was that of Stern (1967). He showed that vertical shear in a water mass with lateral gradients could increase the vertical gradients and conditions favoring salt fingers, and that the resulting convection was probably strong enough to account for the observed downward transport of salt under the Mediterranean outflow. Further, he suggested the process could help maintain the tightness of temperature-salinity (T-S) correlations. Again attacking large problems with broad strokes, Stern (1969b) proposed that salt fingers play a major role in dissipating potential energy put into the ocean thermocline by wind driven Ekman pumping.

Contemporaneous to Stern's conjectures, Tait and Howe (1968) suggested the existence of salt fingers under the Mediterranean outflow, and Cooper and Stommel (1968) suggested the same in the main thermocline near Bermuda. This was based on the presence of thermohaline steps which are the vertical compression of the temperature and salinity gradients into thin layers, and the laboratory demonstrations by Turner (1967) that salt fingers would break up into layers if started from linear gradients. These experiments were soon extended and confirmed by Stern and Turner (1969). More definitive proof for the ocean was obtained several years later. The observations of Williams (1974, 1975) using a shadowgraph technique and Magnell (1976) using a towed conductivity probe confirmed structures of the appropriate horizontal length scale for direct double-diffusive convection

under the Mediterranean outflow. Salt fingering seems almost inevitable when a warm salty current flows over cooler fresher water, when there is a net heat flux into the ocean and an increase in salinity (from evaporation) at the surface, and along intrusion boundaries at many water mass fronts.

As is required on energetic grounds, the density contribution of the energy providing salt flux must be greater than that from the stabilizing heat flux to overcome dissipation. The buoyancy flux ratio (or simply, the flux ratio), χ, has been a source of controversy and is of great oceanographic importance. Turner's (1967) first experiments concentrated on this. He obtained a value of $\chi \approx 0.56$ independent of the two layer density ratio. However Linden (1973) estimated a flux ratio of approximately 0.1 for a heat-sugar system, which should be similar. For salt-sugar fingers both Stern and Turner (1969) and Lambert and Demenkow (1972) obtained constant flux ratios of about 0.9. Stern (1976) predicted constant flux ratios of 0.25 and 0.50 for the heat-salt and salt-sugar cases, respectively. However these values did not correspond well to the values in the work that indicated constant flux ratios. More recently Schmitt (1979b), for the heat-salt case, and Griffiths and Ruddick (1980), for the salt-sugar case, have shown that the flux ratio increases with decreasing density ratio, which seems more reasonable, though the effect is not large. Schmitt found that

$$\chi \approx \begin{cases} 0.7, & \Lambda < 2.5 \\ 0.58, & 2.5 < \Lambda < 4 \\ 0.3, & \Lambda > 6 \end{cases} , \qquad \text{for heat-salt fingers,}$$

while Griffiths and Ruddick found a systematic variation of χ: $0.94 \rightarrow 0.88$ as Λ: $1.2 \rightarrow 2$ for salt-sugar fingers. This is as expected since less powerfully driven systems must be more efficient in extracting potential energy from the salt field while doing little work on the temperature field. For Λ close to 1 the flux ratios

are in excellent agreement with what Schmitt (1979a) predicted for the fastest growing modes in the growth phase, however for larger values of Λ this is no longer the case. There is agreement, though, that the unstable component flux goes as $(\beta \Delta S)^{4/3}$, where ΔS is measured between just above and just below the cellular regime. This was first suggested by Turner (1967), and later by Stern (1969) using different scaling arguments. However, this is quite simplistic since the stabilizing temperature gradient is ignored. Schmitt (1983) claims that the flux is proportional to $(\Delta T)^{4/3}$. This only makes sense if the flux ratio is nearly one for all of the experiments performed.

Though knowledge of some of the flux law dependences is sketchy, the relations obtained in the laboratories do allow rough estimates of expected ocean mixing. These have been applied with some success. Lambert and Sturges (1977) found that the vertical flux divergence of a series of salt finger layers could balance the horizontal divergence of salt in part of the Carribean Sea. Vertical flux divergences are inseparably tied to the vertical structure of the convection modes, as will be discussed in detail below. In the North Atlantic Central Water, where there is little turbulent mixing, Schmitt and Evans (1978) have shown that salt fingers can account for the large observed downward salt fluxes driven by surface evaporation. They further suggest that the process might be intermittent because of internal wave strains and shear effects but do not develop these ideas. Schmitt (1981) has advanced a theory based on nonlinear flux laws and an equation of state with nonconstant coefficients, that salt fingers can account for the observed constant density ratio in these same waters.

Some of the fascination of the phenomenon is that salt fingers provide an efficient mechanism for an upgradient buoyancy flux. This is not entirely mysterious,

since the more fundamental quantities of heat and salt are both transported downgradient. To those familiar only with thermal convection, this may seem counterintuitive if one is used to viewing convection as a rapid method of fluxing buoyancy downgradient. Still in direct double-diffusive convection, the motions are buoyantly driven. Though double-diffusion is relatively fast compared to simple diffusive mixing, it is still much slower than convection with a downgradient buoyancy flux. This is why density inversions are almost never observed in the ocean. To realize how extreme the difference in time scales is we note — all double-diffusive experiments have been run-down experiments but it would be extremely difficult to setup a single component density inversion to observe its run-down.

There are many areas of the ocean where there is a net flux of negative buoyancy into the surface. If an efficient means did not exist for the oceans' center of mass to drop, releasing potential energy, without maintaining an overall negative density gradient, many regions of the central and equatorial waters would have positive or near zero mean density gradient. A major consequence of this would be a decrease in oceanic Richardson numbers, Ri, a ratio of static stability to the square of the shear, and far greater turbulence than is observed. Further there would be a far less powerful tendency for oceanic motions to remain quasi-horizontal. So, the details of double-diffusion can affect not only the details of the ocean microstructure, but also motions strongly affected by the overall density gradient. Additionally, though the temperature and salinity steps are nearly density compensating a stepped density structure still results. This greatly affects the propagation and spectrum of internal waves (see Johns and Cross 1970 and Miles 1972).

Not only are salt fingers a mechanism for rapid mixing, but they also have characteristics quite different from other oceanic mixing events. Like other types of mixing, double-diffusion in the long run makes water with distinctly different properties more homogeneous and mean gradients are weakened. However on the smaller layer scales very sharp gradients of temperature and salinity are created. Normal eddy diffusivity models that treat heat and salt inidividually as passive tracers and look only at the net desity contribution in the dynamics are grossly in error in a region of vigorous salt fingering. Further, in addition to simplistic models underestimating the fluxes, eddy models assuming that the relevant length scales are much larger than the scales appropriate for molecular diffusion yield equal mixing coefficients for heat and salt. In reality the mixing coefficients are greatly different even on fairly large length scales, in the presence of double-diffusion. Garrett (1982) has made some suggestions for this problem in mixing at fronts. Since there is reason to believe that in large parts of the ocean double-diffusive convection is the major source of mixing (*e.g.*: Schmitt and Evans 1978), this is a major problem for establishing heat and salt balances in the general circulation.

There are several features of salt fingers, though not all independent ones, that make them important to large scale ocean structure and dynamics. The heat and salt Nusselt numbers, Nu_T and Nu_S, which are the ratio of the transports with convection to purely molecular diffusion, can be several orders of magnitude greater than one. Any resulting turbulence from the convective motions is on scales of centimeters or tens of centimeters. Salt is transported more efficiently than heat because of its *lower* diffusivity while both are transported downgradient. However, buoyancy, or equivalently density, is transported upgradient, so the density gradient can remain negative with a negative buoyancy flux. The overall

temperature and salinity gradients are concentrated into thin layers so that the thermohaline structure is stepped providing a characteristic signature in regions where direct double-diffusion is a dominant process. An explanation of these steps is the major goal of this work.

3. Basic Equations and Linear Theory

The simplest basic equations appropriate for studying double-diffusion are the Boussinesq equations for an incompressible fluid:

$$\frac{d\mathbf{v}}{dt} = -\frac{1}{\rho_o}\nabla p + \nu\nabla^2\mathbf{v} - \frac{\rho}{\rho_o}g\hat{z} \tag{1}$$

$$\nabla \cdot \mathbf{v} = 0 \quad . \tag{2}$$

Density variations are only kept in the buoyancy term. A detailed discussion of the validity of this approximation, along with some history of it, can be found in Gray and Giorgini 1976. For traditional reasons, an equation of state dependent on just temperature and salinity is chosen, which linearized is

$$\rho = \rho_o(1 - \alpha\theta_* + \beta S_*) \quad . \tag{3}$$

Other systems, such as water with salt and sugar which is frequently studied in the laboratory experiments discussed herein, or hydrogen with helium and heat as might be relevant in stellar dynamics, could be similarly considered. This particular sign convention is chosen because it is familiar to oceanographers. It will be maintained even when referring to systems other than heat and salt where both density coefficients would more naturally be taken to be positive. Fourier-Fick type diffusion laws with constant coefficents are assumed, so

$$\begin{aligned} \frac{d\theta_*}{dt} &= K_T\nabla^2\theta_* \\ \frac{dS_*}{dt} &= K_S\nabla^2 S_* \quad , \end{aligned} \tag{4}$$

which ignore frictional heating and assume that both K_T and K_S are constants.

For constant vertical base gradients of temperature, A_o, and salinity, B_o, and periodic lateral boundary conditions, eq. (1) and eq. (4) with eq. (3), become (see Stern 1975 for details),

$$\frac{d\mathbf{v}}{dt} = -\nabla\phi + \nu\nabla^2\mathbf{v} - [(\beta S - \alpha\theta) + (B_o - A_o)z]g\hat{z} \tag{5}$$

$$\frac{d\theta}{dt} + wA_o = K_T\nabla^2\theta$$

$$\frac{dS}{dt} + wB_o = K_S\nabla^2 S \quad . \tag{6}$$

The linear stability analysis has been done in great detail by Baines and Gill (1969) but the scaling and vertical dependence will be treated here in a different manner to allow for an arbitrary vertical wave number. For expansion in normal modes of the form

$$w = \hat{w}e^{i(k_1x+k_2y+mz)+\lambda t} \quad ,$$

where $\hat{w}$ is the initial perturbation amplitude, and with the linearization ($\frac{d}{dt} \rightarrow \frac{\partial}{\partial t}$) of eqs. (5) and (6), the three dimensional stability problem reduces to a two dimensional probem using $k^2 = k_1^2 + k_2^2$. The vorticity equation in x and z, with the nonlinear terms kept, is then

$$\frac{\partial\varsigma}{\partial t} + J(\psi,\varsigma) = \nu\nabla_2^2\varsigma - g\frac{\partial}{\partial x}(\beta S - \alpha\theta) \quad , \tag{7}$$

where the standard definitions of

$$w = \frac{\partial\psi}{\partial x} \quad , \qquad u = -\frac{\partial\psi}{\partial z} \quad , \tag{8}$$

and,

$$\varsigma = \nabla_2^2\psi \tag{9}$$

apply. The choice of eqs. (8) guarantees that the continuity constraint of eq. (2) is satisfied.

The set of nondimensional parameters used to characterize the double-diffusive problem is traditionally taken as having four members consisting of a salt Rayleigh number (sometimes using the thermal diffusivity in the denominator), a thermal Rayleigh number, Prandtl number, and diffusivity ratio (Lewis number), such as was used by Baines and Gill (1969). Unlike in the one component problem, here the Rayleigh numbers and the mean density gradient are decoupled. A net Rayleigh number $R_* = R_T - R_S$ is useful in characterizing the stability boundaries. For the case of very tall regions, in effect the infinite Rayleigh number limit, the problem can be reduced to three parameters. For these we choose the density ratio, Λ, Prandtl number, σ (or its reciprocal, η), and the diffusivity ratio, τ, similar to the choice of Schmitt (1979a). Note that τ is chosen less than one; the reciprocal convention is frequently used with the same choice of letter. To fully specify the problem the fourth parameter would be a nondimensional height based on the length scale chosen. The length and time scales chosen here are slightly different than Schmitt's in that we use the destabilizing rather than the stabilizing gradient. Schmitt's (1983) later use of the Brunt-Väisälä frequency, N, to give a time scale is problematic because this diverges as the density ratio goes to one, as he notes. The limit $\Lambda \to 1$ is still within the parameter range of interest here. That scaling choice, however, does make clear when growth rates are greater than the buoyancy oscillation frequency, which Schmitt uses as a heuristic to decide if fingers are driven strongly enough to be observed in the ocean. No notational distinction will be made between dimensional and nondimensional versions of growth rates, wave numbers, and other quantities, since the choice is made obvious by the context.

When referring to numerical values quantities are dimensionless unless units of measure are given.

If eqs. (6) and (7) are nondimensionalized by a time scale and a length scale

$$T = \left(\frac{K_T g\beta B}{\nu}\right)^{-1/2} \tag{10a}$$

$$L = \left(\frac{g\beta B}{\nu K_T}\right)^{-1/4} \quad , \tag{10b}$$

respectively, with the salinity perturbation scaled by $S = g\beta BL$ and the temperature perturbation scaled by $\Theta = \Lambda_o S$ the resulting equations are

$$\frac{d\varsigma}{dt} = \sigma\nabla_2^2\varsigma - \sigma\frac{\partial}{\partial x}(S - \Lambda_o\theta) \tag{11}$$

$$\frac{d\theta}{dt} + \Lambda_o\frac{\partial\psi}{\partial x} = \nabla_2^2\theta \tag{12a}$$

$$\frac{dS}{dt} + \frac{\partial\psi}{\partial x} = \tau\nabla_2^2 S \quad . \tag{12b}$$

Equation (9) remains unchanged. The solutions of the linearization of eqs. (9), (11), and (12) are, with $\gamma^2 = k^2 + m^2$,

$$\psi = -\frac{i}{k}w \tag{13a}$$

$$S = -\frac{w}{(\lambda + \tau\gamma^2)} \tag{13b}$$

$$\theta = -\frac{\Lambda_o w}{(\lambda + \gamma^2)} \tag{13c}$$

and require the dispersion relation

$$\eta\lambda^3 + (1+\eta+\eta\tau)\lambda^2\gamma^2 + (1+\tau+\eta\tau)\lambda\gamma^4 + \tau\gamma^6 + (\Lambda_o - 1)\frac{k^2}{\gamma^2}\lambda - (1-\Lambda_o\tau)k^2 = 0 \quad , \tag{14a}$$

or in dimensional form grouped slightly differently,

$$(\lambda+\nu\gamma^2)(\lambda+K_T\gamma^2)(\lambda+K_S\gamma^2)+g(\alpha A-\beta B)\frac{k^2}{\gamma^2}\lambda-gK_SK_T\left(\frac{\beta B}{K_S}-\frac{\alpha A}{K_T}\right)k^2=0 \quad . \tag{14b}$$

The choice of the velocity perturbation as the solution normalization is arbitrary; the solutions could have been expressed just as easily in terms of the temperature or salinity perturbation amplitude. These keep an arbitrary vertical wave number, unlike previous treatments of the linear problem. This nondimensionalization will be referred to as intrinsic since it depends only on fluid parameters and their ratios and not on container scales. The Rayleigh number approach, conversely, is essentially extrinsic. A generalization of eq. (14), including shear and rotation, but with the vertical dependence structure hidden by the use of two Rayleigh numbers, as in Baines and Gill 1969, may be found in Worthem *et al.* 1983.

It can be shown that the principle of the *exchange of stabilities* holds for this problem with $A > 0$ and $B > 0$ as it does in the Bénard problem. This implies several things of importance. There is no subcritical instability at finite amplitude and when the instability is reached λ is real. Further $\lambda = 0$ defines the marginal curve with no complex part, so the initial mode is direct. The instability occurs when the release of potential energy from the buoyancy flux can exceed dissipation. Also for slightly supercritical states the modes saturate at a steady finite amplitude. (For a fuller description of the significance of this see Chandrasekhar 1971 secs. 11–14 and appendix I.) A neccessary condition for instability, with $\Lambda > 1$, then is that $\Lambda\tau < 1$. It is also neccessary that the experimental cell have sufficient height since there is a maximum unstable vertical wave number, so

$$h > \pi\left[\frac{4}{27}(1 - \Lambda_o\tau)\right]^{-1/4} \tag{15}$$

for instability with stress free, perfectly conducting boundaries and unbounded horizontal extent. In more traditional form, using Rayleigh numbers, the minimum condition for direct convection, *i.e.*: real positive growth rate, λ, is (Stern 1960)

$$R_* = \frac{gH^4}{\nu}\left(\frac{\beta B}{K_S} - \frac{\alpha A}{K_T}\right) = R_S - R_T > \frac{27}{4}\pi^4 = R_c \quad ,$$

which is a criterion very similar to that obtained for one component convection.

The complete structure of the growth rate dependence on both vertical and horizontal wave numbers for direct double-diffusive convection can be seen in fig. 2. Though the figure is only for the heat-salt case with $\Lambda = 1.5$, the general appearance of the curves is unchanged for other cases if $\Lambda > 1$. A blow up of the region near the maximum growth rate is in fig. 3a. Going from fig. 3a to 3d the increase in maximum growth rate with decreasing density ratio is evident. The fastest growing modes shift to smaller wave number at lower density ratio. Figures 4 and 5 show growth rates as a function of vertical and horizontal wave numbers for the salt-sugar parameter set. The values of the density ratio match some of those chosen for fig. 3 for comparison.

Though most of the terms of the double-diffusive dispersion relation depend on the square magnitude of the total wave number the differences of the dependence on the vertical and horizontal wave numbers have some important consequences. These features are apparent when studying the marginal curves, $\lambda = 0$, in figs. 2 and 4. In horizontal dependence there is a distinct short wavelength cutoff for the instability. Though there is no clear long wavelength cutoff the growth

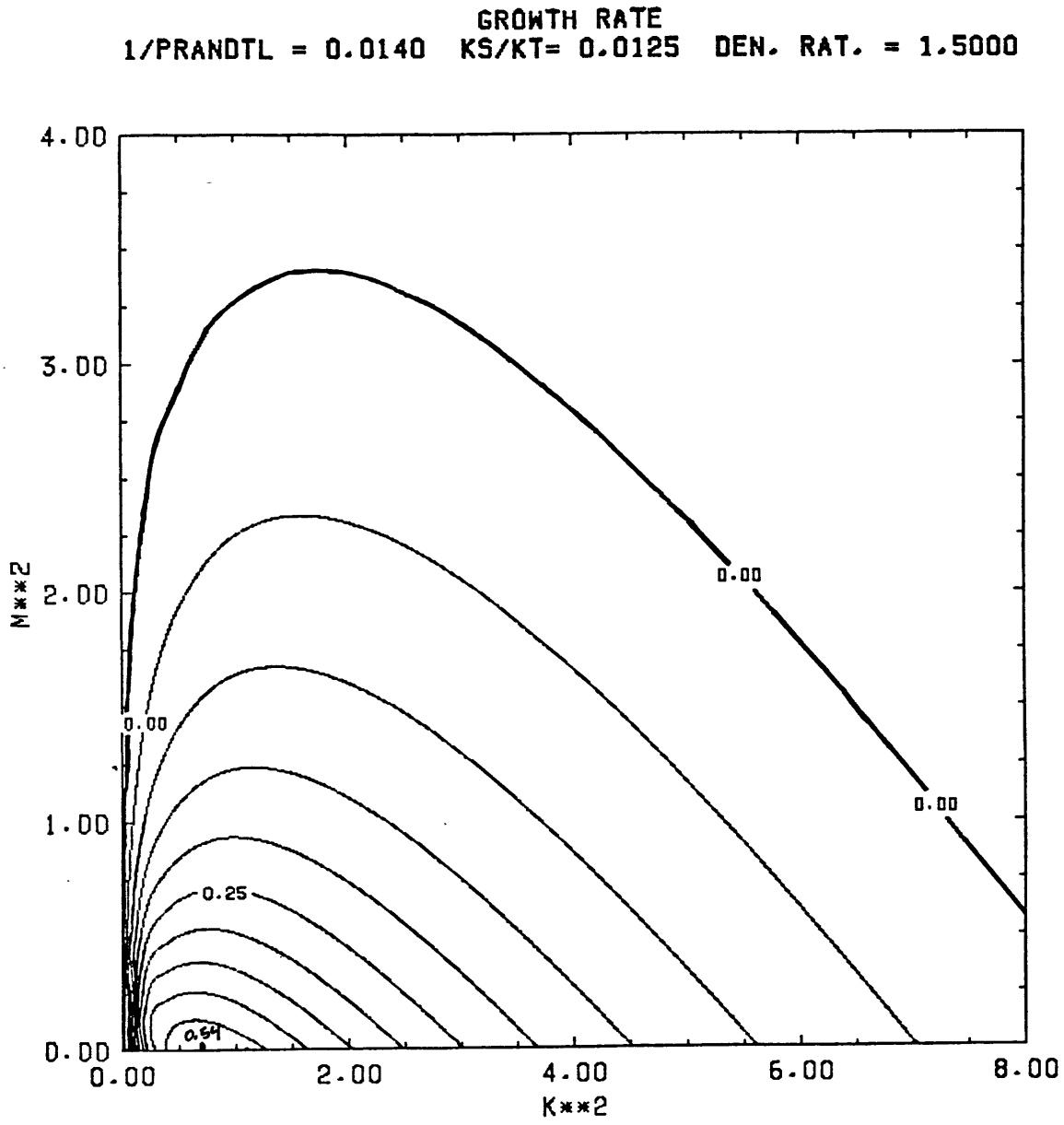

Fig. 2. Growth rate as a function of horizontal and vertical wave numbers for direct double-diffusive convection heat-salt case, density ratio = 1.5. Plots of growth rate are contours of the largest root, λ, of eq. (14a).

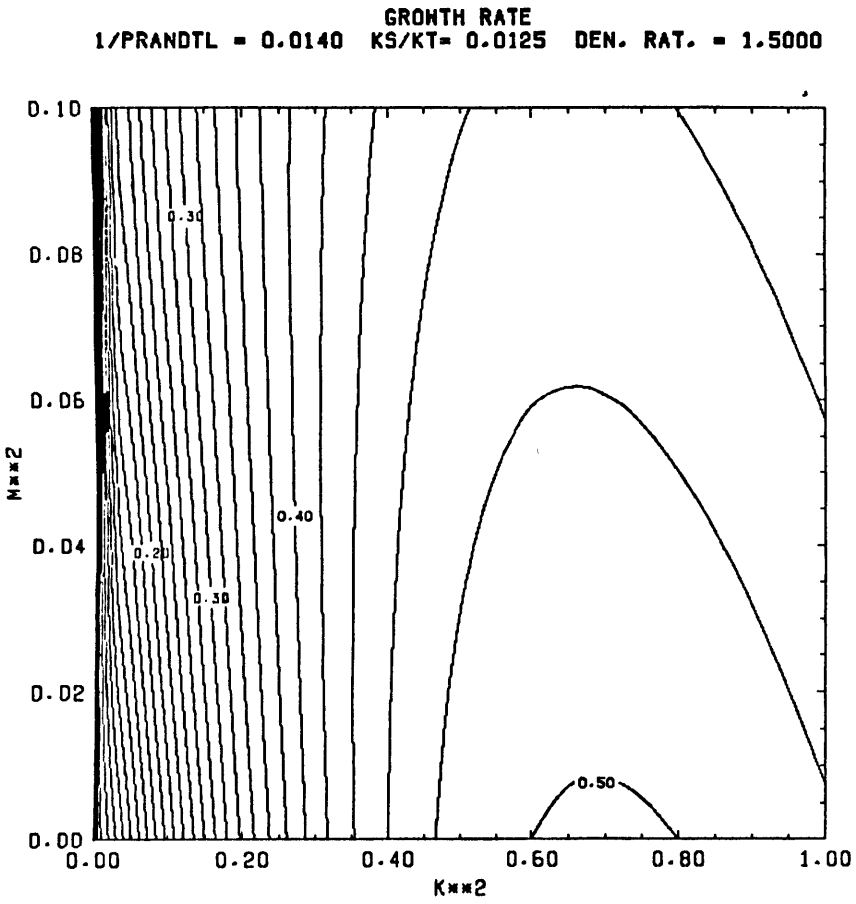

Fig. 3a.

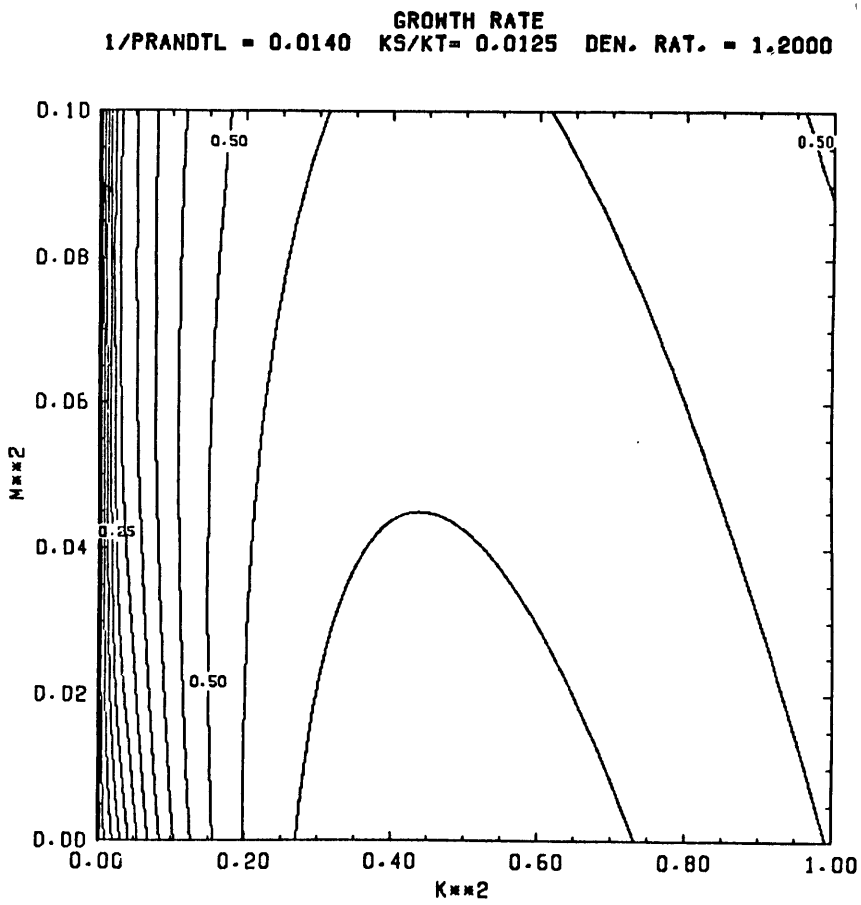

Fig. 3b.

Fig. 3. Growth rate as a function of horizontal and vertical wave numbers for heat-salt case, expanded scale: a) density ratio = 1.5, b) density ratio = 1.2, c) density ratio = 1.1, d) density ratio = 1.0.

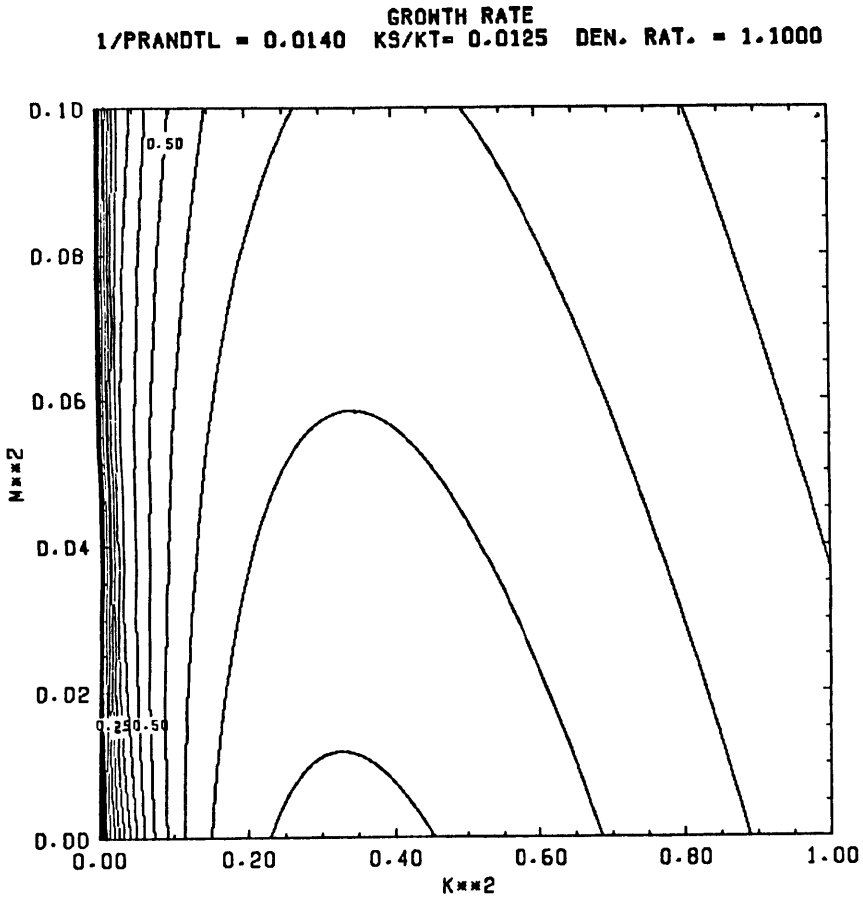
GROWTH RATE
1/PRANDTL = 0.0140 KS/KT= 0.0125 DEN. RAT. = 1.1000
0.10
0.08
0.06
0.04
0.02
0.00
M**2
0.50
0.25
0.50
0.00
0.20
0.40
0.60
0.80
1.00
K**2

Fig. 3c.

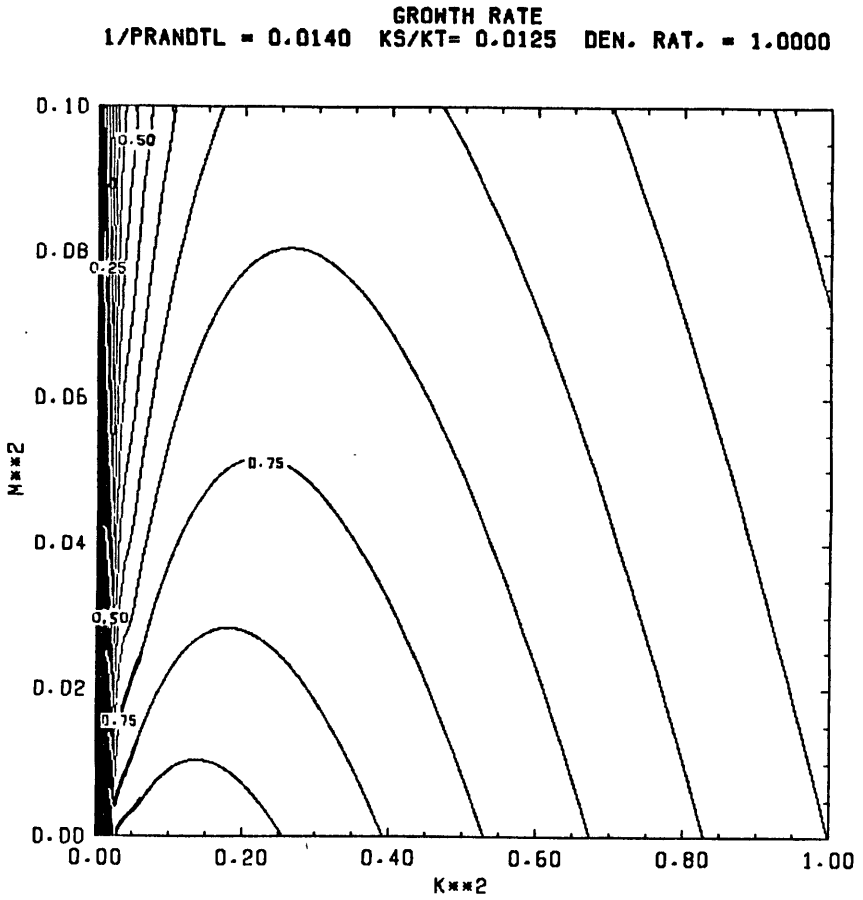
GROWTH RATE
1/PRANDTL = 0.0140 KS/KT= 0.0125 DEN. RAT. = 1.0000
0.10
0.08
0.06
0.04
0.02
0.00
M**2
0.50
0.25
0.75
0.50
0.75
0.00
0.20
0.40
0.60
0.80
1.00
K**2

Fig. 3d.

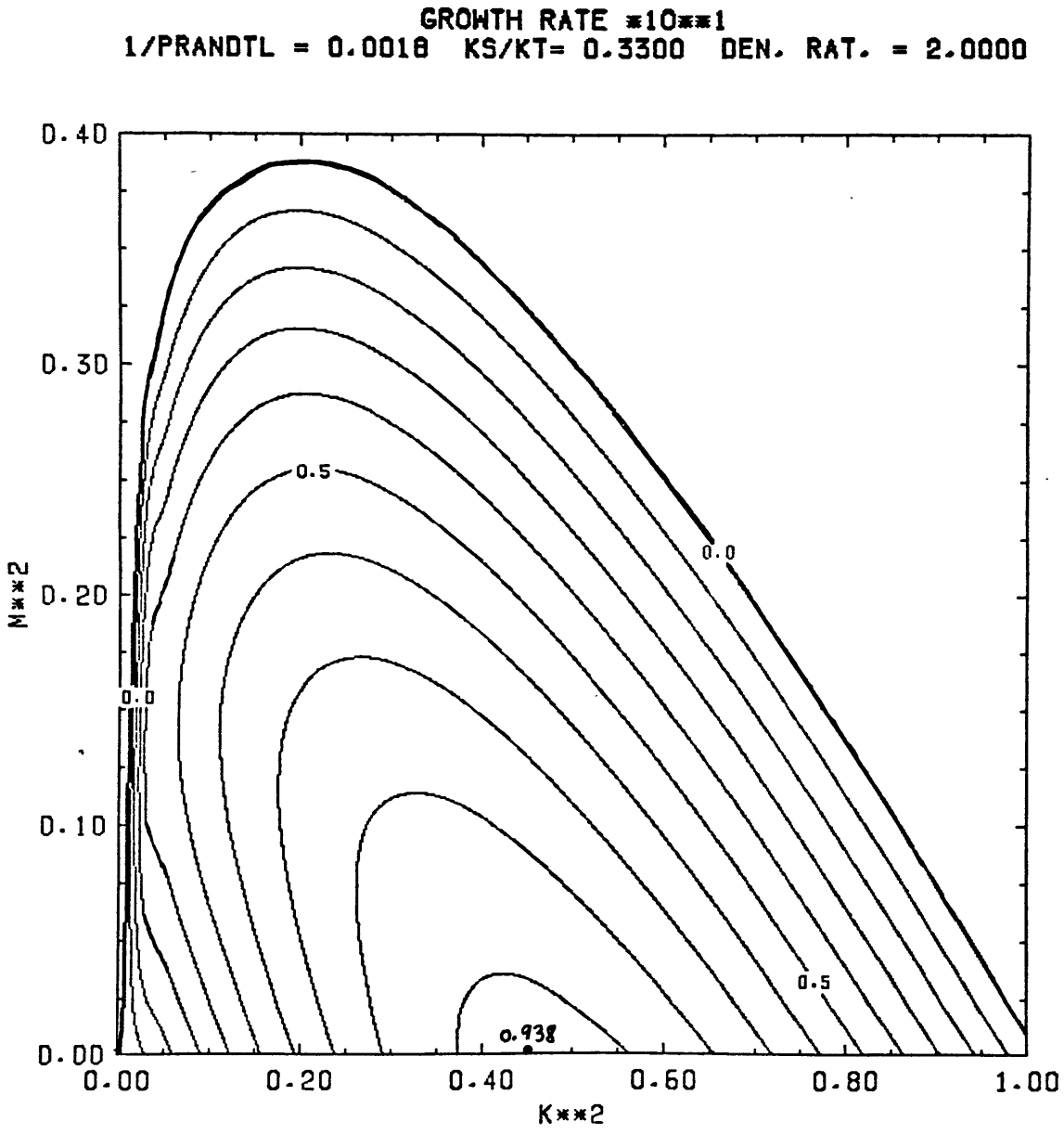

Fig. 4. Growth rate as a function of horizontal and vertical wave numbers for direct double-diffusive convection salt-sugar case, density ratio = 2.0.

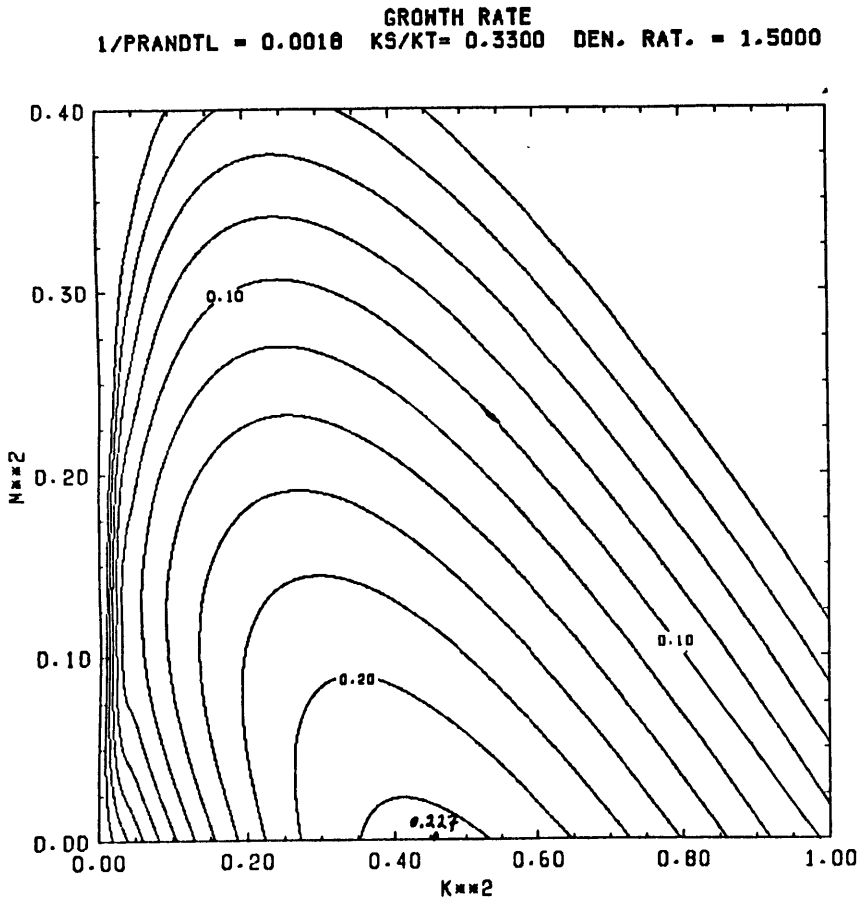

Fig. 5a.

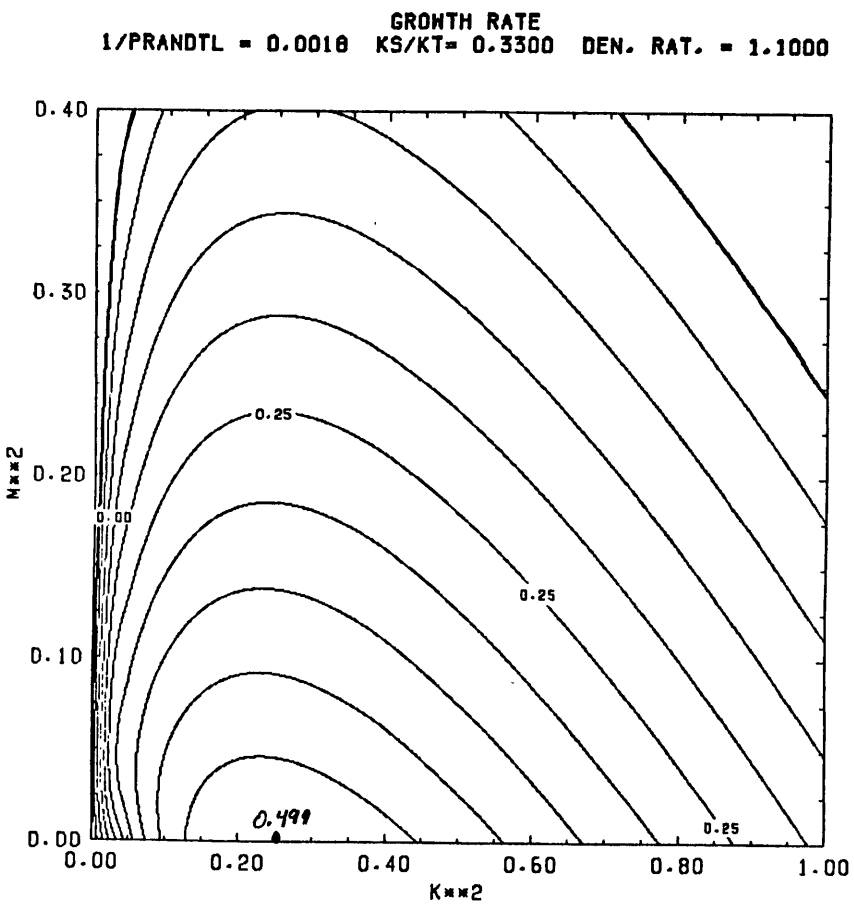

Fig. 5b.

Fig. 5. Growth rate as a function of horizontal and vertical wave numbers for salt-sugar case, expanded scale: a) density ratio = 1.5, b) density ratio = 1.1, c) density ratio = 1.0, d) density ratio = 0.95.

