J. Fluid Mech. (2010), *vol.* 651, *pp.* 443–464.
doi:10.1017/S0022112009994150

Multi-layered diffusive convection. Part 1. Spontaneous layer formation

TAKASHI NOGUCHI† AND HIROSHI NIINO

Ocean Research Institute, The University of Tokyo, Nakano, Tokyo 164-8639, Japan

(Received 25 August 2009; revised 9 December 2009; accepted 18 December 2009; first published online 26 March 2010)

Diffusive convection in an infinite two-dimensional fluid with linear vertical gradients of temperature and salinity is studied numerically and analytically. When the density gradient ratio exceeds a critical value above which diffusive convection grows according to the linear stability analysis, spontaneous layer formation is found to occur. At the first stage nearly steady oscillating motions, the horizontal scale of which is of the order of the buoyancy boundary layer scale δ, arise. After several tens of the oscillation cycle, a transition to the second stage occurs in which overturning convective motions develop and well-mixed regions are formed. The convective motions resemble Rayleigh–Bénard convection at a high Rayleigh number. The well-mixed regions are gradually organized into horizontal layers, a typical thickness of which is of the order of δ. A detailed analysis of the nonlinear process during the layer formation reveals that four modes are responsible for the layer formation: The first mode is the linear fastest-growing mode with wavenumber vector $(k_0, 0)$. The second mode with (k_0, m_0) is weakly growing. The third mode with $(0, m_0)$ is dissipating, and the fourth mode is its higher harmonic having $(0, 2m_0)$. It is shown that a truncated spectral model with the four modes well reproduces the results of the full numerical simulation.

1. Introduction

Step-like vertical structures of salinity and temperature are observed in many regions of the oceans (Swallow & Crease 1965; Hoare 1966; Tait & Howe 1968; Neal, Neshyba & Denner 1969; Neshyba, Neal & Denner 1971). In these structures, nearly neutrally stratified layers are sandwiched by thin interfaces with sharp and stable gradients of density and concentration. Each layer has a typical vertical dimension of about 10 m and has horizontal coherency over several kilometres (Tait & Howe 1968).

One of the most promising candidates to produce these observed layered structures is the double-diffusive convection. There have been a number of laboratory studies on such layering processes, not only from the viewpoint of oceanography (Turner 1965), but also from that of geology (Huppert & Turner 1981). Most of these studies examined the development of convective layers due to some buoyancy flux from the boundaries: e.g. a linear gradient of salinity is heated from the bottom (Turner 1968; Huppert & Linden 1979; Fernando 1987) or linear gradients of salinity and

† Present address: Department of Aeronautics and Astronautics, Kyoto University, Yoshida-Honmachi, Sakyo, Kyoto 606-8501, Japan. Email address for correspondence: noguchi@kuaero.kyoto-u.ac.jp

temperature are sandwiched by uniformly stratified layers (Linden 1976). In both configurations, convective layers are produced one by one from the boundaries. The thickness of the layers decreased with the distance from the boundaries. However, the layered structures observed in the oceans seem to have no such monotonic decrease of layer thickness with depth. Linden (1976) speculated that multi-layered convection may be possible even in the absence of buoyancy forcing from the boundary, if the available potential energy stored in the temperature stratification is released. However, this speculation has never been proved from either a numerical or a laboratory experiment. Thus the mechanism for the formation of uniform layers is not well understood yet.

There are a number of theoretical studies on the stability of a double-diffusive stratification. Baines & Gill (1969) analysed the linear stability of double-diffusive stratification with constant temperature and salinity gradients and summarized the results in a regime diagram. Using a truncated spectral model, Veronis (1965, 1968) and Ahlers & Lücke (1987) studied a nonlinear growth of the diffusive convection in an initially uniform stratification. These theoretical studies considered the double-diffusive problem as an extension of the Rayleigh–Bénard convection, so that they used a configuration in which a layer of the fluid exists between two horizontal parallel plates on which temperature and salinity are specified. Convective motions are driven by the flux through the boundaries.

Linden (1976), on the other hand, speculated that disturbances may be able to develop within a diffusively stratified fluid, even in the absence of boundary fluxes. He anticipated that the disturbances can grow at the expense of the available potential energy stored in the 'top-heavy' temperature stratification. However, this idea has never been pursued until now. Our objectives are to investigate the condition for the growth of small disturbances, their evolution processes and the final state.

The present paper consists of two parts. In Part 1 we investigate the characteristics of the diffusive convection and its nonlinear evolution using a direct numerical simulation (DNS) and an analytical model. In Part 2 (Noguchi & Niino 2010) we analyse more in detail the growth of layers found in the DNS of Part 1 and construct a simple mechanistic model of the layer growth.

Part 1 is organized as follows. In § 2, we formulate the equations of motion in an infinite diffusive stratification. We analyse its linear stability in § 3. In § 4, we outline the numerical model and present the results of numerical simulation in § 5. The results are analysed in terms of nonlinear interaction among several wave modes in § 6. Based on the analysis, we propose a truncated spectral model in § 7. We interpret and discuss the results in § 8. Finally the main findings are summarized in § 9.

2. Formulation of the problem

Consider a fluid stratified with two components T and S, where T denotes the faster diffusing component and S denotes the slower diffusing component. In what follows, we consider a diffusive stratification in which T is stratified unstably ($\mathrm{d}\overline{\rho_T}/\mathrm{d}z > 0$), S is stratified stably ($\mathrm{d}\overline{\rho_S}/\mathrm{d}z < 0$) and the total density $\overline{\rho}$ is also stratified stably ($\mathrm{d}\overline{\rho}/\mathrm{d}z < 0$), where $\overline{\rho_T}$ and $\overline{\rho_S}$ denote contributions of T and S to the total density and the overbar denotes a basic density field (figure 1). In the oceans, $-T$ and S correspond to heat and salt, respectively, and diffusive stratifications tend to be found in high latitudes.

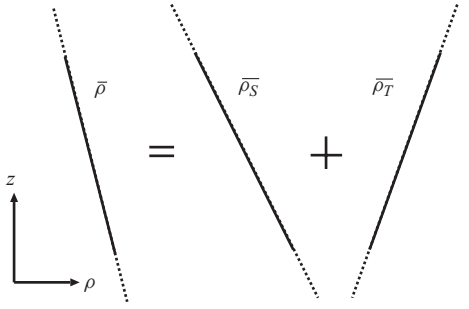

FIGURE 1. Basic stratification.

We will use the Boussinesq approximation. The density is assumed to be a linear function of temperature and salinity:

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

where α and β are the coefficients of expansion for T and S, respectively, and ρ_0, T_0 and S_0 are the reference values for density, T and S, respectively.

In this paper, we will consider only two-dimensional (x, z) motions for simplicity. The governing equations in dimensional form are given as follows:
Equations of motion

$$\frac{\mathrm{D}u_*}{\mathrm{D}t_*} = -\frac{\partial p_*}{\partial x_*} + \nu\nabla_*^2 u_*, \tag{2.2}$$

$$\frac{\mathrm{D}w_*}{\mathrm{D}t_*} = -\frac{\partial p_*}{\partial z_*} - g(\alpha T_* + \beta S_*) + \nu\nabla_*^2 w_*, \tag{2.3}$$

conservation of the faster diffusing component,

$$\frac{\mathrm{D}T_*}{\mathrm{D}t_*} = \kappa_T \nabla_*^2 T_*, \tag{2.4}$$

conservation of the slower diffusing component,

$$\frac{\mathrm{D}S_*}{\mathrm{D}t_*} = \kappa_S \nabla_*^2 S_*, \tag{2.5}$$

and continuity equation

$$\frac{\partial u_*}{\partial x_*} + \frac{\partial w_*}{\partial z_*} = 0, \tag{2.6}$$

where κ_T and $\kappa_S (< \kappa_T)$ are the diffusivities of the faster and the slower diffusing components, respectively, ν is the kinematic viscosity and g is the gravity acceleration. Variables with asterisks are dimensional.

In what follows, we consider a diffusively stratified fluid with an infinite extent. Let the initial stratification be given by

$$T_* = T_{0*} + T_{z*}\, z_*, \quad S_* = S_{0*} + S_{z*}\, z_*, \tag{2.7}$$

where T_{0*} and S_{0*} are constants. The fluid is initially motionless and horizontally homogeneous.