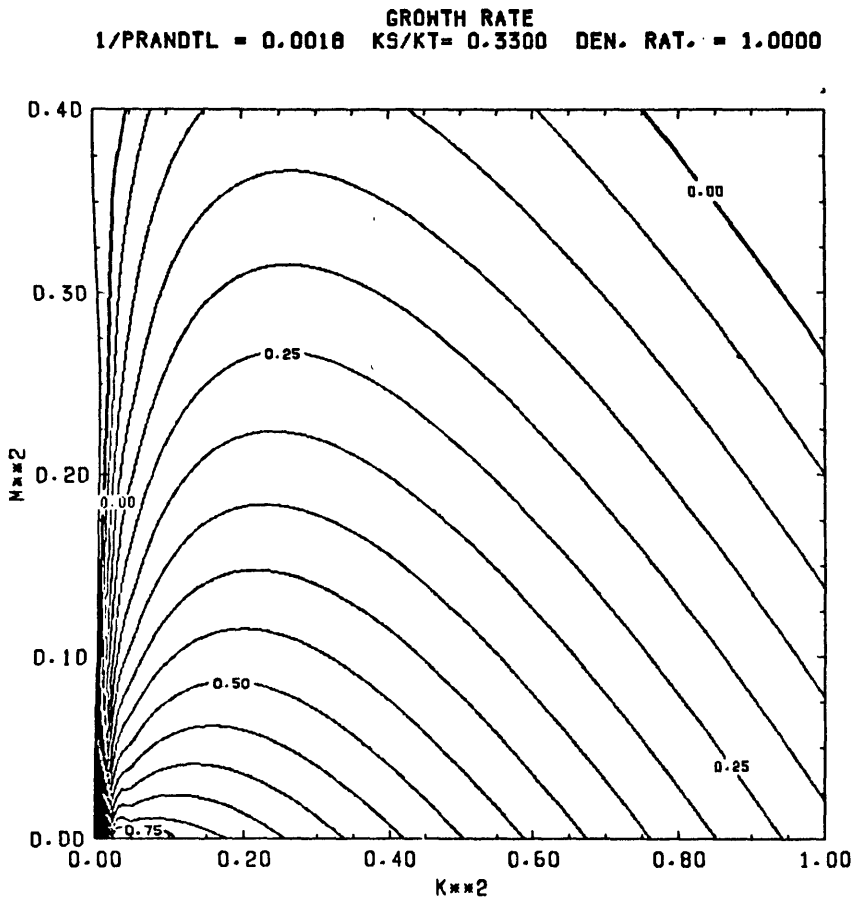
GROWTH RATE
1/PRANDTL = 0.0018 KS/KT= 0.3300 DEN. RAT. = 1.0000
0.40
0.30
0.20
0.10
0.00
M**2
0.00
0.25
0.50
0.75
0.00
0.25
0.00
0.20
0.40
0.60
0.80
1.00
K**2

Fig. 5c.

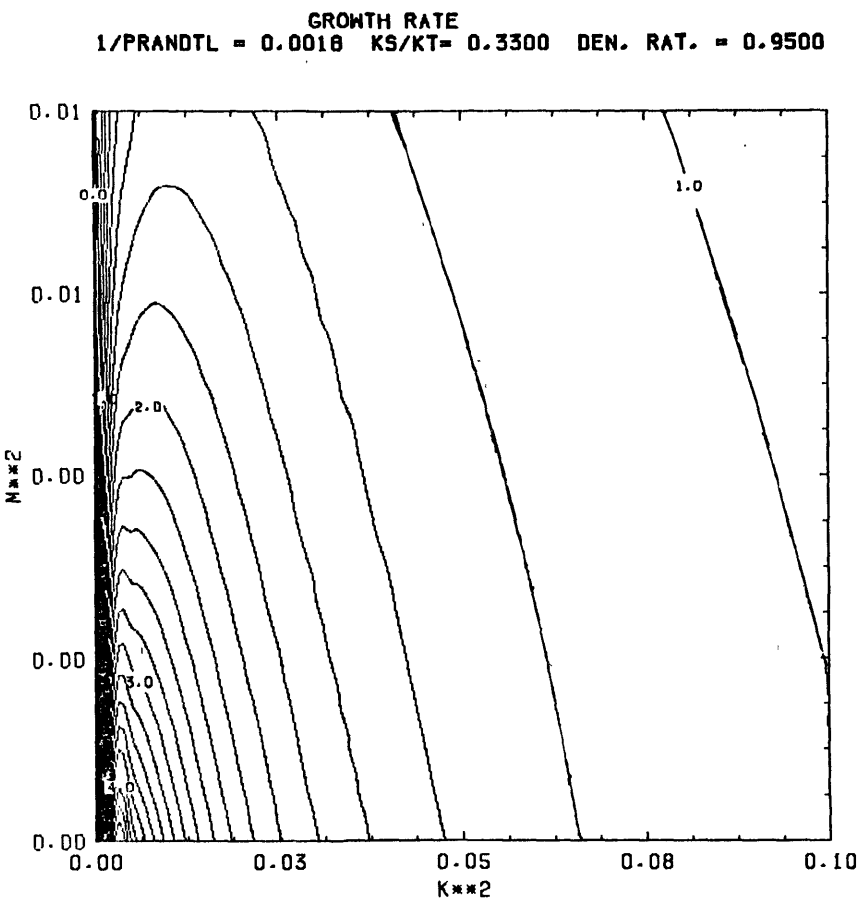
GROWTH RATE
1/PRANDTL = 0.0018 KS/KT= 0.3300 DEN. RAT. = 0.9500
0.01
0.01
0.00
0.00
M**2
0.0
1.0
2.0
3.0
4.0
0.00
0.03
0.05
0.08
0.10
K**2

Fig. 5d.

rate vanishes as the horizontal wave number goes to zero when the vertical wave number is zero. For finite vertical wave number there is a long wave cutoff for maintaining growth. In the vertical dependence there is also a short wave cutoff, however in the long vertical wavelength limit the growth rate remains finite and usually reaches its greatest value. There is little dependence of the fastest growing mode on the vertical mode number as long as the vertical wave number remains small for the first several harmonics. This leads to the preference for tall narrow cells.

Wave numbers will be assumed to be real, though for some types of forced boundary condtions complex wave numbers may be appropriate. The growth rate may in general be complex. For the roots in the direct convection case discussed here it will, however, always be real. Nield (1967) discusses in detail the minimum conditions for direct instability with various other boundary conditions. However, where double-diffusion is observed these conditions are exceeded by at least several orders of magnitude. For constant gradients and a deep layer, a neccessary condition for instability, as was noted, is that the density ratio be less than τ, the ratio of diffusivities; if a system starts with two sharp layers letting diffusion set up gradients at the interface, the criterion for the growth of fingers is (Huppert and Manins 1973)

$$\left(\frac{\alpha\,\Delta T}{\beta\,\Delta S}\right) < \left(\frac{K_T}{K_S}\right)^{3/2},$$

which is a still weaker requirement.

Several other points may be noted from eq. (14). If there is a direct mode, with the gradients of given sign, the other two modes are decaying (Stern 1960).

The fastest growing mode has (Stern 1975)

$$k_f \sim \left[\frac{g(\alpha A - \beta B)}{\nu K_T}\right]^{1/4},$$

or nondimensionally,

$$k_f^2 \sim (1 - \Lambda_o)^{1/2} = \epsilon_o^{1/2} \tag{16a}$$

and

$$\lambda_{\max} \sim 1 - 2\epsilon_o^{1/2} \quad , \tag{16b}$$

in the limits of large Prandtl number, small diffusivity ratio, and tall cells. This clearly does not hold in the limit $\Lambda \to 1$. In this limit the wave number does decrease but does not vanish. The tallest allowed vertical mode for the fastest growing wave number, k_f, is the fastest growing mode. This is proven by Baines and Gill (1969) in their appendix. However if $\frac{\partial \lambda}{\partial k^2} \neq 0$ it is possible to have $\frac{\partial \lambda}{\partial m^2} > 0$ [see eq. (17)]. Examples of this can be seen in the contours of fig.4, amongst others, and explicity in the data of fig. 8b. However this does not occur in the neighborhood of the fastest growing horizontal wave numbers.

A more global picture of the dependences of the maximum growth rate on density ratio was given in Schmitt 1979a and is reproduced here as figs. 6. Note that the time scale is different than the one used here. A very rapid increase in maximum growth rate occurs as Λ approaches one but not until it is fairly close. When $\Lambda < 1$ the increase in maximum growth rate is quite dramatic [see fig. 5d and eq. (20a)]. This is quite clear in both the heat-salt and salt-sugar cases, despite the fact that the heat-salt case is unstable over a much larger range of Λ. Salt finger modes have much smaller growth rates than direct modes with unstable net density

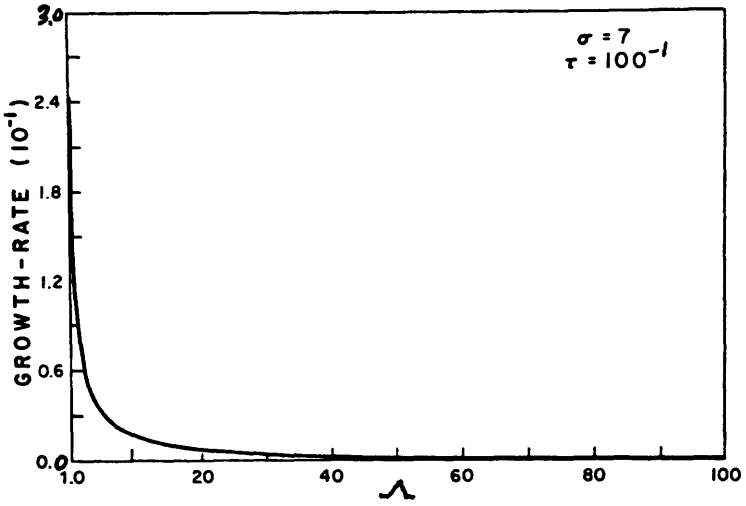

Fig. 6a.

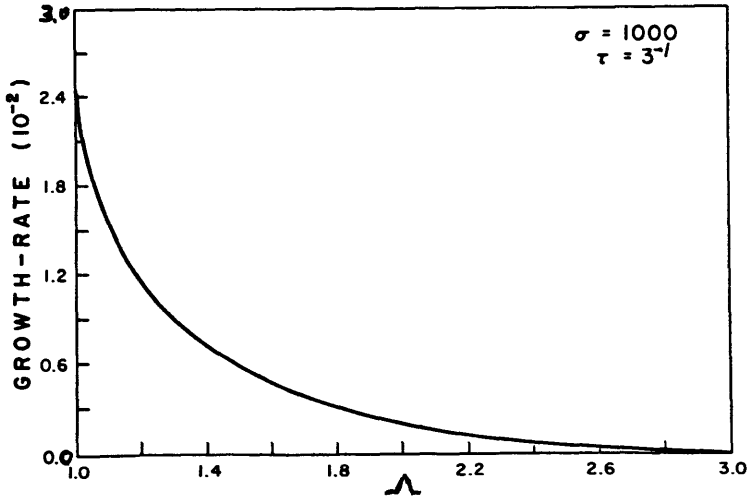

Fig. 6b.

Fig. 6. Maximum growth rate as a function of density ratio: a) heat-salt case, b) salt-sugar case. Time scale is $T = (g\alpha A)^{-1/2}$ allowing direct comparison of growth rates for different systems. Note the order of magnitude difference between the two cases. Figure is adapted from Raymond W. Schmitt, Jr., 1979, *Deep-Sea Res.* **26A**, 23–40, © Pergamon Press.

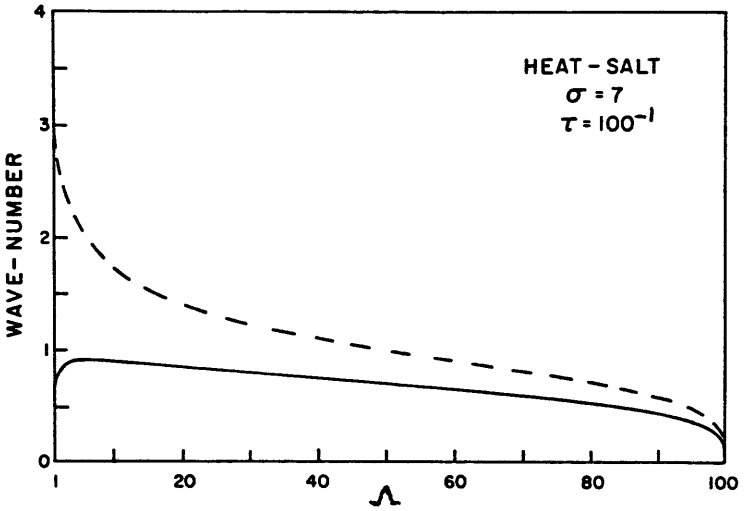

Fig. 7a.

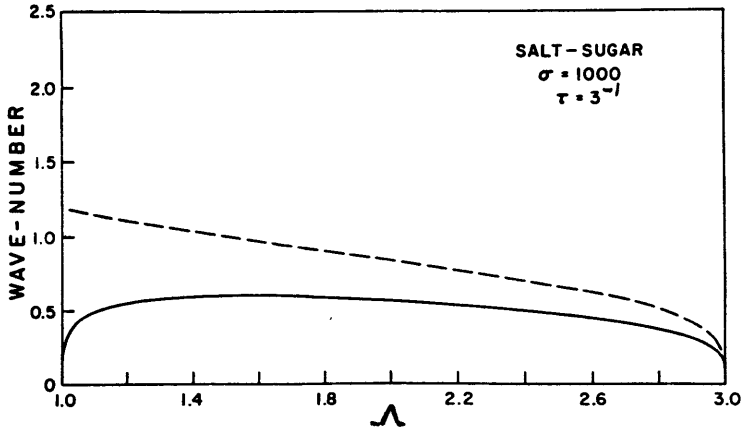

Fig. 7b.

Fig. 7. Wave number of fastest growing mode (solid line) and marginal wave number (dashed line) as a function of density ratio (from Schmitt 1979a): a) heat-salt case, b) salt-sugar case.

gradients. Yet, even if the net density gradient is unstable and the gradients are of the salt finger type, double-diffusion still contributes to the growth. Schmitt also provides a complete picture of the fastest growing wave number for $1 < \Lambda < \tau$, which is reproduced as fig. 7. The decrease in k_f with decreasing Λ as is shown in figs. 2 - 5 does not hold over the entire unstable range, but only as Λ nears one.

From eq. (14a), it can be determined that

$$\frac{\partial \lambda}{\partial m^2} = -\frac{[(1+\eta+\eta\tau)\lambda^2 + 2(1+\tau+\eta\tau)\lambda\gamma^2 + 3\tau\gamma^4 - (\Lambda_o - 1)\lambda k^2/\gamma^4]}{[3\eta\lambda^2 + 2(1+\eta+\eta\tau)\gamma^2\lambda + (1+\tau+\eta\tau)\gamma^4 + (\Lambda_o - 1)k^2/\gamma^2]} \quad (17)$$

(contoured in figs. 8 and 9) and that

$$\frac{\partial \lambda}{\partial k^2} =$$
$$-\frac{[(1+\eta+\eta\tau)\lambda^2 + 2(1+\tau+\eta\tau)\lambda\gamma^2 + 3\tau\gamma^4 + (\Lambda_o - 1)\lambda m^2/\gamma^4 - (1-\Lambda_o\tau)]}{[3\eta\lambda^2 + 2(1+\eta+\eta\tau)\gamma^2\lambda + (1+\tau+\eta\tau)\gamma^4 + (\Lambda_o - 1)k^2/\gamma^2]}$$
$$= \frac{\partial \lambda}{\partial m^2} - \frac{[(\Lambda_o - 1)\lambda/\gamma^2 - (1-\Lambda_o\tau)]}{[3\eta\lambda^2 + 2(1+\eta+\eta\tau)\gamma^2\lambda + (1+\tau+\eta\tau)\gamma^4 + (\Lambda_o - 1)k^2/\gamma^2]} . \quad (18)$$

When taking $\frac{\partial}{\partial k^2}$, m^2 is held constant, while when taking $\frac{\partial}{\partial m^2}$, k^2 is held constant; hence γ^2 is not held constant when taking the partial derivatives. Equations (17), (18), along with the dispersion relation, eq. (14), imply that

$$\frac{\partial \lambda}{\partial k^2} - \frac{\partial \lambda}{\partial m^2} > 0 \quad (19)$$

for growing modes. The inequality holds strictly for $\lambda > 0$, strictly. Note that eq. (19) proves that there is no simultaneous solution to eqs. (14), (17), and (18), for $\frac{\partial \lambda}{\partial k^2} = \frac{\partial \lambda}{\partial m^2} = 0$. This indicates that the fastest growing mode has either m^2 or $k^2 \to \pm\infty$. In this case the sought extremum in λ is at the limit $m^2 \to -\infty$. However, complex or imaginary wave numbers can only match unusually forced boundaries and will not be considered further.

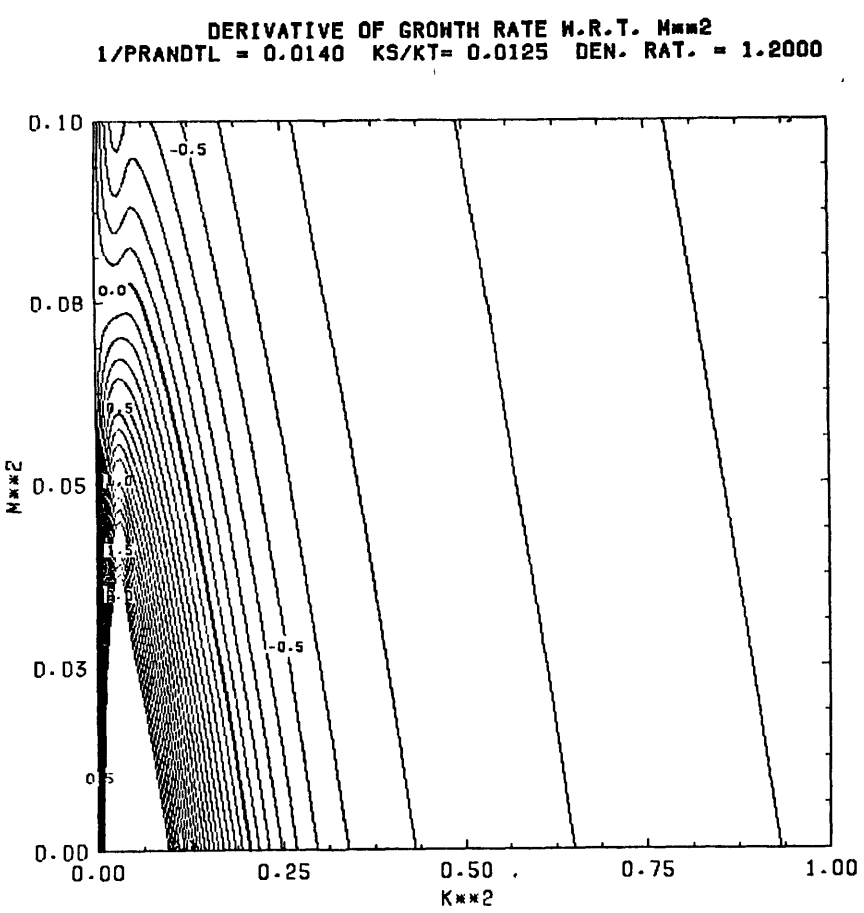

Fig. 8a.

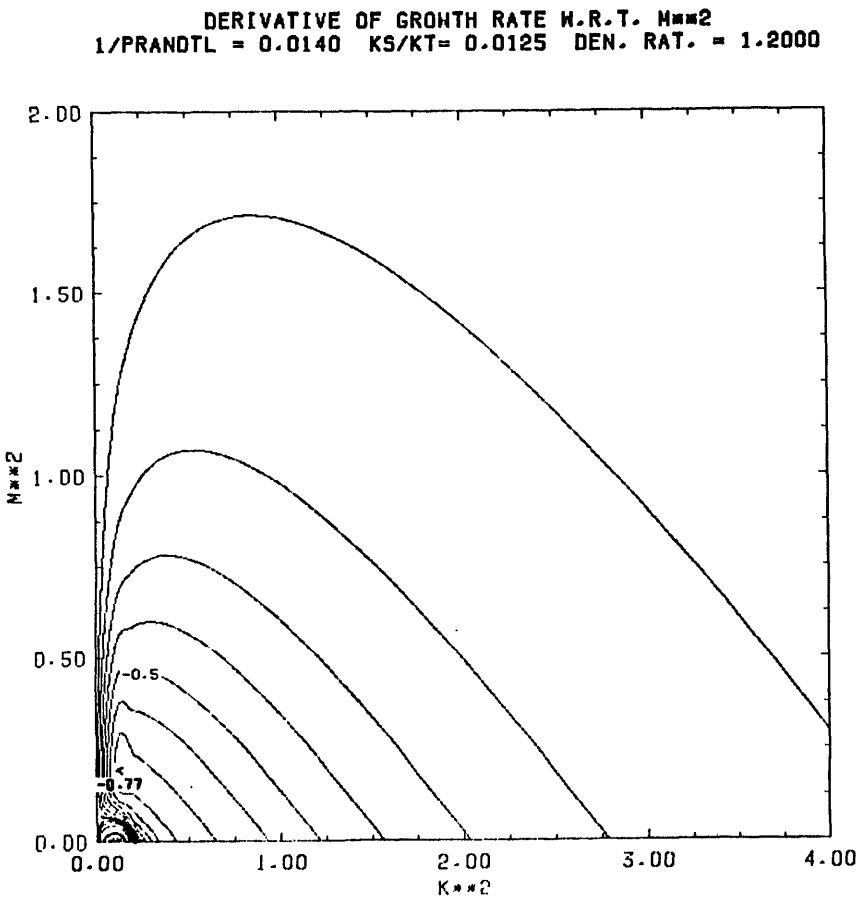

Fig. 8b.

Fig. 8. Derivative of growth rate with respect to square of vertical wave number as a function of horizontal and vertical wave numbers: a) heat-salt case, density ratio = 1.2, b) heat-salt case, density ratio = 1.2, detail. Plots are contours of the values of eq. (17) after choosing the largest root from eq. (14a).

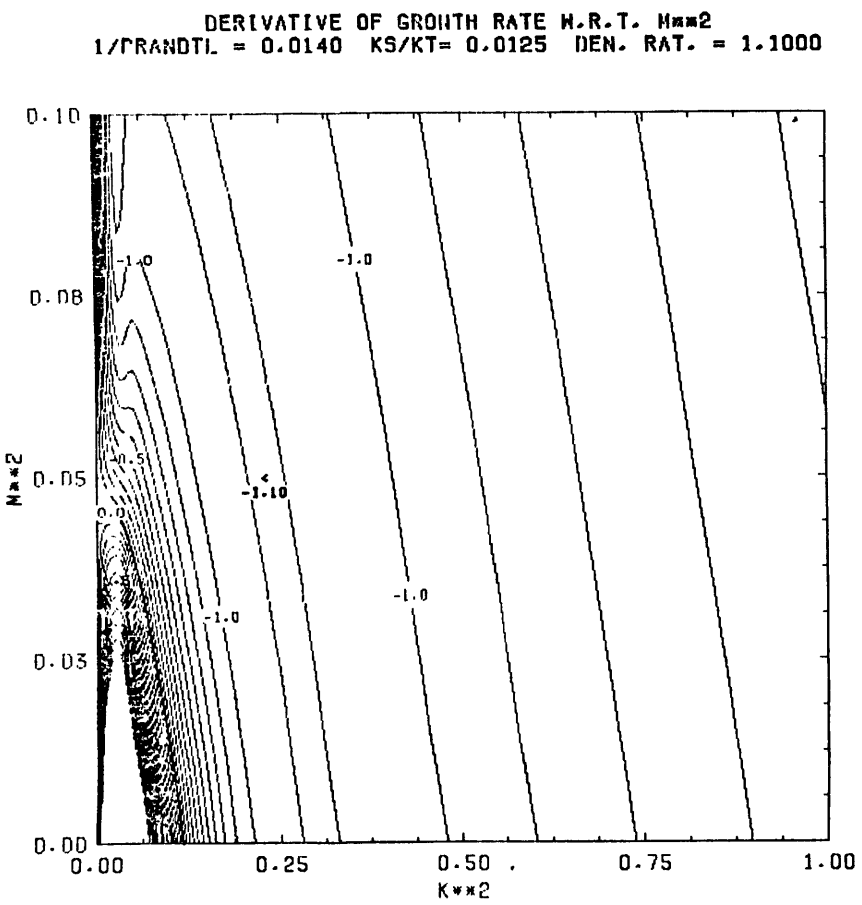

Fig. 9a.

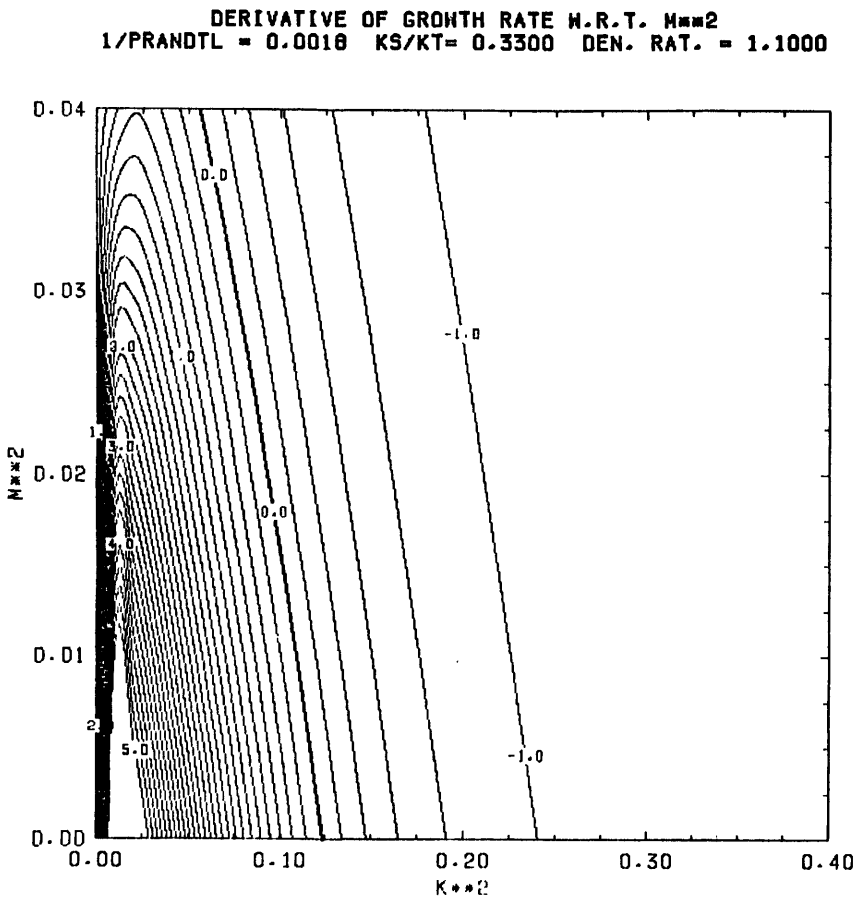

Fig. 9b.

Fig. 9. Derivative of growth rate with respect to square of vertical wave number as a function of horizontal and vertical wave numbers: a) heat-salt case, density ratio = 1.1, b) salt-sugar case, density ratio = 1.1.

When $\Lambda = 1$ the point ($k^2 = 0$, $m^2 = 0$) becomes singular, and remains so for $\Lambda < 1$. For $\Lambda = 1$ the derivative $\frac{\partial \lambda}{\partial k^2}$ blows up at the wave number origin (see fig. 5c). For $\Lambda < 1$ (see fig. 5d.),

$$\lim_{k^2 \to 0} (\lim_{m^2 \to 0}) \lambda = \left(\frac{1 - \Lambda}{\eta} \right)^{1/2} \tag{20a}$$

though

$$\lim_{m^2 \to 0} (\lim_{k^2 \to 0}) \lambda = 0 \quad . \tag{20b}$$

A similar result will be found below for one component convection. There are other roots to be found but care must be taken to follow those are are smoothly connected to the greatest root. The singularity is not simply the result of a mistaken switching of branches. Once a region is created with a density ratio less than one the wave number selection resembles selection in the one component case. The same behavior also occurs in the heat-salt case but is not shown.

For the small values of m^2 that will be of interest here it is useful to have the derivative, $\frac{\partial \lambda}{\partial m^2}$, which is shown in figs. 8 with two scales for the heat-salt case. This derivative for the heat-salt case and the salt-sugar case, both with $\Lambda = 1.1$ is contoured in figs. (9). Using a first order Taylor expansion $\lambda(k^2, m^2)$ can then be obtained to excellent accuracy as

$$\lambda(k^2, m^2) = \lambda(k^2, 0) + m^2 \left. \frac{\partial \lambda}{\partial m^2} \right|_{m^2 = 0} \quad . \tag{21}$$

Near k_f, $\frac{\partial \lambda}{\partial m^2} < 0$, near its largest absolute value, and of order 1. One can see that for some horizontal wave numbers less than the fastest growing one (longer wavelengths), the growth rate increases slightly with increasing vertical wave number. More importantly there is little change in growth rate going from

$m = 0$ to some small finite m since the derivative is of order one. In particular, based on experiments and oceanic observation, we are usually interested in aspect ratios of 10 or more. For example, with $\Lambda = 1.2$ in the heat-salt case the fastest growing modes have nondimensional $k_f^2 \approx 0.42$, and $\lambda_f = 0.63$ so the relevant $m^2 < 0.004$ (see fig. 3b). The derivative $\frac{\partial \lambda}{\partial m^2} \approx -0.72$ (see fig. 8b). Using the approximation of eq. (21) the decrease in λ is less than 0.003 or less than 0.5%. For a 5% decrease down to $\lambda = 0.60$ would require an aspect ratio as small as 3.

The growth rate dependence on wave number for one component convection has some significant qualitative differences. This will be useful in the discussion of why two component convection layers but one component does not. To get the dispersion relation for single component convection quickly, the limits of $\alpha A \to 0$ and $K_T \to 0$ or ∞ (both will work) are taken in eq. (14b), and give

$$\lambda^2 + (\nu + K_S)\lambda\gamma^2 + \nu K_S \gamma^4 - g\beta B \frac{k^2}{\gamma^2} = 0 \quad .$$

The same result is of course obtained from the normal linearization of eqs. (9) – (11), setting $\Lambda = 0$, and with only one diffussion law, eq. (12b). This must be nondimensionalized slightly differently because of the mixing of temperature and salinty parameters in the double-diffusive nondimensionalization previously chosen. Choosing $L = [g\beta B/(K_S \nu)]^{-1/4}$ and $T = (g\beta B K_S/\nu)^{-1/2}$ the nondimensional dispersion relation is

$$\mu\lambda^2 + (1 + \mu)\lambda\gamma^2 + \gamma^4 - \frac{k^2}{\gamma^2} = 0 \qquad (22)$$

where $\mu = K_S/\nu$. Plots of growth rate as a function of vertical and horizontal wave numbers for Prandtl numbers appropriate for heat convection and salt convection are found in figs. (10). By taking derivates of eq. (20) the results

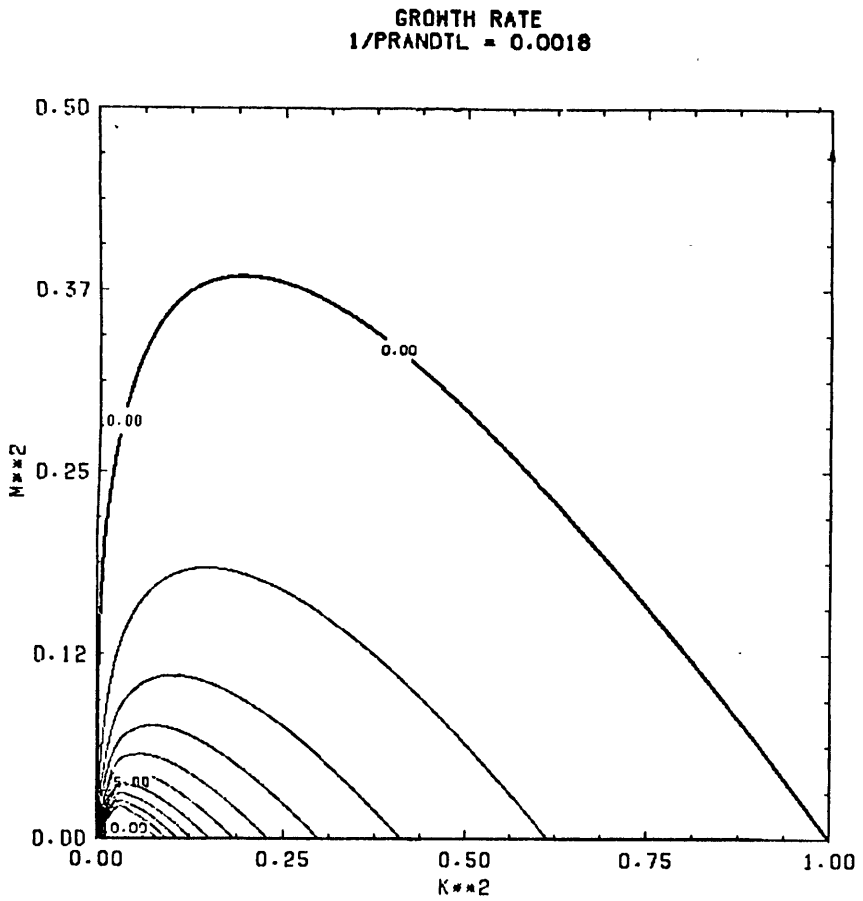

Fig. 10a.

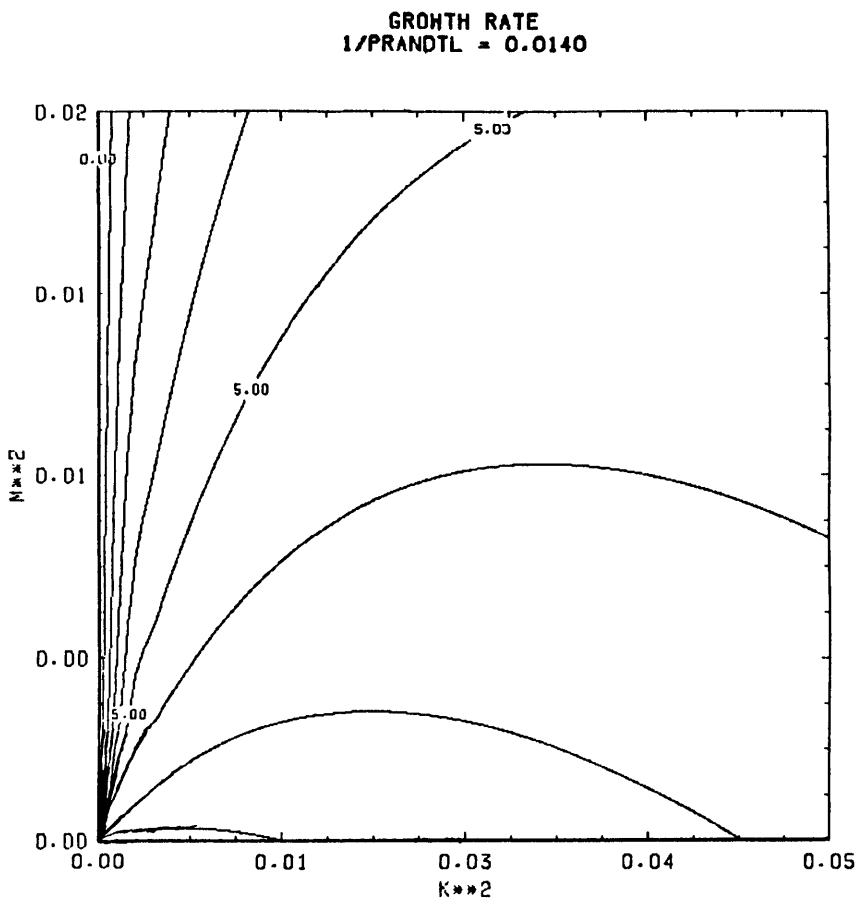

Fig. 10b.

Fig. 10. Growth rate as a function of horizontal and vertical wave numbers for one component convection: a) salt case, b) thermal case. Plots are contours of the largest root, λ, of eq. (22).

$$\frac{\partial \lambda}{\partial k^2} = -\frac{[\lambda(1+\mu) + 2\gamma^2 - m^2/\gamma^2]}{2\mu\lambda + (1+\mu)\gamma^4} \tag{23a}$$

$$\frac{\partial \lambda}{\partial m^2} = -\frac{[\lambda(1+\mu) + 2\gamma^2 + k^2/\gamma^2]}{2\mu\lambda + (1+\mu)\gamma^4} \tag{23b}$$

$$= \frac{\partial \lambda}{\partial k^2} - \frac{1}{[2\mu\lambda\gamma^2 + (1+\mu)\gamma^4]}$$

are obtained. For specified m^2 the maximum of λ occurs at $\frac{\partial \lambda}{\partial k^2} = 0$. Clearly setting $\frac{\partial \lambda}{\partial k^2} = 0$ in eq. (23b) implies $\frac{\partial \lambda}{\partial m^2} < 0$, so as in the two component case the fastest growing modes are as tall as allowed by the boundaries. When $m^2 = 0$, $\frac{\partial \lambda}{\partial k^2} < 0$, the importance of which will soon be apparent.

It is worth noting several differences between the dispersion relation of direct double-diffusion and direct single component convection. For single component convection the point $m^2 = 0$, $k^2 = 0$ is always singular. This can be easily seen since

$$\lim_{k^2 \to 0} \left(\lim_{m^2 \to 0} \lambda \right) = \mu^{-1/2} = \sigma_S^{1/2} \tag{24a}$$

$$\lim_{m^2 \to 0} \left(\lim_{k^2 \to 0} \lambda \right) = 0 \quad . \tag{24b}$$

Further the maximum growth rate possible is obtained by the limits of eq. (24a). All of the contours from 0 to $\mu^{-1/2}$ converge at ($k^2 = 0$, $m^2 = 0$). There is a very strong dependence of k^2 on m^2 for the fastest growing mode unlike the double-diffusive case. Hence the horizontal scale shrinks as the vertical scale is diminished if one follows the fastest growing modes. It still holds that $\frac{\partial \lambda}{\partial m^2} < 0$ for all k^2 of interest and $\frac{\partial k^2}{\partial m^2}\Big|_{\frac{\partial \lambda}{\partial k^2}=0} < 0$, so that the tallest allowed mode is always the fastest growing even for extremely small lateral constraints. This contrasts with the double-diffusive case where $\frac{\partial k^2}{\partial m^2}\Big|_{\frac{\partial \lambda}{\partial k^2}=0}$ can be of either sign depending on

the position in parameter space, but in any case is not particularly large. The behavior of the double-diffusive dispersion relation develops a resemblance to the one component curve as Λ goes to 1 from above by compressing the growth rate contours at small k^2, however the singular point does not occur until $\Lambda < 1$. As was shown in the double-diffusive case, the intrinsic nondimensionalization chosen leads to finite growth rate even with infinite extent for single component convection. This is as it must be since regardless of cell height the gravitational acceleration is finite. The instability condition is then again a minimum required height in terms of the intrinsic length scale:

$$h > \left(\frac{27}{4}\right)^{1/4} \pi \quad . \tag{25}$$

There is a singularity in the maximum growth rate as the Prandtl number goes to inifinity, but this is compensated for by the change in time scale. It is easy to create an intrinsic nondimensionalization which does not have this problem, though it was not chosen here.

The minimum criteria for instability in one component convection, eq. (25), looks almost identical to eq. (15) except that in the one component case $\Lambda = 0$. Using Rayleigh numbers this is equivalent to $R_T = 0$. For the first harmonic to be unstable in either the one or two component case $R_* > 2^4 R_c$; for the second harmonic the requirement is that $R_* > 3^4 R_c$; and so on for higher modes. However once the system has $R_* > R_c$ there is a spectrum of horizontal wave numbers which are unstable. For complex vertical structures to develop it is neccessary for vertical harmonics to be unstable and not just driven by the fundamental through the nonlinear terms of the governing equations. This will be quite important in the analysis of harmonic instabilities.

The experiments examined here all have Prandtl number much greater than one. We are aware of no double-diffusive experiments that have been performed with $\sigma < 1$ and no theory has emphasized such a regime. The roads to chaos are different for the small and large Prandtl number cases (Busse 1981). (The use of the term Prandtl number is not standard if heat is not the convecting component, but is unambiguous.) In the small Prandtl number domains chaotic behavior occurs at supercritical Rayleigh numbers less than $16R_c$, but this is from the interaction of horizontal disturbances (see Golub *et al.* 1975 and their references). Turbulent convection can result form the interactions of many different horizontal modes but the structure is different than what has traditionally been called high Rayleigh number convection. At large Prandtl number the gradients are forced into thin boundary layers as the Rayleigh number is increased and diffusion becomes unimportant in the interior. Time dependent behavior can also occur with R less than $16R_c$ from the interaction of modes of different horizontal wave number but the gradient does not compress to the ends with a nearly isothermal interior. Simplistically, high Rayleigh number convection at large Prandtl number has transport by bubbles and blobs, while low Prandtl numbers lead to convoluted narrow filaments (see Grötzbach 1982). The double-diffusive run-down experiments have large Prandtl number but the convection is diffusion limited in the interior like small Prandtl number convection.

The linear flux ratio can be easily obtained at this point and is simply the ratio of the coefficients of the heat and salt perturbations from eq. (13). So, in nondimensional form

$$\chi = \Lambda \frac{(\lambda + \tau\gamma^2)}{(\lambda + \gamma^2)} \quad . \tag{26}$$

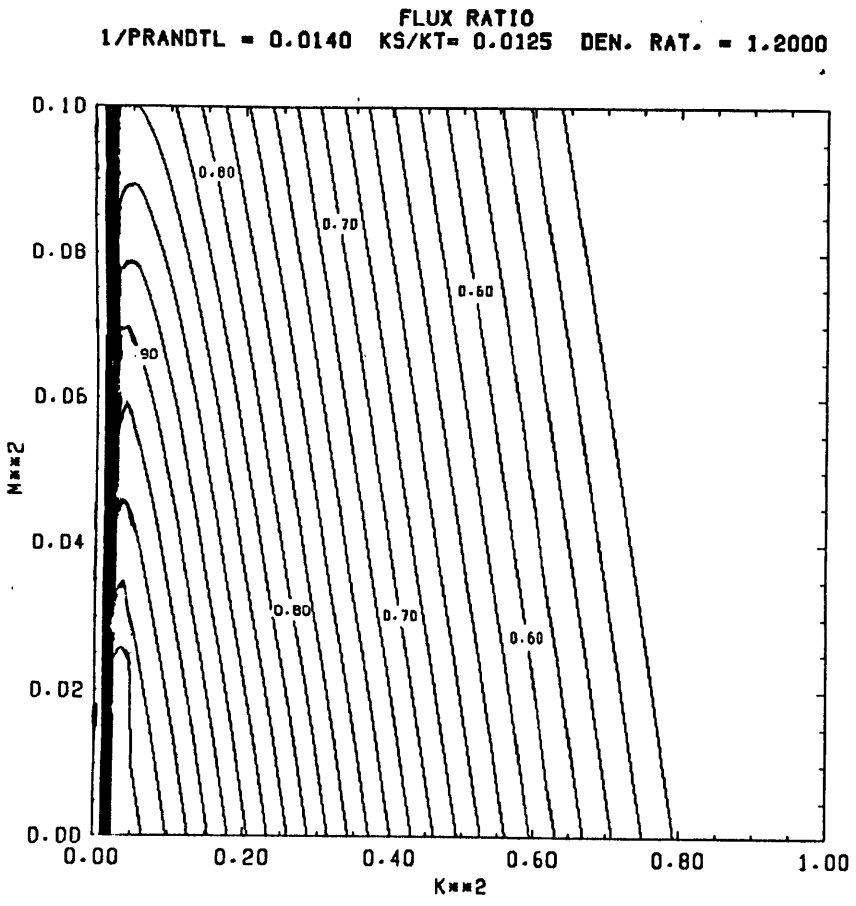

Fig. 11a.

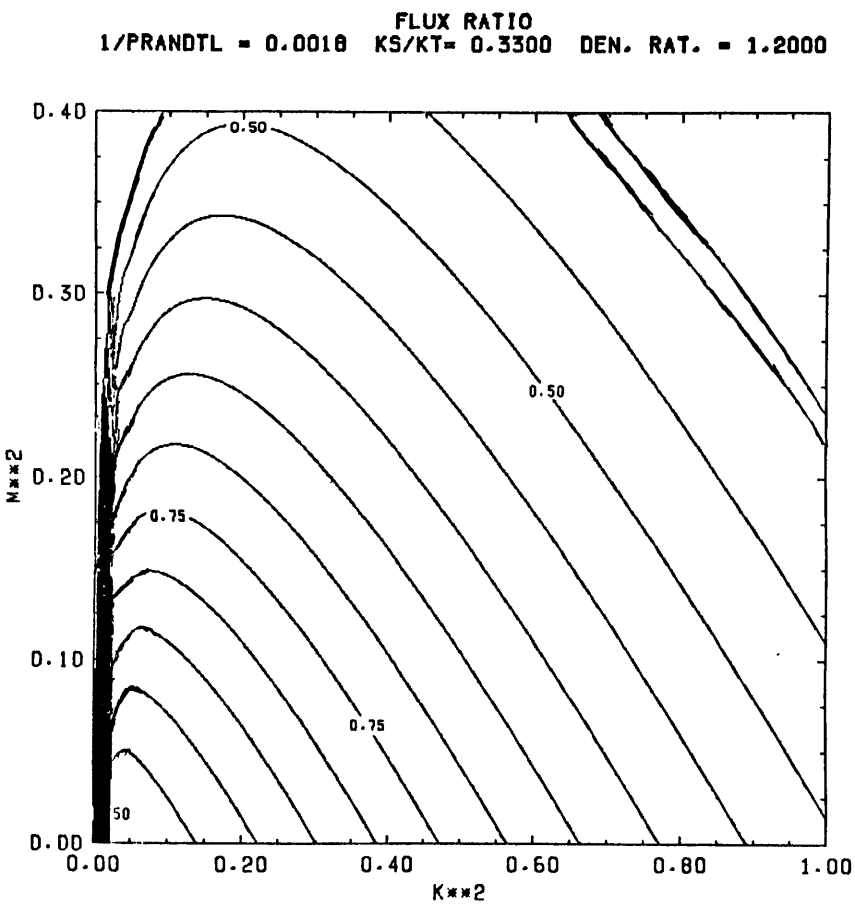

Fig. 11b.

Fig. 11. Buoyancy flux ratio as a function of horizontal and vertical wave numbers: a) heat-salt case, density ratio = 1.2, b) salt-sugar case, density ratio = 1.2. Approximation for growing linear modes from eq. (26).

Contours of χ from eq. (26) are in figs. (11). The flux ratio is a measure of the efficiency of the mode in taking potential energy taken from the reservoirs or from outside of the system. For unstable modes and the gradients as for salt fingers it is bounded by zero and one. As the ratio approaches the value one, most of the potential energy received from the positive buoyancy (negative density) flux of salt is used to drive a negative buoyancy flux of heat. Smaller values indicate more efficient modes in the sense that less energy is used to support a temperature based negative buoyancy flux and hence a larger fraction of the available potential energy can be used to support motion. The net buoyancy flux must of course remain positive. However, when an eventual steady or quasi-steady state is reached the input of potential energy must be balanced by viscous dissipation. When finite amplitude is reached the mode with the greatest buoyancy transport is far in wave number space from the modes that have the smallest flux ratios because of the strong dependence of equilibrium amplitude on wave number.

The dependence of the flux ratio on density ratio and wave numbers is not obvious from eq. (26) alone because of the implied dependences from the dispersion relation. Schmitt (1979a) finds that the observed flux ratios and salt finger widths in the run-down experiments that have been performed correspond well to those calculated for the fastest growing modes. Figure 12, taken from Schmitt's paper, shows χ for $k = k_f$, $m = 0$ as a function of Λ. The preference is not for the modes most efficient at extracting potential energy from the reservoirs, which have smaller flux ratios and larger wave numbers than the faster modes. The small range of χ at k_f in the range of Λ for which run-down experiments have been performed does well to explain why so many investigators concluded that χ was nearly constant. However, since the amplitudes are not known it is still possible that the mode with

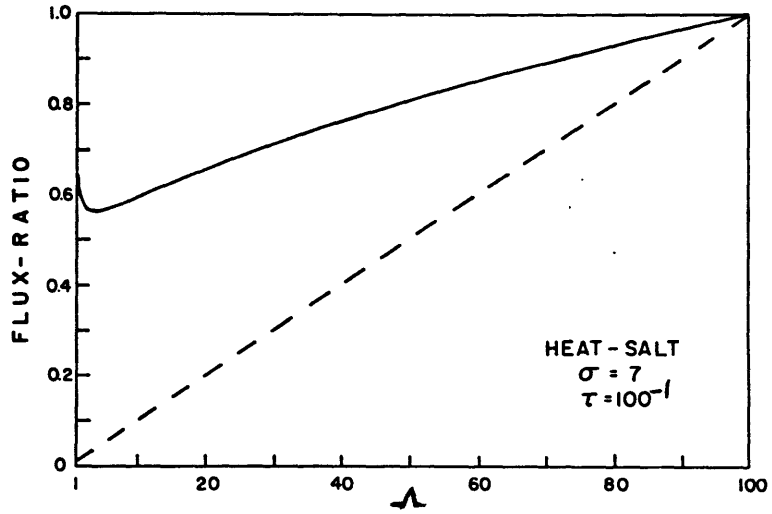

Fig. 12a.

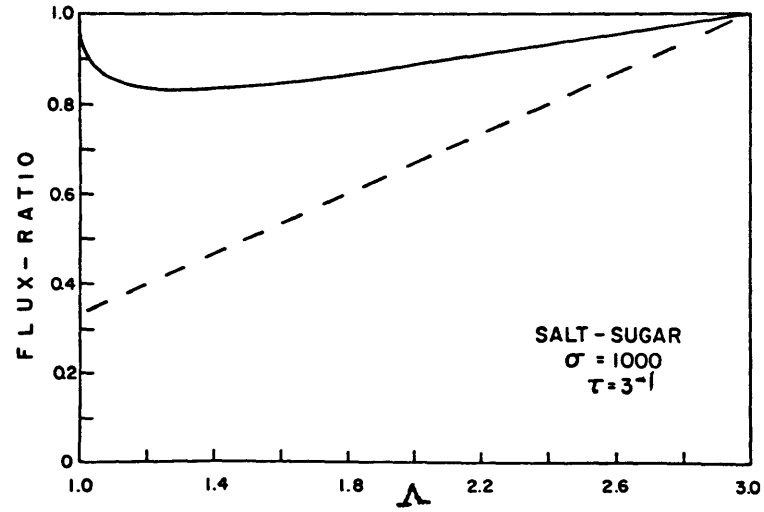

Fig. 12b.

Fig. 12. Flux ratio for fastest growing mode (solid line) and flux ratio for marginal mode (dashed line) as a function of density ratio (from Schmitt 1979a): a) heat-salt case, b) salt-sugar case.

the greatest buoyancy flux is selected, similar to the selection condition that Malkus (1954) recommended for examing thermal convection. Boundary conditions in the ocean may be more realistically viewed as having the fluxes given and hence the flux ratio specified. These linear results and laboratory run-down experiments may not then be as relevant as one might hope them to be.

Nondimensional results must be compared with caution. For example, if for a given convection problem one doubles the cell height but keeps the temperature difference constant the Rayleigh number increases by a factor of eight. The Nusselt number, the nondimensional flux, also goes up; however, the actual flux decreases. Scales must be accounted for carefully. Similar problems were noted above in discussing the behavior of growth rates in the limit of infinte Rayleigh number for the extrinsic scaling, and in the limit of infinite Prandtl number in the chosen intrinsic scaling of the Bénard problem. In retrospect a different intrinsic nondimensionalization that has not been used at all is probably best for comparing results with different Prandtl number and diffusivity ratios. This choice would use $L = [g\beta B/(K_S\nu)]^{-1/4}$, $T = [g\beta B]^{-1/2}$, and keep the same scaling for the temperature and salinity fields. This is similar to Schmitt's choice (1979a, 1983) but using the destabilizing component gradient and diffusivity. The limit of no destabilizing gradient, $A \to 0$, leads to no instability before vanishing. However if the scale using the stabilizing gradient is chosen it is impossible to take the limit of one component convection. Even looking at the effect of a small stabilizing component gradient yields strangely scaled results. After having made our choice in this work it was not worth the effort to rescale all of the data.

It is quite fortuitous that the values for the wavelength and flux ratio for the fastest growing modes correspond so well to those observed. Nonlinear effects at

finite amplitude frequently lead to far more confusing scenarios. If there exist many stable, metastable, or neutral states at finite amplitude that are nearby in phase space, the race to be the observed state may well be to the swiftest. Further, then hysteresis would exist between various states, though this has not yet been studied in the context of double-diffusion. Some experiments have been performed on one component convection that make hysteresis effects obvious (Chen and Whitehead 1968). The linear results that have been obtained will be used extensively to provide clues for finite amplitude effects during the growth phase of direct-double diffusive systems.

4. Nonlinear Theory

Direct double-diffusive convection has proven more resistant to attack by weakly nonlinear theory than the overstable double-diffusive instability. Though the techniques work they do not apply to the parameter ranges of interest where behavior unique to salt fingers occurs. At just slightly supercritical configurations nonlinear theory looks much like the unconstrained Bénard problem with direct modes just saturating. It appears that this problem is also doomed to fall inextricably into the clutches of computers for major headway to be made.

When the gradients are reversed the instabilities get interesting at low Rayleigh numbers because the exchange of stabilities does not hold. Steady, greatly subcritical, finite amplitude instabilities exist (Veronis 1965). Saturation of even a weakly driven mode looks qualitatively different than the initial instability. The presence of a mode frequency as well as a growth rate leads to nontrivial couplings with only a small number of modes, and bifurcation theories can be applied (*e.g.*: Knobloch and Proctor 1981). Chaotic behaviour can also show up. This problem has also been subjected to several numerical studies (*e.g.*: Gough and Toomre 1982).

The earliest first principles nonlinear work on salt fingers [with the possible exception of Stern's (1969) collective instability work which will be held until the next section] was by Straus (1972). His analysis was performed in the limits $\tau \ll 1$, $\sigma \gg 1$ and used a two dimensional numerical Galerkin technique. Using the criteria suggested by Malkus (1954, and also see 1963) that the realized solution maximizes the unstable flux, he finds that the preferred wave number increases dramatically with increasing Rayleigh number, unlike in single component

convection. Straus also tested the stability of his solutions to three dimensional disturbances. As found by Schlüter, Lortz, and Busse (1965) for one component convection, stability is limited by the development of cross-rolls. (In a very long and tedious totally analytical calculation performed in the course of this work, following Schlüter, Lortz, and Busse closely, the structure of the double-diffusive problem was found to be quite similar to the one component problem and gave the same result for the weakly supercritical case. That calculation will not be included here.) The mode with the greatest Nusselt number, the ratio of the heat flux with convection to the heat flux in a purely diffusive state without motion, is always in the stable regime though the marginal mode eventually has too small a wave number to remain stable, providing another distinction with one component convection. (This has recently been challenged for single component convection. See Frick, Busse, and Clever 1983.) The cross-role instabilities lead to square cells as seen in moderate Rayleigh number Bénard type experiments and in salt fingers. The wave number dependences help explain the observed narrow cells. Though several vertical modes had to be included in the expansions the regimes were not sufficiently supercritical to layer.

The weakly nonlinear study by Joyce (1982) was unusual in its use of constant flux boundary conditions but simply shows the saturation of a direct mode, after some overshooting oscillation. Both the height and horizontal wave number were specified; no natural selection of the wave number was allowed. Nothing unusual was predicted and the behaviour just resembled that of a mode saturating as in a Landau equation, though Joyce's equations were more complicated. Decaying oscillations after overshoot of a saturating amplitude were observed in the numerical model presented below for a specified horizontal mode, though the boundary

conditions in this case were fixed properties, not fluxes.

One previous detailed numerical study has been performed by Piacsek and Toomre (1980). Their work models the two layer initial conditions used in the early experiments with heat and salt. The density ratios used varied from 2 to 10. Though two dimensional, the results were quite realistic. They found that the heat variation across the fingers looked nearly sinusoidal while the slower diffusing salt anomaly took on almost a square-wave shape. As the fingers grew the initial sharp salinity gradient split and moved with the ends of the fingers keeping a nearly constant mean salinity in the interior. The thermal gradient remained fairly smooth and the gradient region spread over time extending to the finger ends. The most interesting result however was the development of unstable bulbs at the end of the fingers breaking off into blobs and starting to move through the constant property reservoirs. There was insufficient vertical extent in the model for this process to fully develop, but this was nevertheless encouraging. The bulbs and blobs were unstable with regard to density. Since the model included only a few cells occassional inversions of the horizontally averaged density were observed. These simulation provided the impetus for the numerical modelling done in the course of this work. Zeman and Lumley (1982) have modelled salt fingers using a second-order turbulence closure while reducing the problem to only one, the vertical, spatial dimension. They examined the evolution of an imposed layer for a short period of time. The large number of *ad hoc* constraints limit the usefulness of their model in examining layer evolution though there was good agreement between their calculation of the flux ratio and experimental values.

None of the powerful variational methods has been brough to bear on the direct double-diffusive problem. Stern (1981) has applied such techniques to the

overstable case. The "optimum theory" of Busse (1978b) as was applied to the thermal convection problem could be directly extended to the case of salt fingers. That approach would be able to provide limits for the buoyancy flux, but would not further the study of the layering process.

5. Layer Related Theories and Observations

The history of the association between salt fingers and thermohaline layering has taken some interesting turns. Early theoretical work such as that cited above started with linear mean vertical gradients of both heat and salt. The earliest experiments, however, started with layers such as those by Turner (1967), who argued that in natural oceanic situations salt fingers would be expected at intrusions. Turner (1967) suggested that fingers would maintain sharp interfaces at their ends, and that by fluxing buoyancy would drive the observed larger scale convection above and below the cellular regime. However he did not push this idea very far though it seems correct. He likened this to the interfaces in the oscillatory case started from two layers (Turner 1965) or the sharp interfaces resulting from heating a stable salt gradient from below (Turner 1968), though here the connection is more tenuous. In these cases turbulence driven by the double-diffusive regime helps maintain the sharp gradients. This appears to also apply to the direct case.

As noted earlier, this association of well defined layers with double-diffusive phenomena suggested an explanation for the observations of a thermohaline staircase in some portions of the ocean (Turner 1967, Cooper and Stommel 1968, Tait and Howe 1968). Shortly thereafter it was shown experimentally that with linear gradients as initial conditions sharp layering results (Stern and Turner 1969). Stern (1969a) presented a mechanism for a secondary instability, which he called the collective instability, to explain the breakup by the finger regime into layers. This theory has been recently refined by Holyer (1981). [There is also a process known as collective instability in one component convection (Busse and Whitehead 1974) but this is unrelated.] The mechanism for this secondary instability is the

interaction of the salt fingers with internal wave type modes of wavelength much greater than the cell width. If certain flux conditions are exceeded there is a net transfer of energy from the convection into waves. The waves cause a shear which, they postulate, disrupts the convection cells. The balance of these processes would result in a steady state. Observations are not fully in line with this theory and some of the problems will be discussed below. There is no theory which appears to explain the large horizontal scale of the observed layers both in the laboratory and oceans.

Since the collective instability is the only quantitative theory that has been applied to explain layering for direct double-diffusive convection we will examine some of its assumptions and results. Both Stern (1969a, 1975) and Holyer (1981) assume a basic state that has infinitely tall finite amplitude cells that are marginally stable. These are exact solutions to the nonlinear Boussinesq equations, so it is appropriate to develop perturbation equations linearized about this state. An instability was then sought on a length scale much greater than the finger width. The crucial dynamical assumption made by Stern was that internal waves alter the direction of the fluxes but not their magnitudes, and that the fluxes remain parallel to isohalines. Holyer, by using a multiple scales approach, shows that the flux magnitudes are affected, though their ratio remains unchanged. Besides serious doubts about the relevance of some of the assumptions to layering, both papers have mathematical problems which will be discussed below. In spite of these problems, both get similar results for the onset of collective instability of

$$\frac{\beta F_S - \alpha F_T}{\nu(\alpha A - \beta B)} > a, \qquad a = \begin{cases} 1 & , \text{ Stern 1975} \\ 1/3, & \text{Holyer 1980} \end{cases} \tag{27}$$

Stern (1969) also finds for his fully developed similarity model that

$$H \sim \left[\frac{\nu^8}{g\beta B K_T^3 K_S^3}\right]^{1/4} \tag{28a}$$

$$h \sim \left[\frac{K_S \nu^4}{g\beta B K_T^3}\right]^{1/4} \tag{28b}$$

$$k_f \sim \left[\frac{K_T K_S}{g\beta B}\right]^{1/4}, \tag{28c}$$

where, H is the separation between finger regions, and h the thickness of the finger regions. In a later paper, Stern (1976) calulates a limit on the buoyancy flux with two layers as the initial conditon, and the internal gradients then being dynamically determined, of

$$F \leq \frac{3(1.2)}{16^{4/3}}\left(1 - \frac{K_S}{K_T}\right)^{4/3} (\beta\,\Delta S)^{4/3} (gK_T)^{1/3}$$

in the limit as $(\alpha\,\Delta T)/(\beta\,\Delta S) \to 1$, and also finds

$$k \sim \left\{\frac{4}{3}\left(\frac{K_T \nu}{g\alpha A}\right)\left(\frac{1 + 3K_S/K_T}{1 - K_S/K_T}\right)\right\}^{1/4}$$

$$h \sim \frac{16^{4/3}}{3}(1.2)^2 \nu \left(1 + \frac{3K_S}{K_T}\right)^{-1} [g\beta\,\Delta S(K_T - K_S)]^{-1/3}.$$

Since the basic theory for the collective instability breaks down as $\Lambda \to 1$ it is not clear that this limit should be allowed.

Stern's mathematical problems begin with the description of the basic state. The assumption that $K_S = 0$ requires, as he notes, $\frac{\partial \overline{S}}{\partial z} = 0$ for a steady state to exist. Both the wavelength and the flux ratio are undetermined, though are interdependent, so only one may be treated as a free parameter. The horizontal wavelength is then chosen as being close to the "most unstable disturbance."

However, the basic state chosen has *no* salt finger type instabilities because there is no salinity gradient. It is only through the assumption that $\beta \Delta S \sim \alpha \Delta T$ that a fastest growing mode wave number not dependent on $\beta \Delta S$ is obtained. For this case the assumptions that lead to the scale for the most unstable disturbance do not hold. The choice of a horizontal scale is important and so will come up again in the discussion of Holyer's work below. The collective instability is the result of flux convergences and divergences and the fluxes and their ratios are strongly tied to the horizontal wave number. Despite this, Stern's arguments concerning the "kinematical effects" of internal waves on the convergence of fluxes may still be valid.

While the difficulties in Stern's work are deliberately left out in the open by the author for us to examine and challenge, Holyer's major trouble is far more subtle. The basic state she chooses is infinitely tall marginally stable salt fingers. This is done to have a time independent basic state, as did Stern, but here a mean salinity gradient is included along with a nonzero salt diffusivity. If one wished to find instabilities of this system there are far more powerful and easier to find examples than the collective instabilities; these are salt finger modes that are unstable and closer to, or at, the fastest growing mode. There are three scales used in Holyer's analysis; the mean gradients, a salt finger width, and a collective instability scale much longer than the finger width. Averaging is performed over many finger widths at an arbitrary angle to the finger wave number vector, corresponding to the direction of a possible collective mode wave vector. The unstable collective modes that are found have a nearly vertical wave vector. This limit however violates the scaling assumptions. Once the collective mode is nearly vertical, averaging over its wavelength no longer averages over many finger widths.

It would be hasty to conclude from this brief critisism of collective instability theory that such an instability can not or does not exist. As a cause of layering in direct double-diffusive convection, however, it will not do. Credit must be given to Stern for providing an easy target. His reasoning is made clear even when subject to quick objection. His predictions are powerful, concise, and can be tested experimentally in a straight forward manner. This work will be unable to match his fine example.

There have been several attempts to compare collective instability theory with observations and experiments. Stern (1969) made order of magnitude comparisons with the data of Tait and Howe (1968) using $\beta B = 10^{-8}\,cm^{-1}$ and got a value for H of $10^{1/4} \cdot 10^{3}\,cm$, which is reasonable, and $10^{1/4}$ for h, which is an order of magnitude too small. Hè found the data of Cooper and Stommel (1968) to show more variety in the structure than could be explained by collective instability. Lambert and Demenkow (1972) found the fluxes to be only 10^{-3} of the amount needed for the onset of the instability, but nevertheless quasi-steady layers were present. One might argue that fingers can still exist in a regime where the collective instability is already being driven to some extent but not strong enough to disrupt the cells; however, it is difficult to imagine an instability limiting some phenomenon before the conditions for its onset are met. Though Stern and Turner (1969) and Linden (1973) come within an order of magnitude, there is no particular value of a in eq. (27) about which their values cluster. Linden's experimental values ranged over an order of magnitude from 0.2 to 1.9 for heat-sugar fingers. Also, except for the case of mixing at a front which has clearly different behaivour, oblique motions or disturbances have not been reported. The layers form oriented almost perfectly horizontally, as can be seen in the photographs of figs. (16), exactly where Holyer's

analysis is mathematically inappropriate.

Supercritical salt fingers can survive rather strong motions and shears on length scales greater than the finger scales. Examples of this may be found in Turner and Chen 1974 and Turner 1978. Even Stern and Turner (1969) found that fingers can grow in a strong shear. Linden (1974) found that imposing a steady shear led to two dimensional rolls, not total breakup. Even if fingers pump energy into some larger mode at a finite rate their breakup seems unlikely. It is not sufficient, as Stern (1976) suggests, that internal waves displace the cells of order of a cell width for disruption. The vertical scale of the shear would have to be of the order of the finger width to be destructive. Further, the ends of the fingers have reduced density gradients, though exaggerated thermohaline gradients, so that the shear of internal waves is concentrated elsewhere. These are more reasons that the collective instability doesn't seem to be the mechanism for layering.

If we look at Stern's basic result, eq. (27), it does have two things which give it a qualitative validity; vigorous fluxes are more likely to be associated with an unstable mode, and a steep stable net density gradient is stabilizing. These conclusions are likely to be matched by any secondary instability mechanism and are intuitively reasonable without the collective instability model. With regards to physical details and quantitative validity, the collective instability does not appear to hold up to scrutiny as the cause for the breakup of salt finger convection cells into layers.

Thermohaline layering in the ocean, even when stratified in the salt finger sense may still have multiple causes. Several different forcing situations can exist and theories are not well divided in the literature discussions (though nature may be as

confusing). Three clearly come to mind: 1) a net input of heat and salt from solar heating and excess evaporation, such as in the Atlantic Central waters, 2) a warm salty current over cooler fresher water, such as under the Mediterranean tongue, and 3) at the front of two water masses of different origins and T-S relations. Though the layering processes of the first two are probably similar, the physics of the third may be markedly different. The third case develops alternating layers of fingers and layers started by the oscillatory instability. The setting of vertical scales is very different in this problem. Some theory on this is presented by Toole and Georgi (1981). The references in that paper list most of the other work in this area. Evidence for all three of these in the oceans has already been noted.

Some theoretical work has been done taking the existence of layers for granted and examining the interactions of sets of layers. Using empirical flux laws for the relatively thin double-diffusive regions, separated by larger homogeneous turbulent regions, Huppert (1971) determined the conditions required for a series of layers to be a stable configuration in the oscillatory case. The method is easy to extend and he shows how to apply it to more general flux laws and the salt finger case. Though Huppert restricted the dependencies of the flux laws to density ratio, and power law dependence on the temperature difference across each double-diffusive layer, generalizations are only a practical, and not conceptual, difficulty. Schmitt (1981), applying a method strongly related mathematically to Huppert's, shows that the known semi-empirical flux laws for salt fingering should lead to a constant density ratio in the ocean. He notes observations that appear to confirm this in several ocean regions with density ratios ranging from 1.15 to 1.9.

The observations of Williams (1975) demonstrated that the layered structure of the ocean could be quite intricate. In the Tyrrhenian Sea the fine structure steps had

about a $30\,m$ scale with a mean salinity gradient of $\approx 5 \cdot 10^{-7}\,cm^{-1}$ concentrated into gradient regions of about $6\,m$, with nearly compensating temperature gradient. Within these $6\,m$ layers multiple salt finger layers were found with the gradient concentrated into thinner layers. For a strong interface under the Mediterranean outflow a well defined finger region about $40\,cm$ tall with a fairly smooth internal temperature gradient, the structure of the salinity field within the fingers could not be resolved but had a ΔS of $30 \cdot 10^{-6}$, by mass, for a mean gradient of just under $10^{-6}\,cm^{-1}$ and the mean density ratio can be calculated as about 1.2. Based on Linden's (1973) laboratory work, that salinity gradient is probably further concentrated to a few centimeters at each end of the cells. This is beyond the usual measuring capability of ocean going instruments.

Possibly the most fascinating finger experiments were those of Linden (1978) who chronicled the development of layers in a salt-sugar system. The experiments started with a layer of linear gradients sandwiched between two constant property reservoirs, each about $12\,cm$ deep. A small discontinuity in concentration was initially present at the boundaries between the gradient and reservoir regions, though we do not feel that this was strongly related to the evolution of the systems. The photographs were produced with a shadowgraph technique. This produces light and dark regions on a screen by the differential focusing from the variations of the index of light refraction with density. The method is sensitive to the second spatial derivative of the index of refraction (Shirtcliffe and Turner 1970). The first sequence shown in fig. 14 does not exhibit layering. Initially vigorous convection shrinks the layer thickness, as can be seen going from fig. 14a to fig. 14b. Eventually as the system ran down the layer thickened (not shown).

The second sequence shows more structure. The initial pattern is salt fingers

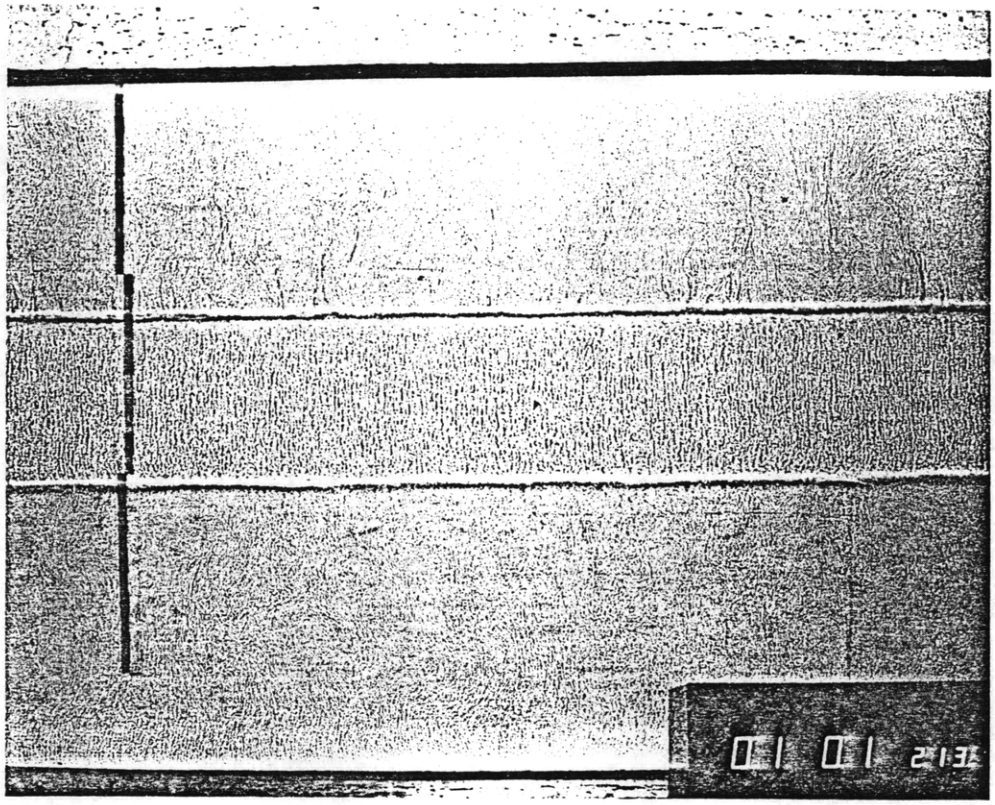

Fig. 14a.

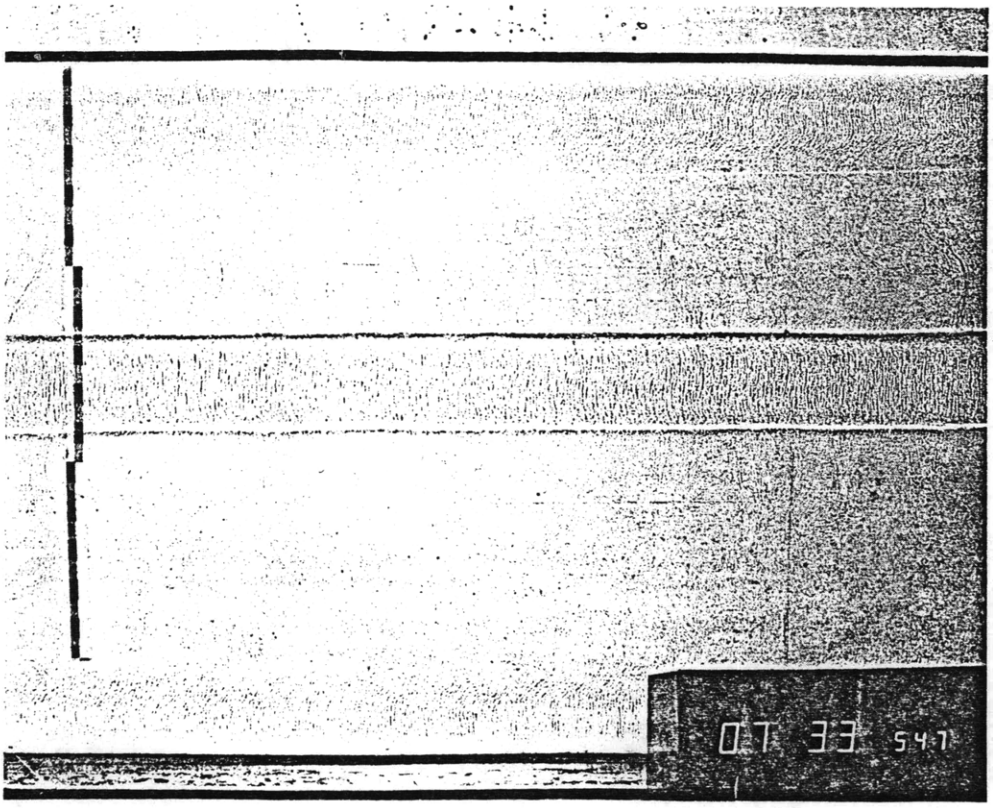

Fig. 14b.

Fig. 14. Time lapse photographs of salt-sugar fingers without layering. The full depth of the tank is visible, the depth being shown by the tape on the left-hand side which has alternating light and dark regions each 1 *cm* long and offset every 10 *cm*. The digital clock in the lower right-hand corner shows the time in hours, minutes, seconds and tenths since the beginning of the addition of the top layer. $\Lambda_v = 1.16$, $\beta \Delta S / h_v = 8.0 \cdot 10^{-3} cm^{-1}$. Note the decrease in height of the layer and that the layer remains continuous. Photgraphs and parts of the captions for figs. 14 – 16 are from P. F. Linden, *J. Geophys. Res.* 83 C6, 2902–2912, © American Geophysical Union.

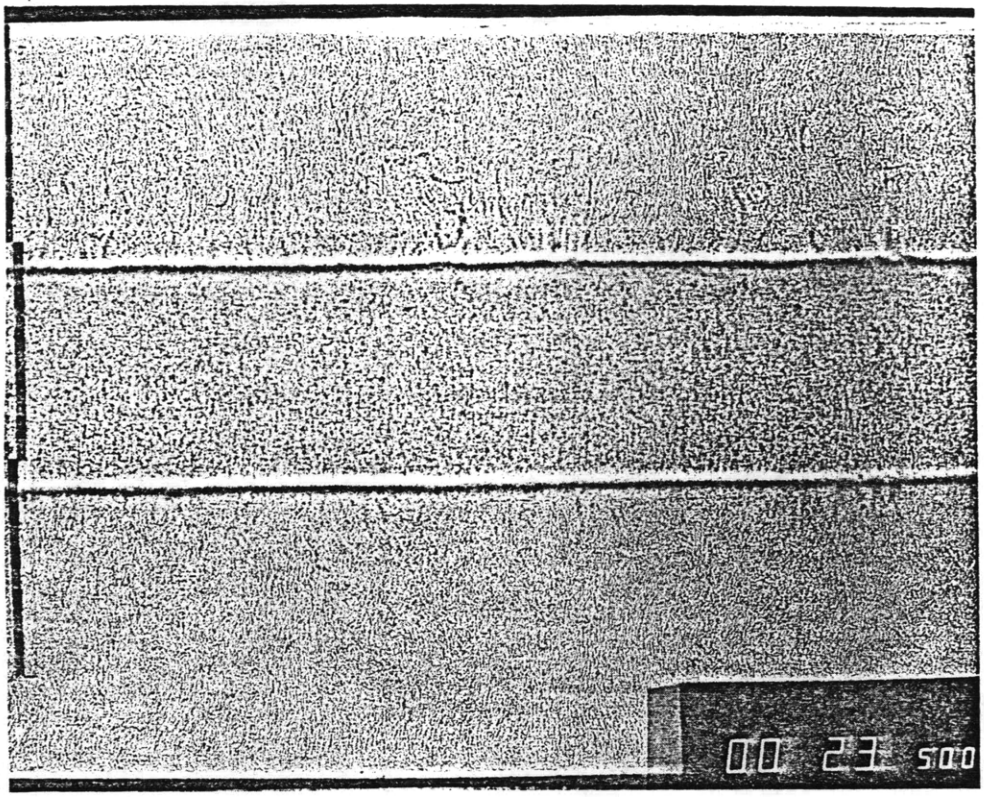

Fig. 15a.

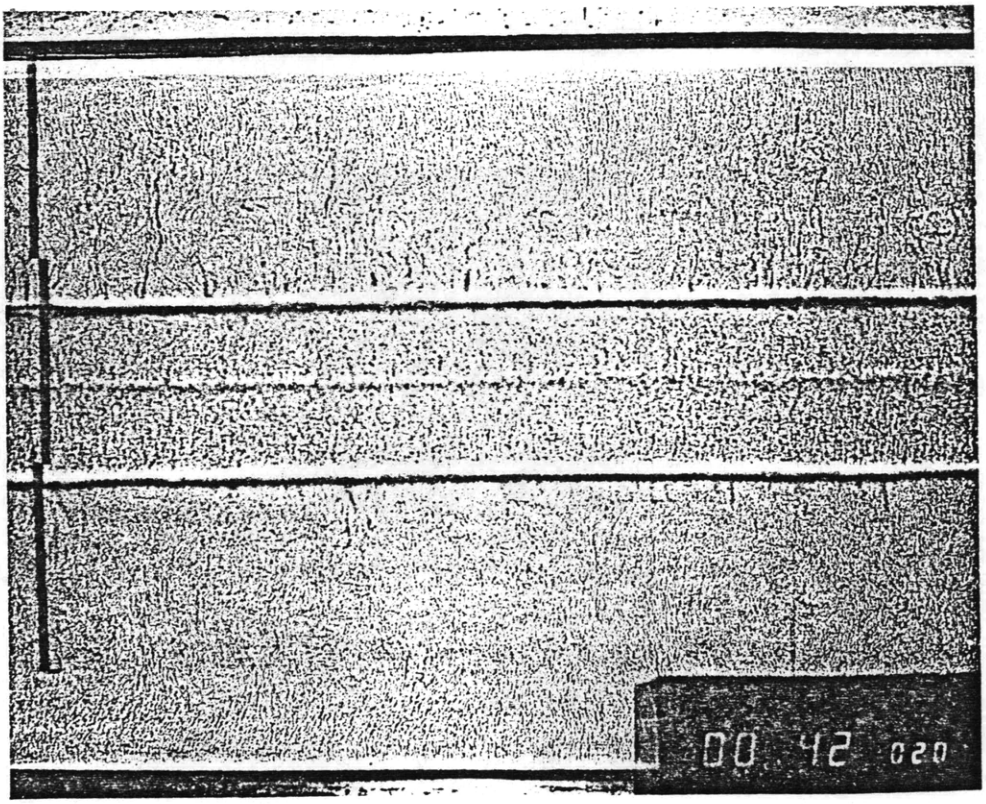

Fig. 15b.

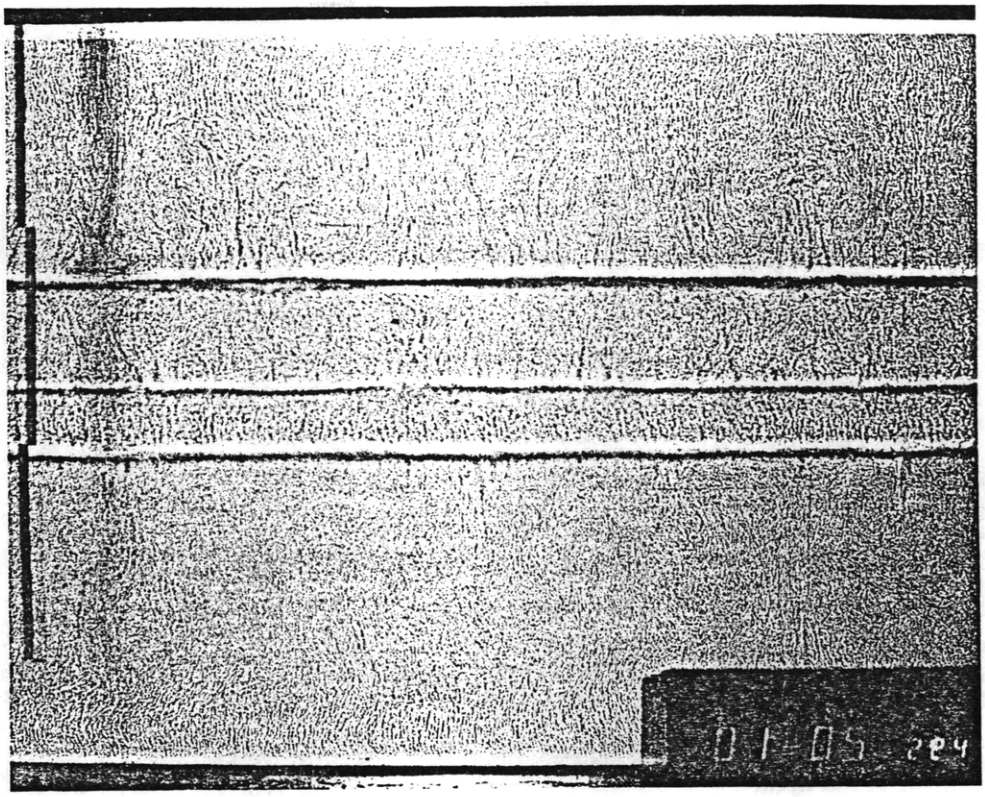

Fig. 15c.

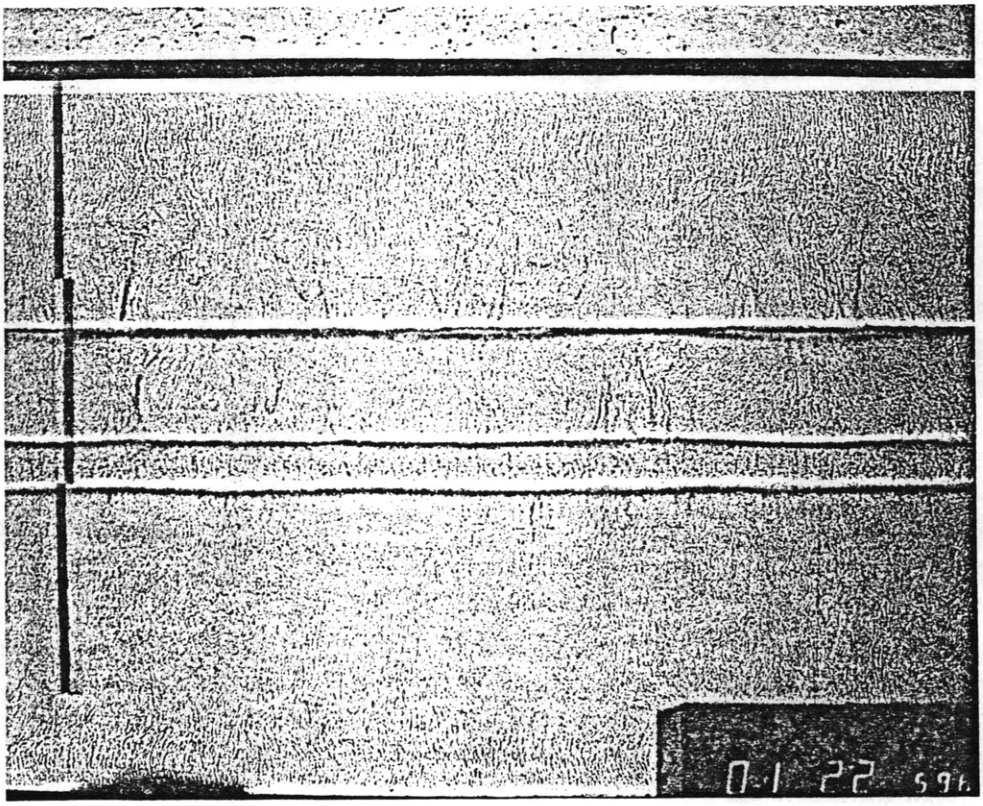

Fig. 15d.