Although there is no external length scale, the present configuration has an internal length scale based on diffusivity of the faster diffusing component, kinematic viscosity, gravity acceleration and vertical density stratification due to the gradient of the faster

diffusing component:

$$\delta \equiv \left| \frac{g\alpha \overline{T}_z}{\kappa_T \nu} \right|^{-1/4}. \tag{2.8}$$

This length scale is equivalent to the horizontal thickness of the wall boundary layer in the heat-up problem of a stable stratification from a sidewall (Prandtl 1952) and is sometimes called a 'buoyancy layer thickness'. It is also identical to the scale introduced by Stern (1969) for the salt-finger problem. The length scale δ also gives a layer thickness for which a Rayleigh number based on the temperature difference is equal to unity in the linear temperature stratification.

Let us now non-dimensionalize the governing equations by scaling length by δ, time by δ^2/κ_T and the velocity by κ_T/δ. The faster and the slower diffusing components are scaled by their concentration difference over the vertical distance of δ. The resulting vorticity equations derived from (2.2)–(2.6) are:

$$\left(\frac{\partial}{\partial t} - Pr\nabla^2\right)\nabla^2\psi = -J(\psi, \nabla^2\psi) - Pr\left(\frac{\partial T}{\partial x} - \frac{1}{\gamma}\frac{\partial S}{\partial x}\right), \tag{2.9}$$

$$\left(\frac{\partial}{\partial t} - \nabla^2\right)T = -J(\psi, T) - \frac{\partial \psi}{\partial x}, \tag{2.10}$$

$$\left(\frac{\partial}{\partial t} - \tau\nabla^2\right)S = -J(\psi, S) - \frac{\partial \psi}{\partial x}, \tag{2.11}$$

where $J(a, b)$ is the Jacobian operator

$$J(a, b) \equiv \frac{\partial a}{\partial x}\frac{\partial b}{\partial z} - \frac{\partial a}{\partial z}\frac{\partial b}{\partial x},$$

and ψ is the streamfunction, $u = -\partial\psi/\partial z$, $w = \partial\psi/\partial x$.

Equations (2.9)–(2.11) show that the present problem is governed by three non-dimensional parameters, Prandtl number, $Pr = \nu/\kappa_T$, diffusivity ratio, $\tau = \kappa_S/\kappa_T$ and density gradient ratio, $\gamma = \alpha T_z/\beta S_z$. Once we have chosen solute components for T and S and solvent, Pr and τ are determined. Thus, γ is the only parameter we can vary. In the present paper, we consider a heat-salt system so that $Pr = 7$ and $\tau = 0.01$.

3. Linear stability analysis

Based on (2.9)–(2.11), we first examine the response of the fluid to infinitesimal disturbances superposed on the initially motionless fluid.

Let us consider an infinitesimal disturbance in the form of a plane wave whose wavenumber vector is $\boldsymbol{k} = (k, m)$, where k and m are horizontal and vertical wavenumbers, respectively. Any variable, say η, is assumed to be expressed as

$$\eta = \tilde{\eta}\, \mathrm{e}^{\mathrm{i}(kx+mz)}\mathrm{e}^{\sigma t},$$

where $\tilde{\eta}$ is a constant amplitude and $\sigma \equiv \sigma_\mathrm{r} + \mathrm{i}\sigma_\mathrm{i}$ is a complex growth rate. Dropping the second-order terms with respect to the disturbance, and assuming that a non-trivial solution exists, we obtain the dispersion relation, which is a cubic equation of σ:

$$[\sigma + (k^2+m^2)][\sigma + Pr(k^2+m^2)][\sigma + \tau(k^2+m^2)]\left(1 + \frac{m^2}{k^2}\right) - Pr\,[(\sigma + \tau(k^2+m^2)) - \gamma^{-1}(\sigma + (k^2+m^2))] = 0. \tag{3.1}$$

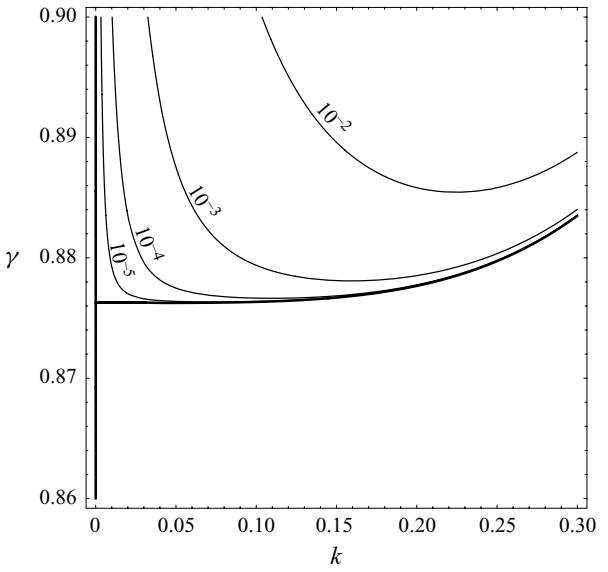

FIGURE 2. Growth rate (the largest real part of σ) in the γ–k plane, for heat-salt system. The thick curve denotes neutral curve. Only contours of positive values are shown.

For a given combination of the horizontal and vertical wavenumbers (k, m), the growth rate σ can be calculated from (3.1). Since the equation generally has three roots, we choose the root which has the largest real part among the three as the growth rate for the disturbance with wavenumbers (k, m).

It is found that growth rate has a maximum at $m = 0$ for all γ. Figure 2 shows the dependency of the growth rate on k. It is seen that the uniformly stratified basic state is linearly unstable when γ is greater than a critical value γ_{cr} of 0.876. It is interesting to note that this critical value for the infinite layer is same as that for a layer between horizontal free-slip boundaries in the limit of infinite thermal and solutal Rayleigh numbers (Baines & Gill 1969, (2.13)). When γ exceeds γ_{cr}, the instability occurs at zero vertical wavenumber, although its growth rate is infinitesimal. When we increase γ further, the fastest growing wavenumber becomes finite and increases as γ.

The growth rate at slightly supercritical stratification, say $\gamma = 0.88$, is shown as a function of (k, m) in figure 3. All modes inside the outermost semicircle have positive growth rates. Note that all the growing modes have growth rates with non-zero imaginary part. Thus, the disturbance grows with oscillation. The figure also shows that the fastest growing mode (FGM) is at $(k, m) = (0.187, 0)$, and has a growth rate of 2.7×10^{-3} and frequency of 0.98. In the physical space, this mode takes the form of infinitely long vertical columns, whose horizontal width is about 33 δ. The vertical velocity inside the columns repeats oscillations while increasing its amplitude at a very slow rate compared to the oscillation frequency.

4. Numerical simulation

The linear stability analysis in the previous section describes well the initial growth of the disturbances of small amplitude. When the amplitude becomes large, however, the linear analysis is unable to predict the evolution of the disturbances including the dominant wavenumber. In order to study the nonlinear evolution of the disturbances, we have to rely on a nonlinear numerical experiment. In the rest of this section,

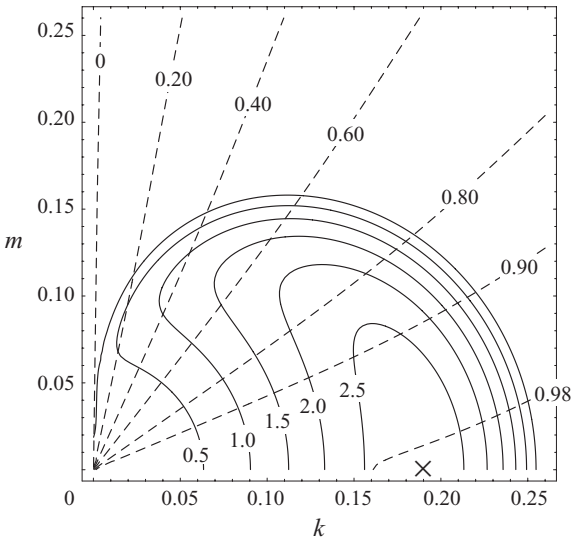

FIGURE 3. Complex growth rate in the k–m plane. The solid line shows the growth rate (real part), and the dashed line shows the frequency (imaginary part). Note that the value of the real part is multiplied by 10^3. The cross mark on the k-axis denotes the fastest growing mode (FGM).

we will focus on a diffusive stratification of heat-salt system with $\gamma = 0.88$, which is weakly unstable with respect to diffusive convection according to the linear theory.