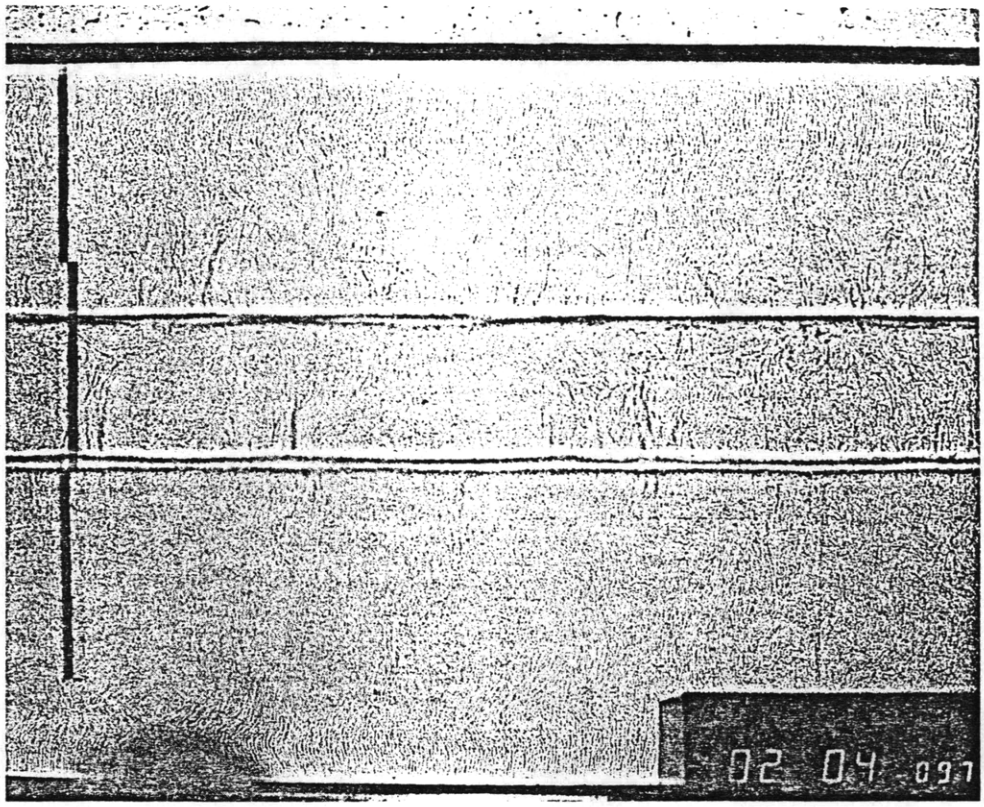

Fig. 15e.

Fig. 15. Time lapse photographs of salt-sugar fingers with layering (from Linden 1978). $\Lambda_o = 1.1$, $\beta \Delta S / h_o = 8.3 \cdot 10^{-3} cm^{-1}$. Note that the first harmonic clearly develops and is then followed by the dominance of higher harmonics.

throughout the gradient region (fig. 15a). However the convection is so vigorous that the clear cellular pattern is not evident. It is possible that for a short while the whole layer is turbulent. By the second photograph, fig. 15b, a sharp gradient evenly divides the cells. We attribute this to the modification of the mean field by the first harmonic. In the next step another sharp gradient developed just below the upper gradient region boundary. The strong asymmetry evident at this time eventually vanishes. The final photograph of this sequence shows two finger layers with each having sharp gradients visible at both ends. Since the initial configuration was so extremely supercritical, we believe that fairly high harmonics that were unstable dominated the system.

The third sequence, fig. 16, layers more gracefully and it is easier to distinguish the laminar cells from the chaotic regions. Thin, but clearly cellular layers develop at two levels. Eventually the cells lengthen with the weakening of the gradients as in the first experiment. No large scale lateral motion was reported. The development of horizontal layers associated with a sharp density gradients at their top and bottom is evident. A theory for the formation of these layers, by the mean field modifications from modes with high vertical wavenumbers and the accompanying change in the stability of some regions, is presented in the next chapter.

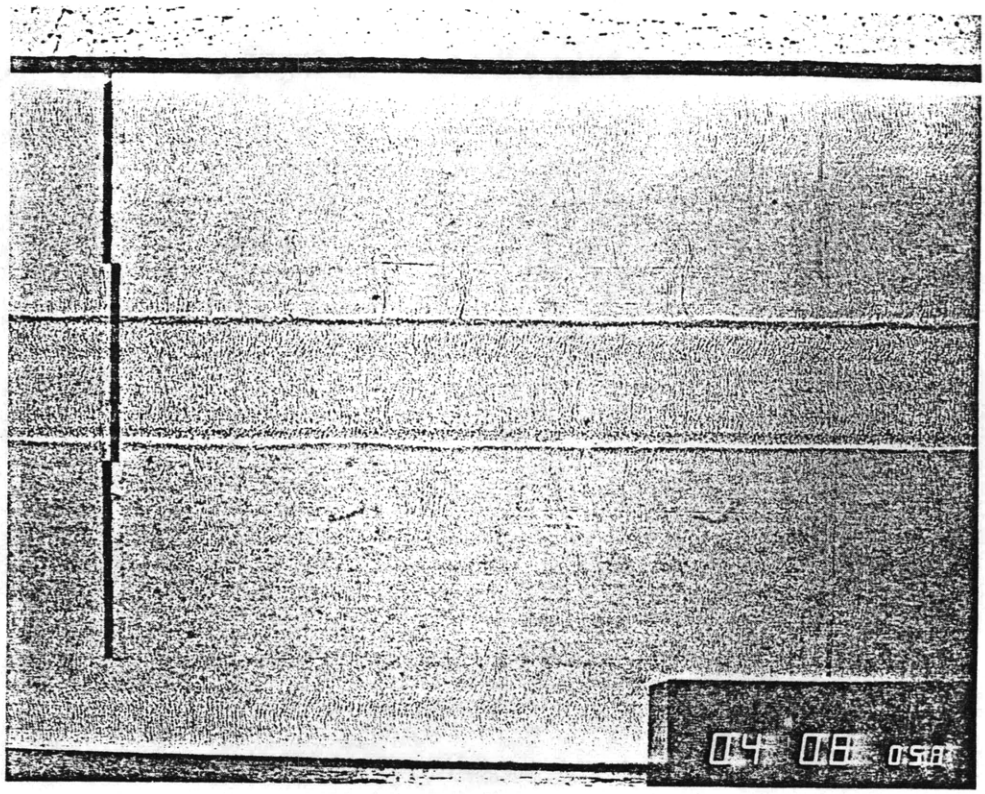

Fig. 16a.

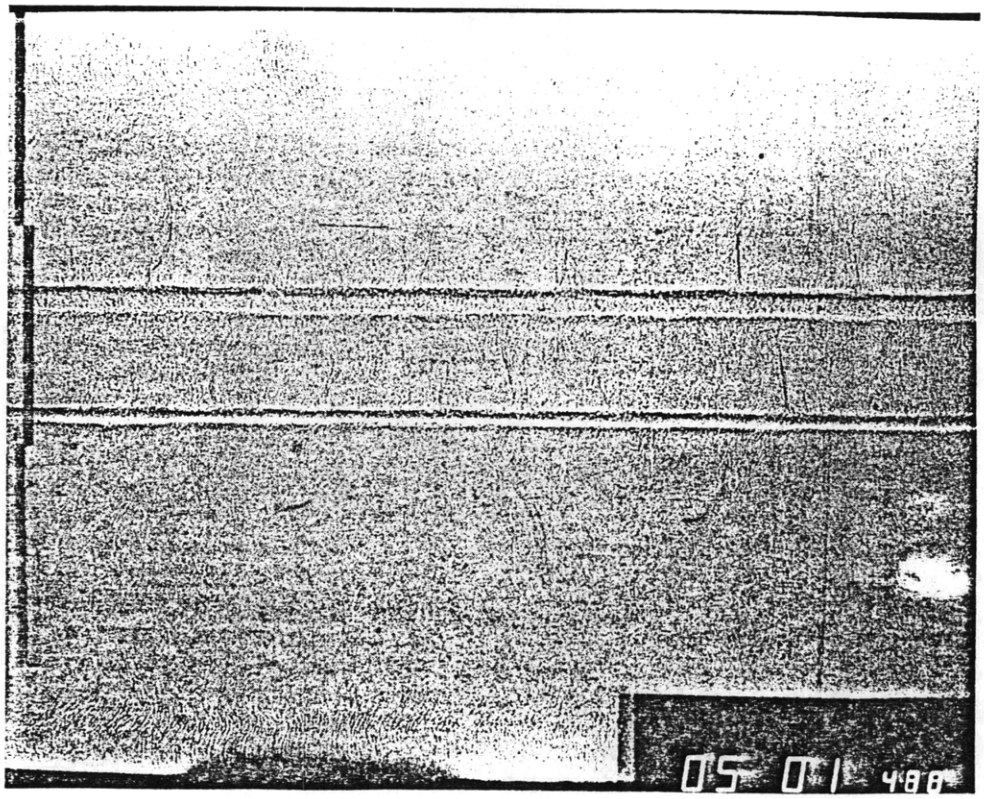

Fig. 16b.

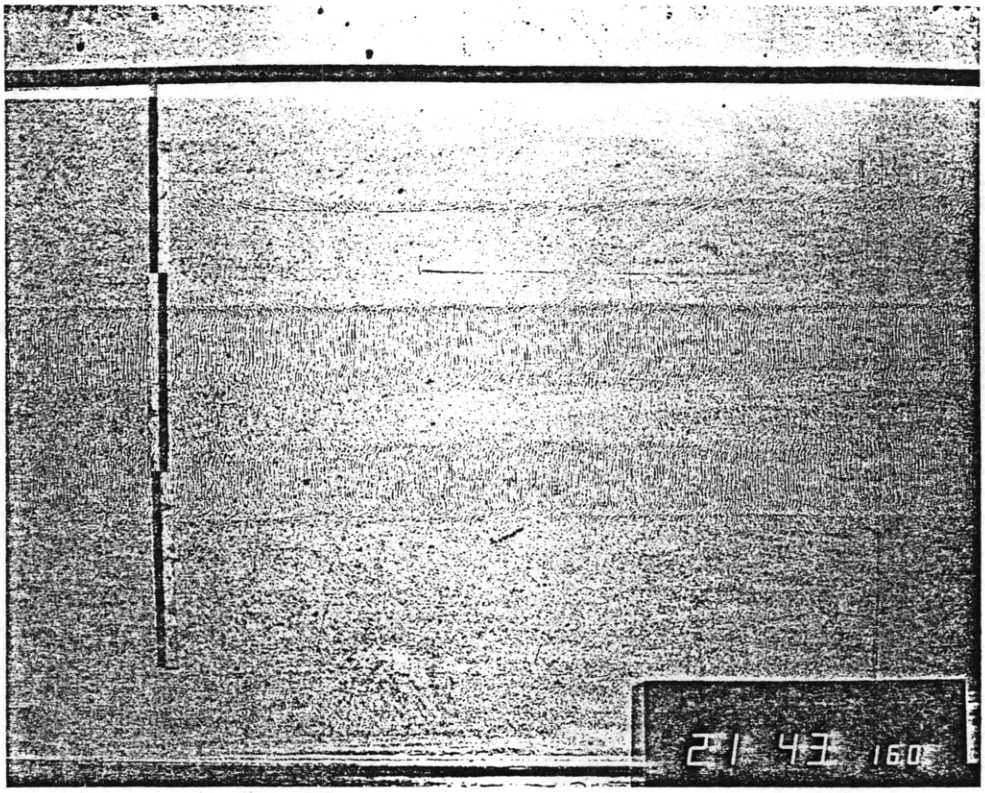

Fig. 16c.

Fig. 16. Time lapse photographs of salt-sugar fingers with layering (from Linden 1978). $\Lambda_o = 1.05$, $\beta \Delta S / h_o = 4.4 \cdot 10^{-3}$. Note the establishment of a central convecting region during the time between 16a and 16b and the subsequent growth of the fingers in the two separate interfaces in fig. 16c.

6. Harmonic Instabilities

For a highly supercritical basic state there is a large spectrum of unstable modes. This applies to the ocean where salt fingers are observed and almost all of the laboratory work that has been performed. Not only is the effective Rayleigh number enormous because of the large vertical extent, but also the density ratio is small leading to large growth rates. In weakly nonlinear theory saturation of the vertical fundamental is sufficient to stabilize the system and the fundamental weakly drives the first vertical harmonic through the nonlinear terms. As the systems become more supercritical the higher vertical modes are themselves unstable and their dynamics become more loosely coupled to the fundamental. Further, the fundamental alone when saturated will be unable to stabilize the system to instabilities on smaller scales.

One class of instabilities takes on a very different role in the parameter ranges of layering in double-diffusion; these are the vertical harmonics. They have the ability to greatly alter the base conditions and control where turbulent layers will occur. We will demonstrate that even with initial white noise, the fastest growing mode is not the mode that has the greatest effect on a system, at least by some important criteria. The approach here is based on examining flux divergences and has not been previously applied to salt fingers. Collective instability theory looks at modification of fluxes by internal waves. The analysis that follows concentrates on the effects of the convective instabilities themselves on the stability of the system with regard to other direct convective instabilities. The system we will examine is unbounded in the vertical, initially with linear positive gradients of temperature and salinity, and net stable density gradient. Equations (12) are still the full nonlinear

diffusion equations. Horizontally averaged they become

$$\frac{\partial \overline{\theta}}{\partial t} - \frac{\partial^2 \overline{\theta}}{\partial z^2} = -\frac{\partial}{\partial z}\mathcal{F}_T \quad , \qquad \mathcal{F}_T = \overline{w\theta}$$
$$\frac{\partial \overline{S}}{\partial t} - \tau\frac{\partial^2 \overline{S}}{\partial z^2} = -\frac{\partial}{\partial z}\mathcal{F}_S \quad , \qquad \mathcal{F}_S = \overline{wS} \quad . \tag{29}$$

Until the disturbance amplitudes reach a point where the nonlinear terms in the basic equations become significant, it is valid to look at the forced solutions to eqs. (29) by determining and using the Reynolds fluxes of a mode from linear theory as the forcing. This approach will provide a tendency calculation for the early evolution of the mean field without feedback to the linear growing modes. A horizontally in phase linear combination of solutions, eqs. (13), to the linearized problem of eqs. (9), (11), and (12), yields

$$-\frac{\partial}{\partial z}\mathcal{F}_T = \Lambda_0 \frac{\hat{w}^2 m e^{2\lambda t}\sin 2mz}{(\lambda+\gamma^2)}$$
$$-\frac{\partial}{\partial z}\mathcal{F}_S = \frac{\hat{w}^2 m e^{2\lambda t}\sin 2mz}{(\lambda+\tau\gamma^2)} \tag{30}$$

while it is still required that the dispersion relation, eq. (14a), be satisfied. The initial value problem for eqs. (29), with the forcing as in eqs. (30), and initial conditions

$$\overline{\theta}(t=0) = \overline{S}(t=0) = 0 \tag{31}$$

has solutions

$$\overline{\theta} = \Lambda_0\left[\frac{\hat{w}^2 m\left(e^{2\lambda t} - e^{-4m^2 t}\right)\sin 2mz}{2(\lambda+2m^2)(\lambda+\gamma^2)}\right]$$
$$\overline{S} = \left[\frac{\hat{w}^2 m\left(e^{2\lambda t} - e^{-4\tau m^2 t}\right)\sin 2mz}{2(\lambda+2\tau m^2)(\lambda+\tau\gamma^2)}\right] \quad . \tag{32}$$

For very small initial noise amplitudes the decaying transients are unimportant since only at times after $e^{2\lambda t} \gg 1$ do the solutions have significant amplitudes. This is equivalent to not strictly requiring the initial conditions of eq. (31). Dropping the transients in eq. (32), the vertically local, horizontally averaged density ratio is then

$$\Lambda(z) = (\alpha\frac{\partial\bar{\theta}_*}{\partial z})/(\beta\frac{\partial\bar{S}_*}{\partial z}) = \left(\Lambda_0 + \frac{\partial\bar{\theta}}{\partial z}\right)\left(1 + \frac{\partial\bar{S}}{\partial z}\right)^{-1} \tag{33}$$
$$= \Lambda_0\left[1 + \frac{\hat{w}^2 m^2 e^{2\lambda t}\cos 2mz}{(\lambda + 2m^2)(\lambda + \gamma^2)}\right]\left[1 + \frac{\hat{w}^2 m^2 e^{2\lambda t}\cos 2mz}{(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)}\right]^{-1}.$$

The layers of greater instability will occur at regions of reduced density gradient so we look for the minima of $\Lambda(z)$ which occur where $\cos(2mz) = -1$. If the growing terms are still much less than one, the expression may be Taylor expanded. Equation (33) then simplifies to

$$\min[\Lambda(z)] \approx \Lambda_o\left(1 - \hat{w}^2\Gamma\right) \tag{34}$$
$$\Gamma = m^2 e^{2\lambda t}\left[\frac{1}{(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)} - \frac{1}{(\lambda + 2m^2)(\lambda + \gamma^2)}\right].$$

For an initial disturbance consisting of spatial white noise, it is not the fastest growing mode that most greatly affects the mean fields and $\Lambda(z)$. This can be easily understood with the following heuristic. The fastest growing mode is the fundamental, which has the smallest allowed vertical wave number with $k = k_f$. For an unbounded fluid this is $m = 0$, and for a deep region m is very small. However, this mode has no effect on the mean field since it has no flux divergence because $m = 0$ and little flux divergence if m is extremely small. On the other hand, there is some maximum vertical wave number above which there is no growth, so

$$m^2 < \left[\frac{4}{27}(1-\Lambda\tau)\right]^{1/2} \tag{35a}$$

or in dimensional form,

$$m^2 < \left[\frac{4}{27}\frac{g}{\nu}\left(\frac{\beta B}{K_S} - \frac{\alpha A}{K_T}\right)\right]^{1/2} . \tag{35b}$$

This is determined in a manner similar to determining the marginal instability criteria for a fixed cell, with conducting, stress free boundaries and hence is essentially the same result as eq. (15).

To find the mode that has had the greatest effect of the density ratio at a given time, set

$$\frac{\partial}{\partial m^2}\min[\Lambda(z)] = \frac{\partial}{\partial k^2}\min[\Lambda(z)] = 0 \quad .$$

Taking these derivatives of eq. (34), we therefore require

$$\begin{aligned} 0 = &\left(1 + 2m^2 t\frac{\partial\lambda}{\partial m^2}\right)\left[\frac{1}{(\lambda+2m^2\tau)(\lambda+\gamma^2\tau)} - \frac{1}{(\lambda+2m^2)(\lambda+\gamma^2)}\right] \\ &- \frac{m^2\left[\left(\frac{\partial\lambda}{\partial m^2}+2\tau\right)\left(\lambda+\gamma^2\tau\right)+\left(\lambda+2m^2\tau\right)\left(\frac{\partial\lambda}{\partial m^2}+\tau\right)\right]}{(\lambda+2m^2\tau)^2(\lambda+\gamma^2\tau)^2} \\ &+ \frac{m^2\left[\left(\frac{\partial\lambda}{\partial m^2}+2\right)\left(\lambda+\gamma^2\right)+\left(\lambda+2m^2\right)\left(\frac{\partial\lambda}{\partial m^2}+1\right)\right]}{(\lambda+2m^2)^2(\lambda+\gamma^2)^2} \end{aligned} \tag{36}$$

and

$$0 = 2t\frac{\partial\lambda}{\partial k^2}\left[\frac{1}{(\lambda+2m^2\tau)(\lambda+\gamma^2\tau)} - \frac{1}{(\lambda+2m^2)(\lambda+\gamma^2)}\right]$$
$$- \frac{\left[\frac{\partial\lambda}{\partial k^2}\left(\lambda+\gamma^2\tau\right)+\left(\lambda+2m^2\tau\right)\left(\frac{\partial\lambda}{\partial k^2}+\tau\right)\right]}{(\lambda+2m^2\tau)^2(\lambda+\gamma^2\tau)^2}$$
$$+ \frac{\left[\frac{\partial\lambda}{\partial k^2}\left(\lambda+\gamma^2\right)+\left(\lambda+2m^2\right)\left(\frac{\partial\lambda}{\partial k^2}+1\right)\right]}{(\lambda+2m^2)^2(\lambda+\gamma^2)^2} \quad . \tag{37}$$

The simultaneous solution of Eqs. (14), (36), and (37), still has a time dependence and holds for $e^{2\lambda t} \gg 1$ or, approximately, $\lambda t > 1$. A criterion for the closure of this system must still be specified, or determined.

How much effect can the flux divergences have before the dynamics are surely nonlinear? Admittedly for mathematical convenience, and fortuitously for oceanographic relevance, take $\Lambda = 1+\epsilon$, $\epsilon \ll 1$. In the areas of the ocean where salt fingers are observed to be vigorous $\epsilon \sim 0.2$. This is not totally fortuitous since layering can only occur if Λ is not much greater than one. Further this is also required for vigorous convection and would naturally result in an area with a large negative buoyancy flux into the surface (with standard conventions for direction this is a positive buoyancy flux). The acceleration term in the vorticity equation is largest for the fastest growing mode; then it is of order $\lambda\hat{w}e^{\lambda t}$. The nonlinear terms are of order $i\gamma^2\hat{u}\hat{w}e^{2\lambda t}$ or $i\gamma^2(m/k)\hat{w}^2e^{2\lambda t}$, so their ratio, q, is

$$q \sim \frac{m\hat{w}e^{\lambda t}}{\lambda} \quad . \tag{38}$$

The ratio of the viscous term to the nonlinear terms is $m\hat{w}e^{\lambda t}/(\sigma\gamma^2)$ while the ratio of the density gradient driving term to the nonlinear terms is $m\hat{w}e^{\lambda t}\gamma^2(\lambda+\tau\gamma^2)/(\sigma k^2) \sim m\hat{w}e^{\lambda t}\lambda/\sigma$ which is even less than q for $\sigma > 1$. As long as $q \ll 1$,

the approach taken is still valid. From eq. (33) for heat-salt fingers, so $\tau \ll 1$, assuming $\gamma^2/\lambda \sim 1$, and $m^2 \ll \gamma^2$,

$$\min[\Lambda(z)] - 1 \sim \epsilon - \frac{m^2 \hat{w}^2 e^{2\lambda t}}{2\lambda^2} \sim \epsilon_o - \frac{q^2}{2} . \tag{39}$$

So if

$$\epsilon \ll \frac{1}{2} \tag{40}$$

regions are created that have a mean statically unstable density gradient, while all the dynamics are still essentially linear.

It is not the case that the nonlinear terms are in this ratio to all the linear terms. However, for the dynamics to remain linear at least two linear terms from each of the equations (11) and (12) must be much greater than the nonlinear terms. In particular the diffusive terms in the salt conservation equation, (12b) is quite small if $\tau \ll 1$, and then the time derivative term balances the advection term. In the salt conservation equation the ratio of the nonlinear terms to the time derivative term is again q as in the vorticity equation. The nonlinear terms over the linear advection of the mean salinity field are also of order q if $\tau \ll 1$, otherwise they are less than q and the ratio of the nonlinear terms to the diffusion term is $q/(\tau\gamma^2) \gg q$ for $\tau \ll 1$. This does not present a problem for the reason noted that there are still two large linear terms to maintain a consistent dominant balance that does not include the nonlinear terms. A similar analysis can be performed on the heat conservation equation (12a) and yields the same required condition.

The results of the previous calculations are at least suggestive even when eq. (40) does not strictly hold. The nonlinear terms do not force saturation of

the normal modes before the stability characteristics of some regions are changed enormously. This is crucial for all of the tendency calculations presented here. If the fundamental does reach large amplitude before breakdown occurs, the process is not stopped but may be significantly slowed down. As long as the fundamental can not fully stabilize the system harmonics will still develop.

A search for the mode that first creates levels that are statically unstable in the mean for a given initial density ratio and noise level, provides conditions that close the system of eqs. (14), (36), and (37). The criterion of the creation of regions with positive density gradient was used by Turner (1968), and also by Linden (1976), for theoretical examinations of the limitations of layer thickness and the initiation of additional layers, when the gradients are as in the oscillatory double-diffusive case. Demonstration that regions exist that have a mean positive density gradient is not sufficient to generate an additional type of instability. However, it is suggestive that regions can be created that have such large accelerations that the resulting motions will be turbulent and not cellular, since the growth rate grows rapidly with decreasing density ratio. Modifications of the mean field in one component convection cannot create regions with significantly greater accelerations than in the initial state.

Since it is not clear how to pick an initial perturbation amplitude and spectrum it is not yet possible to analytically solve the set of eqs. (14), (36), and (37). Modes with the greatest effect at early times have a much larger m^2 and smaller k^2 than observed. However the time dependence of the pattern is strong. After long amounts of time, *i.e.*: $\lambda t \gg 1$, small vertical wave numbers become more important, as is observed and indicated in the contours of Γ at later times, and the

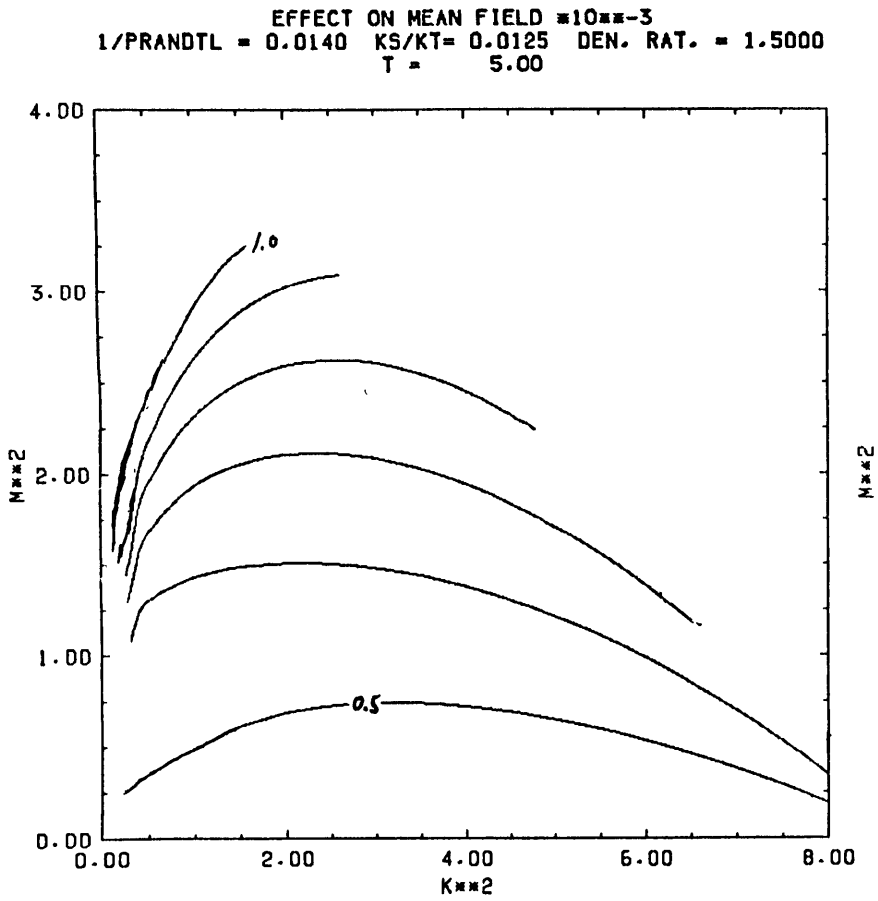

Fig. 17a.

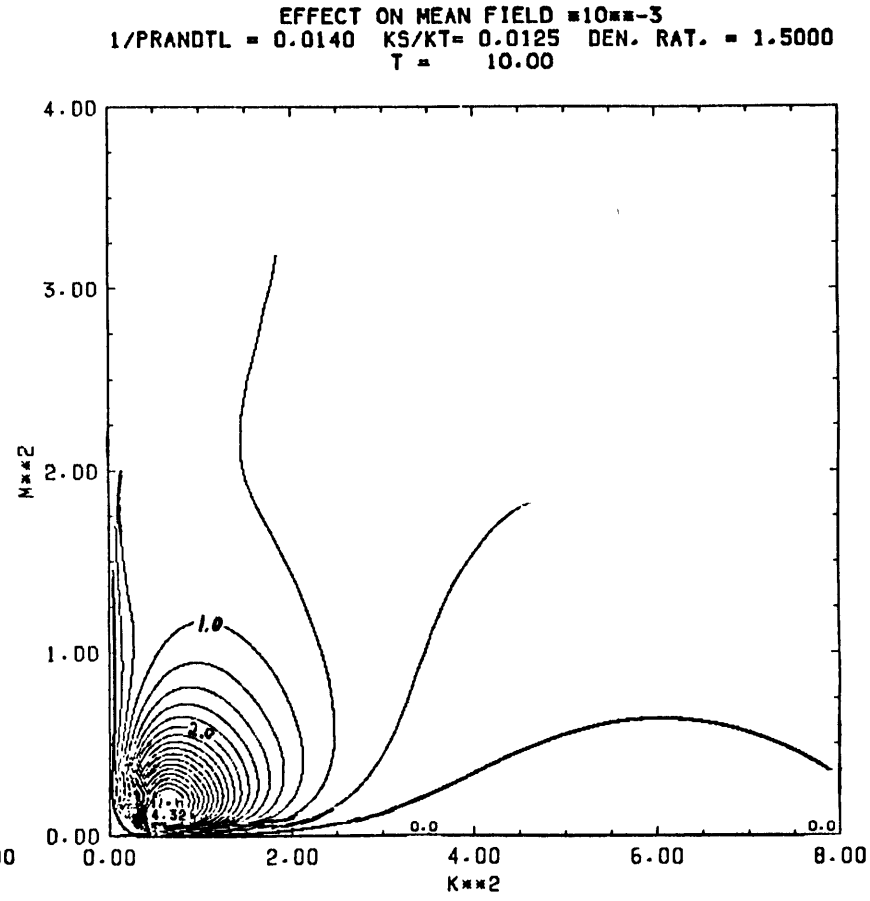

Fig. 17b.

Fig. 17. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for heat-salt case, density ratio = 1.5: a) $t = 5$, b) $t = 10$, c) $t = 20$, d) $t = 25$. Contours are plots of Γ from eq. 34.

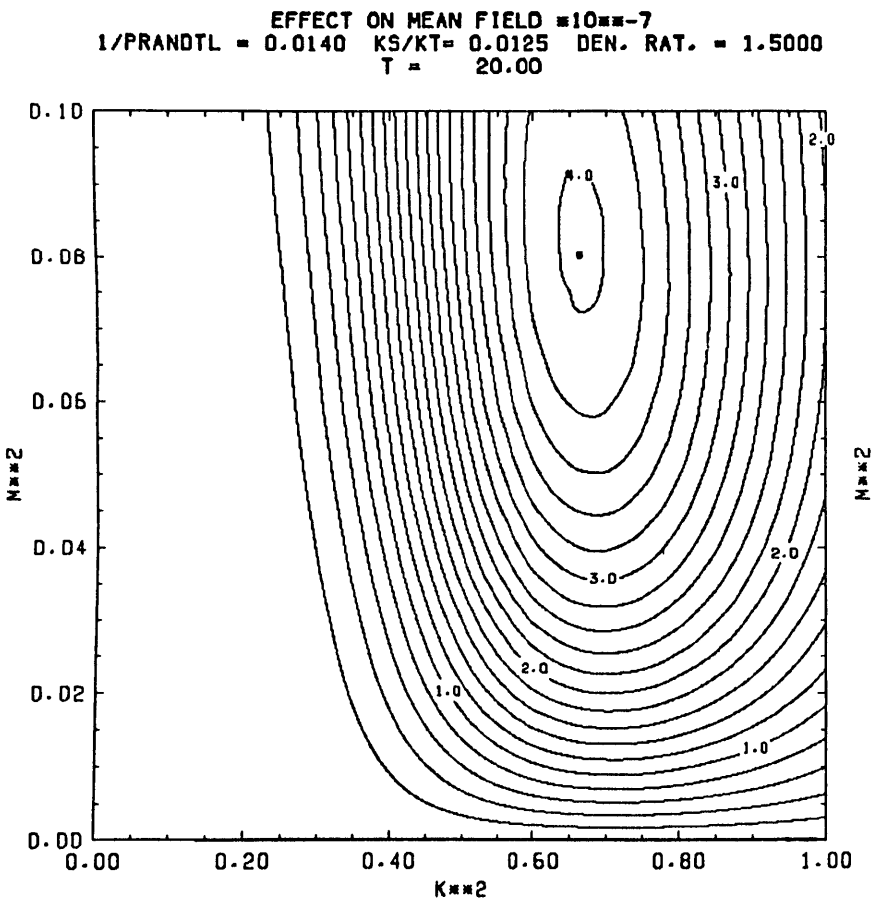

Fig. 17c.

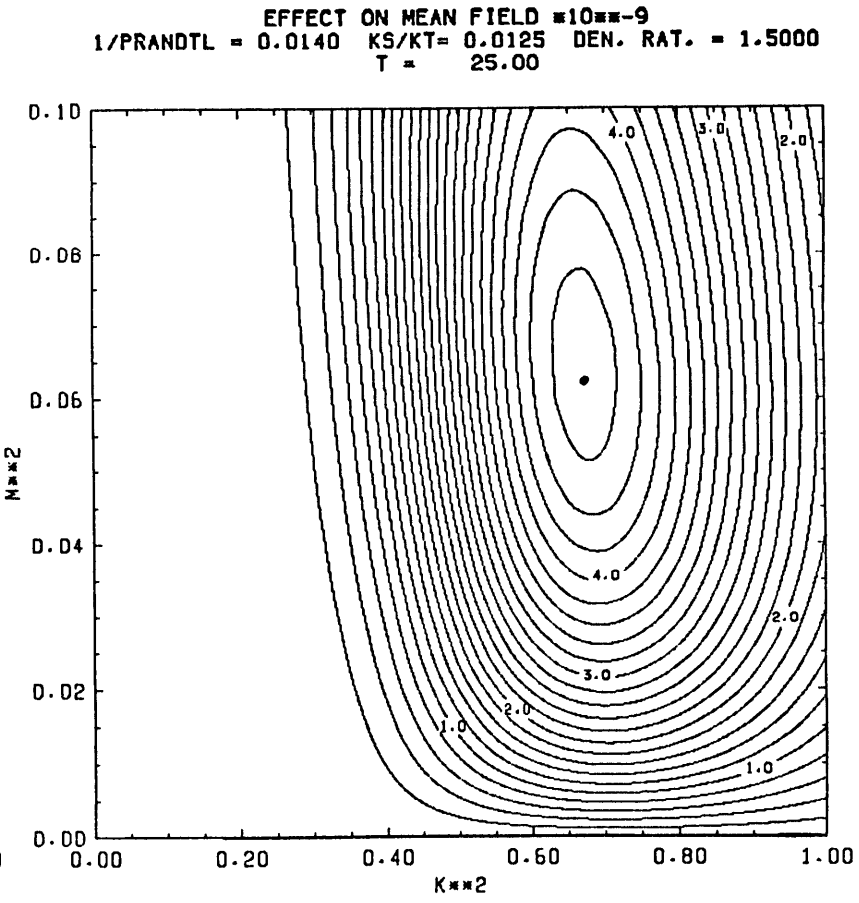

Fig. 17d.

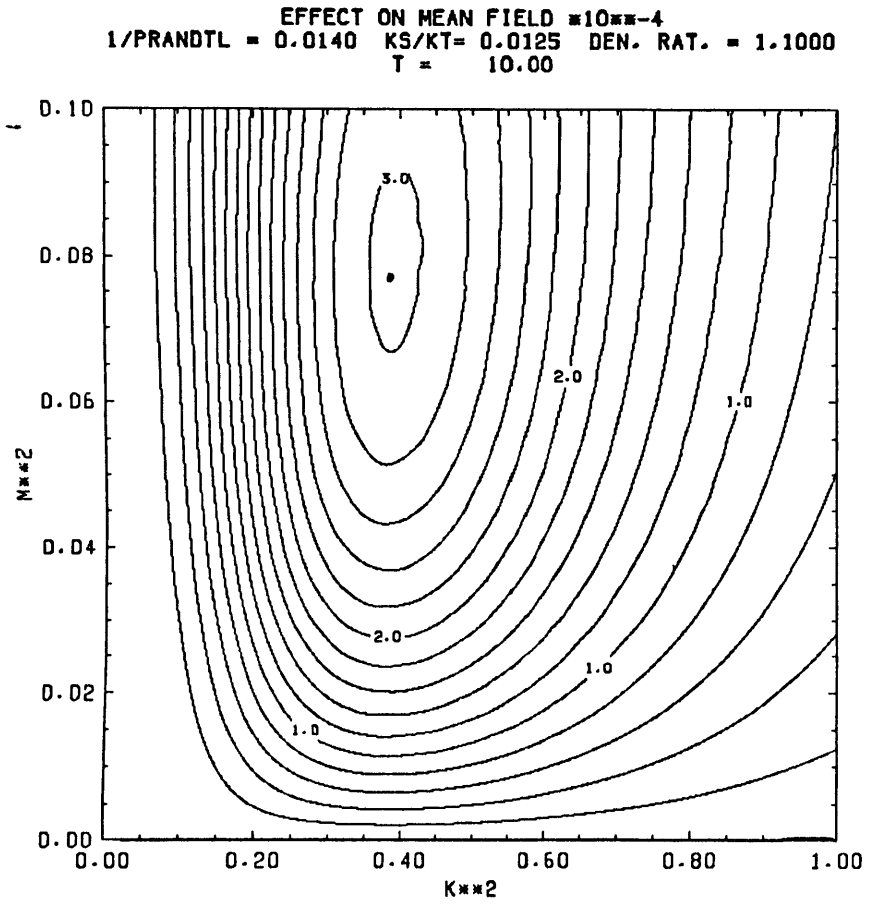

Fig. 18a.

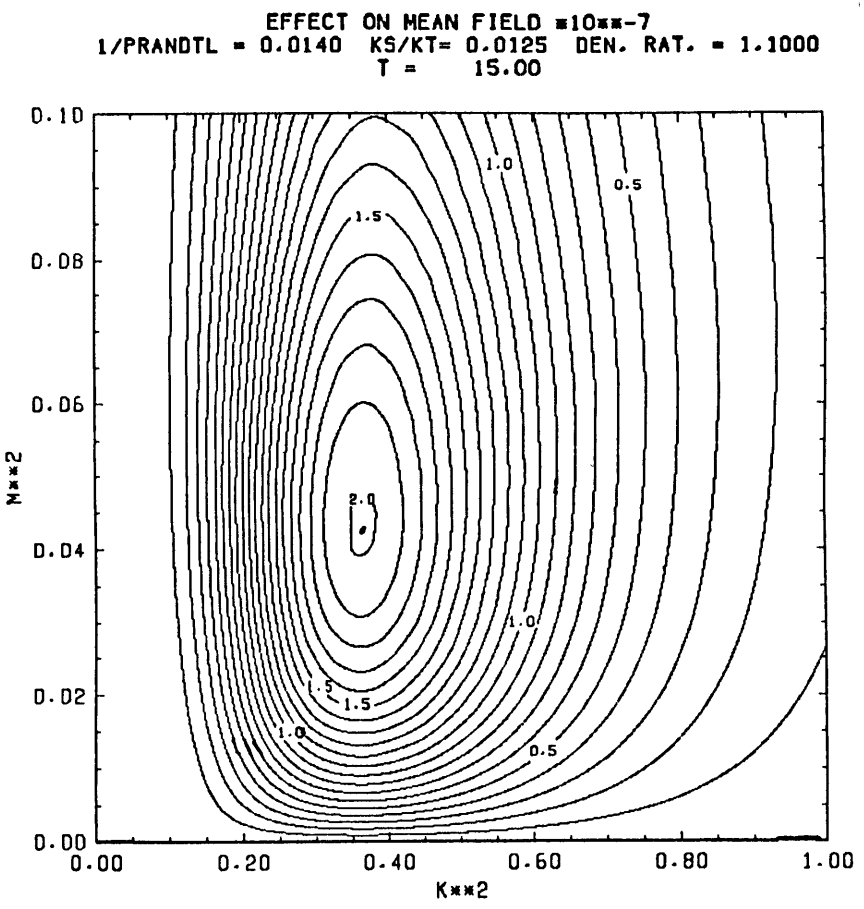

Fig. 18b.

Fig. 18. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for heat-salt case, density ratio = 1.1: a) $t = 10$, b) $t = 15$, c) $t = 20$, d) $t = 25$.

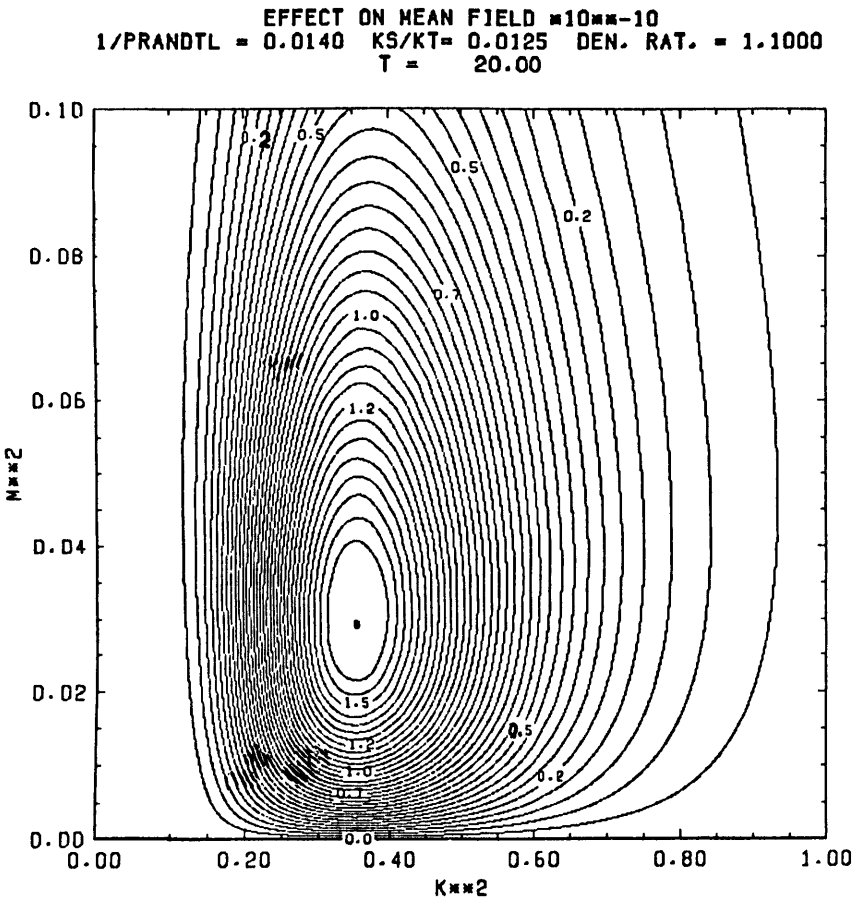

Fig. 18c.

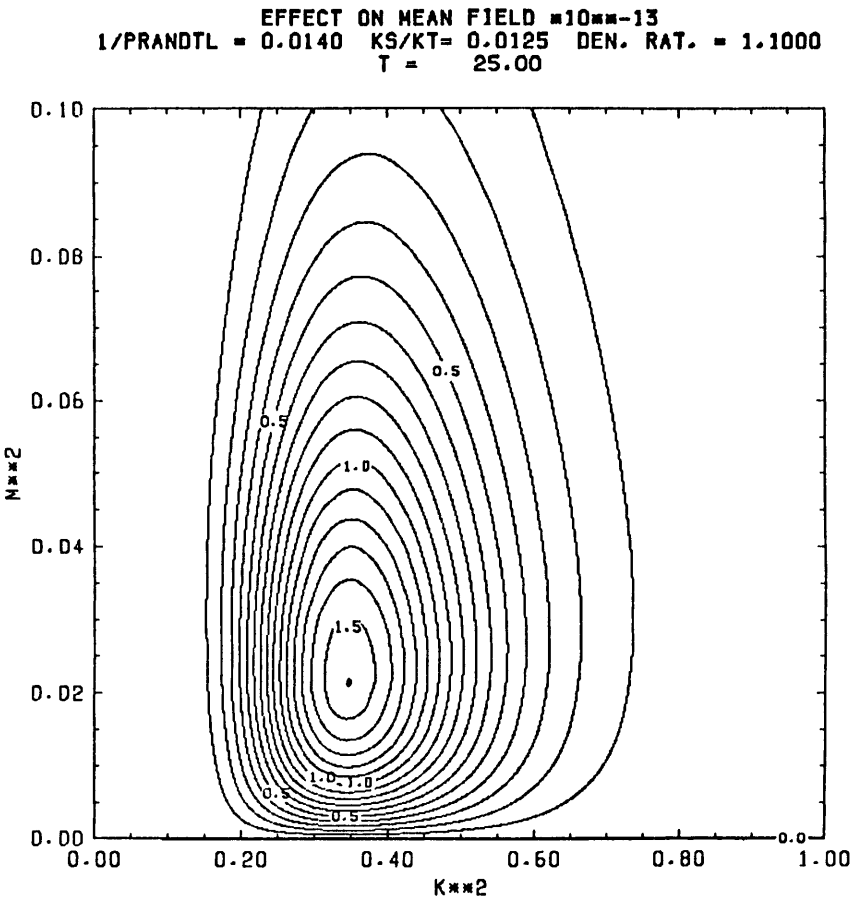

Fig. 18d.

most important horizontal wave number approaches k_f.

Figures 17 - 21 provide contours of Γ from eq. (34) at various times later than $t = 0$ but within a range where exponential growth might still be occurring. With this approach we bypass the search for a particular closure, at least for the moment. As time passes the mode with the greatest effect evolves toward the fastest growing mode at $(k_f^2,\ m^2 = 0)$. However $\Gamma_{\max}(t)$ always occurs at finite m^2. To allow a perturbation to grow from Angstroms to centimeters is an amplification of 10^8 which is more than 18 e-folding periods. This is an amplification of 10^{16} for quantities proportional to the square of the amplitudes. If $\lambda_{\max} \sim 0.5$ this criterion is reached at $t \sim 37$. Though the resulting amplification ratios are enormous they are not unreasonable given the probable small size of the initial perturbations, which may be well less than an Angstrom in displacement at the appropriate horizontal and vertical wavelengths.

The values of the density ratio which are presented in figs. 17 - 21 were chosen to be relevant for oceanic situations and laboratory experiments where layering has been observed, particularly Linden's 1978 experiments. At the earliest time shown, fig. 17a, $t = 5$ and $\lambda_{\max} t$ is only about 2.5. The behavior is swamped by the dependence on m^2 and is pushed up to the marginal curve, beyond which no data was plotted. By the the time $\lambda_{\max} t \approx 5$ the result is more along the lines of our expectations. There is little qualitative development of the pattern later on, though the scale of the figures is changed. The continual evolution of the location of $\Gamma_{\max}$ in wave number space toward k_f^2 and smaller m^2 is evident.

Figures 17 and 18 show the development of Γ for the heat-salt case while figs. 19 - 21 apply to the salt-sugar case. The qualitative behavior is identical for the

different parameter sets though the numerical values are different. The width of the peak at Γ_{max} is fairly broad but narrows as time passes in absolute magnitude. When normalized by the value of m_l^2 at Γ_{max} the width becomes fairly constant after a few e-folding periods. For example in the heat-salt case with $\Lambda = 1$ the full width at half maximum goes 0.12, 0.073, 0.054, as t goes from 10 to 25 by increment 5, while the width normalized by m_l^2 goes 1.6, 2.6, 2.4, 2.7. The large relative width indicates that the criterion of choosing m_l is not sharply selective and includes several harmonics. The plots of Γ provide only relative amplification. An initial spectrum is still required to generate relative magnitudes.

As an example, we take a $10\,m$ high region with $\Lambda = 1.1$, $\beta B = 10^{-6}\,cm{-}1$ for the heat-salt case using the nominal parameter values from the introduction, except for convenience take $K_T \sim 10^{-3}$. This example has a slightly astronomical Rayleigh number of order 10^{15}. The scales from eqs. (11) are $T \approx 100\,sec$, $L \approx 10^{-1/2}\,cm$. From fig. (3c) we get $k_f^2 \approx 0.32$, so the fastest growing wavelength is $3.5\,cm$ and $\lambda \approx 0.715$. The fundamental has $m_1^2 \approx 10^{-6}$. The first *hundred* harmonics all have $\lambda > 0.70$. It would take over 30 e-folding periods, or more than one hour, before the fundamental had a factor of 2 more amplification over the slowest of these. About the 140^{th} harmonic has $m^2 \approx 0.02$ and a vertical wavelength of $14\,cm$, while λ has only dropped to 0.69 requiring 20 e-folding periods for a relative factor of 2 in amplification. Their relative effects on the mean field however are quite different. At $t = 25$, $e^{\lambda_{max} t} \sim 10^8$ and the effect of the 140^{th} harmonic on the mean density ratio is twice that of the hundredth harmonic and orders of magnitudc greater than the effect of the fundamental.

For a laboratory example we take a $10\,cm$ height region with $\Lambda = 1.1$, $\beta B =$

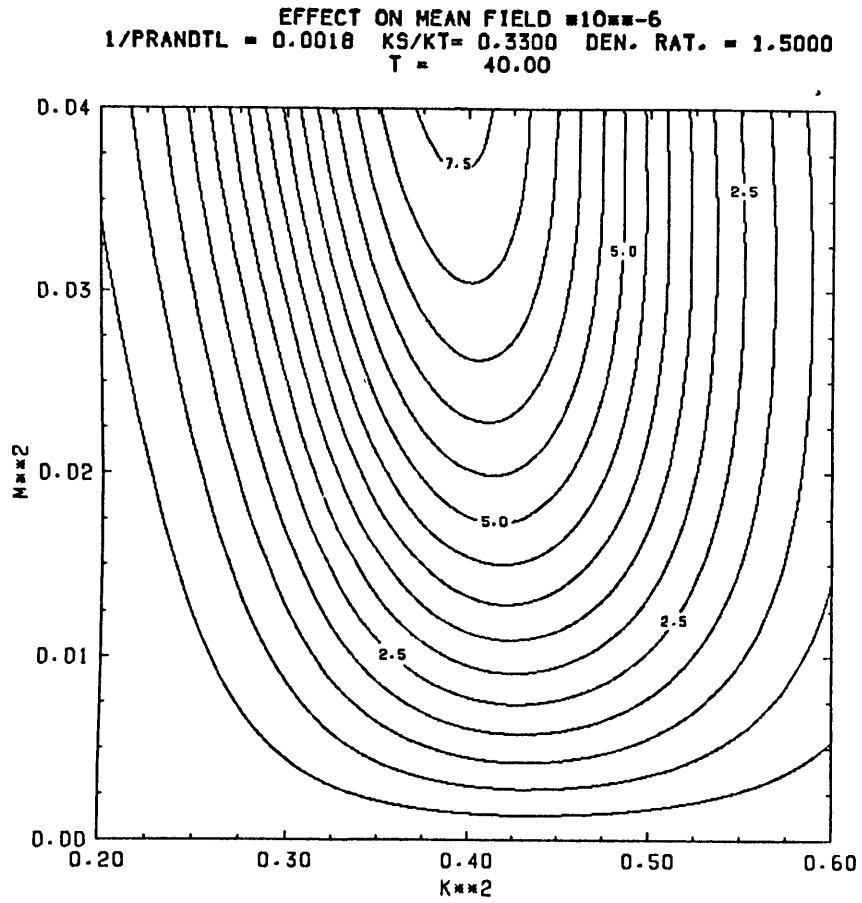

Fig. 19a.

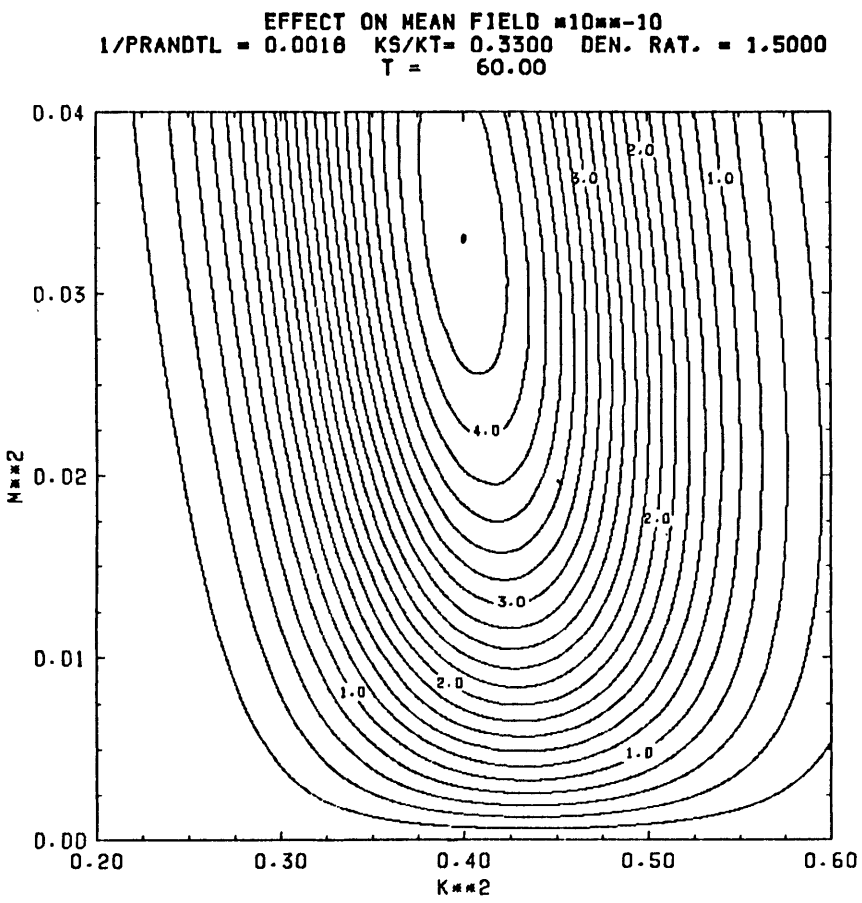

Fig. 19b.

Fig. 19. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar case. density ratio = 1.5: a) $t = 40$, b) $t = 60$, c) $t = 80$, d) $t = 100$.

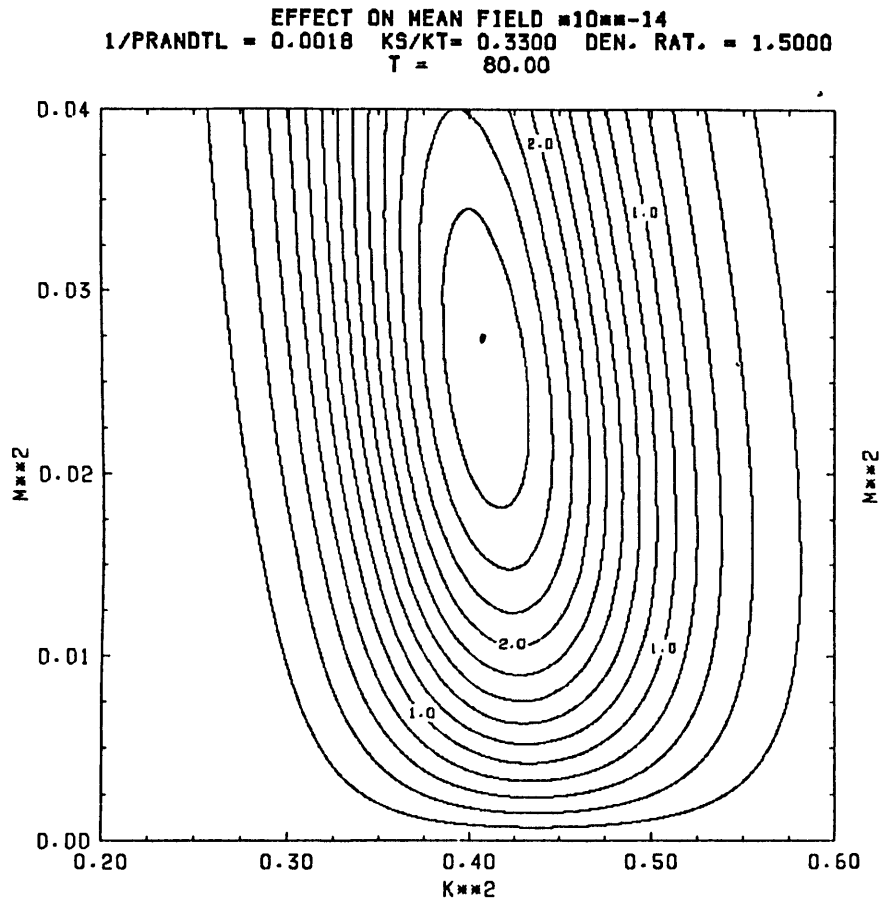

Fig. 19c.

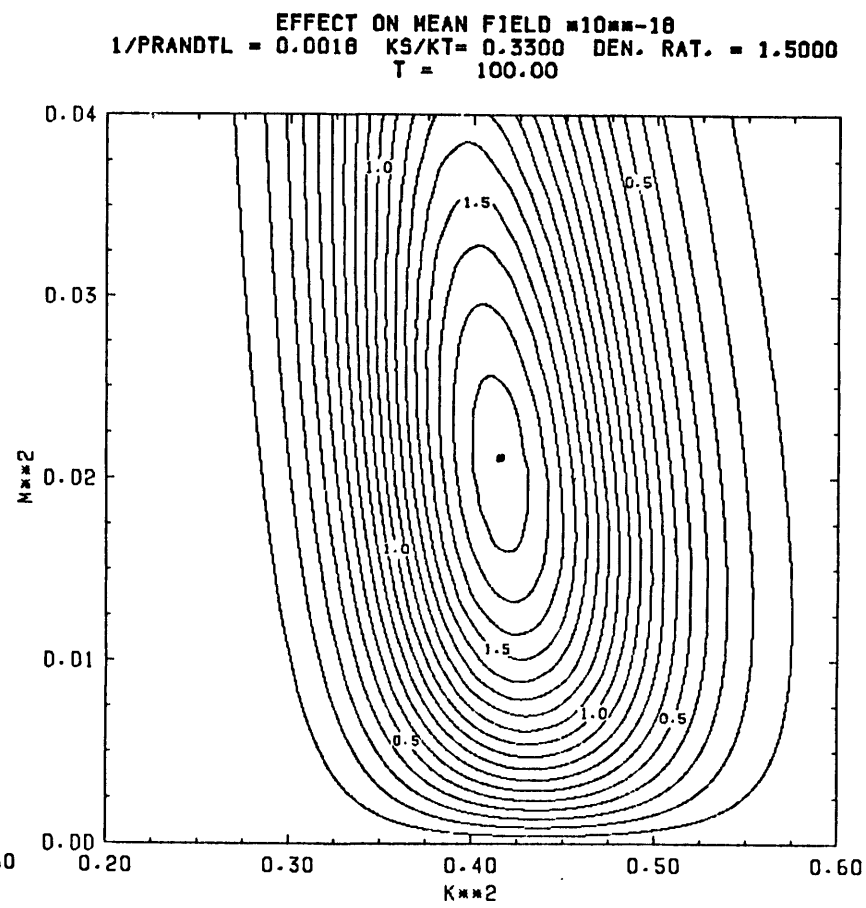

Fig. 19 d.

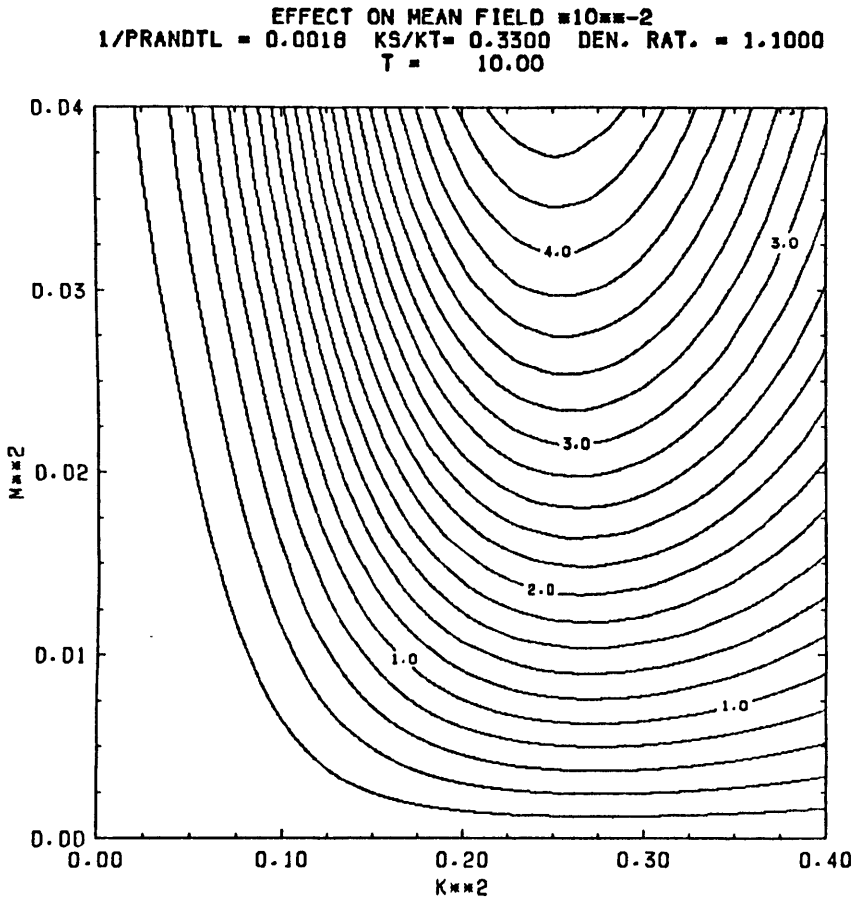

Fig. 20a.

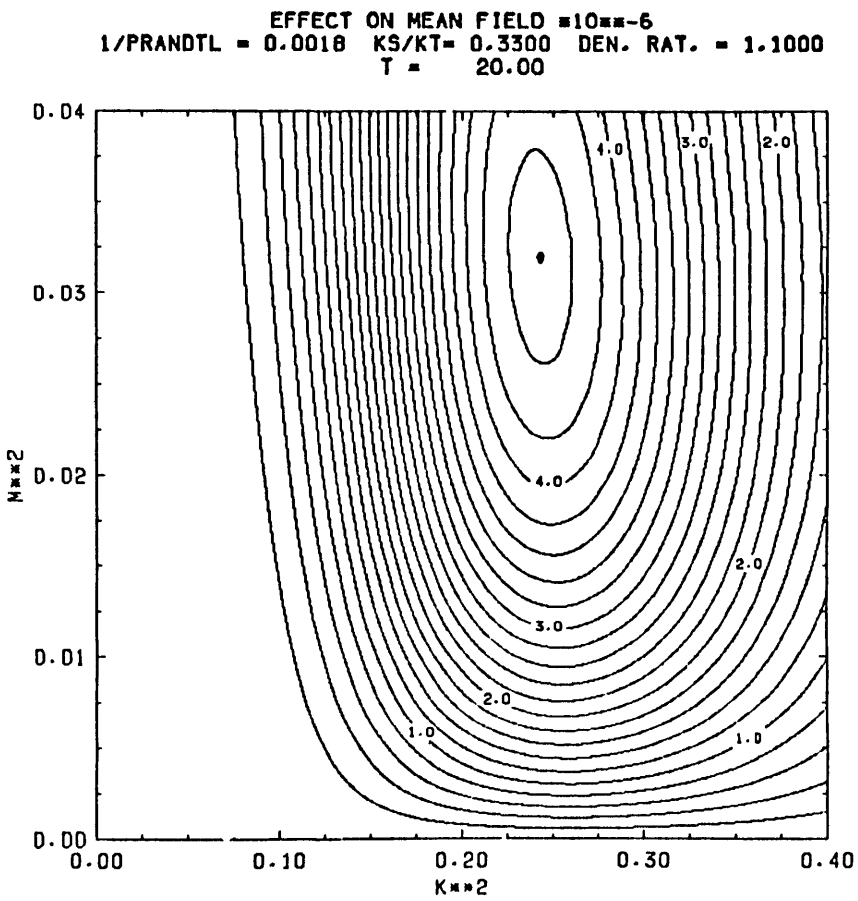

Fig. 20b.

Fig. 20. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar case, density ratio = 1.1: a) $t = 10$, b) $t = 20$, c) $t = 30$, d) $t = 40$, e) $t = 50$, f) $t = 60$.

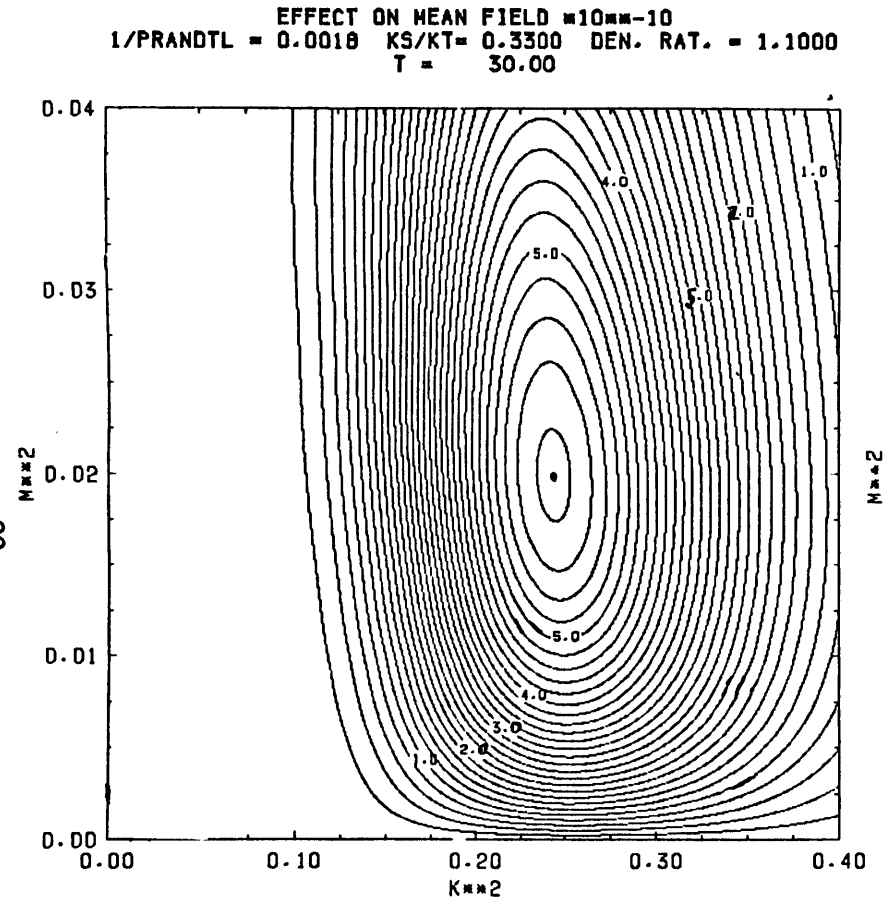

Fig. 20c.

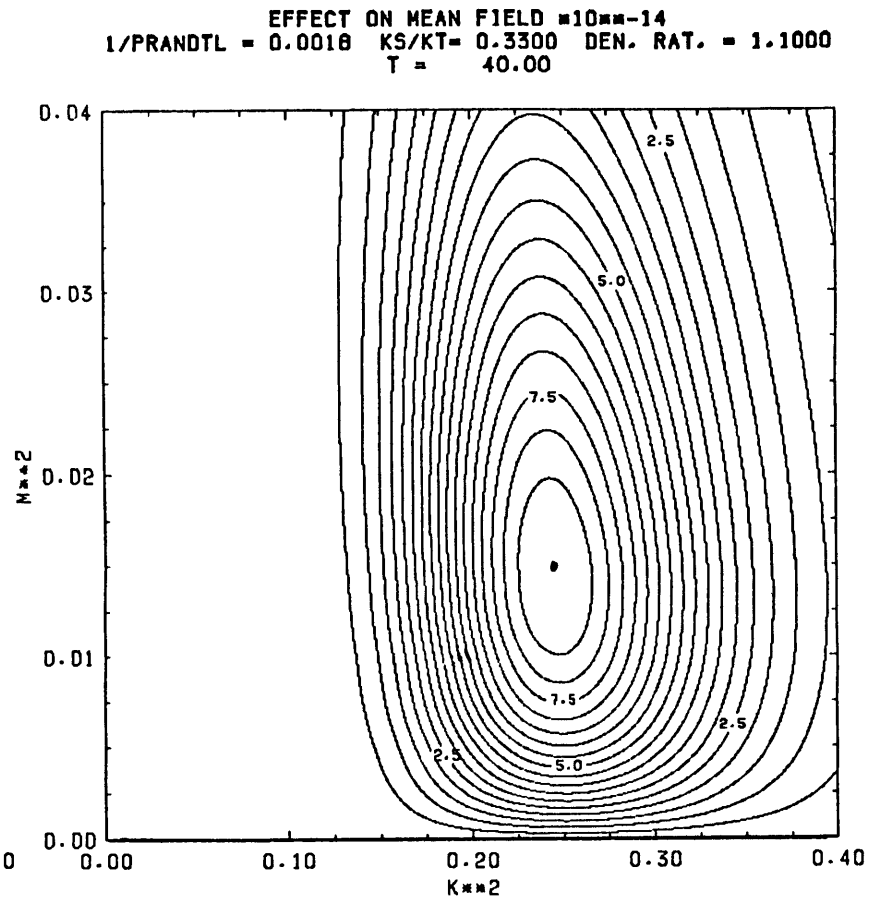

Fig. 20d.

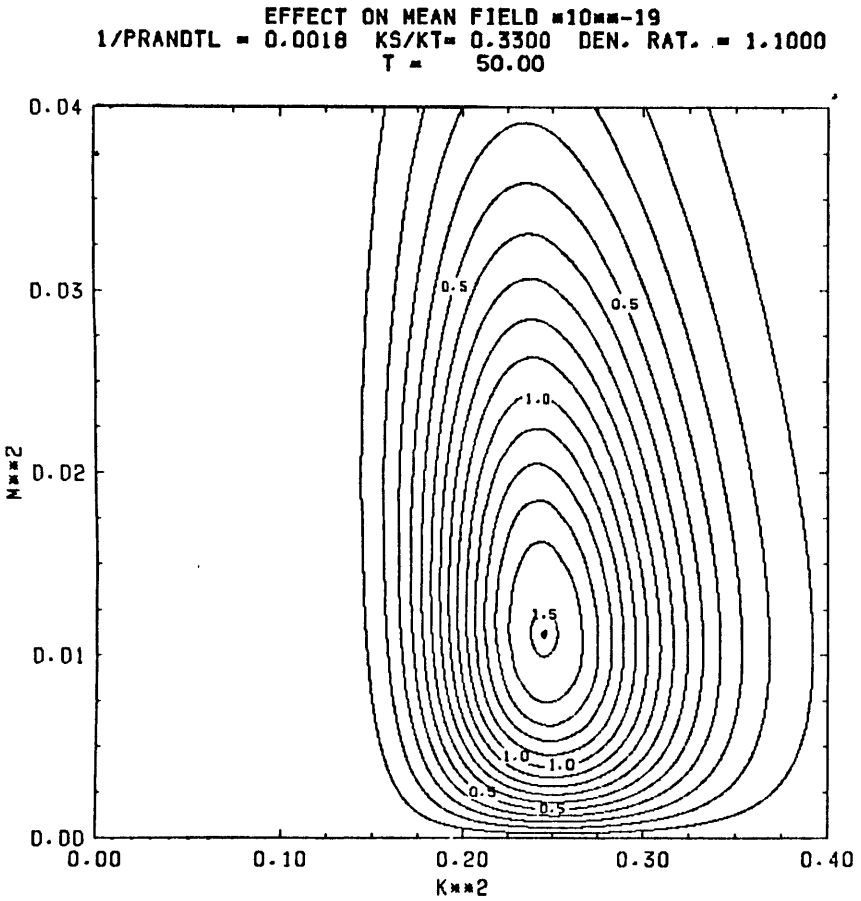

Fig. 20e.

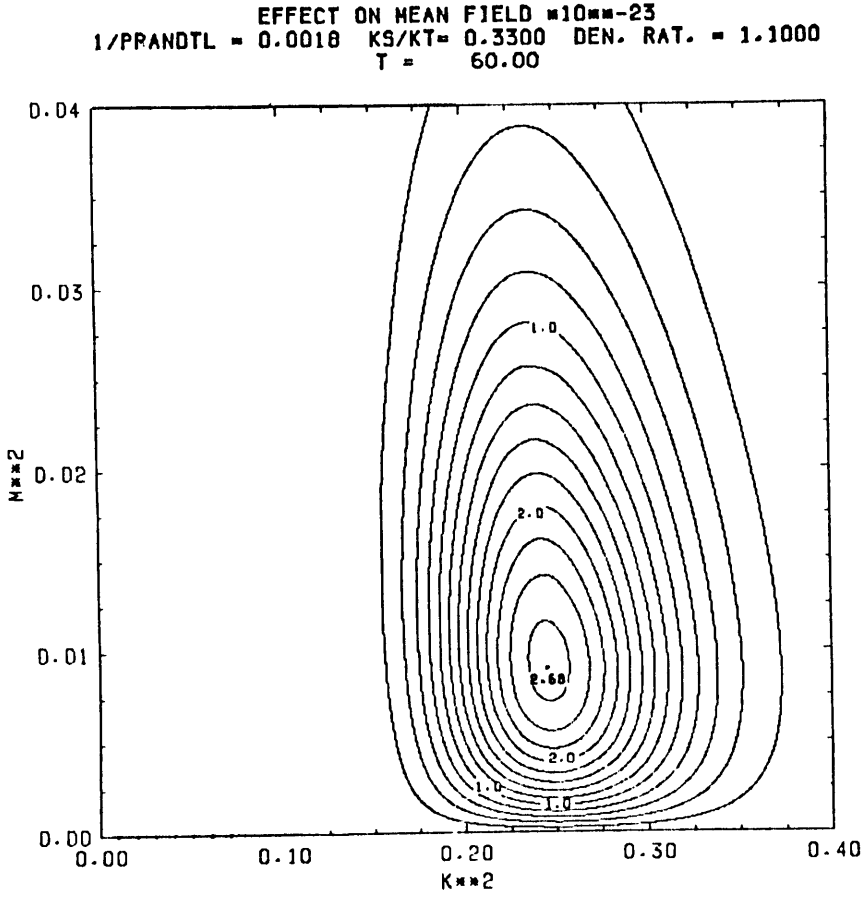

Fig. 20 f.

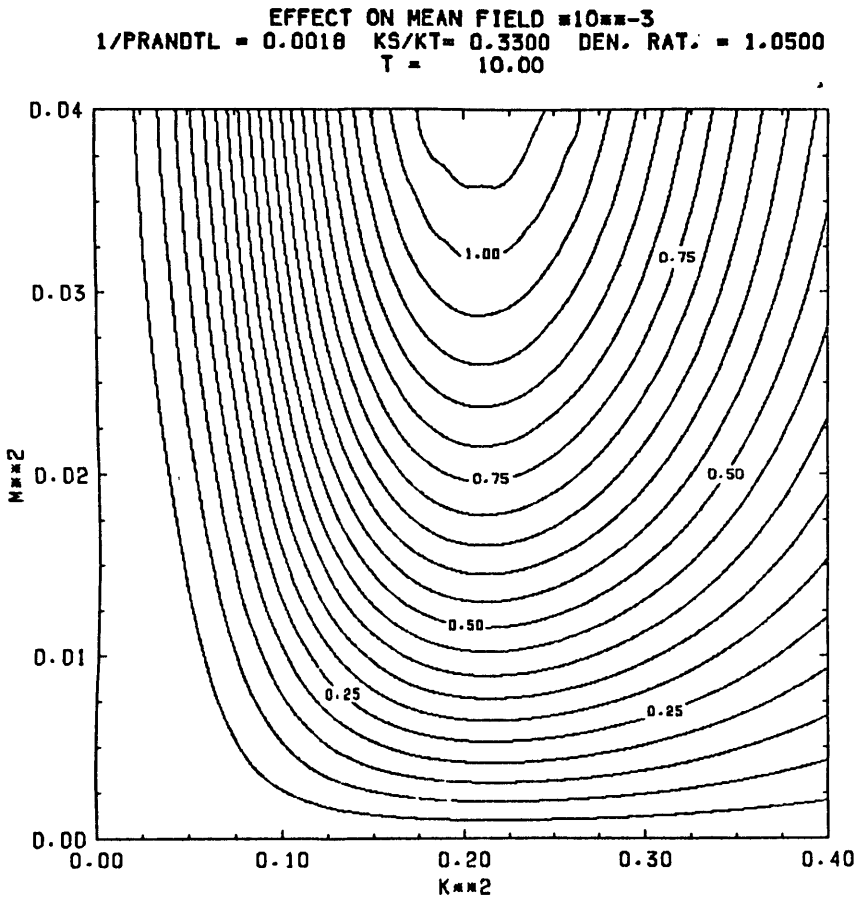

Fig. 21a.

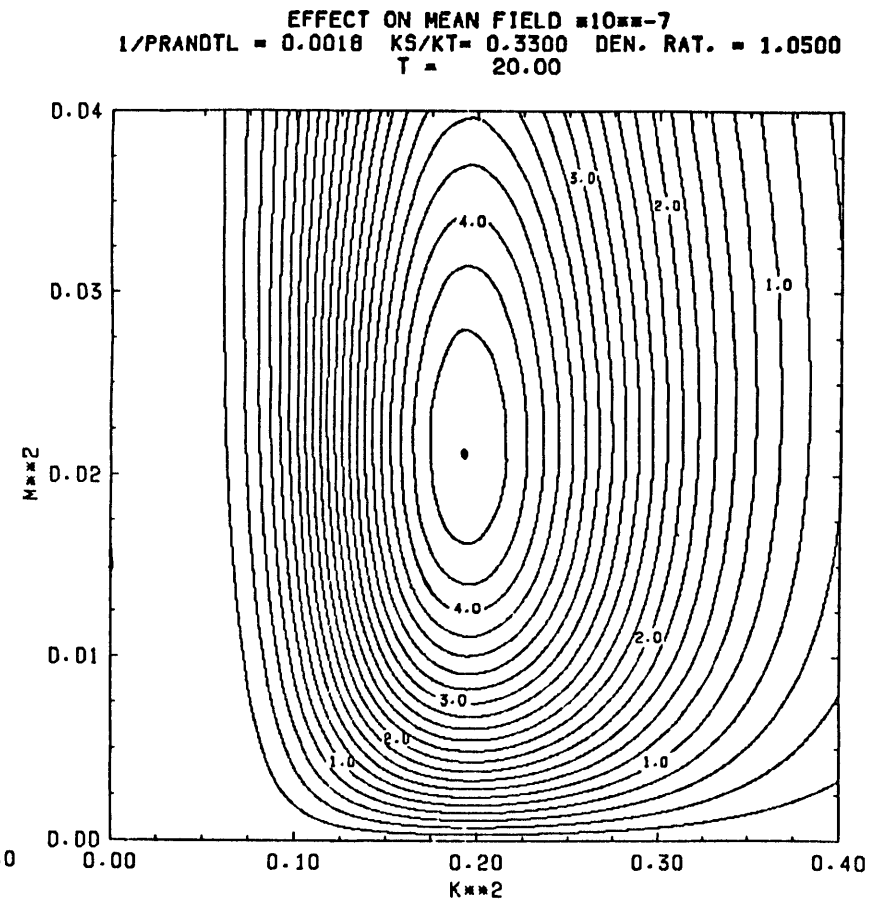

Fig. 21b.

Fig. 21. Effect on mean density ratio by growing modes as a function of horizontal and vertical wave numbers for salt-sugar case, density ratio = 1.05: a) $t = 10$, b) $t = 20$, c) $t = 30$, d) $t = 40$.

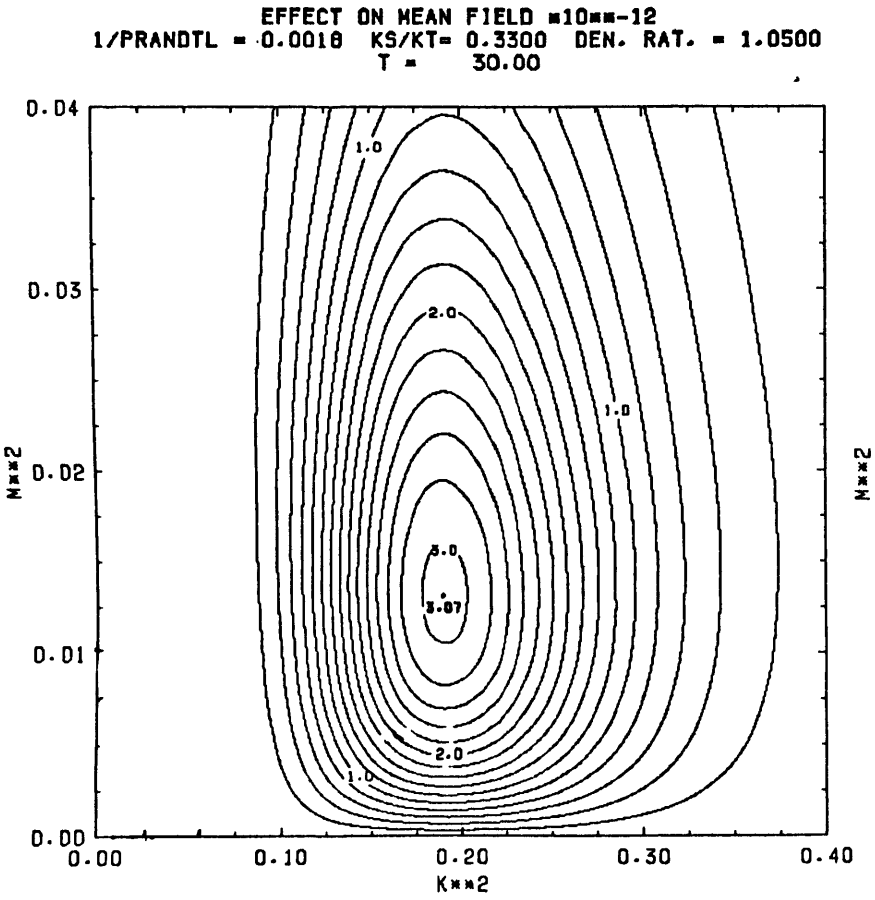

Fig. 21c.

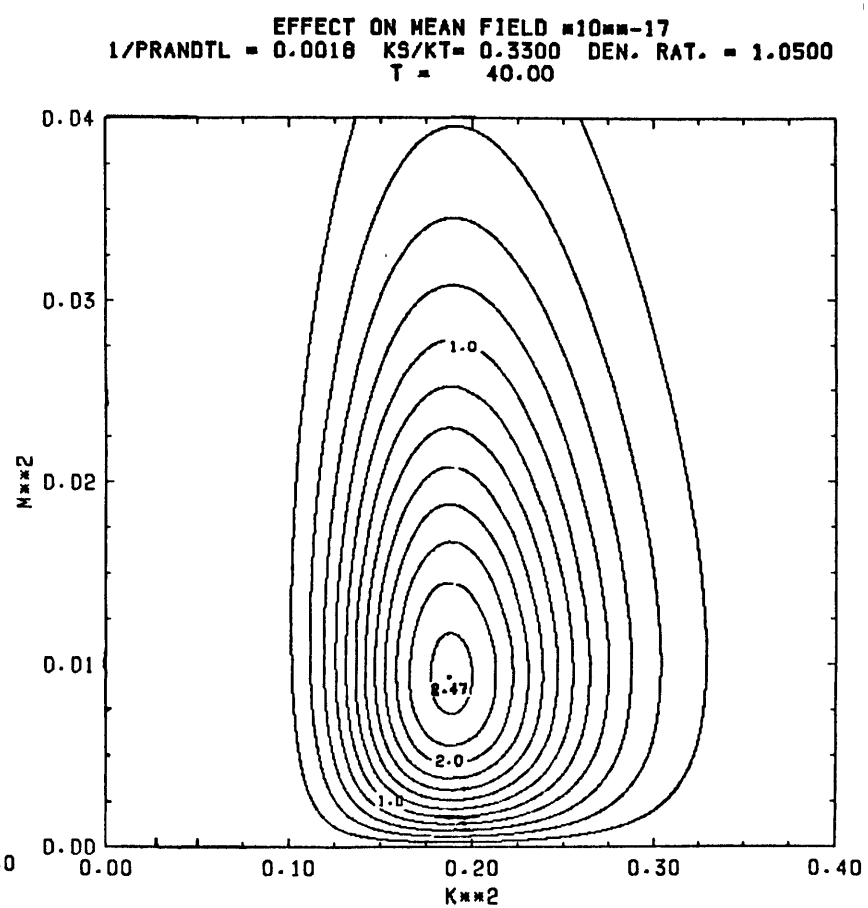

Fig. 21d.

$10^{-2}\,cm^{-1}$. The resulting values as above are $T \approx 8\,sec$, $L \approx 0.011\,cm$, $k_f^2 \approx 0.25$, $l_f \approx 0.14\,cm$, and $m_1^2 \approx 10^{-5}$. the growth rate of the fastest growing allowed mode is 0.499. Over sixty harmonics have a growth rate over 0.45. The relative effects on the mean field are similar to the heat-salt case. At $t = 50$, taken from fig. 20e, $m_l^2 = 0.011$. The derivative $\frac{\partial\lambda}{\partial m^2}\Big|_{k=k_f} = -1.05$. The relative amplification is $e^{(m_l^2 - m_1^2)\frac{\partial\lambda}{\partial m^2}t} \approx e^{0.6} \approx 2$, while $\Gamma_{\max}$ is orders of magnitude greater than $\Gamma(k_f, m_1)$, and the vertical wavelength corresponding to m_l is $0.7\,cm$.