In the numerical experiment, the governing equations (2.9)–(2.11) are cast into a finite-difference form. The time evolution of variables are calculated by 'leap-frog' scheme except for the viscosity and diffusion terms for which the Euler forward difference scheme is used. Heun scheme is used as a numerical filter at every 20 steps. Variables are configured on a square staggered grid system: i.e. vorticity and streamfunction are placed at the vertex, while T and S are placed at the centre. The third-order upstream scheme (Roache 1972) is used for the advection terms, since it is suitable for handling very sharp internal interfaces. Although the scheme is known to have spurious numerical diffusion, it turns out that the model succeeded to make quite a realistic prediction of the flow characteristics as will be seen below. No artificial parameterization of turbulent mixing were used. Only the molecular viscosity and molecular diffusivities were considered. The Poisson equation which calculates streamfunction from vorticity is solved by the Fourier transform method.

The calculation is made on a square domain of $632\delta \times 632\delta$ in the x–z plane. This size is about 20 times larger than the wavelength of the FGM predicted by the linear stability theory. Grid points of 256×256 are used for the calculation, so that the grid interval is sufficiently small to resolve the FGM.

Periodic boundary conditions are imposed on all physical variables at the horizontal and vertical boundaries. For temperature and salinity, which have basic linear gradients in the vertical direction, only their deviations from the basic states are explicitly predicted while imposing periodic boundary conditions on the deviations (figure 4).

A small white noise with random phase is imposed as the initial temperature disturbances. Thus every mode has same amplitude over the whole wavenumber space, initially. The amplitude of the initial disturbances is about 10^{-4} times the

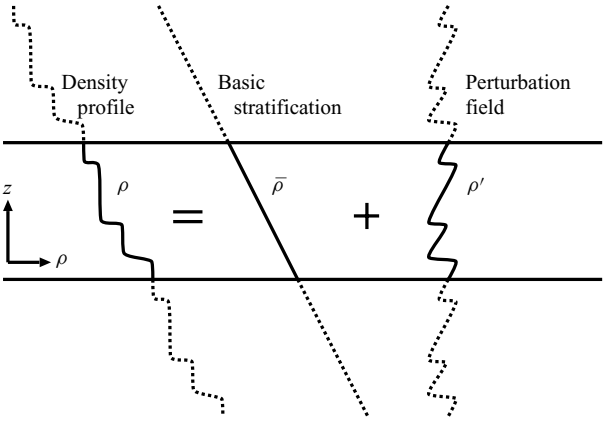

FIGURE 4. Schematic of boundary conditions for ρ, T and S. Only the perturbation field, which is obtained from original stratification by subtraction of the constant basic stratification, is assumed to be cyclic.

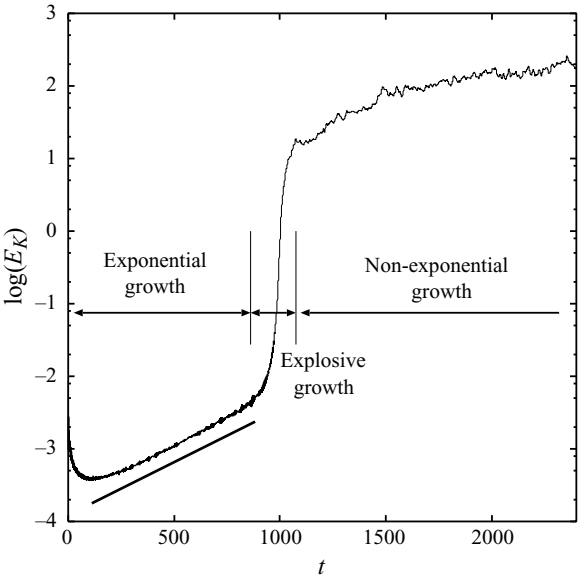

FIGURE 5. Time evolution of kinetic energy (E_K) per unit volume, averaged over the whole calculation domain, for $\gamma = 0.88$. The slope of the straight line gives the growth rate of the FGM predicted by the linear stability analysis.

temperature difference between the top and bottom of the calculation domain. The time step for the integration is $\Delta t = 5 \times 10^{-3}$, and the calculation is continued until $t = 4000$.

5. Numerical results

5.1. *Overview*

Figure 5 shows the time evolution of the kinetic energy density averaged over the whole calculation domain. After a brief period of total energy decrease, which is

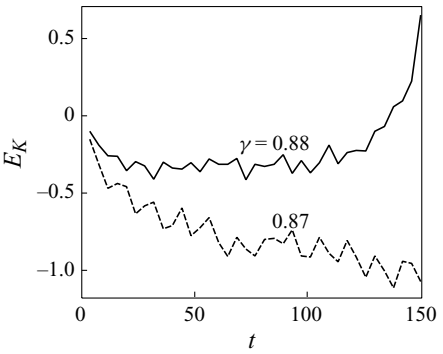

FIGURE 6. Time evolutions of kinetic energy averaged over the whole calculation domain, for supercritical and subcritical values of $\gamma = 0.88$ (solid) and $\gamma = 0.87$ (dashed), respectively.

presumably due to damping of a class of disturbances with negative growth rates, the initial random perturbations first start to grow exponentially with time, at a rate comparable to that of the FGM predicted by the linear theory ($100 \lesssim t \lesssim 900$). Following this exponential growth phase, a sudden explosive growth of kinetic energy density takes place around $t \approx 1000$.

After the kinetic energy per unit volume exceeds $O(1)$ at around $t \approx 1100$, the growth rate slows down. The kinetic energy exhibits much larger fluctuations than before. In this stage, the growth is no longer exponential.

For the convenience of the following analysis, the time evolution of the disturbance will be divided into three stages, which will be called exponential growth stage, explosive growth stage and non-exponential growth stage, respectively, as shown in figure 5.

5.2. *Exponential growth stage*

5.2.1. *Linear growth rate*

A preliminary experiment has been made with two different density ratios: one is supercritical ($\gamma = 0.88$) and the other is subcritical ($\gamma = 0.87$). Figure 6 shows the time evolutions of kinetic energy for these two experiments. For $\gamma = 0.87$ the energy decreases monotonically with time, while for $\gamma = 0.88$ it first decreases slightly because of damping of modes with negative growth rates, but soon increases with time exponentially. These results are consistent with the prediction of the linear theory and support the accuracy of the present numerical model.

5.2.2. *Growth in wavenumber space*

Figure 7 shows the time evolution of the amplitude of temperature perturbation in the (k–m) wavenumber space. During the initial period ($t \lesssim 200$), the amplitude of each mode on the wavenumber space evolves in a way very similar to that predicted by linear theory (figure 7*a*, *b*; cf. figure 3). Modes of large horizontal and vertical wavenumbers decay rapidly, while those inside the semicircle of the neutral curve retain their amplitudes. The mode with the largest amplitude is the FGM predicted by the linear theory for $t \lesssim 200$.

As time further elapses, however, the modes along the m-axis start to grow (figure 7*c*, *d*), which contradicts the linear theory. All of the modes with $m = 0$ have negative growth rate. By $t \approx 700$, the FGM ceases to be dominant and a mode whose wavenumber is about $(k, m) = (0.19, 0.09)$ starts to have the largest amplitude (figure 7*c*). Note that this new dominant mode has the same horizontal wavenumber as the FGM. The modes along the m-axis are still growing at this time.

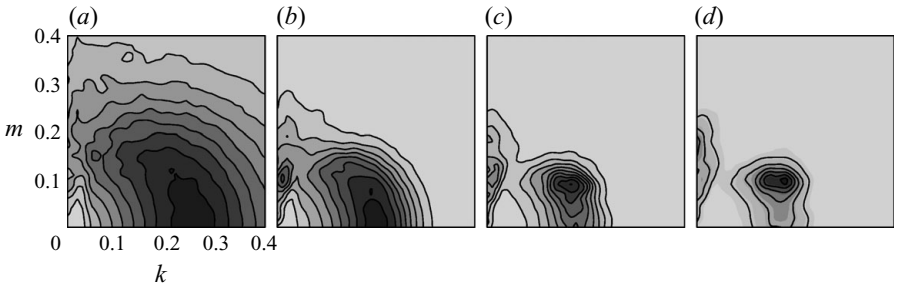

FIGURE 7. Amplitude of Fourier modes of the temperature perturbation in the wavenumber-plane, at four different times: (*a*) $t = 50$ (soon after the initiation), (*b*) $t = 200$, (*c*) $t = 700$ and (*d*) $t = 900$. The contours are normalized by the largest amplitude at each time.

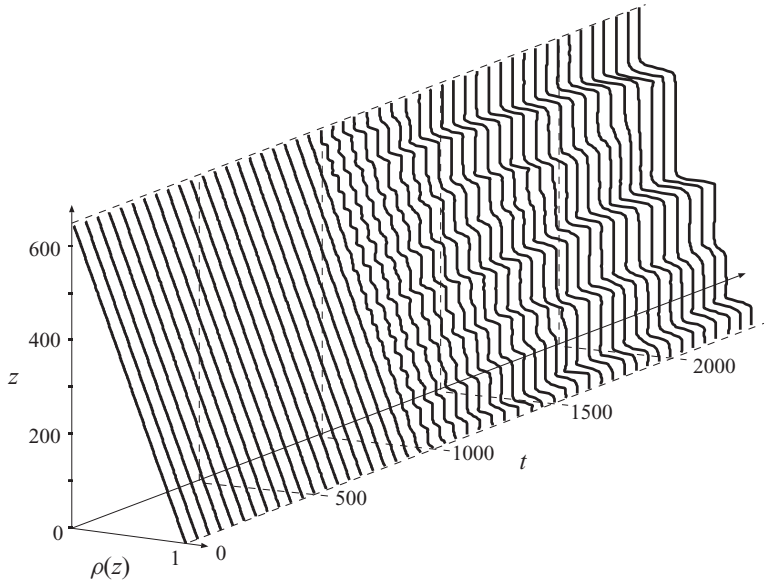

FIGURE 8. Time evolution of vertical density distribution. The density is averaged in horizontal direction, and is displayed every 50 non-dimensional time units.

The new dominant mode continues to grow by $t \approx 900$ (figure 7*d*). At the same time, the amplitude that is maximum along the m-axis shifts to larger m.

These complex behaviours are not explained by the linear theory and certainly reflect nonlinear interactions among various modes of the diffusive convection.

5.3. *Explosive growth stage*

At the explosive growth stage, many aspects of the flow change drastically. The kinetic energy increases by 3 orders of magnitude. The most striking event in the explosive growth stage is a sudden formation of layered structures in the density field. This is rather unexpected since the vertical density profile changes very little from basic state throughout the linear growth stage, and the linear theory predicts growth of vertically uniform columnar convection.

Figure 8 shows the time evolution of the horizontally averaged non-dimensional density $\overline{\rho}$ as a function of height. The vertical profile of $\overline{\rho}$ remains almost linear until $t \approx 900$. After $t \approx 900$, however, it starts to show small modulations. Those modulations rapidly develop into remarkable step-like structures. When the step-like

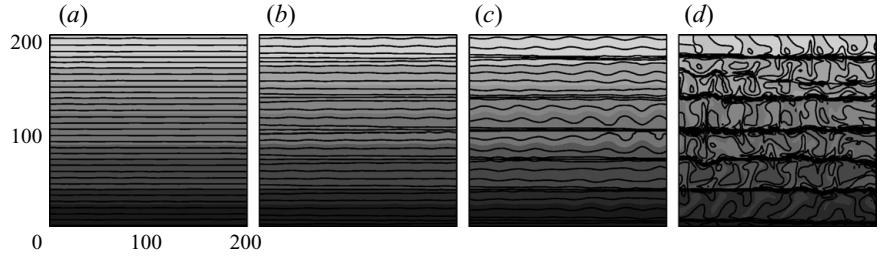

FIGURE 9. Snapshots of net density distribution at four different times before the discrete layers are formed: (*a*) $t = 800$, (*b*) $t = 950$, (*c*) $t = 980$ and (*d*) $t = 1010$. Only $(1/3) \times (1/3)$ of the entire calculation domain is displayed.

structures are first formed at $t \approx 1000$, the thickness of the steps is about 35 in non-dimensional length, which is close to the horizontal wavelength of the FGM. The fact that the thickness of the steps is considerably uniform in the vertical direction appears to imply an existence of some kind of mutual adjusting mechanism.

Figure 9 shows snapshots of density field at four different times. At $t = 800$, the deformation of the density stratification is very small and only random modulations of isopycnals (isolines of density) are visible, indicating that the fluid motions are of small amplitude. By $t = 950$, however, inhomogeneity in the density stratification becomes evident (figure 9*b*). Alternating layers of denser and sparser isopycnals emerge, although the amplitude of modulation of each isopycnal still remains small.

The inhomogeneity in the vertical stratification is slowly enhanced and at the same time wavy disturbances in each layer become visible by $t = 980$ (figure 9*c*). The horizontal wavelength of the disturbances is about 33, which is identical to that of the linear FGM.

The amplitude of the wavy disturbances grows further and eventually results in overturning (figure 9*d*). Strong mixing is initiated instantly and nearly homogeneous layers separated by strongly stratified interfaces are formed. During the processes from figure 9(*b*) to 9(*d*), the regions with dense and sparse isopycnals do not show any migrations in the vertical direction and their vertical intervals remain same.

5.4. *Non-exponential growth stage*

As time elapses further, these steps gradually increase their thickness through mergers with adjacent steps. The vertical profile seen after $t = 1000$ (figure 8) resembles very much that observed in the polar oceans (Neal *et al.* 1969). After the initial layers are established at $t \approx 1200$, the layers evolve in a self-similar manner. Each homogeneous layer is sandwiched by sharp interfaces and its thickness increases with time by merging with its adjacent layers, with their shape unchanged. Though it is not shown in figure 8, the layers continue to grow with time continuously after $t \approx 2500$ and eventually a single layer fills the whole depth of the calculation domain. It is likely that such layer mergings would continue if the calculation domain is expanded.

Figure 10 shows a snapshot of the density stratification together with the streamlines at $t = 1500$. There are eight layers of nearly homogeneous density sandwiched by thin interfaces with sharp and very stable density gradients. The thickness of the interfaces is only about 10 in non-dimensional length scale and the vertical density gradient at the interface is about 20 times of that of the basic stratification. The interface is

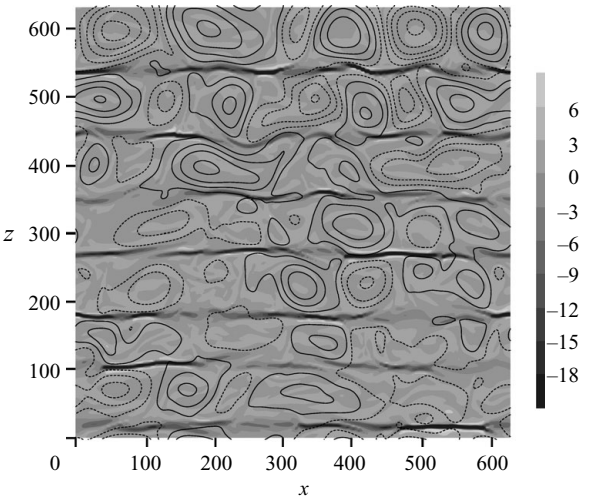

FIGURE 10. Vertical density gradient (shade) and streamlines (contours) at $t = 1500$.

neither flat nor uniform in the horizontal direction and the vertical density gradient at the interface shows large horizontal variations.

Each layer of nearly homogeneous density contains active convective cells, whose aspect ratio is approximately unity. These convective motions contribute to mix the density thus maintaining the homogeneous layer. The convective motions are driven by the density flux of the fastest diffusing component through the interface. Sharp interfaces tend to transport more T than S, because of the difference in the molecular diffusivity, and result in unstable stratifications right above and below the interfaces. The convective motions in each layer resemble Rayleigh–Bénard convection cells at a high Rayleigh number. Convective mixed layers sandwiched by diffusive interfaces are characteristic of the diffusive convection.

It is of great interest to understand how the layer structure is formed and how the layers merge to grow in size. In the following section, the mechanism of the layer formation is examined in detail. The mechanism for the growth of layers will be investigated in Part 2 (Noguchi & Niino 2010).