The flux divergence ratio is defined as $\partial_z F_T / \partial_z F_S$. It is only relevant for the exponential growth period. At steady steady the vertical flux divergences must vanish. Taking the vertical derivatives of the solutions from eqs. (13)

$$Q = \frac{\partial_z F_T}{\partial_z F_S} = \Lambda \frac{(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)}{(\lambda + 2m^2)(\lambda + \gamma^2)} \quad . \tag{41}$$

The flux divergence ratio is a measure of a mode's efficiency in extracting potential energy from the mean fields within the cell. This is distinct from the flux ratio which is a measure of a mode's effectiveness in obtaining potential energy from the reservoirs at the ends of the cells or at the boundaries. It is strongly related to the formula for the potential energy change calculated in eq. (44b). Contours of the flux divergence ratio are in figs. 22. The values of Q are bounded by 0 and 1. Low values are indicative of a mode that is efficient in extracting energy from the salt field but does not lose much of it doing work on the temperature field. There does not appear to be any preference for modes with small flux divergence ratios in realized double-diffusive convection.

It is not difficult to determine the relative efficacy of different modes in extracting potential energy from the system. The gravitational potential energy

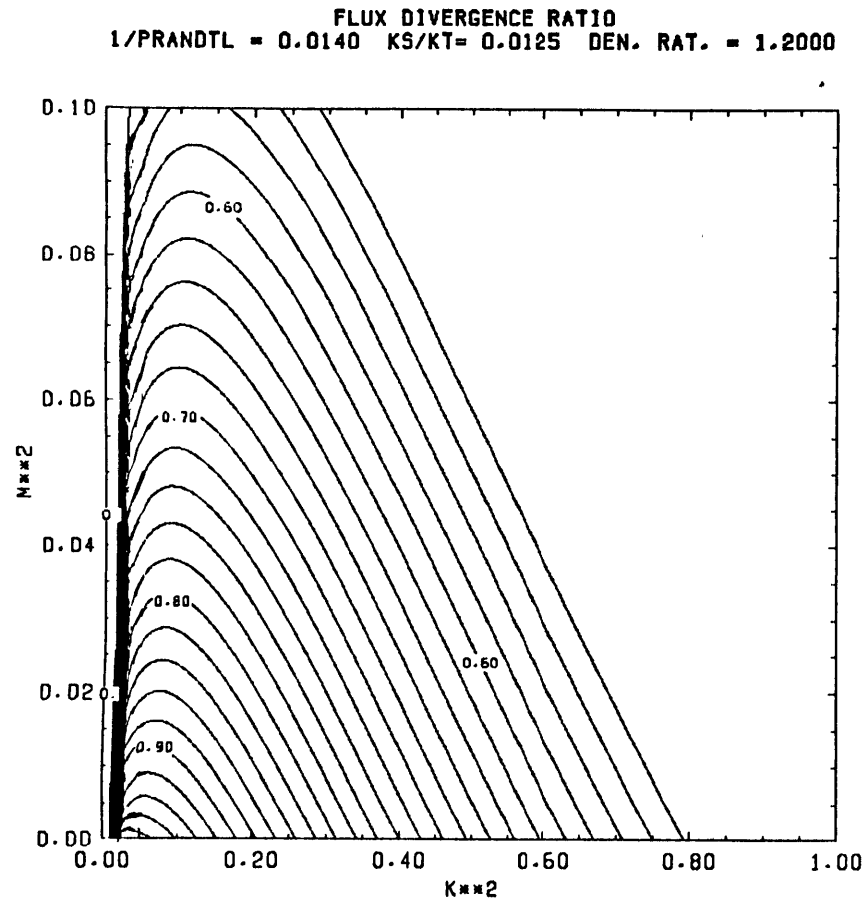

Fig. 22a.

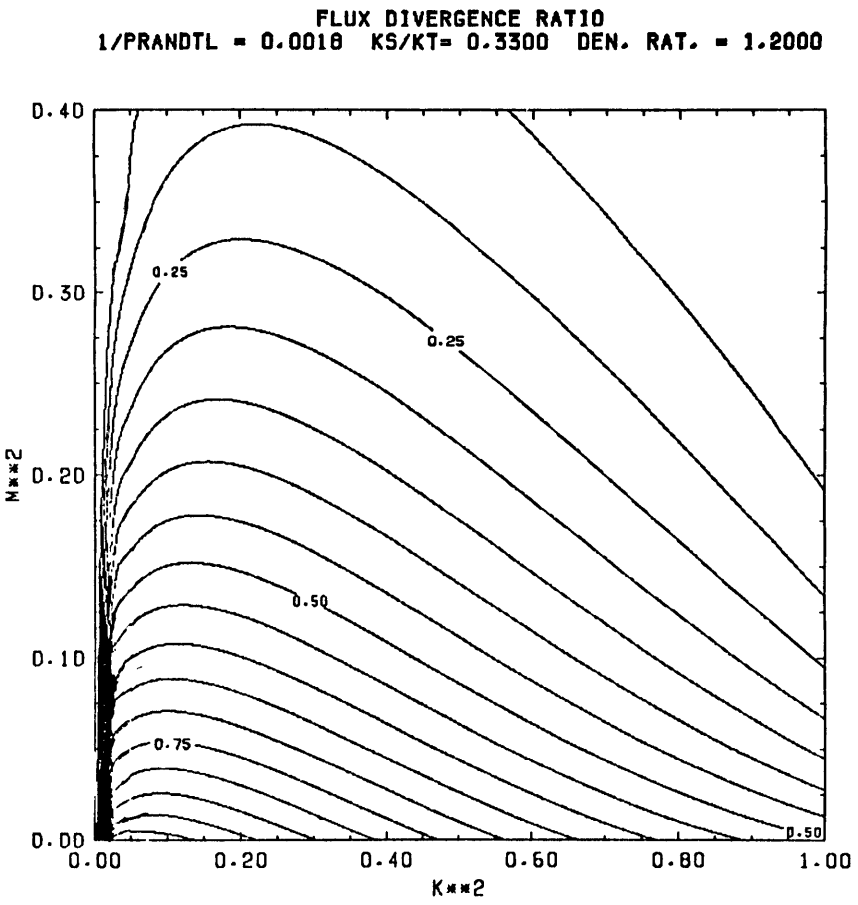

Fig. 22b.

Fig. 22. Flux divergence ratio a function of vertical and horizontal wave numbers: a) heat-salt case, density ratio = 1.2, b) salt-sugar case, density ratio = 1.2. Contours are plots of Q from eq. (41).

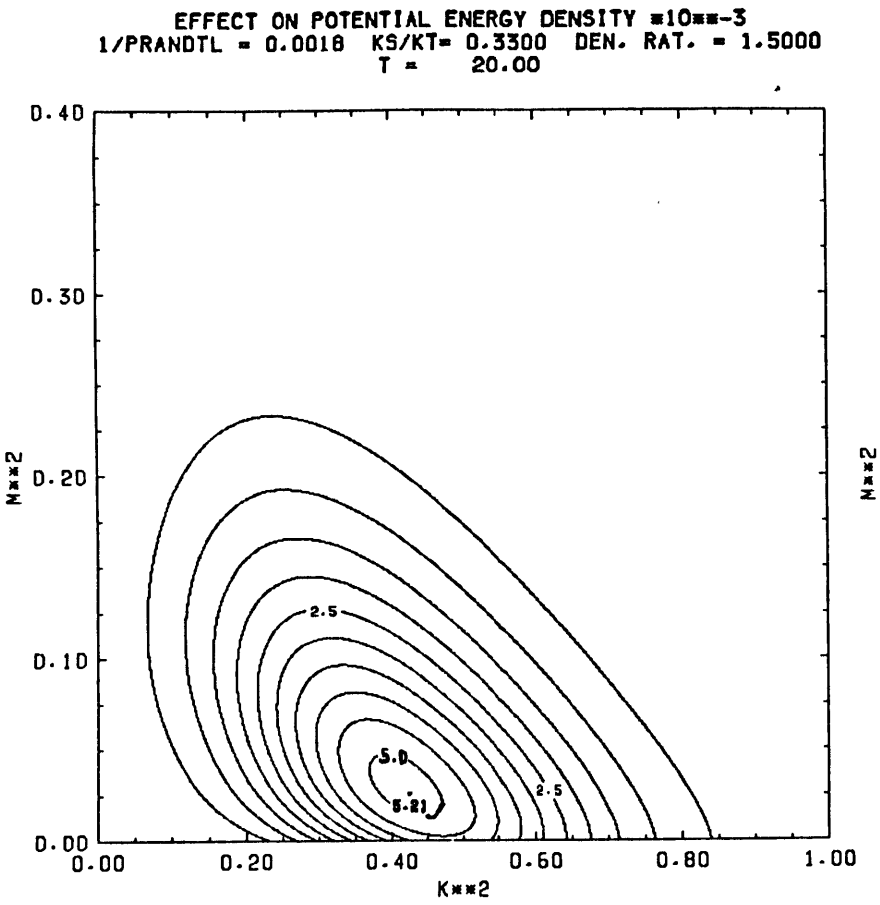

Fig. 23a.

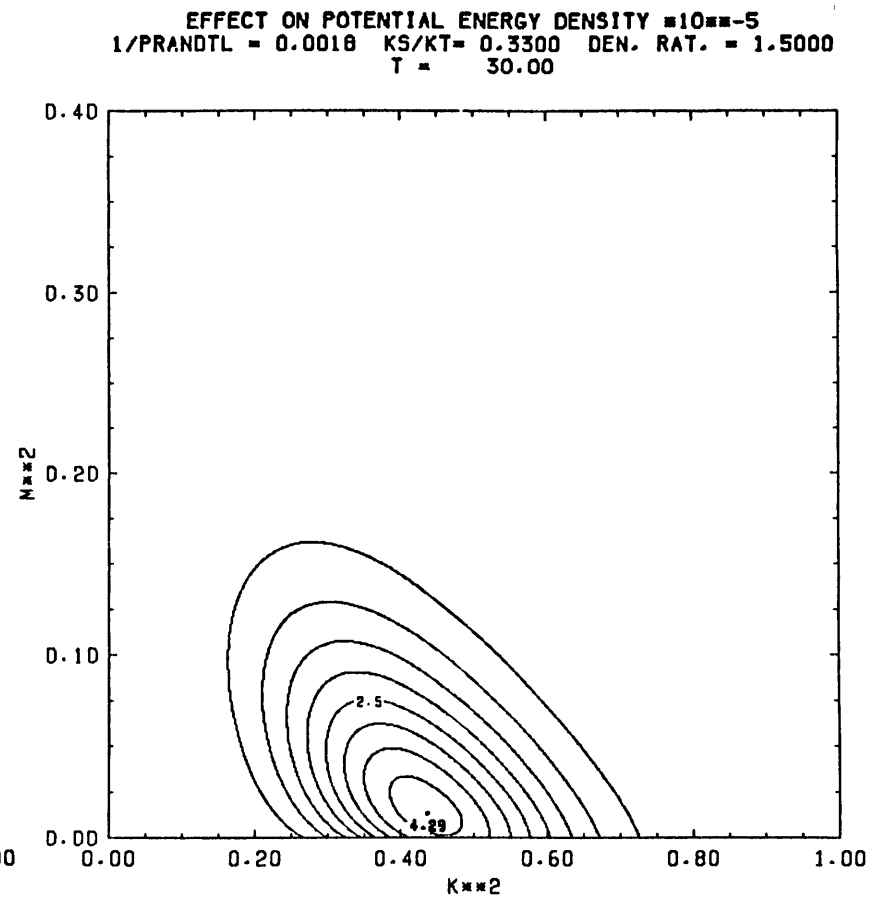

Fig. 23b.

Fig. 23. Effect on mean potential energy density by growing modes as a function of horizontal and vertical wave numbers for salt-sugar case, density ratio = 1.5: a) $t = 20$, b) $t = 30$, c) $t = 40$, d) $t = 50$. Contours are plots of δU from eq. (44b).

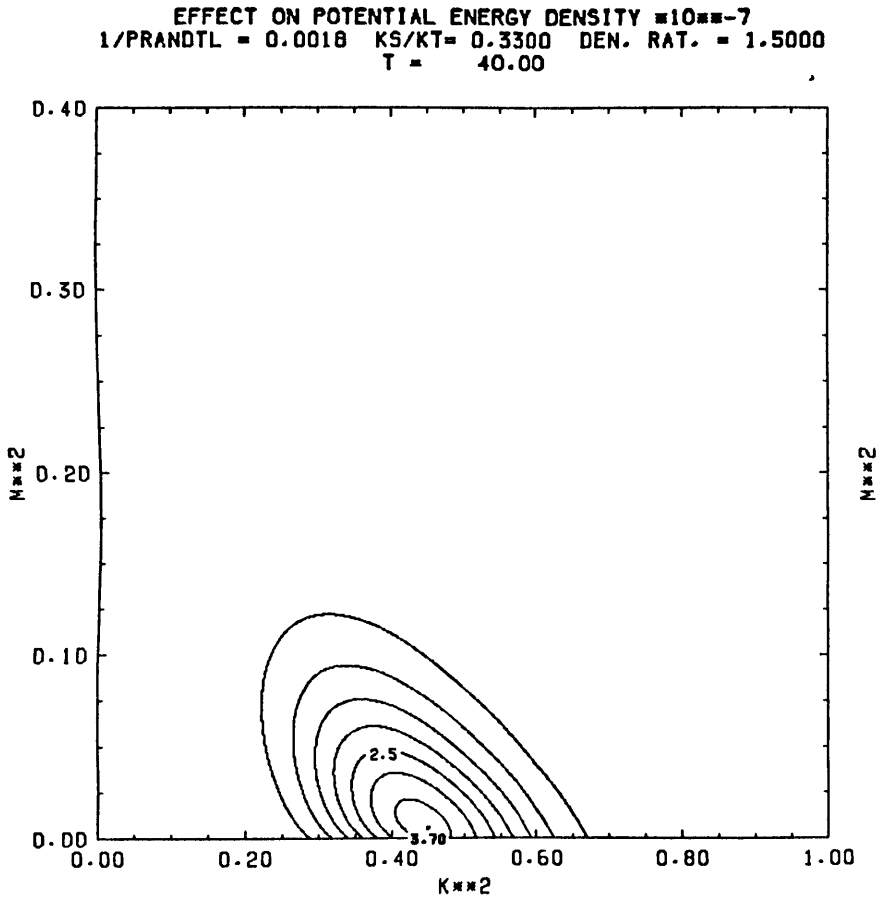

Fig. 23c.

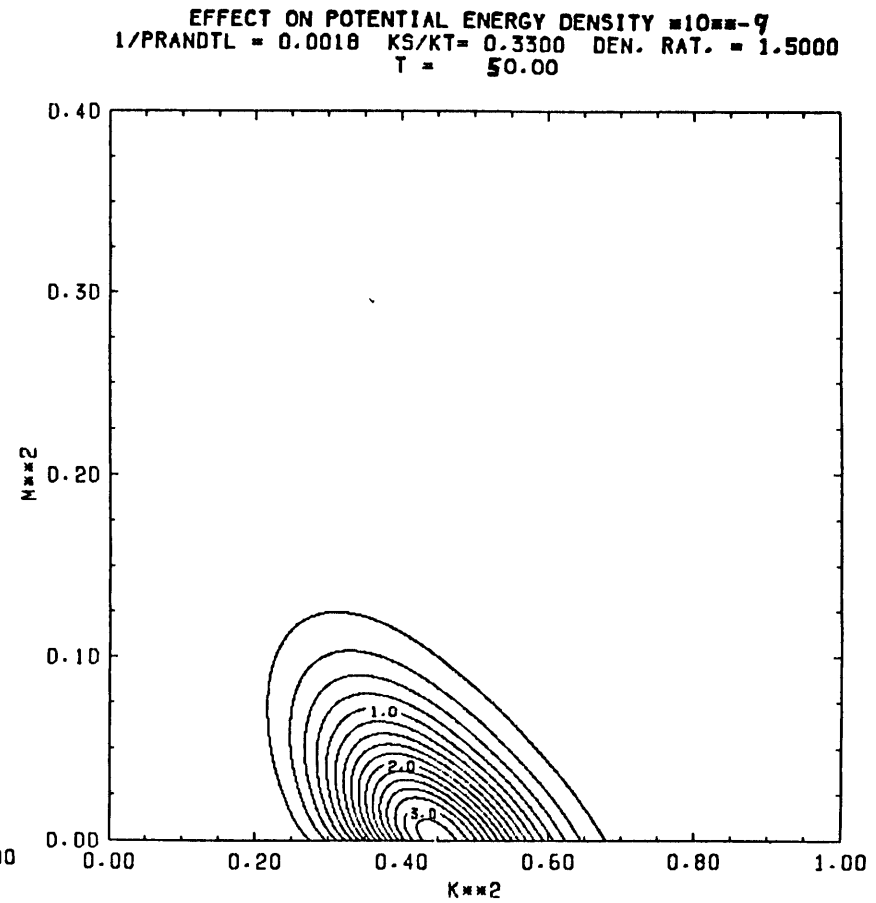

Fig. 23d.

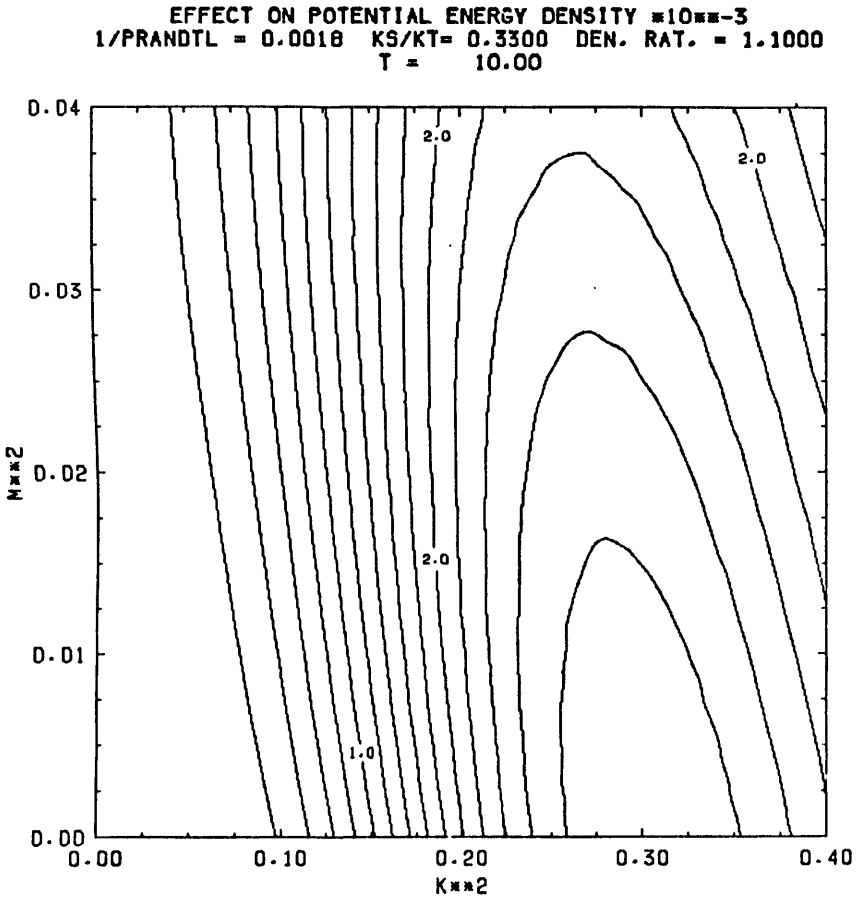

Fig. 24a.

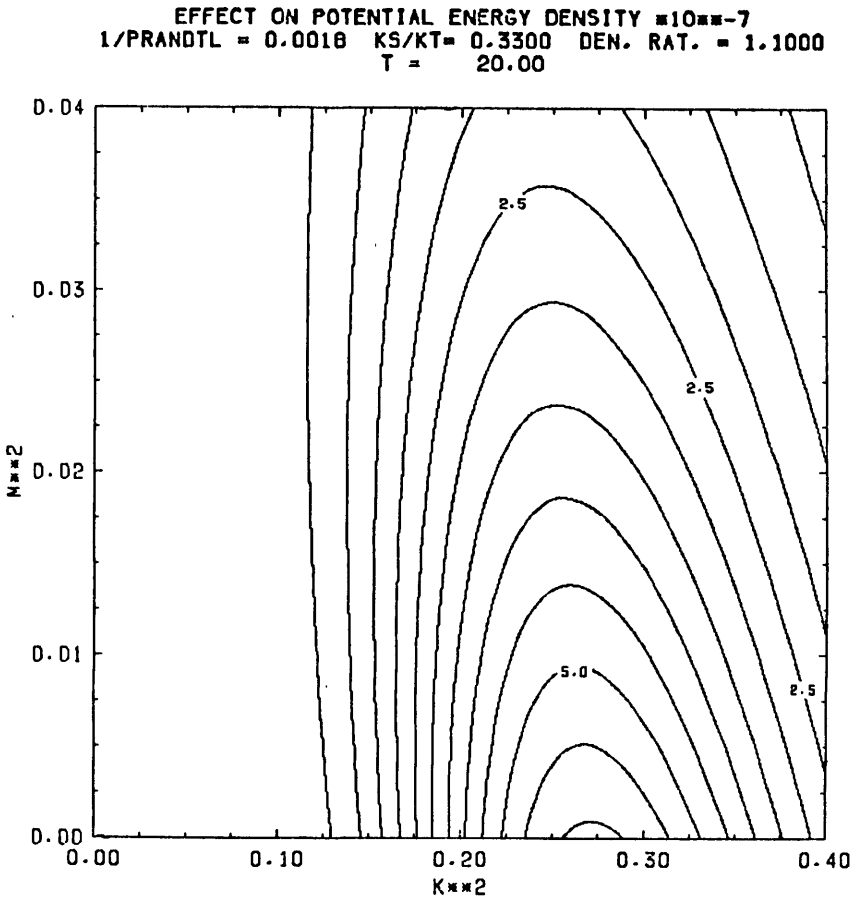

Fig. 24b.

Fig. 24. Effect on mean potential energy density by growing modes as function of horizontal and vertical wave numbers for salt-sugar case, density ratio = 1.1: a) $t = 10$, b) $t = 20$, c) $t = 30$, d) $t = 40$.

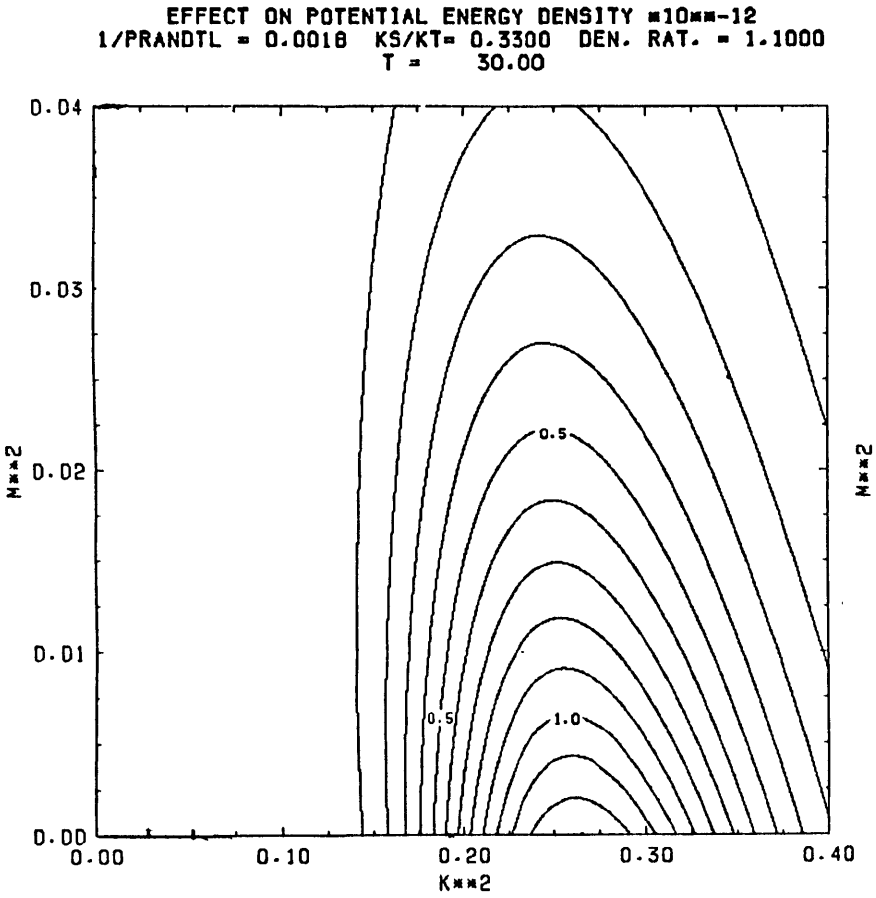

Fig. 24c.

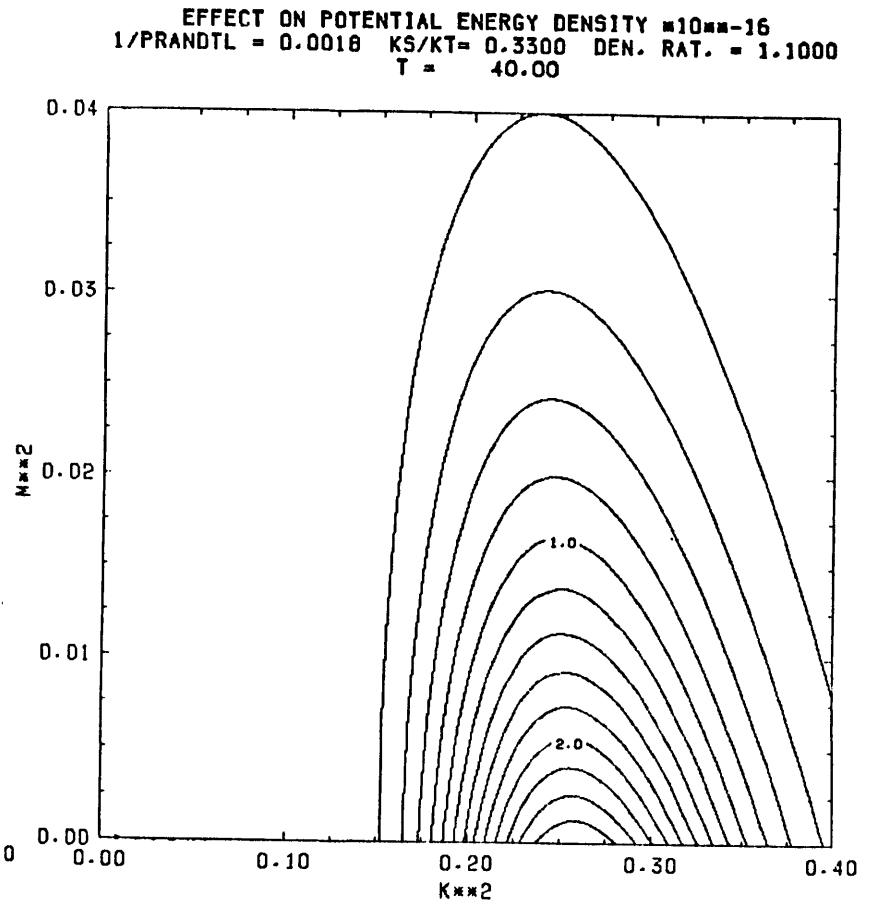

Fig. 24d.

may be defined as

$$U = \int_0^h \int_y \int_x gz\rho(z)\, dx\, dy\, dz \quad .$$

Referenced to the initial state, the amount of energy extracted is

$$\delta U = C \int_0^h gz(\beta \overline{S} - \alpha \bar{\theta})\, dz \quad , \tag{42}$$

where C is the lateral area of the integration and horizontal averaging. Putting in the effects of the growing normal mode solutions from eqs. (13) and vertically averaging as well as dividing by the volume considered gives a change in the average potential energy density in nondimensional form of

$$\delta \mathcal{U} = \frac{1}{h} \int_0^h z \sin(2mz) \frac{\hat{w}^2 m}{2} e^{2\lambda t} \left[\frac{1}{(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)} - \frac{\Lambda}{(\lambda + 2m^2)(\lambda + \gamma^2)} \right] dz \quad . \tag{43}$$

Integrating over one wavelength regardless of its length, so $h = \pi/m$, gives a vertical average of

$$\delta \mathcal{U} = \left[\frac{1}{(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)} - \frac{\Lambda}{(\lambda + 2m^2)(\lambda + \gamma^2)} \right] \frac{\hat{w}^2 e^{2\lambda t}}{4} \tag{44a}$$

$$= \frac{[1 - Q(k^2, m^2)]}{[(\lambda + 2\tau m^2)(\lambda + \tau\gamma^2)]^{-1}} \frac{\hat{w}^2 e^{2\lambda t}}{4} \quad . \tag{44b}$$

This can be calculated easily as a function of (k^2, m^2) for various parameter sets as a function of time was as done above for the analysis of the nonlinear change of the basic stability characteristics. Some results are shown in figs. 23 and 24.

The limit of zero vertical wave number should be examined carefully. In this limit the amplitude of the mean field adjustment vanishes linearly to first order.

However, averaging over a full period of a $\sin(2mz)$ shaped perturbation for an arbitrary potential energy spectrum results in a linearly diverging singularity as m^{-1} times the amplitude spectrum. These cancel to give a finite result with a nonsingular finite limit as $m \to 0$. Taking the derivative of the potential energy density with respect to the square of the vertical wave number it becomes clear that $\frac{d(\delta U)}{dm^2}$ is small and can be of either sign for small m^2 at $t = 0$. Analytically

$$\frac{\partial \delta U}{\partial m^2} = \frac{w^2}{4} e^{2\lambda t}\Bigg(- \frac{[(\frac{\partial \lambda}{\partial m^2} + 2\tau)(\lambda + \tau\gamma^2) + (\lambda + 2\tau m^2)(\frac{\partial \lambda}{\partial m^2} + \tau)]}{(\lambda + 2\tau m^2)^2(\lambda + \tau\gamma^2)^2} + \Lambda_o \frac{[(\frac{\partial \lambda}{\partial m^2} + 2)(\lambda + \gamma^2) + (\lambda + 2m^2)(\frac{\partial \lambda}{\partial m^2} + 1)]}{(\lambda + 2m^2)^2(\lambda + \gamma^2)^2} + 2t \frac{\partial \lambda}{\partial m^2} \delta U \Bigg) \quad . \tag{45}$$

So, eventually $\frac{\partial \delta U}{\partial m^2} < 0$ everywhere in the parameter range of the growing modes since $\frac{\partial \lambda}{\partial m^2} < 0$ for k near k_f.

One could perform some of these calculation on single component convection. In particular, calculation of the effect on the mean field and potential energy change could be made. However the change of the mean field only changes the growth rates of other modes by changing the scaling time which is proportional to the square root of the gradient. A dramatic dependence such as is on the density ratio in the double-diffusive problem is unavailable.

Some further approximations can be made to the results obtained by eqs. (36) and (37), though it might be difficult to have them met. However the resulting calculation emphasizes the problems of the time dependence. For m^2, τ, $(\Lambda - 1)$ all much less than one

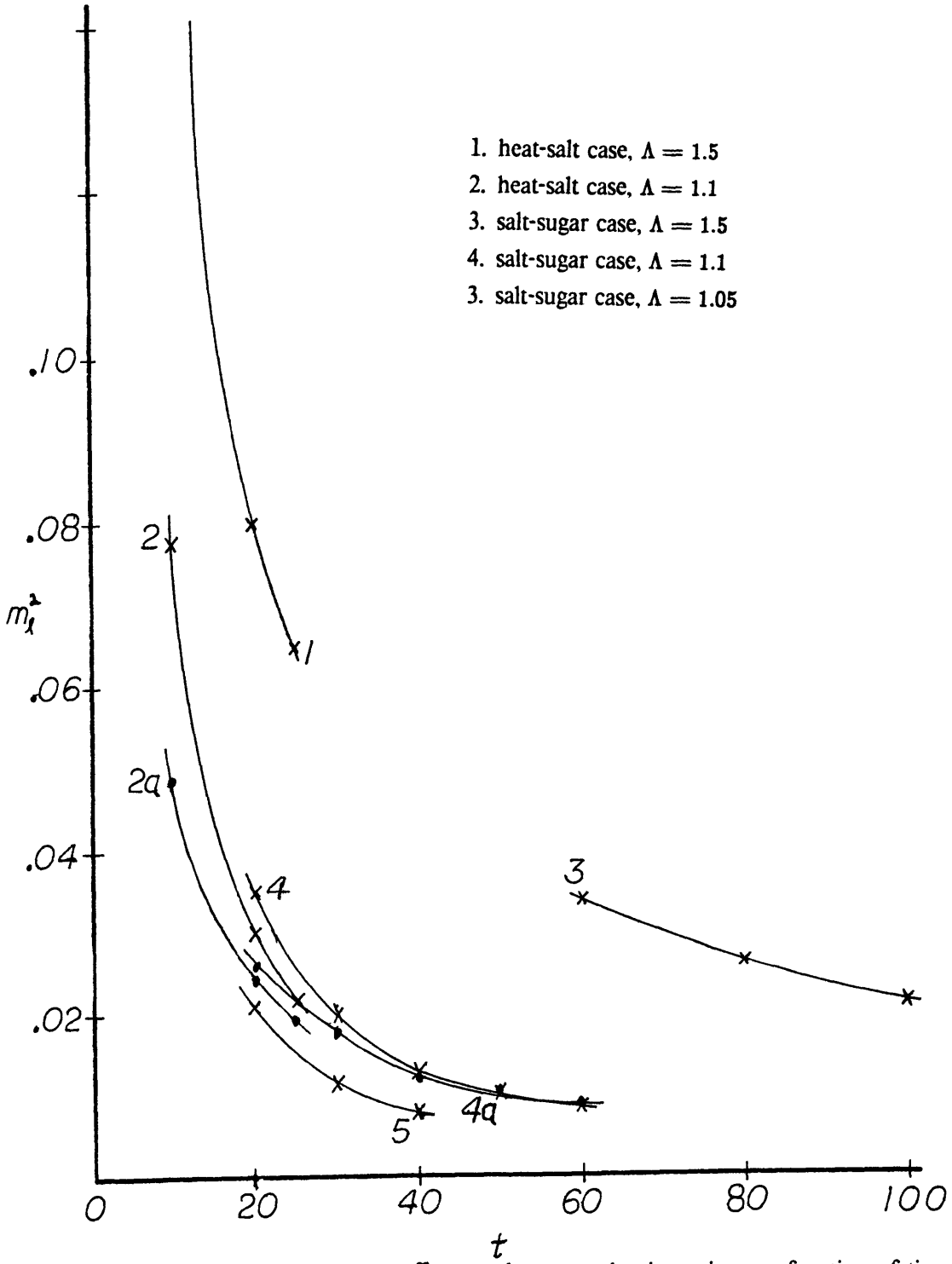

Fig. 25. Vertical mode with the greatest effect on the mean density ratio as a function of time. Points for lines 1 – 5 are obtained from the location of Γ_{max} in figs. 18 – 22, respectively. Lines 2a and 4a are calculated from eq. (48) with data from figs. 9.

$$\Gamma \sim \frac{m^2 k e^{2\lambda t}}{\lambda^2(\lambda + k^2)} \quad . \tag{46}$$

To search for the strongest vertical mode by this criterion take $\frac{\partial}{\partial m^2}$ and divide out common factors to get

$$0 = 1 + 2tm_l^2 \frac{\partial \lambda}{\partial m^2} - m_l^2 \frac{\partial \lambda}{\partial m^2}\left[\frac{3\lambda^2 + 2\lambda k^2}{\lambda^2(\lambda + k^2)}\right] \quad . \tag{47}$$

In eq. (47) the ratio of the third term to the second is $(3\lambda + 2k^2)/[2\lambda t(\lambda + k^2)]$, which is of order $1/(\lambda t)$, while the ratio of the third term to the first is clearly just equal to the third term, so when $\lambda t \gg 1$,

$$m_l^2 \sim \frac{1}{2t\frac{\partial \lambda}{\partial m^2}} \quad , \tag{48}$$

where the derivative is evaluated at $(k_f, m = 0)$. Some curves of this are plotted in fig. 25 and compare very well with the corresponding curves from figs. 19 and 21 also included in fig. 25. The same result can be obtained by making the same approximations in eqs. (26) and (27) which makes the balance of terms more intuitively clear, but this path of derivation is shorter. [Without being very careful, eq. (26) is dropped and k is taken as k_f. Assuming $m^2 \ll 1$ then leaves only the first line of eq. (27) which then requires the vanishing of the first factor.]

Still further approximations can be made. In the same limits used above with $k = k_f$, so $\frac{\partial \lambda}{\partial k^2} = 0$, approximating eq. (18) yields

$$\frac{\partial \lambda}{\partial m^2} \sim \frac{-1}{k^2(2\lambda + k^2)} \quad . \tag{49}$$

Now substituting from eq. (49) into eq. (48) the result obtained is

$$m_l^2 \sim \frac{k_f^2}{t}\left(\lambda_f + \frac{k_f^2}{2}\right) \quad . \tag{50}$$

Further substitution of the approximations, also made in the same limits, from eqs. (16) while keeping only the terms lowest order in ϵ leads to the simple result that

$$m_l^2 \sim \frac{\epsilon^{1/2}}{t} \quad . \tag{51}$$

The result of eq. (51) that has m_l independent of the fluid parameters in the limits of small τ and η is consistent with the result that each pair of lines 1 and 3, and lines 2 and 4, from fig. 25, are nearly on the same curve and have the same density ratios. The actual values obtained with eqs. (50) and (51) are considerably smaller than those from eq. (49), are much less acurate, and are not plotted.

As time goes by there is a continual shift to preferring taller vertical modes for disrupting the system. The smaller the initial noise spectrum the longer the amount of time needed for any mode to reach sufficient amplitude for layers to become distinct. After long times, modes with a growth rate larger by even a small amount gain an advantage over modes with large vertical wave number since the growth rate increases with decreasing vertical wave number in the parts of wave number space of interest.

As noted earlier while discussing fig. 6, there is a dramatic increase in growth rate as $\Lambda \rightarrow 1$. The maximum growth rate continues to grow when $\Lambda < 1$, as Λ decreases, which is still a well defined problem. If localized density ratios approach one in regions with vertical extent much greater than the finger width but much smaller than the fully system height, the results should be quite dramatic. In the course of this process the local mean gradient will also be altered so that the time

and length scales shift proportional to the reciprocals of the square root and the fourth root of the salinity gradient, respectively, but this effect is small compared to the dependence of the growth rate on the density ratio. This points out a major distinction with one component convection where the latter scaling effect is the only one that might lead to fast growing instabilities, but this is more than countered by the smaller vertical restrictions.

The harmonics will not take over as quickly as the calculations for figs. 18 – 22 indicate. The fundamental though unable to stabilize the system will still try to do so; it increases the density ratio in the interior as it decreases it at the extremes, the vertical harmonics will sense a weaker driving potential and hence have slower growth rates. To quantify this would require a nonlinear time dependent growth rate calculation and is probably not worth the effort, since simple sinusoidal vertical structure functions could not be used. In a highly supercritical system it will still be unable to succeed in preventing the growth of the higher modes. Just how poor this assumtion is is demonstrated by Linden's (1978) experiments. In the experiment shown in fig. 16 an e-folding period is 20 sec and about 100 of them elapse before the first harmonic is visually obvious. For the experiment of fig. 17 an e-folding period is about 30 sec and harmonics higher than the first take over after about 500 e-foldings. Higher modes also grow in one component convection but in that case lead to the characteristic chaotic high Rayleigh number convection.

In single component convection modes with small vertical extent disrupt the entire laminar cell structure not just regions of it. This difference in behavior is because of the difference in the relation of the horizontal and vertical scales as well as the creation of regions with larger exponential growth than the primary modes in the double-diffusive case. The shorter cells for single convection also have

smaller horizontal wavelengths. Hence, vertical harmonics in this case are generally disruptive of the cellular structure. The vertical harmonics in double-diffusion have almost no difference with the fundamental in the preferred horizontal wavelength, so the harmonics do not disrupt the cellular planform.

There is some horizontal scale shift in the direct double-diffusive problem that is indicated in this theory and in the observations reported. In layers where the density ratio has been driven down to near one, vertically local modes that have the same wavelength as the fastest growing modes which have altered the mean field, will have their growth rates increased. However, the fastest growing modes in the layers of reduced density ratio have a larger wavelength than the initially important modes. This is only the case when Λ is near one as is clear in fig. (7) taken from Schmitt (1979a). This development of larger scale motions in the disordered layers has been noted in the literature (Linden 1973) and by this author's observations of motion picture films by Schmitt (provided through personal communication). Further; even though the modes with larger wavelengths were not the dominant modes of the entire system they probably have significant amplitude by the time the low density ratio regions develop so that when they have the largest growth rates in the system they become rapidly significant.

The regions of increased density ratio will remain cellular. Not only is there little change in the fastest growing horizontal wave number with increase in vertical wave number as seen in figs. 2 – 5, but there is also little shift in the fastest growing wave number with increase in density ratio as seen in fig. 7. It is encouraging that the shifts in density ratio that are part and parcel of harmonic instability theory do not even suggest the need for a wholesale restructuring of the cellular planform in the regimes where it should remain intact.

The poorest assumption in this calculation is the choice of the initial noise spectrum. Neither the ocean not the laboratory are likely to have a spatial white noise spectrum in the parameter region and spectrum section of interest. The question of the shape of the noise spectrum is probably only relevant to the laboratory work where the base state initial condition are so well ordered. At the spatial scales of interest only two sources of perturbations are occur, thermal noise and fossil turbulence. Thermal noise is always present and salt fingers are very long wavelength phenomenon in the realm of the random thermal acoustic modes. In this range the energy per mode is essentially independent of k and there is little distinction in the behavior of liquids and solids. The smoothed mode density is the Debye spectrum (for an elementary discussion see Kittel 1976) to a very good approximation which decays as k^2 in the limit $k \rightarrow 0$, so that the energy density per wave number also decays as k^2 in the limit $k \rightarrow 0$. Assuming that maximum displacement is proportional to potential energy as in a simple harmonic oscillator, the maximum displacement will decrease proportionally. These are the only motions that cannot be eliminated in a maximally quiet system. If there is any motion imposed on the system or driven at larger scales a dissipation spectrum is also present. At what is very small scale in the dissipation process a Kolmogorov spectrum decaying in energy as $k^{-5/3}$ and $m^{-5/3}$ holds (see Tenekes and Lumley 1972 for an elementary discussion of this spectrum). Even with the most carefully prepared experiment some fossil turbulence remains from the set up which must damp out. It is impossible from first principles to assign an amplitude to the Kolmogorov spectrum.