6. Theoretical analysis on the layer formation mechanism

6.1. *Spectral equations*

The results of the numerical simulation suggest that some kind of nonlinear interactions are necessary to generate the layers. The linear theory predicts that the FGM has zero vertical wavenumber. After $t \approx 900$, however, rapid formation of layers, which corresponds to zero horizontal wavenumber and finite vertical wavenumber, occurs. The purpose of this section is to clarify the mechanism for the layer formation through nonlinear interactions using a truncated spectral model. As for the choice of modes, our approach here is semi-empirical. First, we analyse the nonlinear interactions in the DNS to find a set of modes that play principal roles. Next, we construct a highly truncated spectral model and reproduce the observed layer formation in the DNS.

To analyse the results of the DNS, the governing equations (2.9)–(2.11) are Fourier-transformed with respect to the spatial variables x and z:

$$\left(\frac{d}{dt} + Pr\,K_0^2\right) K_0^2 \widetilde{\psi}(k_0, m_0, t) = \sum_{\boldsymbol{k}_1+\boldsymbol{k}_2=\boldsymbol{k}_0} (k_1 m_2 - k_2 m_1) K_2^2\, \widetilde{\psi}(k_1, m_1, t)\, \widetilde{\psi}(k_2, m_2, t) + \mathrm{i}k_0 Pr \left[\widetilde{T}(k_0, m_0, t) - \frac{\widetilde{S}(k_0, m_0, t)}{\gamma}\right], \tag{6.1}$$

$$\left(\frac{d}{dt} + K_0^2\right) \widetilde{T}(k_0, m_0, t) = \sum_{\boldsymbol{k}_1+\boldsymbol{k}_2=\boldsymbol{k}_0} (k_1 m_2 - k_2 m_1)\, \widetilde{\psi}(k_1, m_1, t)\, \widetilde{T}(k_2, m_2, t) - \mathrm{i}k_0 \widetilde{\psi}(k_0, m_0, t), \tag{6.2}$$

$$\left(\frac{d}{dt} + \tau K_0^2\right) \widetilde{S}(k_0, m_0, t) = \sum_{\boldsymbol{k}_1+\boldsymbol{k}_2=\boldsymbol{k}_0} (k_1 m_2 - k_2 m_1)\, \widetilde{\psi}(k_1, m_1, t)\, \widetilde{S}(k_2, m_2, t) - \mathrm{i}k_0 \widetilde{\psi}(k_0, m_0, t), \tag{6.3}$$

where tilde denotes an amplitude of a Fourier-transformed variable, $\boldsymbol{k}_i \equiv (k_i, m_i)$ and $K_i^2 \equiv k_i^2 + m_i^2$. The effects of the nonlinear interactions are expressed by the first terms on the right-hand side of (6.1)–(6.3). Note that, if these terms are neglected, the equations reduce to those used by the linear theory to derive the dispersion relation (3.1).

The nonlinear interaction terms consist of products of two different modes, (k_1, m_1) and (k_2, m_2). To generate the mode (k_0, m_0) through nonlinear interactions, the modes (k_1, m_1) and (k_2, m_2) must satisfy the relations $k_0 = k_1 + k_2$ and $m_0 = m_1 + m_2$ simultaneously. Since the nonlinear terms contribute to exchanging energy among modes, they will be henceforth called as energy transfer functions. The energy transfer function for mode (k_0, m_0) will be denoted as

$$N_{\widetilde{X}}(k_0, m_0; k_1, m_1) = \sum_{\boldsymbol{k}_1} [(k_0 - k_1)m_1 - k_1(m_0 - m_1)]\widetilde{\psi}(k_0 - k_1, m_0 - m_1)\widetilde{X}(k_1, m_1),$$

where $\widetilde{X}$ can be $\widetilde{\psi}$, $\widetilde{T}$ or $\widetilde{S}$. Note that $k_2 = k_0 - k_1$ and $m_2 = m_0 - m_1$ are automatically determined once (k_0, m_0) and (k_1, m_1) are given.

6.2. *Spectral analysis of the simulation results*

In order to identify the nonlinear interactions that lead to the layer formation, we will examine the energy transfer function just before the explosive growth stage.

Since there are infinite points on the k–m plane, we start from looking at the point (k_0, m_0) corresponding to the FGM in the linear theory, where $k_0 = 0.19$ and $m_0 = 0$. Then we sweep the interacting counterpart (k_1, m_1) over the k–m plane and plot the magnitude of energy transfer function $N_{\widetilde{T}}(k_0, m_0; k_1, m_1)$, averaged over $650 < t < 750$, based on the results of the numerical simulation. The resulting energy transfer function is shown in figure 11(a). In the following, the modes corresponding to (0.19,0), (0.19, 0.09) and (0, 0.09) will be called modes A, B and C, respectively. The first quadrant of figure 11(a) ($k_1 > 0$, $m_1 > 0$) shows that the FGM (mode A) has strong interactions with mode B and mode C. There is another pair of modes, which interacts with mode A, in the fourth quadrant of figure 11(a). This is simply showing the symmetric property of $N_{\widetilde{T}}(k_0, m_0; k_1, m_1)$ with respect to m_1 and the processes associated with these modes are essentially identical to those in the upper half-plane. From now on,

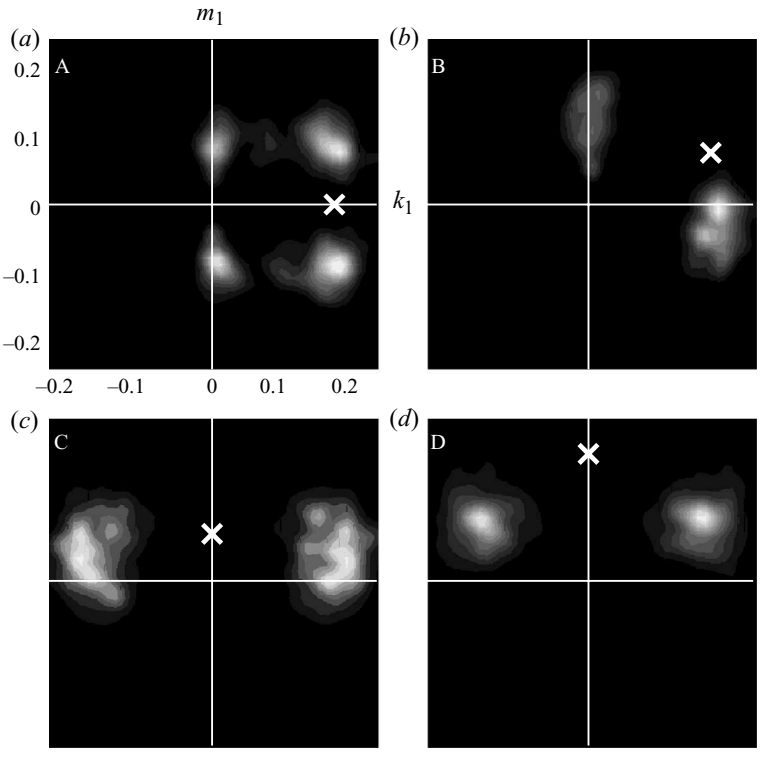

FIGURE 11. The magnitude of nonlinear interaction terms (energy transfer functions) $N_{\bar{T}}(k_0, m_0; k_1, m_1)$ for modes (*a*) A, (*b*) B, (*c*) C and (*d*) D. The white cross mark in each panel denotes the corresponding mode (k_0, m_0). The absolute value is plotted on k_1–m_1 plane. Values are averaged over $650 < t < 750$. Gray scales are in arbitrary unit for each panel.

no discrimination between mirror-image modes ($\pm k, \pm m$) will be made when we refer to the types of the modes, i.e. the mode with $(0.09, \pm 0.19)$ will be called mode B and that with $(0, \pm 0.19)$ will be called mode C.

It is curious to examine how modes B and C are produced. First, let us examine the energy transfer function for mode B (figure 11*b*). It is seen that mode B interacts strongly with mode A and weakly with mode C. A band of modes on m-axis also appears to interact with mode B. Note that the mode (0, 0.18), which is at the high wavenumber end of this band, corresponds to the vertical interval of the layered structure to emerge (see figure 8). This mode will henceforth be called mode D.

A similar examination of mode C shows that it strongly interacts with modes A and B (figure 11*c*). Thus it seems that modes A, B and C could constitute a closed triad. However, this closed triad alone does not explain the generation of mode D. Figure 11(*d*) shows that mode D interacts only with mode B. The vertical wavenumber of mode D is twice as large as that of mode B. It is important to note that the mode D cannot interact with its spatial subharmonic mode C, because the nonlinear transfer function for modes with $k = 0$ vanishes.

The above analysis suggests that the nonlinear interactions among the four modes A, B, C and D are essential for the formation of the layered structure. In the following subsection, it will be demonstrated that this is in fact the case.

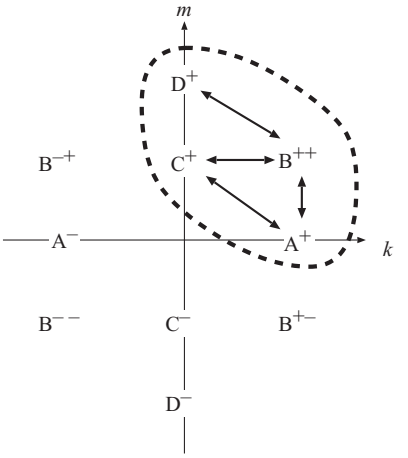

FIGURE 12. Configuration of the four-mode model. Encircled by broken line are the independent four modes and others are their mirror-image modes. Arrows denote the interactions between the four modes.

7. Four-mode model

In this section we will consider a simple truncated nonlinear spectral model in which only the four modes identified in the analysis in the previous section are considered. The relationship among the selected four modes are summarized in figure 12. The modes that can directly interact with each other are connected by arrows. Note that interactions only in the first quadrant are shown for simplicity, but there are similar interactions in each quadrant as guessed from a consideration of symmetricity.

The construction of the present model are summarized as follows:

(*a*) Only the four modes (and their mirror-image modes) are considered.

(*b*) The principal mode A is the FGM of the linear stability theory.

(*c*) Mode B has the same horizontal wavenumber as A.

(*d*) Vertical wavenumbers of modes B and C are the same, and are taken to be 0.09, based on the analysis of the DNS (§ 6.2).

(*e*) Modes A, B and C constitute an interaction triad.

(*f*) Mode D, which has a vertical wavenumber twice as large as that of mode C, interacts only with mode B.

According to the linear theory (figure 3), modes A and B have positive growth rates, while modes C and D have negative growth rates. The growth rate of mode B is about a half of that of mode A. Note that modes C and D are non-oscillatory.

Now we consider only the above four-mode set into (6.1)–(6.3) and integrate with respect to time. The initial amplitude of the modes are taken to be $|A| = 10^{-1}$, $|B| = 10^{-3}$ and $|C| = |D| = 0$, by considering the amplitude of modes in the DNS at $t = 50$ (figure 7*a*).

Figure 13 shows the time evolution of the magnitude of the streamfunctions in the truncated model. Modes A and B grow exponentially nearly at the growth rates predicted by the linear theory until $t \approx 1300$, whereas modes C and D, which are decaying modes according to the linear theory and do not exist at the initial time, are excited through nonlinear interactions with modes A and B and grow faster than both A and B. This rapid growth can be explained as follows. The nonlinear production of mode C from modes A and B is expected to be $O(|A|\,|B|)$. If $|A|$ and $|B|$ grow as $\exp(\sigma_A t)$

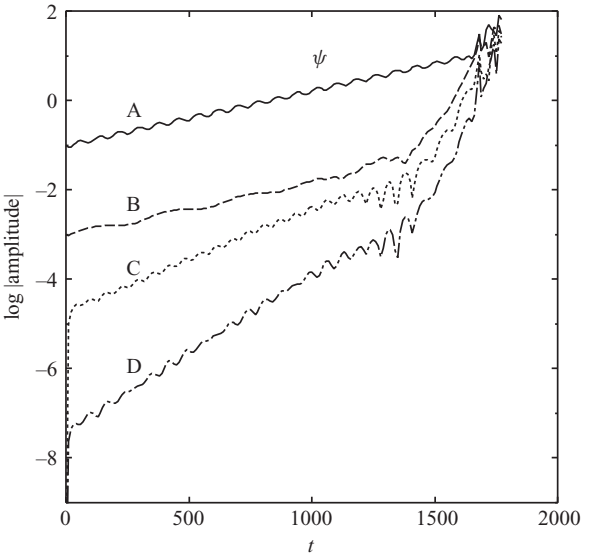

FIGURE 13. Time evolution of the amplitude of the streamline perturbation for modes A–D. The calculation is started from the initial state in which $|A| = 10^{-1}$, $|B| = 10^{-3}$ and $|C| = |D| = 0$.

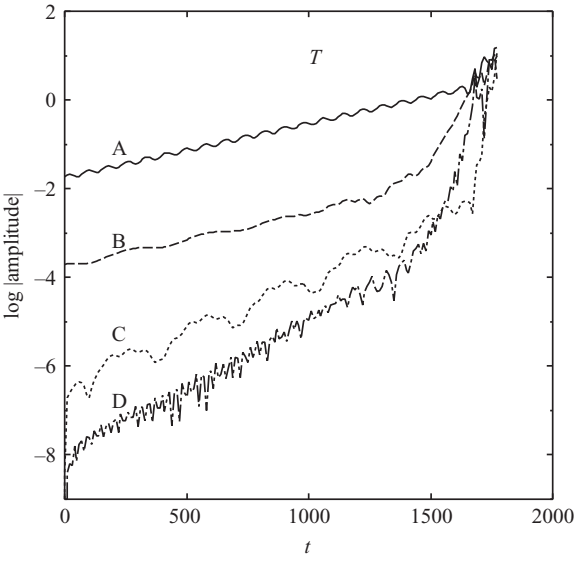

FIGURE 14. Time evolution of the amplitude of the temperature perturbation for modes A–D.

and $\exp(\sigma_B t)$, where σ_A and σ_B are the growth rates of modes A and B, respectively, the amplitude of C would grow as $\exp[(\sigma_A + \sigma_B)t]$, which is faster than both A and B.

After $t \approx 1300$, modes B, C and D start to grow even faster than before, and by $t \approx 1700$, all the four modes have nearly the same amplitude of $O(10^1)$. Finally at $t \approx 1750$, mode D, which has been of the smallest amplitude, overtakes the rest of the three modes. Figure 14 shows the time evolution of the amplitude of temperature perturbations. The growth of temperature amplitude of mode D is even more dramatic than that of the streamfunction.

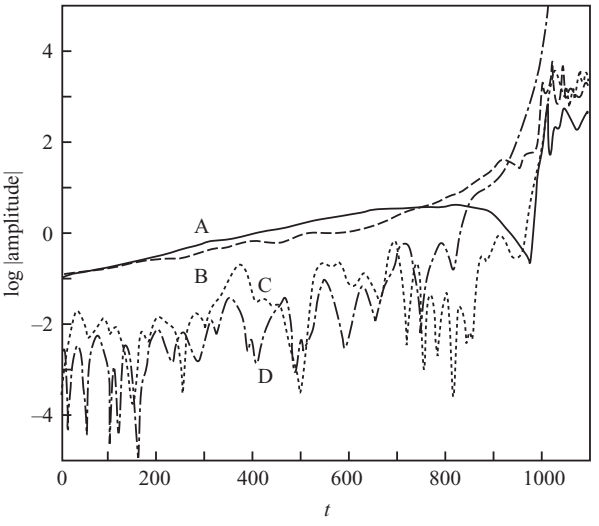

FIGURE 15. Time evolution of temperature perturbation of the four modes in the DNS.

Figure 15 shows the evolution of the four modes simulated by the DNS for comparison. The amplitude of the temperature perturbations for modes A and B grow nearly exponentially, with little fluctuations until $t \approx 650$. After this time, the growth of mode A slows down and mode B starts to show faster growth than before. By $t \approx 750$, mode B becomes dominant over mode A. Modes C and D have small amplitudes of $O(10^{-2})$ until $t \approx 300$, but then they start to grow at a rate somewhat larger than that of modes A and B. About $t \approx 750$, when the amplitudes of mode B overtakes the mode A, mode C starts to show a decrease until $t \approx 860$. On the other hand, mode D continues to grow and overtakes mode A at $t \approx 850$ and then mode B at $t \approx 930$. After $t \approx 930$, mode D becomes the dominant mode. It is also noteworthy that the amplitudes of modes A, B and C show remarkable increases just before $t = 1000$ (cf. figure 5).