The relevant perturbation spectra are an interesting set since one is red with a cutoff and one is blue with a cutoff. By adjusting the relative magnitudes the peak

could be put wherever it matched the data. From Linden's (1978) experiments we can get a crude estimate of the time scales of the growth of the fastest growing mode and its perturbations which appears to be tens of e-folding times and longer. This correspond to enormous amplifications well in excess of 10^{10}, the ratio of a meter to an Angstrom. Because of the care taken in filling the tank and the extremely long time needed for finite amplitude effects to develop, it seems reasonable to assume that the initial noise amplitudes at the wavelengths of rapidly growing double-diffusive instabilities are minuscule. No explicit calculation of the disturbance amplitudes has been made.

Until a region is no longer unstable to finger growth on vertical scales smaller than the region height but of the same order, it seems reasonable that harmonic instabilities will disrupt larger scale motions by altering the mean field with their flux divergences. A cellular layer isolated by two turbulent regions, if unstable to perturbations that are smaller but not as small as the first harmonic would expel gradient towards the cell ends, making the interior less unstable. If instabilities shorter than the first harmonic are possible, there are two possibilities. Should all the harmonics not be strong enough to create regions in the interior with strong enough gradients for turbulence, the situation would resemble the one just discussed above. However, if the harmonics are strong enough to generate an interior turbulent region, a split would occur into two or more finger regions, each shorter and less supercritical.

These scenarios and the behavior photographed in Linden's (1978) experiments, shown in figs. 14 – 16 are in excellent qualitative concordance with the theory presented here. Layers form where strong density gradients, visible on the shadowgraph occur. There are no regions of noticably strong, broad scale shear. In

fig. 14 a mode with larger vertical wavenumber than the fundamental dominates the system initially shrinking the finger layer thickness. However the wavenumber is not sufficiently large to create multiple layers. As the gradients weaken this trend is eventually reversed and the finger layer thickens.

The development of the first harmonic is clear in fig. 15b. Though the harmonic creates a visible high gradient region it does not cause an immediate complete breakdown. Actually the convection in the finger region of both figs. 15a and 15b is so vigorous as to be fairly disordered. Finally higher harmonics create thin cellular regions bounded by strong gradients. Initially the two finger regions are asymmetric in figs. 15c and 15d, probably due to some asymmetry in the initial conditions. However this is unstable on the larger scale and a symmetric arrangement develops.

The layering process can be seen more clearly in figs. 16. In this case the gradients are weaker and the evolution slower. Even in the first photograph, fig. 16a, the beginning of layer formation is evident along with the sharpening of the gradients at the finger ends. This becomes quite pronounced by fig. 16b. Some evidence of the first harmonic can be faintly seen though it is weak and the central region is quite disordered. The last state shown in fig. 16c is taken much later and there is extremely sharp contrast between the laminar fingers and the separating chaotic layer. By this time the overall gradients have been greatly weakened. The system is still supercritical enough to maintain layers. The harmonic instability theory also predicts a decrease in the vertical wavenumber of the most powerful mode over time, but after layers have formed we hesitate to apply this.

If shear was the primary cause of the layers, tilts of the fingers should be

visible in the early stages before breakdown stabilizes the system. Further there would be no need for strong gradients to be associated with the layer boundaries so quickly, though they would eventually develop at the finger ends. However it is just where the sharp gradients occur before layers are distinct that the layers divide. This corresponds to the harmonic instability mechanism.

7. Numerical Simulations

Numerical simulations were performed with a code similar to that used by Piacsek and Toomre (1980). However the scheme was pushed further into the supercritical unstable regime than previously and the initial conditions were qualitatively different. Equations time stepped were the vorticity equation (7), the heat equation, and the salt equation (6), with a solution of a Poisson equation (9) obtained at each time step. The grid spacings were constant in both the $\hat{x}$ and $\hat{z}$ directions, though the spacings in the two directions were different. The temperature and salinity grids were staggered with the stream function and vorticity grids. Boundary conditions in the $\hat{x}$-direction were periodic. The top and bottom boundaries had constant temperature and salinity, and were of the no slip type.

The time step procedure used was a Dufort-Frankel leap-frog scheme (see Roache 1972) for the diffusion terms of all three conservation equations. Simple averaging was performed at regular intervals to cancel leap-frog divergences of the alternating time steps. An Arakawa Jacobian (Arakawa 1966) was used on the vorticity advection term while the salt and heat equations had simple second order differencing corresponding to J_{++} in Arakawa's notation for their advection terms. The density term of the vorticty equation also used simple second order differencing by use of the staggered grid. The Poisson equation was solved using a spectral method outlined by Williams (1969). Much of the code was provided by Steve Piacsek including a Poisson solver written by D. A. Moore and some output routines by Glyn O. Roberts. The code was cleaned up, vectorized and altered for the Cyber-205, and new initialization routines were written.

Some runs were attempted using Arakawa Jacobians on the advection terms

of all three conservation equations and without a staggered grid. These developed numerical instabilities at parameter values that for the other scheme were stable. The all-Arakawa scheme was tried because of explosive numerical instabilities of the first scheme even though the Courant-Friedrichs-Lewy (CFL) condition requiring that advection not exceed a gridpoint spacing in a time step was always satisfied. Evidence of possible numerical instability in the codes used by Piacsek and Toomre (1980) had already been noticed even at much less supercritical conditions when near the ends of runs started with a sharp interface, salinities were observed that were beyond the boundary values (Piacsek, personal communication). Though the Arakawa Jacobian was originally used in the vorticity equation because of a previous use of the codes, it is the diffusion equations that are more likely to develop a Phillips type instability (Phillips 1959) because of their smaller diffusion coefficients, so the experiments were performed using Arakawa Jacobians there. This scheme proved to be unstable even in parameter ranges where the first scheme was stable, which was surprising. The instabilites were finally eliminated by a decrease in grid size and a corresponding decrease in time step. Even though a single diffusion equation is unconditionally stable using a Dufort-Frankel, scheme if the CFL condition is met, the set of three coupled equations with a leap-frog appears not to be. The nature of these numerical instabilities has not been fully analyzed yet, though either a diffusive of leap-frog type instability is expected. We know of no studies of the computational stability of finite difference schemes of such complicated coupled equations. The problems are exacerbated by the fact that the chosen base state is physically unstable.

The primary hope before these simulations were begun was to be able to examine in detail the evolution of the layering process, by whatever mechanism it

occurred. This goal was not achieved. It was not possible even using a Cyber–205 computer operating at about 20 million floating point operations per second for nearly an hour to run for a sufficient number of time steps, with a fine enough grid, and large enough vertical extent. Initial runs attempting to model one of Linden's (1978) experiments developed explosive numerical instabilities probably related to the inability to resolve the salinity gradients on the course grid spacing. The numerical experiments that remained stable are presented here and appear to have had sufficient resolution in the interior but were probably too coarse near the top and bottom boundaries.

Three simulations will be described all of which were numerically stable. All have infinitely conducting constant property boundaries with no slip at the top and botton. Periodicity was imposed as a lateral boundary condtion. The runs vary only in the initial conditions and total box height. The first simulation was started with constant linear gradients and was given perturbations with a white vertical specturm over several modes and a single horizontal mode. A spectrum of several modes in the horizontal was added in the second simulation. The third simulation imbedded a region with linear gradients between two initially constant property reservoirs and had a limited white noise spectrum in both the vertical and horizontal.

All of the simulations had several parameters and conditions identical. Parameters were chosen to correspond to the salt-sugar case. This reduced a problem that Piacsek and Toomre had of the grid being finer than needed for the faster diffusing component and yet still unable to resolve the gradients of the driving component. The kinematic viscosity, ν, was taken as $0.01\ cm^2/sec$, $K_T = 1.8 \cdot 10^{-5}\ cm^2/sec$, and $K_S = 6 \cdot 10^{-6}\ cm^2/sec$. Hence the Prandtl number was 556 and the diffusivity

ratio was 3. The vertical grid size was $0.04\,cm$, the horizontal grid spacing $0.02\,cm$, and the time step $0.1\,sec$. The solute differences were measured directly in density units with ΔS set at $5 \cdot 10^{-3}\,g/cm^3$ and ΔT at $6 \cdot 10^{-3}\,g/cm^3$, giving $\Lambda = 1.2$. The grid box was 66 points laterally by 130 points vertically, in the gradient region, and so was $1.28\,cm$ wide by $5.12\,cm$ tall. In the third simulation the gradients were unchanged and the reservoirs were added by putting 64 more grid rows both above and below for a total height of 258 points or $10.24\,cm$.

The initial perturbations for simulation A were all at a single horizontal wavelength corresponding to the fastest growing allowed mode which fit four waves across the grid. The first ten allowed modes in the vertical were included. Each mode was made to vanish at the top and bottom boundaries and had random lateral phase. Perturbation amplitudes were set to $10^{-6}g/cm^3$, equivalent to physical displacements of $10^{-3}cm$. The sixty-four bit accuracy of the computations was sufficient to maintain evidence of only wave number four across the grid in the two dimensional profiles (not shown) after more than 30,000 steps.

Exponential growth lasted about $600\,sec$ as can be seen in fig. 26. At that time the salinity profile appears as a single mixed layer with sharp boundary layers. By $t = 700\,sec$ growth has slowed and the mode amplitude has become so large that the interior gradient is inverted. After this time the amplitude overshoot is corrected by a slight decrease while the profile becomes smoother, indicating the decay of some unstable, short vertical scale, transients. At $t = 2000\,sec$ there is evidence of two mixed regions each associated with a physical boundary. In the last profile the slow extension of the mixed regions into the interior is seen. We will return to this last point when discussing the salt Reynolds flux.

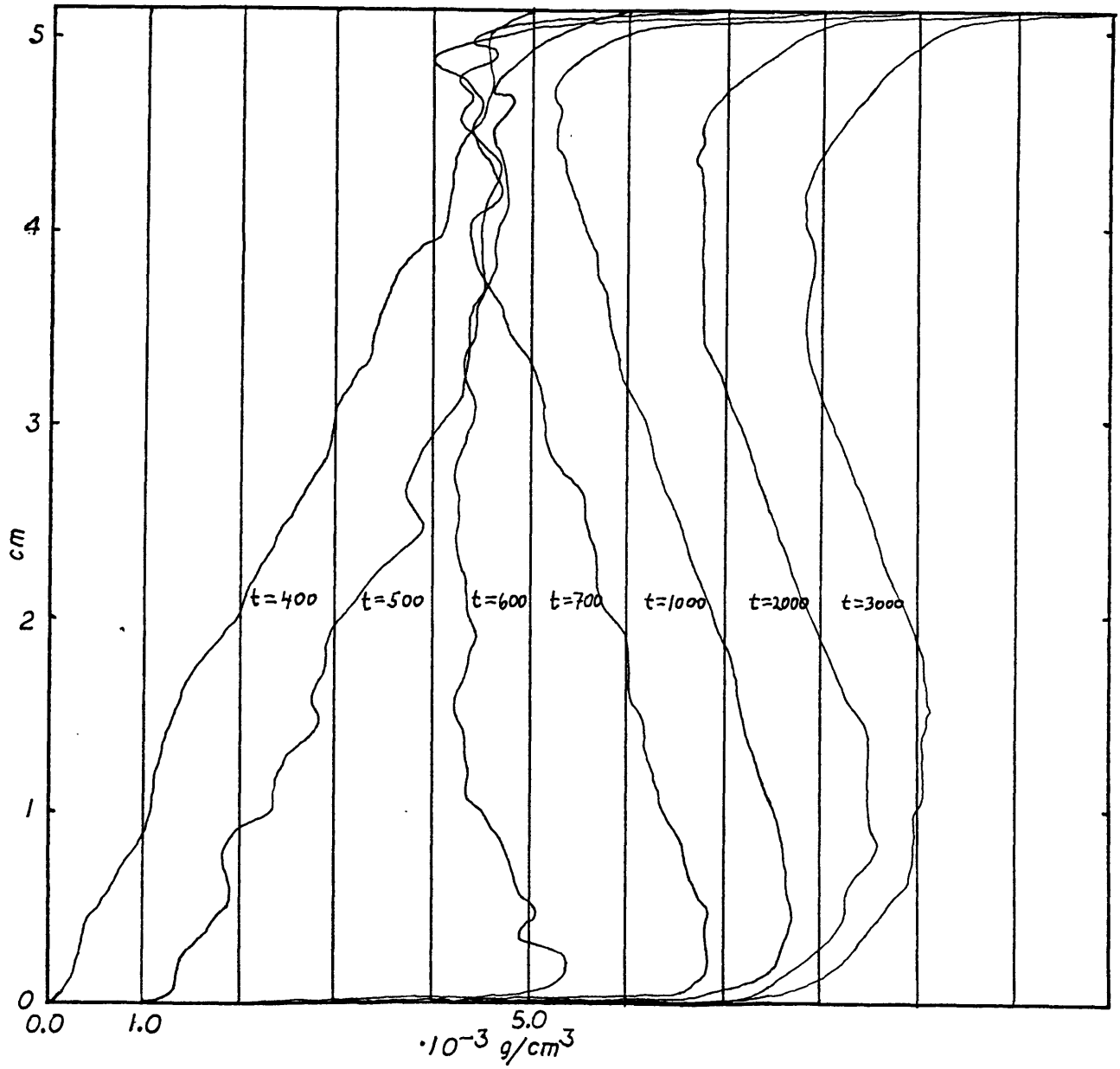

Fig. 26. Development over time of mean salinity field in numerical simulation *A*. See text for description of simulations.

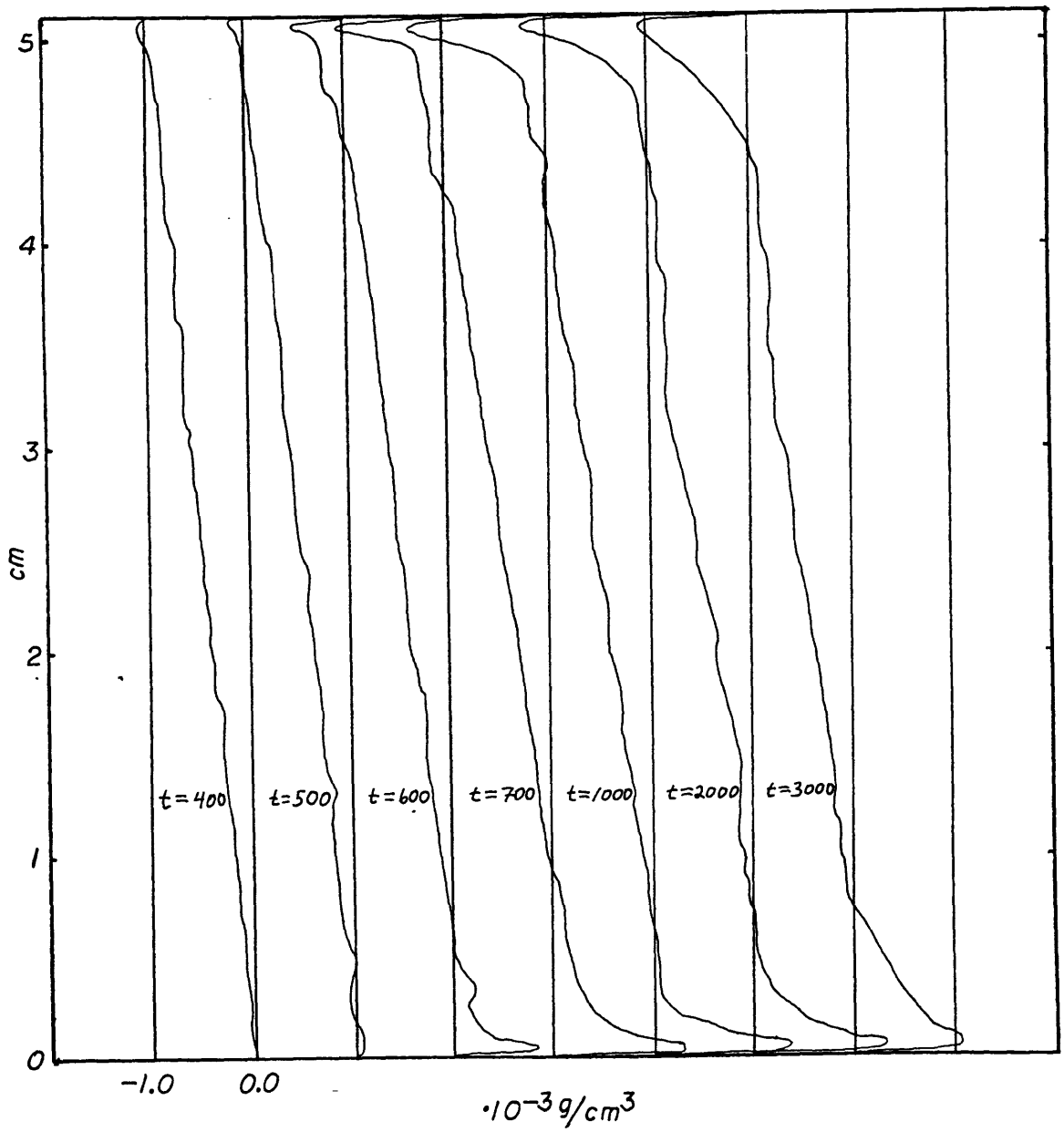

Fig. 27. Development over time of mean density field in numerical simulation *A*.

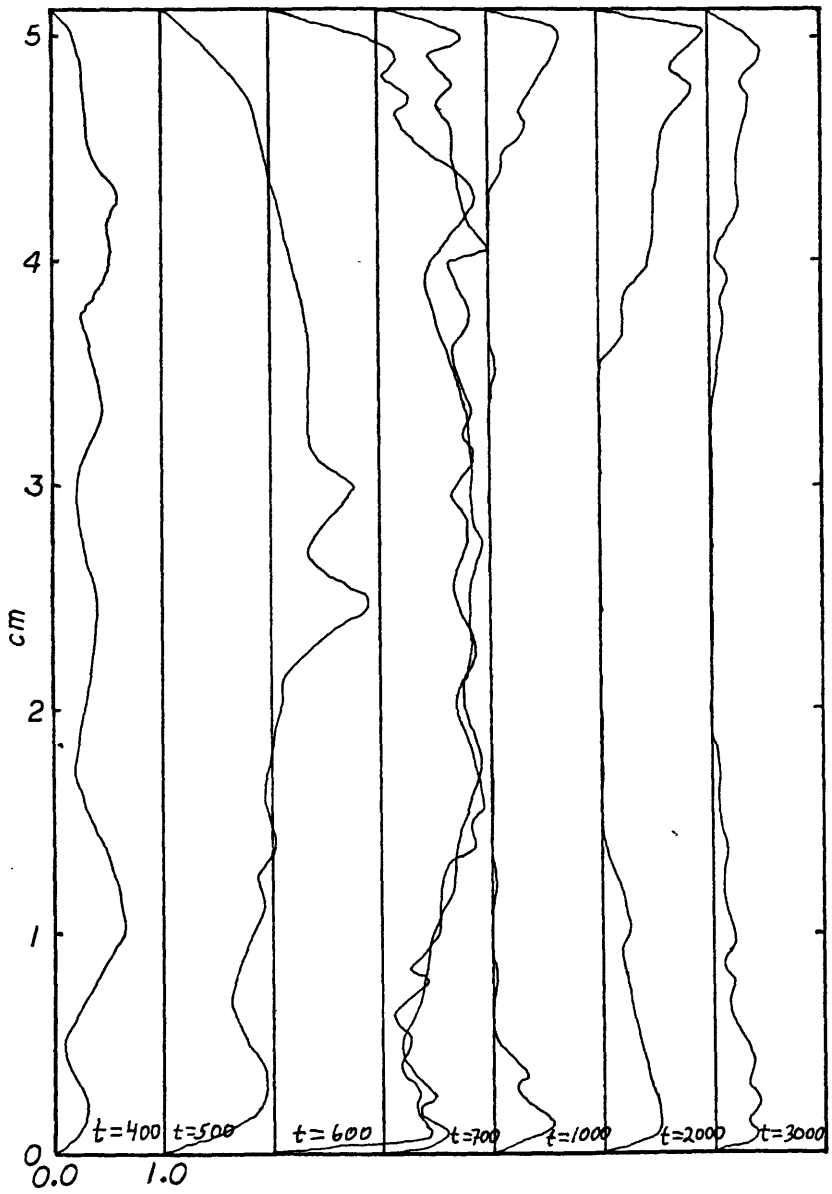

Fig. 28. Development over time of salt Reynolds flux in numerical simulation *A*.

The density profile evolution for simulation A is in fig. 27. At the boundaries a strong density gradient inversion developed as expected. The exponential growth follows the same pattern as the salinity profile development. Eventually growth slows at $t = 700\,sec$ and then decays slightly. In the later profiles the density boundary layer thickens. The salt Reynolds flux profiles in fig. 28 provide information on the intensity of convection as a function of vertical coordinate. Initially the convection is vigorous throughout, growing exponentially and predominated by the fundamental, as indicated by the broad central peak, with some harmonics superimposed. At $t = 600\,sec$ the convective motion is nearly independent of z, except at the boundaries, and by $t = 700\,sec$ the convection has begun to decay, though the salinity and density profiles continued their development as noted above. In the last three profiles convection in the interior has been thoroughly quenched and the time evolution shows the slow extension of the convecting layers toward the interior.

The continuing destabilizing buoyancy flux forces the convective regions to encroach on the quiescent center. Though the fundamental has created an interior that is absolutely stable, it has not been able to fully stabilize the system. The simulation was stopped before convection inevitably extended between the boundaries. This does represent the development of two separated convecting layers but this was not exactly what was anticipated. The last configuration is far from steady state because of the vertical divergence of the salinity flux, yet individual modes have saturated and are in a quasi-steady state. Whether a slow oscillation of this pattern would have occurred or a steady state developed is unknown.

Simulation B started with multiple modes in both the vertical and horizontal

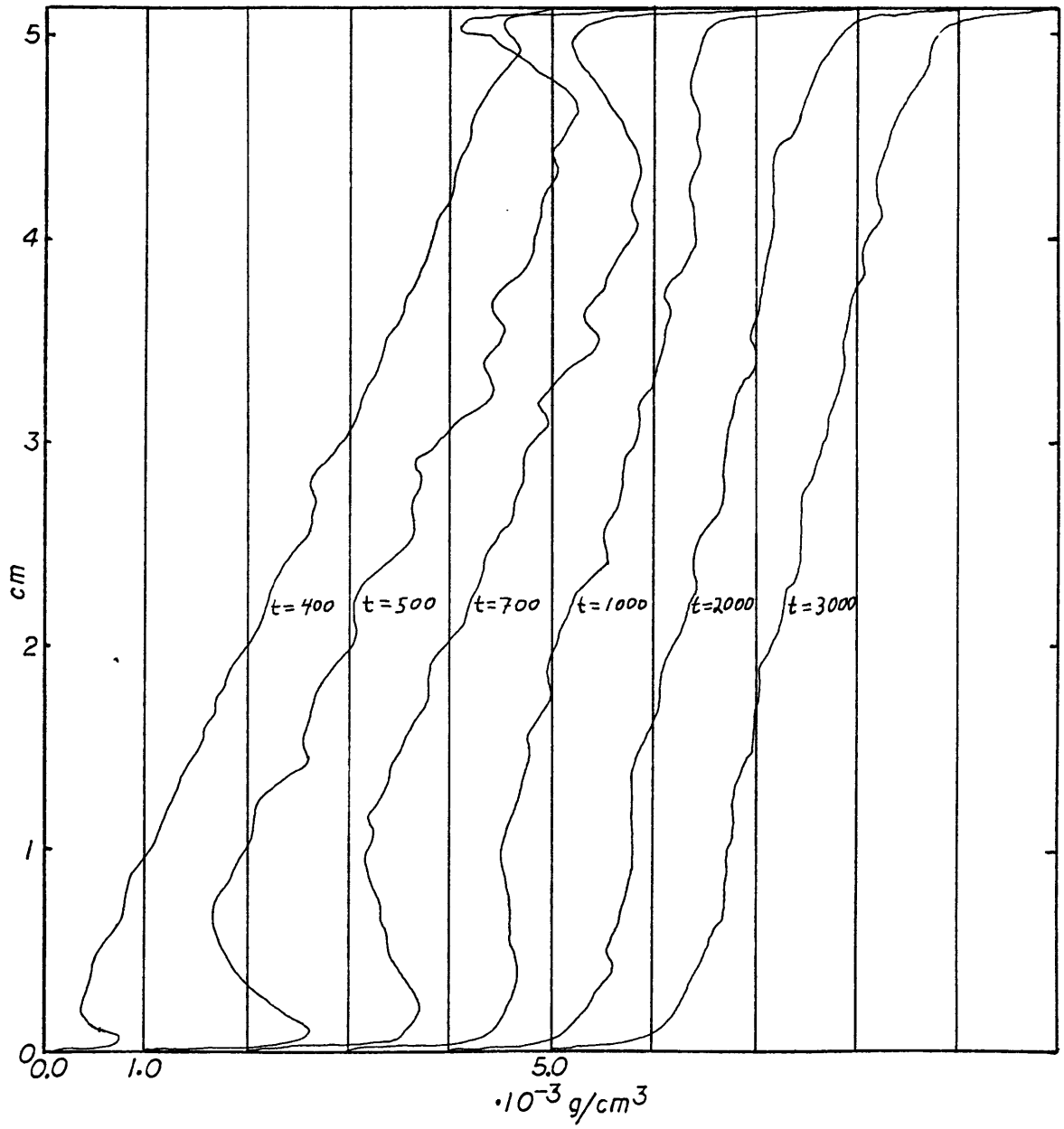

Fig. 29. Development over time of mean salinity field in numerical simulation *B*.

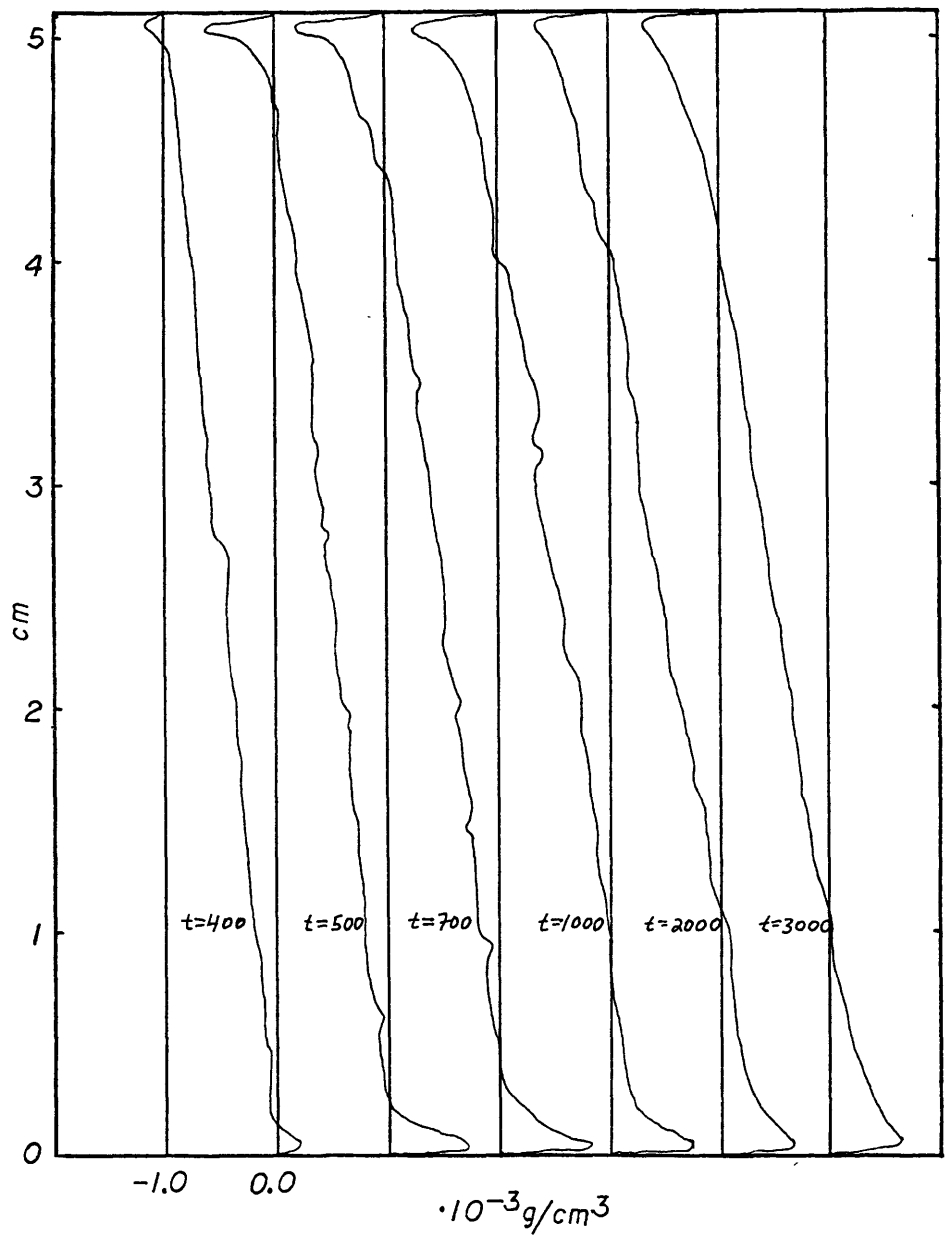

Fig. 30. Development over time of mean density field in numerical simulation *B*.

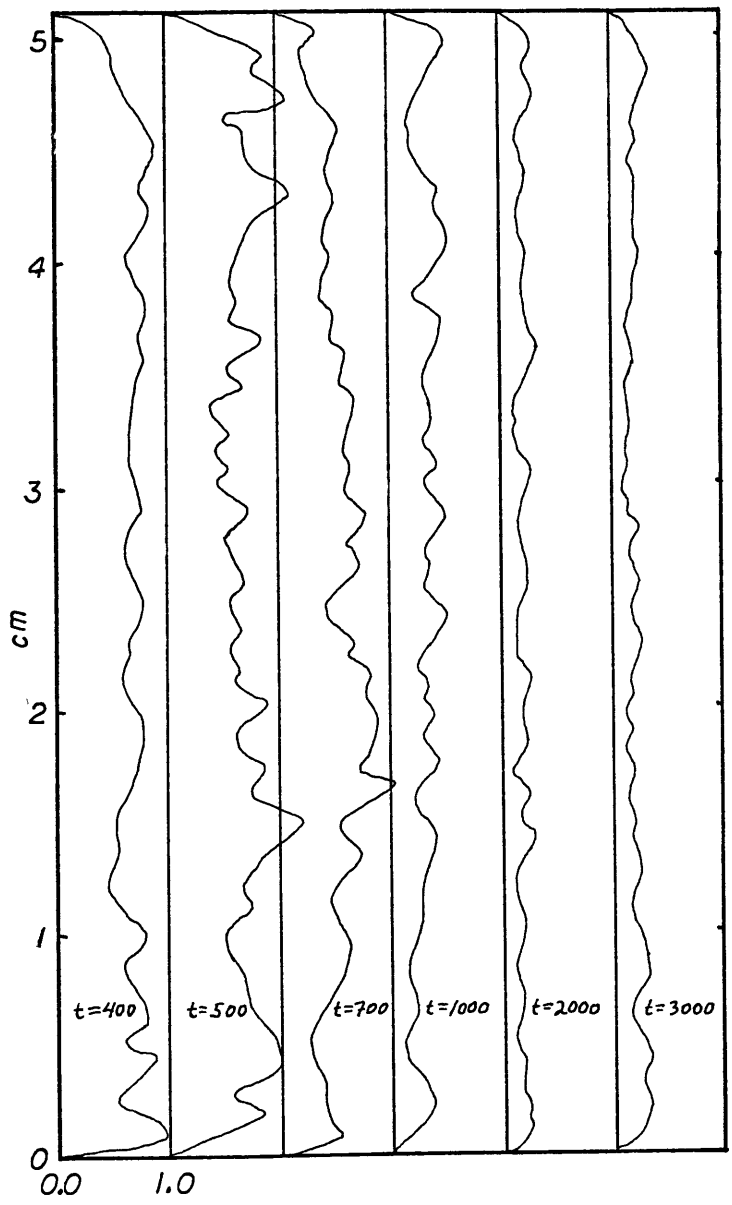

Fig. 31. Development over time of salt Reynolds flux in numerical simulation *B*.

directions. Again the modes were arranged to individually match the end boundary conditions and have random phase in the $\hat{x}$-direction. Each mode had an amplitude of $10^{-6} g/cm^3$ as before. Profiles of salinity, density, and salt Reynolds flux are presented in figs. 29, 30, and 31, respectively. The structure in the two dimensional profiles (not shown) is not easily discernable to the eye. The results were greatly different than the previous simulation. Salinity inversions developed just outside the boundary layer. The inversion layers thickened over time and eventually disappeared. The final configuration had about half of the total salinity difference concentrated in the two boundary layers and the remainder in a nearly linear interior gradient. Density inversions at the boundaries quickly developed. The evolution of the density profile was quite similar to that in simulation *A*. Growth and eventual decay of the convective motion can be seen in the evolution of the salt Reynolds flux. Throughout all of the profiles here the Reynolds flux is nearly depth independent except at the boundaries. The final result at $t = 3000\, sec$ is close to a consistent steady state.

The difference in the results between simulations *A* and *B* indicate that the interactions of various lateral modes is quite complex and strongly affects the vertical structure. A broad specturm of vertical modes was insufficient for the development of a high Rayleigh number configuration with gradients compressed to the boundaries and a nearly homogeneous interior. The interaction of multiple modes in more than one spatial dimension appears to be required for the mean fields associated with turbulent convection. It is conceivable that many vertical modes are sufficient for aperiodicity but the simulations were not run long enough to determine this. Since simulation *A* is constrained in a similar was as a single mode calculation (*e.g.*: Gough and Toomre 1982) the results obtained here may

indicated that such a single mode approach, which is considerably less expensive than the modelling done here, might not be effective for analysing direct double-diffusion. We are not aware if similar effects have been noticed in numerical work on high Rayleigh number thermal convection. Simulation *B* appears to reasonably model the double-diffusive analog of high Rayleigh number, parallel plate bounded, thermal convection, an experiment not yet performed.

The last numerical experiment, simulation *C*, added reservoirs to the previous one at both ends. Perturbations again had spectra in both directions which were extended into the constant property reservoirs. The salinity profiles in fig. 32 again show the initial salinity inversion associated with the boundaries thickens and finally vanishes. The earlier profiles inidcate a good deal of noise in the vertcal which smooths out over time. The linear gradient region shortens to a height of about $4\,cm$. When the system settles down two clear mixed region are again present. Two fifths of the salinity gradient is concentrated at the boundaries and the remainder is in the interior linear gradient region. Sharp density inversions at the boundaries occur as expected (fig. 33). The inner boundary layers, where the density inversion is compensated, thicken with the passage of time and the central gradient region shrinks. From the salt Reynolds flux profiles (fig. 34) the driven convection can be seen to peak and then decay slightly. An examination of the fluxes from the convection and diffusion indicates that the system is still far from steady state. However, from the standpoint of individual modes finite amplitude saturation has been reached.

The reservoirs in simulation *C* did not break down. They maintained significant homogeneity even at the end of the simulation and even though the solute concentrations had drifted. With the exception of boundary effects this model

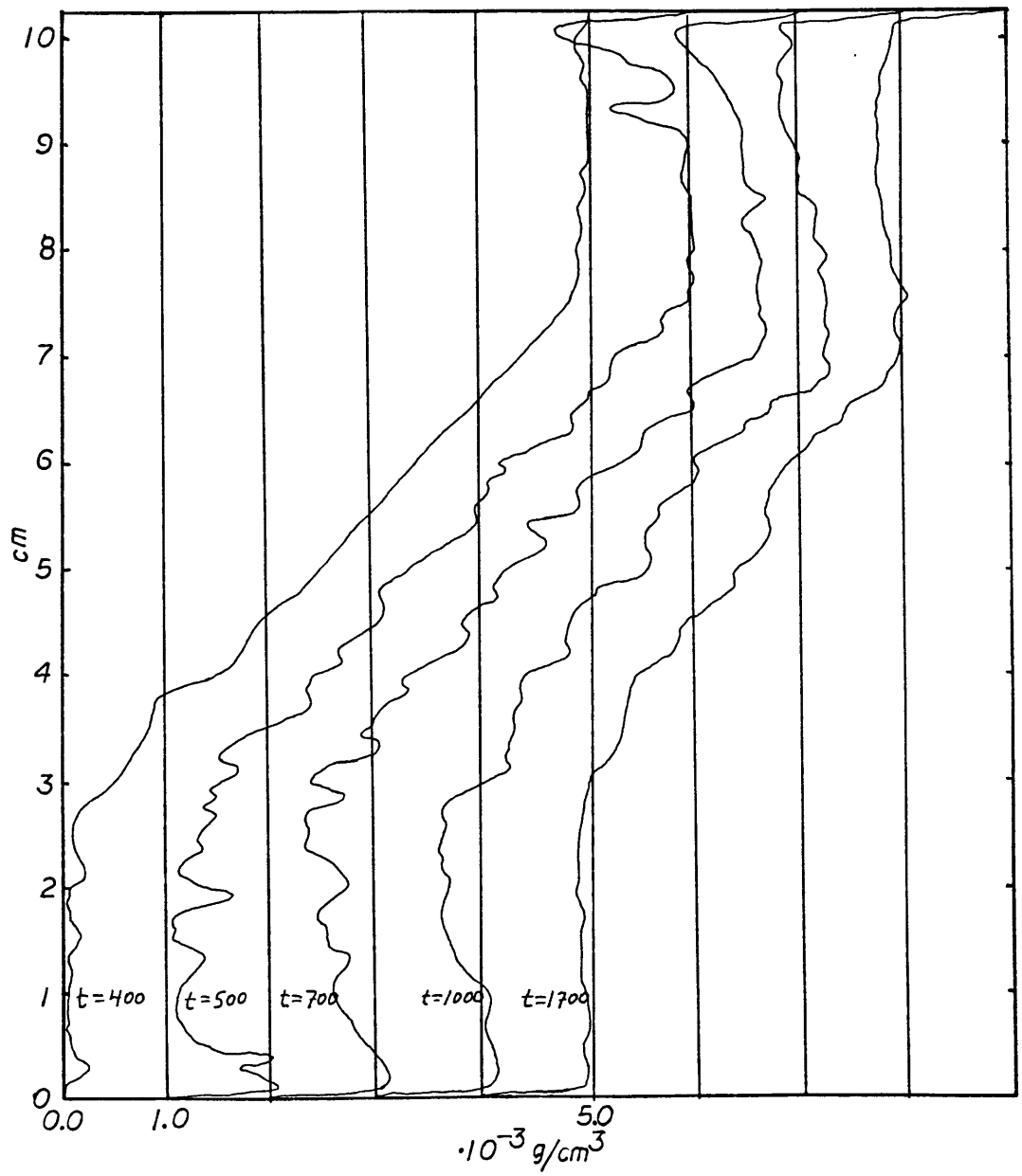

Fig. 32. Development over time of mean salinity field in numerical simulation *C*.

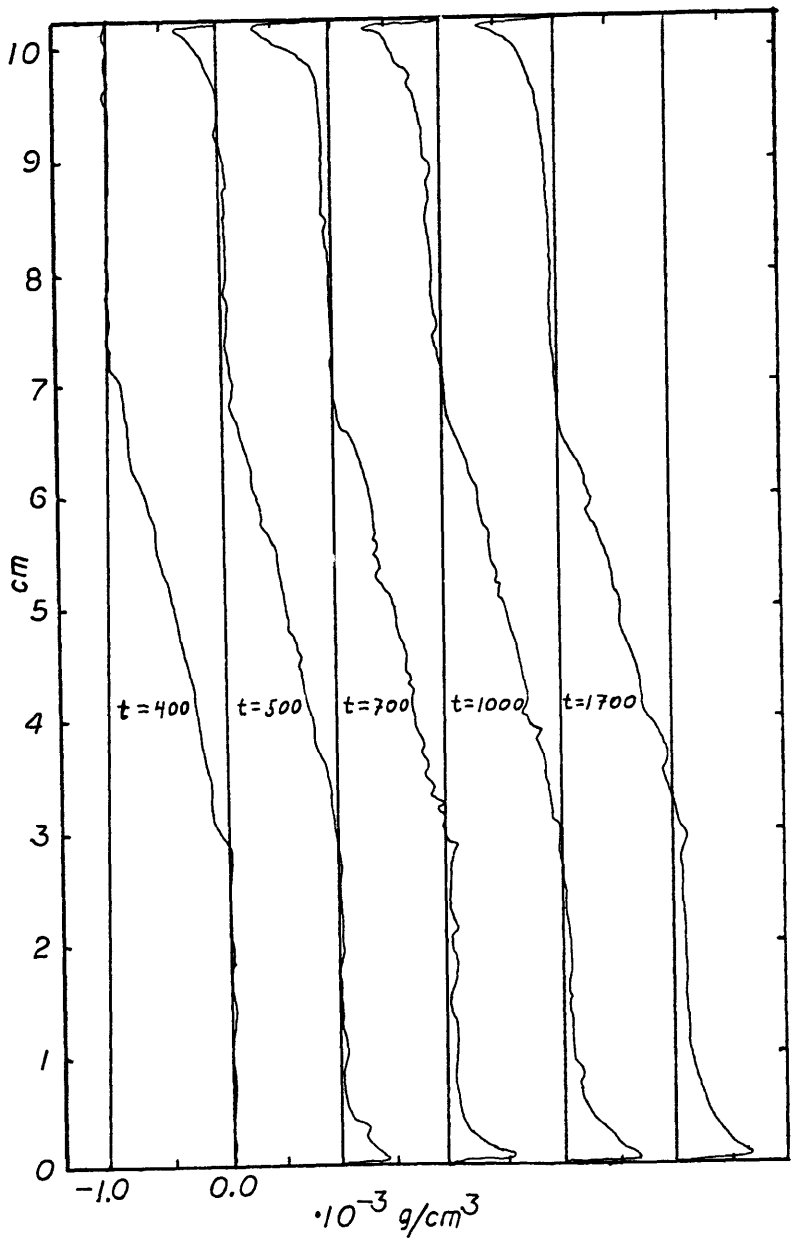

Fig. 33. Development over time of mean density field in numerical simulation *C*.