The above comparison shows that the behaviour of the modes reproduced in the four-mode model (figures 13 and 14) is reasonably similar to that found in the DNS (figure 15). At the initial stage, modes A and B show exponential growths, drawing energy from the basic stratification as predicted from the linear stability theory. On the other hand, modes C and D, which are decaying modes according to the linear theory, are produced from modes A and B through nonlinear interactions until their amplitudes attain $O(1)$ and catch up with those of modes A and B. Although the time required for the layer structures to develop appears to be slightly different between the four-mode model and the DNS, this timing is considered to depend on the initial amplitudes of the four modes.

The behaviour of the fully nonlinear system, after mode D obtains an amplitude of $O(1)$ (i.e. after the layer is formed), is beyond the scope of the four-mode model. However, it is also noted that the results of the DNS (figure 15) show that, after obtaining the largest amplitude among all the modes ($t \gtrsim 1000$), mode D grows further while the other modes eventually saturate at almost a constant amplitude of $O(10^2)$.

8. Discussion

8.1. *Physical interpretation of the results of the four-mode model*

The nonlinear interactions illustrated by the four-mode model can be physically interpreted as follows. First, mode A as well as mode B grows due to the linear

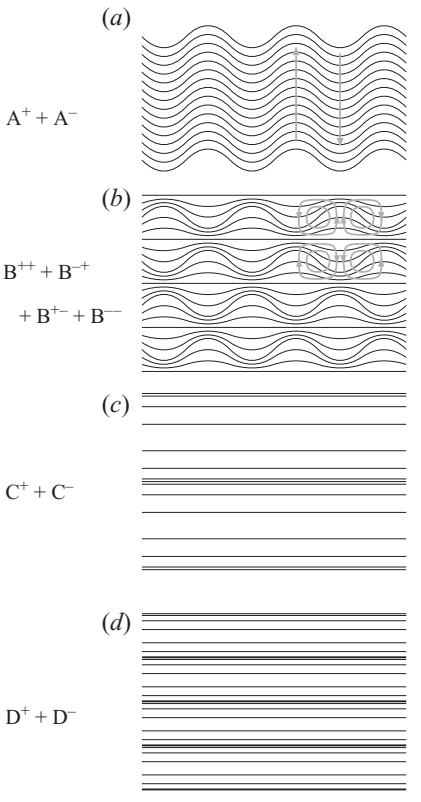

FIGURE 16. Sketches of the deformation of the basic density stratification due to the four modes. Arrows show motions accompanied with modes A and B. For superscripts + and −, see figure 12.

instability mechanism. As modes A and B develop, mode C is generated through the nonlinear interaction between modes A and B, while mode D is generated through that between mode B and its mirror-image mode. Figure 16 illustrates schematically the deformations of the basic stratification by modes A, B, C and D. The deformations due to modes C and D take the form of alternating layers of dense and sparse isopycnals except that the vertical length scale of layers for mode C is two times as large as that for mode D. Until modes C and D obtain $O(1)$ amplitude, modes A and B continue exponential growth to feed modes C and D through nonlinear production.

When the amplitude of mode D reaches $O(1)$, the vertical diffusion starts to enhance mode D itself in the following way (see figure 17). The deformations of the density stratification for modes C and D are horizontally homogeneous and thus can be considered as one-dimensional stratification. Let us assume that the initial deformations of the vertical profiles for T and S are similar and the density stratification is stable at any height. In a double-diffusive system, as time elapses, the vertical diffusion relaxes the deformation of T faster than that of S. This results in an enhancement of deformation of density stratification. Therefore, the vertical motion of mode B, whose antinode is at the level of weakest stratification in mode D, becomes enhanced and the amplitude of mode B increases. Conversely, this amplification of mode B also contributes to increase the amplitude of mode D, through the mode B–mode B interaction. It should be noted that the self-enhancing mechanism due to vertical differential diffusion is decoupled with fluid motions and is only effective

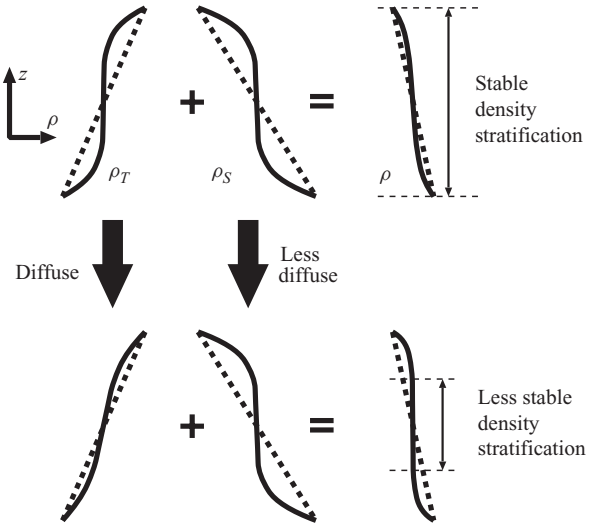

FIGURE 17. The self-enhancement mechanism of layer modes C and D due to vertical differential diffusion.

when the amplitude of the modulation of stratification is of the order of unity. Thus the mechanism is essentially a nonlinear process and beyond the scope of the linear stability analysis.

It is worth noting that the self-enhancing mechanism also applies to linearly stable ($\gamma < \gamma_{cr}$) stratification, if vertical modulation in stratification is given externally. This possibility will be explored in a separate paper.

8.2. *Vertical scale of the layer*

In the present four-mode model, we have used an empirical value 0.09 as the vertical wavenumber m_B of mode B on the basis of the analysis of the DNS. The wavenumber m_D of the resulting layer is given by $m_D = 2m_B$. It may be difficult to give an explanation for why this empirical value of m_B is preferred. In what follows, however, we will try to answer this question at least partially.

In order to examine how the value of m_B affects the layer formation, additional calculations have been made for the four-mode model with different values of m_B and $m_D = 2m_B$. In these calculations, the horizontal wavenumber of modes A and B, k_A and k_B, respectively, were taken as the same value as those in the previous section and m_B was varied. The results of the calculations are shown in figure 18, where the times required for the amplitude of mode B, C and D to overtake that of mode A is plotted as functions of $m_D = 2m_B$. It turns out that the four-wave model predicts growth of all modes for any value of m_D between 0.04 and 0.4. However, mode D as well as modes B and C seems to grow fastest when m_D is between 0.14 and 0.18, which is consistent with the range of m_D realized in the DNS.

In the DNS in which white noises are used as initial disturbances, the layer thicknesses are not strictly constant and show some variability for runs. For such a nonlinear system, it is very likely that the flow evolution would depend strongly on the initial condition, so that the exact value of the wavenumber of mode B may vary from experiment to experiment.

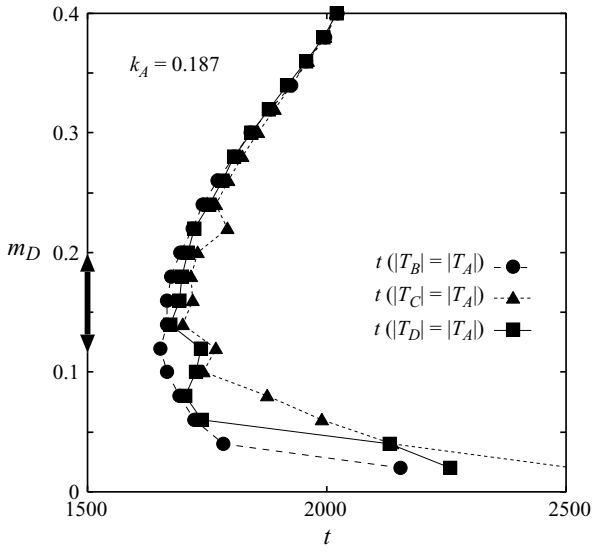

FIGURE 18. Time for mode A to be overtaken by modes B (circle), C (triangle) and D (square), as the function of m_D. The arrow on the vertical axis indicates the range of m_D that are frequently observed in the DNS initiated from random noise.