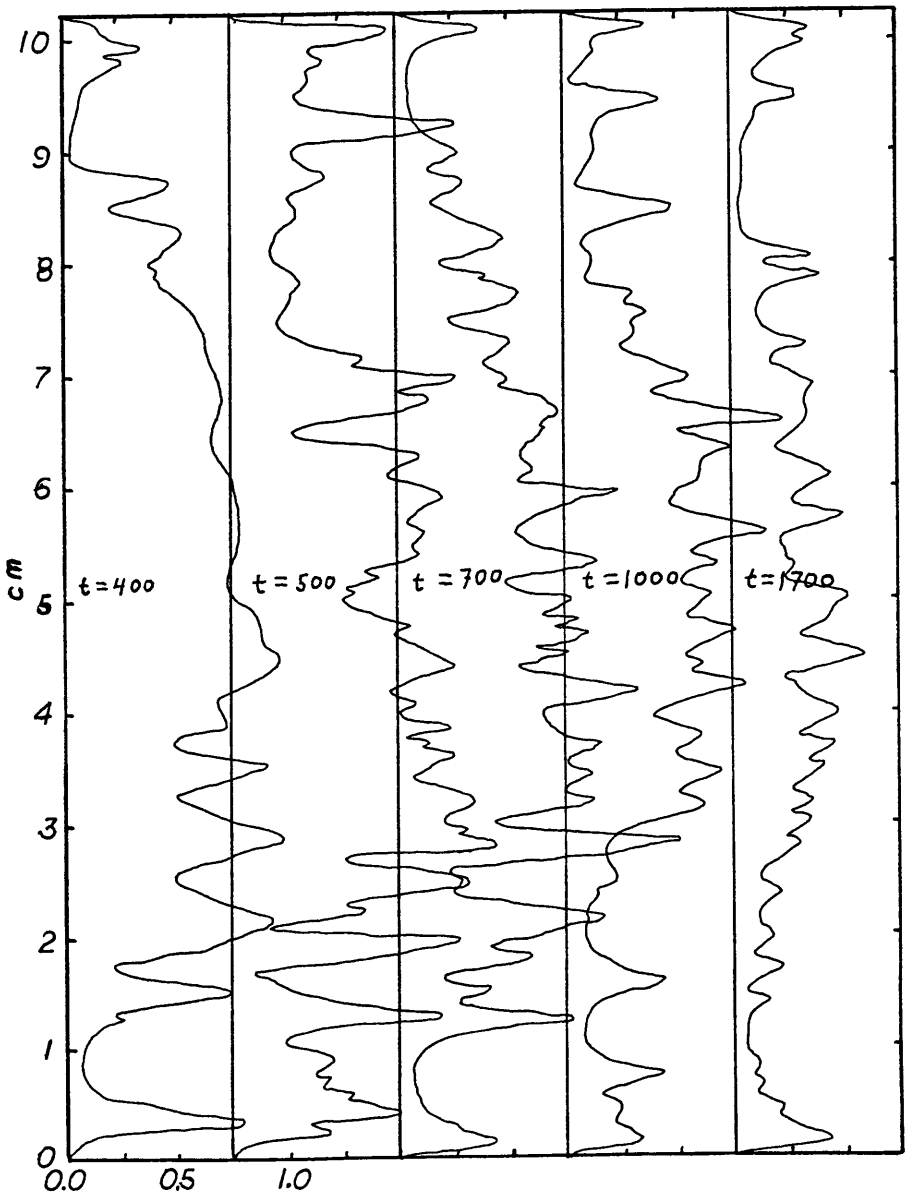

Fig. 34. Development over time of salt Reynolds flux in numerical simulation *C*.

corresponds well to the development of Linden's experiment shown in fig. 14. Again the search for layers was frustrated, but the simulation did capture the essential physics of a complex laboratory experiment, which is encouraging.

Significant improvements in computational efficiency are still possible with available techniques. These will come at the expense of additional programming. One of the simplest would be the dynamical determination of the time step to keep it close to the CFL limit. More complete vectorization of the code can be accomplished but the resulting programs would be less transportable. At greater programming cost some dynamic changes in grid size could be implemented. It is quite clear that the salinity gradients at the boundaries need a much finer grid for proper resolution than was used here. If a simulation does develop layers a finer grid would be needed at some interior levels.

Enough of the physics of double-diffusive convection has been captured in the simulations to recommend simulations of longer time and greater height. The choice of boundary conditions, however, will continue to present a complex problem and will be discussed further in the next section. The boundary conditions used do not correspond to any actual experiment. It is possible to match the boundary conditions of the rundown experiments but the eventual steady state is one of no motion. Numerical boundary conditions that are appropriate for oceanic layering are not obvious.

8. Discussion and Collected Thoughts

The strong time dependence of the vertical wave number that most strongly affected the system, as is contained in eqs. (36) and (37), and made plain in eq. (48), is a distressing result. It does not hurt the basis of the theory but makes it impossible to determine, from what we have done so far, a vertical length scale. One must have some idea of the initial noise spectrum to do that if layering occurs significantly before saturation. After a mode creates chaotic regions it locks-in and the time scale needed to change the configuration lengthens enormously. Whether this reflects the reality of the system, that there is a range of possible vertical scales based on the initial noise spectrum, or whether there is a unique scale associated with finite amplitude effects, we are as yet unable to say.

A system that has the property that the final configuration has little or no dependence on the details of the initial conditions is called robust. If a system is not robust it may be difficult, or impossible, to predict its final state. There may exist a large number of accesssible steady states, which may or may not be contiguous in phase space. Examples of this have been found in direct convection problems. Startin with Chen and Whitehead (1968) and followed by Busse and Whitehead (1971) it has been shown that by using an initial perturbation to force a particular planform a stable configuration may be created that would not otherwise result. Further, if the forced array is not stable the breakdown of the cellular pattern in favor of a new one can be quite slow. At fairly low Rayleigh numbers a range of stable rolls is predicted by Busse (1967) for thermal convection and by Straus (1972) for double-diffusive convection. It seems quite possible that a given double-diffusive problem may admit multiple stable state with different planforms,

as thermal convection, and multiple stable vertical structures.

The fact that rolls are unstable to cross-rolls is evidenced the same way in double-diffustion as in one component convection. Square cells occur in supercritical regimes that are not sufficiently driven to be chaotic. Fine distinctions have been made between cross rolls, bimodal convection, and square cells (see Busse 1981), but they are of little importance here. The result of Schlüter, Lortz and Busse (1965) that cross rolls are the strongest instability to limit some rolls in thermal convection, an unpublished result of ours for direct double-diffusion leading to the same conclusion, the work of Straus (1972), and work by Busse (1967) giving the stability regime called the "Busse bubble," are consistent with this. The highest Rayleigh numbers in thermal convection that still allow stable, time independent behavior, exhibit a square planform (Busse and Whitehead 1971). In double-diffusive convection the layering regimes also take this pattern, we believe, for similar stability considerations. The aspect ratio in double-diffusion is of course far from one. Whether or not the various instabilites that Busse (1978a) has noted in his reviow on thermal convection can occur in double-diffusion will be unknown until more careful observations are made on better controlled experiments.

An interesting set of *gedanken* experiments results if one looks at various fixed boundary problems in the steady state. They should help clarify the broad issue of the stability problem. Take top and bottom boundaries separated by a given large distance, H. At each of these boundaries we fix the temperature and salinity. This is a well posed problem that can, at least in principle, have a steady solution. Any steady solution also has vertically constant mean fluxes. The possibility that there will be no stable state with stationary horizontal means cannot be dismissed. Schmitt and Evans (1978) suggest that fingering is an intermittent process in the

ocean because of effects from internal wave strain. However, though wave strain can disrupt the laminar cells it can not easily prevent the convective flux. It is possible that different levels are active at different times but this would require that the sharp gradient regions migrate. If this is the case a Huppert (1971) type analysis would show such behavior if realistic flux laws are used. The individual layers, though, would still have to be stable in to smaller scale instabilities. If no steady, stable solution exists for the imposed vertical scale, the result must be either layering, time dependence, or both. Running the gamut between easily analyzed extreme cases, may serve to illustrate some of the similarities between one and two component convection.

For constant positive temperature difference, measured from the bottom, there is only one steady solution, which has constant gradient and is stable. For constant positive salinity difference, the standard one component convection problems are returned, depending only on the salt Rayleigh number, R_S, and the salt Schmidt number, σ_S. As R_S increases from zero, there is eventually a transition from a purely conductive state to barely supercritical cellular convection state. Further increase brings about greater degeneracy of the the solutions, different cellular structures, and eventually turbulence. For any given large R_S sufficient for turbulence, there is a thermal Rayleigh number, R_T, great enough to return the system to a purely conductive state. Just short of that, a full height cellular mode would be expected. Up to some R_T, the entire interior would remain turbulent. Somewhere in between the last two described states alternating layers appear to occur.

In double-diffusion there appears to exist a range of Rayleigh numbers, R_S and R_T, for which neither full height cellular or full height turbulent regimes

are stable. There is nothing, *a priori*, that precludes a similar situation occuring for one component convection, but we have no reports vertical wave numbers greater than the first allowed predominating. Single component convection has however, provided a recent surprise related to layering. Hele-Shaw convection, which is extremely constrained horizonally to a vertical slot, does exhibit two time dependent stacked cells (Frick and Müller 1983). In the range studied these cells remained laminar; there was no development of a front, just the strong presence of the first allowed harmonic. The prevention of cells of width comparable to height, as is usual in one component convection, is clearly forced in this problem leading to the increased vertical structure. Direct double-diffusive convection has narrow cells without imposed constraints so there may be a connection, particularly if the cell Reynolds numbers are supercritical for shear instabilities, discussed below.

The question remains why at ultra-high Rayleigh numbers at which most double-diffusive experiments are performed is not turbulent convection from top to bottom observed as in the Bénard problem. This does not have a simple answer. First it is important to note that the double-diffusive analogy of the Bénard experiment has never been performed. The numerical experiments presented here did contain only one vertical cell for the length of time they were run, which appeared to be sufficient to reach at least a quasi-steady state. It is considerably more difficult to maintain constant solute concentrations at a boundary than constant temperatures. The numerically generated solution with an initial broad noise specturm resulted in a configuration with hydrostatically unstable steep gradient boundary layers and an interior not all that dissimilar to the simulation of Grötzbach (1982) on low Prandtl number convection. Low Prandtl number convection shares double-diffusion's property that diffusion remains important in

the interior. However, diffusion is less important when convection is driven from the boundaries, as in our numerical simulations, than in Linden's (1978) and other run-down experiments, which are driven from the interior as much as the reservoirs.

In a convective steady state there must be a positive buoyancy flux (negative density flux) to counteract dissipation. Therefore at the boundaries of the cell where the vertical velocities vanish the density gradient must be positive, *i.e.*: in the unstable sense. It may well be that the numerical simulations accurately reflect what would happen in the so-far-unperformed steady state salt fingering experiment with constant temperature and salinity boundaries. Whether or not layers can be a steady state phenomenon remains a serious unanswered question. So far all of the laboratory experiments have been of the run-down variety. Lambert and Sturges (1977) labeled observed layers in the ocean "steady" because they persisted for four days. However the run-down times of lab experiments which are clearly not truly steady state stretch into several days even with more vigorous convection, much smaller water volumes, and smaller vertical extents. Still the time scales of layer life times are considerably longer than the transport times for a parcel of fluid to move through the system.

Some possible boundary conditions that are more realistic for the ocean and the laboratory but difficult to make mathematically precise might avoid the extreme boundary layers and allow layered convection. An example would be stirred layers with constant properties and a turbulent velocity spectrum. Numerically this might be accomplished by specifying boundary property values at the top and bottom and imposing a turbulent velocity profile within a restricted layer. This however is assuredly analytically inaccessible. The presence of homogeneous reservoirs at the ends of the fingers in both laboratory and ocean situations seems strongly tied to

the lack of development of a single high Rayleigh number type cell. The presence of a physical boundary at the finger ends that forces the vertical velocity to vanish totally further requires a strong density inversion. At the softer boundaries that form internally the velocity need not vanish for all time. Yet, the end of the of the fingers is qualitatively different than the interior and the internal finger velocities are much greater than the movement of the internal boundary layers. A possibility is that the velocities at the end are small and time average to zero. This would require at least some small time dependence in a neighborhood of the finger ends. The ends of the fingers do drive what appears as a time dependent high Rayleigh number cell. However, there is great difficulty in separating the tail, the dog, and the wag. The third simulation crystallized this confusion since it is clear that the convection is driven from the conducting boundaries, yet the interior gradient region along with the surrounding mixed retions were long lived. The rapid increase in the accelerations of parcels if the salinity gradient gets very compressed puts limits on the amplitude of any given long vertical mode.

Layering in the ocean, particularly in the central waters, may be encouraged by the time dependence of the forcing. In regions driven by heating and evaporation the forcing has a strong diurnal component, which is sufficiently long for modes to approach saturation. The layering on two different scales reported by Williams (1975) we suggest is the result of a combination ot time dependent forcing and harmonic instabilities. Layering on the scale of meters is caused by a strong positive buoyance flux at the surface forced by air-sea interactions, *i.e.* surface heating and evaporation, driving penetrative turbulent convection on time scales faster than strictly diffusion driven convection. The cells resulting from this process are still unstable to harmonic instabilities on scales of tens of centimeters.

The question posed on layering in one and two component convections can also be inverted a bit. Since greatly supercritical regimes in the Bénard problem also have unstable harmonics why is layering not observed there? The most important difference is how the finite amplitude effects of a mode affect the stability of the fluid. Growth of normal modes does increase the unstable gradient in some regions of the one component problem. However, growth rates only grow as the inverse of the scaling time, proportional to $B^{1/2}$, and the restiction on the vertical extent of these regions will not allow faster growing modes. In the double-diffusive case the growth rate is strongly dependent on the density ratio. So, even regions restricted vertically can support local instabilities with faster growth rates than the taller disturbances.

It is interesting that many vertical modes were not sufficient to lead to the expected profile for turbulent convection in numerical simulation *A*. In the model of turbulence put forward by Ruelle and Takens (1971) the interaction of many modes, all unstable because of the supercriticality of a system and having incommensurate time constants leads, to turbulence. Whether or not the long time behavior of the simulation would have proved to be aperiodic, and therefore "turbulent" in some sense, is not known. However the *vertical* spatial structure traditionally associated with turbulent convection appears to require the interaction of many *lateral* modes. We have not yet developed an understanding of this process.

Both the experiments of Linden (1978) and our numerical work show that once a sharp gradient layer has been created it is fairly long lived. Constant reservoirs above and below gradient regimes with fingers rapidly become like turbulent high Rayleigh number convection. When the fingers invert the density gradients at their ends, plumes will travel through the entire reservoir because there is no stabilizing

density gradient to slow them down and require further driving from diffusion. This is consistent with both experiment and simulation. Why the interface migrates so slowly is not fully understood.

For fingers growing between two homogenous layers the interior velocities were much greater than the front velocity lengthening the fingers (Linden 1973). This indicates a nearly closed cell and requires boundary layers to maintain the fluxes. Coming up with a finite amplitude theory, even at the level of scaling arguments, that finds this scale separation is an open challenge. On the spatial scale of finger widths and the boundary layer thickness a steady state appears to be reached. Typical vertical velocities of $0.5\,cm/sec$ with a width of $0.23\,cm$ gives a finger Reynolds number ~ 1 for salt-sugar. Glucose-sucrose fingers had typically $w \sim 0.003\,cm/sec$ and $L \sim 0.006\,cm$ for $Re \sim 2 \cdot 10^{-3}$ (Linden 1973). The lower of these Reynolds numbers indicate disruption of cells is not likely to be from shear type instabilites. However,the flow is full of inflection points in the velocity profile. The critical Reynolds number for a viscous sinusoidal flow is $\sqrt{2}$ (Green 1974). A square wave profile is always unstable (Beaumont 1981). The velocity profile in fingers is probably not exactly sinusoidal, but somewhere between sinusoidal and square and may have a critical Reynolds number even less than one. However, it is not know how effective negative density gradients are in stabilizing a flow above the critical Reynolds number for the pattern. This is a well defined, tractable problem with available techniques and requires a modification of the Orr-Somerfeld equation. Whether finite amplitude instabilities at even lower Reynolds numbers exist, as the result of simple subcritical finite amplitude instabilities or from the more complex interactions leading to instability as demonstrated by Orszag and Patera (1983) in bounded shear flows, is an open question. The possibility that

shear flow instabilities contribute to cellular disruption is definitely worth further study, in particular since it is long wavelengths that are first unstable (Beaumont 1981) once the critical Reynolds number is exceeded.

The experiments and observations make it clear that direct double-diffusive convection leads to at least two types of regions with qualitatively different physics. The configuration could be taken as being made up of alternating layers of laminar cellular convection and turbulent convection with steep gradients at their common boundaries. Hence, we have a model composed of two somewhat familiar models, that have been subject to extensive investigation for one component convection. The cellular regime is probably in a state not far from the marginal curve, so that a weakly nonlinear theory can be applied. The turbulent regime probably fluxes buoyancy downgradient, has statically unstable boundary regions, and a neutral interior.

Several things appear to belong in such a theory. Even without fixed boundaries forcing a hydrostatically unstable net density gradient, as in some of the simulations, an organized convective regime strong enough to invert the density gradient at its ends is probably needed to drive the turbulent regime. The limitation on the height of the cellular regime is that in its interior, not counting the boundary layer gradients, it is unable to support growth of the first harmonic, *i.e.*: the Rayleigh number of the interior alone is of order $16\,R_c$. Linden's (1973) data show that the solute gradient becomes so weak as to be unmeasureable accurately in the laminar interior of heat-sugar fingers. How close the resulting configuration is to marginal stability can hence not be determined. Limitations on the height of the turbulent regime are tied to a parcel's ability to move with velocities on a time scale set by the density inversion at the finger ends rather than the slower diffusive time scale. The

boundary layers themselves are probably near marginal based on their thinness, as proposed by Howard (1964) for high Rayleigh number thermal convection. These conditions are not sufficient to close the problem and give unique length scales. When restrictions can be made to configurations that match the heat and salt fluxes in the laminar and turbulent regions that might well close the problem. Fig. 2 in Schmitt 1979 that shows two density ratios with the same flux ratios for small Λ at the fastest growing mode for each Λ is intruiging in looking at this.

Recently, density inversions have been found at the edges of layers stratified to allow direct double-diffusiton in the equatorial Pacific Ocean. Worthem and Ostapoff have found inversions of about $40 \cdot 10^{-6}$, over about $5\,cm$ (E. Mollo-Christensen, personal communication). The root mean square vertical velocity vanishing over the thickness of this layer needed to support this is of order $0.5\,cm/sec$. This is more than sufficient to provide oceanic magnitude buoyancy fluxes with turbulent convection which Schmitt and Evans (1978) estimate of the order $10^{-8}\,cm/sec$ in areas of the North Atlantic where salt fingers are probable.

In the course of this work it would have been useful to have results on single mode finite amplitude calculations, somewhat in the spirit of Joyce (1982). Particularly of interest are saturation amplitudes of different modes as a function of horizontal and vertical wavenumbers, and density ratio for a given fluid with solutes. Accompanying information of note is the flux ratio with a comparison to the growth phase value, and the individual fluxes. As a numerical calculation it is clearly easier than the ones reported herein, but would require many calculations to determine the functional dependences. Possibly results could be obtained by integral methods, at least approximately, as can be found in Chandrasekhar (1961) for thermal convection. Should such calculations show that the most vigorously

fluxing modes are not the fundamental then it would be easier to conclude that layers are a steady state phenomenon and also be possible to suggest a preferred vertical scale. We do not hazard a prediction on the outcome of such work.

The strong horizontal coherence of the layers is the result of two things. Harmonics must be standing modes and are hence locked in place. If the initial conditions then have no broad lateral variation the horizontal translational symmetry must be basically maintained. The second reason applies to all layered phenomenon. The equations for convective motion, in particular in vorticity equation form, are very intolerant of horizontal variations of density. It is essentially horizontal symmetry breaking which is required in continuum mechanics to allow vertical motions, though the forcing is vertical. If the forcing requires horizontal divergences on a scale larger than the finger widths a balance requires vertical flux divergences to satisfy conservation. Even experiments with extreme horizontal divergences, such as the experiments by Ruddick and Turner (1979) involving vertical fronts lead to nearly level layers of cells. The vertical scale in that work is probably not strongly tied to interactions of harmonics since alternating regimes include both direct and overstable convections.

Worthem *et al.* (1983) also take the position that flux divergences lead to layering. However the fact alone that a mode with vertical structure will have flux divergences is insufficient for layering to occur. There must be some mechanism to force coherence of modes. If the finite amplitude modifications of a mode make it likely that another mode will end up spatially out of phase, layering will not result. Alternatively as proposed herein, if the modified state leads to more localized instabilities positioned selectively, layers will develop. Worthem *et al.* show an example of this also where internal waves increase shears so that shear

instabilities will result.

The issue of boundary conditions brings up the problem of the nearly constant density ratio that is observed in the laboratory. This result may be essentially tied to the fact that these are run-down experiments. One still has a mathematically well defined problem if one chooses to fix the fluxes and only one end value. So clearly one can regulate a system to have a density ratio far from Schmitt's (1979) curve in his fig. 2 of χ for the fastest growing modes as a function of Λ. However in a run-down experiment that one observes the fastest growing modes is not surprising. The weakness of the dependence of χ on Λ for the range of Λ used in most experiments explains the lab results. If oceanic fingering is limited by essentially flux boundary conditions, or more realistically mixed boundary conditions, results could be markedly different.

Whether or not one chooses to consider harmonic instabilies true secondary instabilites is probably based more on semantics than physics. The harmonics growing contemporaneously with the faster growing fundamental are in some sense also primary instabilities. The high harmonics whose large growth rates depend on the modification of the mean fields longer wavelength modes could well be called secondary instabilites, though they are still direct double-diffusive convective modes. If one started with a base state of a saturated fundamental, not as was done by the calculations herein, there may be harmonics that are still unstable because the fundamental was unable to fully stabilize the system. However, these harmonics are not drawing on the modification of the linear gradient state as their *primum mobile* and in fact have smaller growth rates than in the original state. After fairly long harmonics create regions of reduced density gradient not up against the boundaries, as the fundamental creates, more rapidly growing harmonics can

occur. We consider modes that have greater growth rates after the modification of the mean field than they did initially, though this rapid growth is localized, to be sceondary instabilities in the spirit of the term. It is these modes that finally disrupt the laminar cellular motions.

9. Summary and Conclusions

A new mechanism for layer formation in direct double-diffusive convection has been proposed. This theory is labeled harmonic instability theory because higher harmonics of the fastest growing convective mode, still themselves convective modes, are shown to be able to disrupt the system. The vertical flux convergences and divergences of these harmonics create layers with density ratio reduced from the mean and even sometimes creating density inversions. These regions of reduced density ratio are unstable to additional small scale convective instabilities with larger growth rates than the primary instabilities.

The differences in the growth rates of the fastest growing fundamental and its first several harmonics are negligible in greatly supercritical systems. The mode that is most rapidly changing the mean field can have a vertical wavenumber significantly greater than the fundamental. However, which mode has affected the mean field the most at any given time is highly time dependent. Conceivably then, the resulting layer scale is strongly dependent on the initial noise spectrum. The model is, therefore, explicative but not predictive, which is disappointing. This system is an interesting demonstration of the fastest growing mode not being preferred because smaller slightly slower growing modes can condition a disruption of it. The only other theory of layering, the collective instability, has been shown to have problems in its fit with available data, mathematical procedures, and the basic physics. It does not appear to be able to provide an appropriate disruption for layer formation.

We conclude that the vertical structure in double-diffusion is the result of interactions of only convective modes as is the case with one component convection.

Several differences between the dispersion relations of one and two component convection lead to qualitatively different configuations. Harmonics in single convection do create regions with stronger gradients, but the more restricted vertical extent does not allow markedly faster growing modes; however, double-diffusive convection by only slightly changing the mean fields can significantly alter the density ratio allowing much faster growing modes. Higher modes in the one component system are fully disruptive of the lateral structure because the strongest horizontal scales rapidly decrease with the decrease of the vertical scales, giving the characteristic chaotic high Rayleigh number cell. Salt finger harmonics have wavelengths similar to the taller disturbances so the cellular pattern is only broken in the thin layers where the secondary instabilities based on low density ratio take over and create restricted turbulent convection regions. Once this configuration has developed, with the gradients concentrated into very thin layers and laminar cell alternating with less organized transports, we believe that there is not much that can disrupt it; *i.e.*: no other instabilities have significant growth rates. That is not to say that the system is globally stable with respect to the boundary conditions. However, the system may be locally stable or any remaining instabilities have negligible growth.

The expense of running a straight forward numerical model with sufficient height to allow layering, with a small enough grid to capture motion of plumes and bulbs, over sufficiently long time for layers to develop, is still prohibitive. A more sophisticated code using different grid sizes for the salt, heat, and vorticity fields along with a dynamically determined grid sized to resolve internal fronts but remaining coarse elsewhere could possibly bring computation costs down to a manageable level. Dynamically altering the time step would be fairly easy way

to increase efficiency. The basic numerical techniques to do this already exist but would still require significant labor to apply them to double-diffusion. It is also possible that a stirred layer numerical boundary region might give greatly different results than the fixed property boundaries used here. A numerical experiment up to a steady state with such a configuration might prove interesting.

Several important problems remain outstanding, the most conspicuous of which is a determination of whether or not a series of layers can remain as a steady state. This remains unsolved on all of experimental, theoretical and numerical levels. It seems likely that at least two laminar finger regions would exist each associated with a boundary, as resulted in Linden's (1978) experiments. However once a layer has formed its lifetime is of the order of the time it takes the fingers to equilibrate the turbulent resevoirs on each side of it. In both the lab and ocean this can be several days or more. To develop a full understanding of the problem a manageable approach might be to separately model the laminar and turbulent regimes and then piece them together with an approach akin to Huppert (1971).

The possibility of instabilities that are not essentially convective should still be considered. Though shear flow instabilities may not provide the initial impetus for layering they may provide an upper bound for layer heights when the finger Reynolds number is supercritical for an appropriate viscous periodic shear flow. Both the numerical simulations performed in the course of this work and some observations imply Reynolds numbers that are close to critical or even exceed it so such an investigation is worthwhile.

Though there is reason to believe from our numerical studies that single mode (laterally) calculations would not accurately depict the full vertical structure of

salt fingers, they would still provide a first order estimate of mode amplitudes at saturation. Amplitude estimates are of particular importance for oceanographic modes requiring heat and salt fluxes. A related problem of interest is the transport of "passive" tracers. Ocean constituents that occur in insufficient quantity to noticably affect the equation of state, and hence drive any motion, still have increased fluxes because of the presence of double-diffusion. The transport of such constituents suffers from the interesting paradox that lower molecular diffusivities lead to greater fluxes. Since the transport of passive chemical species are used for diagnosing the general circulation as well as determining nutrient availability for biological processes and constituent availability in chemical reaction, and since double-diffusive may be the chief mixing process in large regions of the oceans, this is a problem of major oceanographic importance.

Finally, we note that the work here somewhat unifies the theories applied to double-diffusive convection with the more studied problem of single component convection. We propose that the resulting configurations of alternating layers in highly supercritical states of direct double-diffusive convection are the result of the interactions of many convective instabilities over a range of vertical scales and not instabilities of another nature. This is consonant with the generally accepted belief that the planforms and vertical structures in thermal convection stem from nonlinear interactions of multiple convecting modes.

References

Arakawa, Akio, 1966: *J. Comp. Phys.* **1**, 119–143, Computational design for long-term numerical integrations of fluid motions: two dimensional incompressible flow. Part I.

Arons, Arnold B., 1981: in **Evolution of Physical Oceanography**, Bruce A. Warren and Carl Wunsch, eds., *xiv – xviii*, MIT Press, Cambridge, Massachusetts, The scientific work of Henry Stommel.

Baines, P. G. and A. E. Gill, 1969: *J. Fluid Mech.* 37 (2), 289–306, On thermohaline convection with linear gradients.

Beaumont, D. A., 1981: *J. Fluid Mech.* **108**, 461–474, The stability of spatially periodic flows.

Busse, F. H., 1967: *J. Math. and Phys.* **42** (2), 140–150,On the stability of two-dimensional convection in a layer heated from below.

Busse, F. H., 1978a: *Rep. Prog. Phys.* **41**, 1929–1967, Non-linear properties of thermal convection.

Busse, F. H., 1978b: *Adv. in Appl. Mech.* **18**, 77–121, The optimum theory of turbulence.

Busse, F. H., 1981: in **Hydrodynamic Instabilities and the Transition to Turbulence**, H. L. Swinney and J. P. Gollub, eds., ch. 5, Transition to turbulence in Rayleigh-Bénard convection.

Busse, F. H. and J. A. Whitehead, 1971: *J. Fluid Mech.* **47** (2), 305–320, Instabilities of convection rolls in a high Prandtl number fluid.

Busse, F. H. and J. A. Whitehead, 1974: *J. Fluid Mech.* **66**, 67–79, Oscillatory and collective instabilities in large Prandtl number convection.

Calman, Jack, 1977: *Dynam. of Atmos. and Oceans* **1**, 277–297, Experiments on high Richardson number instability of a rotating stratified shear flow.

Chandrasekhar, S., 1961: **Hydrodynamic and Hydromagnetic Stability**, Clarendon Press, Oxford.

Chen, M. M. and J. A. Whitehead, 1968: *J. Fluid Mech.* **31**, 1–15, Evolution of two-dimensional periodic Rayleigh convection cells of arbitrary wave-numbers.

Cooper, J. W. and H. Stommel, 1968: *J. Geophys. Res.* 73 (18), 5849–5854, Regularly spaced steps in the main thermocline near Bermuda.

Copley, S. M., A. F. Giamei, S. M. Johnson, and M. F. Hornbecker, 1970: *Metallurgical Trans.* **1**, 2193–2204, The origin of freckles in unidirectionally solidified castings.

Frick, H., F. H. Busse, and R. M. Clever 1983: *J. Fluid Mech.* **127**, 141–153, Steady three-dimensional convection at high Prandtl numbers.

Frick, H. and U. Müller, 1983: *J. Fluid Mech.* **126**, 521–532, Oscillatory Hele-Shaw convection.

Garrett, Christopher, 1982: *J. Phys. Oceanogr.* **12**, 952–959, On the parameterization of diapycnal fluxes due to double-diffusive intrusions.

Gollub, J. P., S. L. Hulbert, G. M. Dolny, and H. L. Swinney, 1975: in **Photon Correlation Spectroscopy and Velocimetry**, H. Z. Cummins and E. R. Pike, eds.. NATO Advanced Study Institute on Photon Correlation Spectroscopy, Plenum Press, New York, 425–439, Laser Dopler study of the onset of turbulent convection at low Prandtl number.

Gough, D. O. and Juri Toomre, 1982: *J. Fluid Mech.* **125**, 75–97, Single-mode theory of diffusive layers in thermohaline convection.

Gray, Donald D. and Aldo Giorgini, 1976: *Int. J. Heat Mass Trans.* **19**, 545–551, The validity of the Boussinesq approximation for liquids and gases.

Green, J. S. A., 1974: *J. Fluid Mech.* **62** (2), 273–287, Time-dependent turbulence near the viscous limit.

Griffiths, R. W., 1979: *Deep-Sea Res.* **26A**, 383–397, The transport of multiple components through thermohaline diffusive interfaces.

Griffiths, R. W. and B. R. Ruddick, 1980: *J. Fluid Mech.* **99** (1), 85–95, Accurate fluxes across a salt-sugar finger interfaces deduced from direct density measurements.

Grötzbach, Günter, 1982: *J. Fluid Mech.* 119, 27-53, Direct numerical simulation of laminar and turbulent Bénard convection.

Holyer, Judith Y., 1981: *J. Fluid Mech.* **110**, 195–207, On the collective instability of salt fingers.

Howard, L. N., 1964: *Proc. 11th Int. Cong. Theoret. Appl. Mech.*, 1109–1115, Convection at high Rayleigh number.

Huppert, Herbert E., 1971: *Deep-Sea Res.* **18**, 1005–1021, On the stability of a series of double-diffusive layers.

Huppert, Herbert E. and P. F. Linden, 1979: *J. Fluid Mech.* **95**(3), 431–464, On heating a stable salinity gradient from below.

Huppert, Herbert E. and Peter C. Manins, 1973: *Deep-Sea Res.* **20**, 315–323, Limiting conditions for salt fingering at an interface.

Jevons, W. S., 1857: *Phil. Mag. and J. Sci.* **XIV - Fourth series**, London, 22–35, On the *cirrous* form of cloud.

Johns, Bryan and Michael J. Cross, 1970: *J. Mar. Res.* **28** (2), 215–224, The decay and stability of internal wave modes in a multisheeted thermocline.

Joyce, Terrence M., 1982: *J. Mar. Res.* **40 Supplement**, 291–306, Marginally unstable salt fingers: limits to growth.

Kittel, Charles, 1976: **Introduction to Solid State Physics, fifth ed.**, John Wiley and Sons, New York.

Knobloch, E. and M. R. E. Proctor, 1981: *J. Fluid Mech.* **108**, 291–316, Nonlinear periodic convection in dobule-diffusive systems.

Lambert, Richard B., Jr., 1974: *Deep-Sea Res.* 21, 529–546, Small-scale dissolved oxygen variations and the dynamics of Gulf Stream eddies.

Lambert, R. B. and J. W. Demenkow, 1972: *J. Fluid Mech.* **54** (4), 627–640, On the vertical transport due to fingers in double-diffusive convection.

Lambert, Richard B., Jr. and Wilton Sturges, 1977: *Deep-Sea Res.* **24**, 211–222, A thermohaline staircase and vertical mixing in the thermocline.

Linden, P. F., 1973: *Deep-Sea Res.* **20**, 325–340, On the structure of salt fingers.

Linden, P. F., 1974: *Geophys. Fluid Dyn.* **6** (1), 1–27, Salt fingers in a steady shear flow.

Linden, P. F., 1976: *Deep-Sea Res.* **23**, 895–908, The formation and destruction of fine-structure by double-diffusive processes.

Linden, P. F., 1978: *J. Geophys. Res.* **83** (C6), 2902–2912, The formation of banded salt finger structure.

Linden, P. F. and J. E. Weber, 1977: *J. Fluid Mech.* **81** (4), 757–773, The formation of layers in a double-diffusive system with a sloping boundary.

Magnell, Bruce, 1976: *J. Phys. Oceanogr.* **6**, 511–523, Salt fingers observed in the Mediterranean outflow region (34°N, 11°W) using a towed sensor.

Malkus, W. V. R., 1954: *Proc. Roy. Soc.* **A225**, 196, The heat transport and spectrum of thermal turbulence.

Malkus, W. V. R., 1963: in **Theory and Fundamental Research in Heat Transfer**, 203–212, Outline of a theory of turbulent convection.

Mason, D. W., 1976: *Biophys. J.* **16**, 407-416, A diffusion driven instability in systems that separate particles by velocity sedimentation.

McDougall, Trevor J., 1983: *J. Fluid Mech.* **126**, 379–397, Double-diffusive convection caused by coupled molecular diffusions.

McIntyre, Michael E., 1970a: *Geophys. Fluid Dyn.* **1**, 19–57, Diffusive destabilization of the baroclinic circular vortex.

McIntyre, Michael E., 1970b: *Geophys. Fluid Dyn.* **1**, 58–89, Role of diffusive overturning in nonlinear axisymmetric convection in a differentially heated rotating annulus.

Merceret, Francis J. 1977: *Boundary-Layer Meteorology* **11**, 121–123, A possible manifestation of double diffusive convection in the atmosphere.

Miles, John W., 1972: *J. Fluid Mech.* **53** (3), 557–573, Internal waves in a sheeted thermocline.

Munk, Walter H., 1966: *Deep-Sea Res.* **13**, 707–730, Abyssal recipes.

Nield, D. A., 1967: *J. Fluid Mech.* **29** (3), 545–558, The thermohaline Rayleigh-Jeffreys problem.

Orszag, Steven A. and Anthony T. Patera, 1983: *J. Fluid Mech.* **128**, 347–385, Secondary instability of wall-bounded shear flow.

Pearlstein, A. J., 1981: *J. Fluid Mech.* **103**, 389–412, Effect of rotation on the stability of a doubly diffusive layer.

Phillips, Norman A., 1959: in **The Atmosphere and the Sea in Motion**, Rossby Memorial Volume, Bert Bolin, ed., Rockefeller Inst. Press, New York, 501–504, An example of non-linear computational instability.

Piacsek, S. A. and J. Toomre, 1980: in **Marine Turbulence**, J. C. J. Nihoul, ed., 193–219, Nonlinear evolution and structure of salt fingers.

Roache, Patrick J., 1972: **Computational Fluid Dynamics**, Hermosa Publishers, Albuquerque, New Mexico.

Ruddick, B. R. and J. S. Turner, 1979: *Deep-Sea Res.* **26A**, 903–913, The vertical length scale of double-diffusive intrusions.

Ruelle, David and Floris Takens, 1971: *Commun. Math. Phys.* **20**, 167–192, On the nature of turbulence.

Schaefer, Joseph T., 1975: *J. Atmos. Sci.* **32**, 2278–2284, Nonlinear biconstituent diffusion: a possible trigger of convection.

Schlüter, A., D. Lortz and F. Busse, 1965: *J. Fluid Mech.* **28**, 223–239, On the stability of steady finite amplitude convection.

Schmitt, Raymond W., Jr., 1979a: *Deep-Sea Res.* **26A**, 23–40, The growth rate of super-critical salt fingers.

Schmitt, Raymond W., Jr., 1979b: *J. Mar. Res.* **37** (3), 419–436, Flux measurements on salt fingers at an interface.

Schmitt, Raymond W., 1981: *J. Phys. Oceanogr.* **11** (7), 1015–1026, Form of the temperature-salinity relationship in the Central Water: evidence for double-diffusive mixing.

Schmitt, Raymond W., 1983: *Phys. Fluids* **26** (9), 2373–2377, The characteristics of salt fingers in a variety of fluid systems, including stellar interiors, liquid metals, oceans and magmas.

Schmitt, Raymond W., Jr. and David L. Evans, 1978: *J. Geophys. Res* **83** (C6), 2913–2919, An estimate of the vertical mixing due to salt fingers based on observations in the North Atlantic Central Water.

Schmitt, Raymond W. and Richard B. Lambert 1979: *J. Fluid Mech.* **90** (3), 449–463, The effects of rotation on salt fingers.

Scorer, Richard, 1972: **Clouds of the World**, Lathian Publishing, Melbourne, Australia, 54–56.

Shirtcliffe, T. G. L. and J. S. Turner, 1970: *J. Fluid Mech.* **41** (4), 707–719, Observations of the cell structure of salt fingers.

Stern, Melvin E., 1960: *Tellus* **12** (2), 172–175, The "salt-fountain" and thermohaline convection.

Stern, Melvin E., 1967: *Deep-Sea Res.* **14**, 747–753, Lateral mixing of water masses.

Stern, Melvin E., 1969a: *J. Fluid Mech.* **35**, 209–218, Collective instability of salt fingers.

Stern, Melvin E., 1969b: *Deep-Sea Res.* **16** (3) Supplement, 263–267, Salt finger convection and the energetics of the general cirulation.

Stern, Melvin E., 1975: **Ocean Circulation Physics**, Academic Press, New York, ch. 9–11.

Stern, Melvin, 1976: *J. Mar. Res.* **34** (1), 95–110, Maximum buoyancy flux across a salt finger interface.

Stern, Melvin E., 1981: *J. Fluid Mech.* **114**, 105–121, Inequalities and variational principles in double-diffusive turbulence.

Stern, Melvin E. and J. Stewart Turner, 1969: *Deep-Sea Res.* **16**, 497–511, Salt fingers and convecting layers.

Stommel, Henry, Arnold B. Arons, and Duncan Blanchard, 1956: *Deep-Sea Res.* **3**, 152–153, An oceanographic curriosity; the perpetual salt fountain.

Stong, C. L., 1971: *Sci. Amer.* **224** (6), June, 124–128, The amateur scientist: Experiments with salt fountains and related instabilities of water.

Straus, Joe M., 1972: *J. Fluid Mech.* **56** (2), 353–374, Finite amplitude doubly diffusive convection.

Tait, R. W. and M. R. Howe, 1968: *Deep-Sea Res.* **15**, 275–280, Some observations of thermo-haline stratification in the deep ocean.

Tennekes, H. and J. L. Lumley, 1972: **A First Course in Turbulence**, MIT Press, Cambridge, Massachusetts.

Toole, John M. and Daniel T. Georgi, 1981: *Prog. Oceanog.* **10**, 123–145, On the dynamics of double-diffusively driven intrusions.

Turner, J. S., 1965: *Int. J. Heat Mass Transfer* **8**, 759–767, The coupled turbulent transports of salt and heat across a sharp density interface.

Turner, J. S., 1967: *Deep-Sea Res.* **14**, 599–611, Salt fingers across a density interface.

Turner, J. S., 1968: *J. Fluid Mech.* **33** (1), 183–200, The behaviour of a stable salinity gradient heated from below.

Turner, J. S., 1978: *J. Geophys. Res.* **83** (C6), 2287–2901, Double-diffusive intrusions into a density gradient.

Turner, J. S. and C. F. Chen, 1974: *J. Fluid Mech.* **63** (3), 577–592, Two-dimensional effects in double-diffusive convection.

Turner, J. S. and Henry Stommel, 1964: *Proc. Natl. Acad. Sci.* **52**, 49–53, A new case of convection in the presence of combined vertical salinity and temperature gradients .

Ulrich, Roger K., 1972: *Astrophys. J.* **172**, 165–177, Thermohaline convection in stellar interiors.

Veronis, George, 1965: *J. Mar. Res* **23** (1), 1–17, Finite amplitude instability in thermohaline convection.

Walin, Gösta, 1964: *Tellus* **16** (3), 389–393, Note on the stability of water stratified by both salt and heat.

Williams, Albert J., 3rd, 1974: *Science* **185**, 941–943, Salt fingers observed in the Mediterranean outflow.

Williams, Albert J., 3rd, 1975: *Deep-Sea Res.* **22**, 811–829, Images of oceanic microstructure.

Williams, Garreth P., 1969: *J. Fluid Mech.* **37** (4), 727–750, Numerical integration of the three-dimensional Navier-Stokes equations for an incompressible flow.

Worthem, Sylvia, Erik Mollo-Christensen, and F. Ostapoff, 1983: *J. Fluid Mech.* **133**, 297–319, Effects of rotation and shear on doubly diffusive instability.

Wunsch, Carl, 1970: *Deep-Sea Res.* **17**, 293–301, On oceanic boundary mixing.

Zeman, Otto and John L. Lumley, 1982: *J. Mar. Res.* **40** (2), 315–330, Modeling salt-fingering structures.