8.3. *Unimportance of resonant interactions*

Although a classical resonant mechanism such as a parametric subharmonic instability often plays an important role in nonlinear interactions (McComas & Bretherton 1977), the present four-mode mechanism does not satisfy the resonant interaction condition: the matching condition with respect to the frequencies is not required as seen in (6.1)–(6.3). In that sense, we can judge that the resonant interaction is unrelated to the present layer formation mechanism. However, near the m-axis, some modes that satisfy the resonant condition (figure 19) are indeed excited as is seen in figures 7(b) and 7(c). Thus, it may be possible that some kind of resonant interactions occur before the explosive growth stage. Such interactions, however, do not seem to be important during the layer formation as shown in figure 7(d).

There are several reasons that exclude resonant interactions in the present problem. Firstly, mode D (and also mode C) corresponding to the layer structure is not oscillatory according to the linear theory. Since the resonant interactions require that the matching condition is satisfied among three modes not only for wavenumber vectors ($\boldsymbol{k_1} + \boldsymbol{k_2} = \boldsymbol{k_0}$) but also for frequencies ($\omega_1 + \omega_2 = \omega_3$), non-oscillatory modes cannot interact with the FGM to draw energy (see figure 19). Secondly, mode D as well as mode C deforms the vertical stratification. This shifts the oscillation period of mode B longer, since the antinode of mode B is located where the vertical density gradient is weakened by mode D. This would disturb the temporal resonant condition between modes B and D. Thirdly, the subharmonic of the FGM is a decaying mode according to the linear stability analysis. A subharmonic resonance requires the subharmonic wave to be at least linearly neutral. Otherwise, the subharmonic wave would decay out before it becomes large enough to interact with the principal wave.

8.4. *Role of mode D*

It is of interest to understand why the nonlinear interactions do not close among the triad of modes A, B and C, but require an addition of mode D. If mode D were not

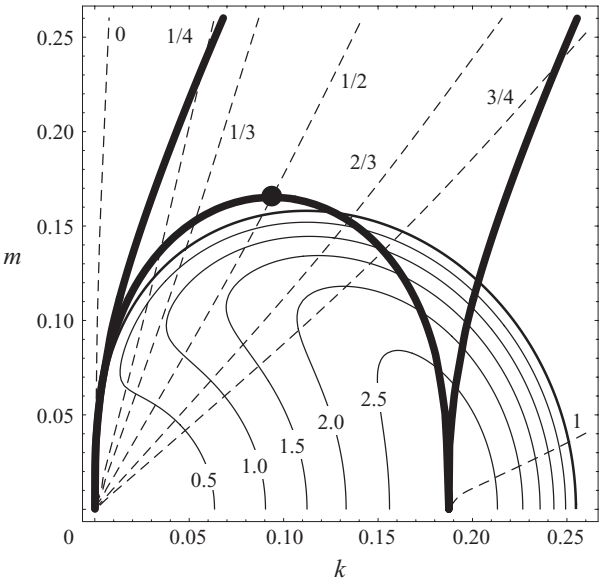

FIGURE 19. A possible criterion for a parametric instability. Any oscillatory mode on the thick line satisfies the parametric resonance condition with the FGM. Frequency (dashed line) is scaled by that of the FGM. The solid circle on the thick curve shows the mode which can draw energy from the FGM through parametric subharmonic resonance.

included, mode C would be the only mode that contributes to the layer formation. There are at least two reasons to believe that mode C would be less effective than mode D to grow. Firstly, a deformation of the density stratification due to mode C has a vertical scale twice as large as that of mode B, so that the position of the weakest stratification due to the deformation of mode C does not match with the antinode of mode B. Secondly, since the vertical scale of mode C is twice as large as that of mode D, the self-enhancing mechanism due to vertical diffusion (§ 8.1) would operate four times less effectively.

9. Conclusions

A spontaneous layer formation in unbounded diffusive stratification in a two-component fluid has been studied numerically and analytically. In the DNS, the spontaneous layer formation is found to occur when the density gradient ratio exceeds a critical value above which diffusive convection grows according to the linear stability analysis. When the stratification is supercritical, nearly steady oscillating motions, the horizontal scale of which is of the order of the horizontal wavelength of the fastest growing mode (FGM) in the linear stability analysis, are initiated. The disturbance energy at this stage grows nearly exponentially in accord with the prediction of the linear stability analysis.

After several tens of the oscillation cycle, a transition to the second stage, in which overturning convective motions develop and well-mixed regions form, occurs. The convective motions resemble Rayleigh–Bénard convection at a high Rayleigh number. The well-mixed regions are then gradually organized into horizontal layers, a typical thickness of which is of the order of the horizontal wavelength of the FGM in the linear stability analysis. The disturbance energy at this stage exhibits an

explosive growth. At the third stage after the layer structures are formed, the layers repeat merging with the adjacent layers and increase their thickness with time. The disturbance energy at this stage increases rather slowly.

A detailed analysis of the nonlinear interaction during the layer formation reveals that four modes are responsible for the layer formation. Mode A is the FGM with zero vertical wavenumber, and mode B is a weakly growing mode which has the same horizontal wavenumber k_0 as mode A and a finite vertical wavenumber m_0. Mode C has zero horizontal wavenumber and the same vertical wavenumber as mode B. The mode D also has zero horizontal wavenumber but has a vertical wavenumber twice as large as mode C. Both modes C and D are decaying modes according to the linear stability analysis. A severely truncated spectral model with the four modes is found to well describe the results of the DNS.

The physical mechanism for the layer formation has been proposed based on the results of the DNS and the truncated spectral model. The explosive growth of the disturbance energy is found to occur after the amplitudes of modes C and D attain $O(1)$. This is shown to be caused by the self-enhancing mechanism of the layer structure due to vertical differential diffusion. What determines the vertical scale of layers was also discussed.

The physical mechanisms investigated here may be important in many geophysical situations, such as polar oceans and magma chambers, where double-diffusive processes play important roles. As we have shown, formation of layers leads to vigorous Rayleigh–Bénard type convection in multilayers. The convective motion and associated transports of physical quantities are enforced as layer thickness increases through merging. The dynamics of the layer merging is an attractive and important topic in a variety of geophysical flows and will be treated in detail in Part 2.

The authors thank Professor Ryuji Kimura for many fruitful discussions and insights. T. Noguchi would also like to acknowledge the financial assistance of cooperative research program of Ocean Research Institute, The University of Tokyo.

REFERENCES

Ahlers, G. & Lücke, M. 1987 Some properties of an eight-mode Lorenz model for convection in binary fluids. *Phys. Rev.* A **35**, 470–473.

Baines, P. G. & Gill, A. E. 1969 On thermohaline convection with linear gradients. *J. Fluid Mech.* **37**, 289–306.

Fernando, H. J. S. 1987 The formation of a layered structure when a stable salinity gradient is heated from below. *J. Fluid Mech.* **182**, 525–541.

Hoare, R. A. 1966 Problems of heat transfer in Lake Vanda, a density stratified Antarctic lake. *Nature* **210**, 787–789.

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

Huppert, H. E. & Turner, J. S. 1981 A laboratory model of a replenished magma chamber. *Earth Planet. Sci. Lett.* **54**, 144–172.

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

McComas, C. H. & Bretherton, F. P. 1977 Resonant interaction of oceanic internal waves. *J. Geophys. Res.* **82**, 1397–1412.

Neal, V. T., Neshyba, S. & Denner, W. 1969 Thermal stratification in the Arctic ocean. *Science* **166**, 373–374.

Neshyba, S., Neal, V. T. & Denner, W. 1971 Temperature and conductivity measurements under upper Ice Island T-3. *J. Geophys. Res.* **76**, 8107–8120.

Noguchi, T. & Niino, H. 2010 Multi-layered diffusive convection. Part 2. Dynamics of layer evolution. *J. Fluid Mech.* doi:10.1017/S0022112010994160.

Prandtl, L. 1952 *Essentials of Fluid Dynamics*. Blackie.

Roache, P. J. 1972 *Computational Fluid Dynamics*. Hermosa Publishers.

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

Swallow, J. C. & Crease, J. 1965 Hot salty water at the bottom of the Red Sea. *Nature* **205**, 165–166.

Tait, R. I. & Howe, M. R. 1968 Some observation of thermohaline stratification in the deep ocean. *Deep-Sea Res.* **15**, 275–280.

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

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

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

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