JOURNAL OF COMPUTATIONAL PHYSICS 103, 320–335 (1992)

A Chebyshev Collocation Algorithm for 2D Non-Boussinesq Convection

P. LE QUÉRÉ, R. MASSON, AND P. PERROT

LIMSI-CNRS, BP 133, 91403 Orsay Cedex, France

Received October 30, 1990; revised November 14, 1991

A Chebyshev collocation algorithm is developed to integrate the time-dependent Navier–Stokes equations for natural convection flow with large temperature differences. The working fluid is assumed to be a perfect gas and its thermophysical properties vary with temperature according to Sutherland laws. The governing equations do not allow for acoustic waves. The generalized Helmholtz and Uzawa operators which arise from time discretization are solved iteratively and the performances of several types of preconditioners and iterative schemes are examined. The algorithm is validated by computing almost Boussinesq flows and by comparing with previous results obtained with a finite difference algorithm. We investigate the effects of the temperature difference and of total mass contained within the cavity on the transition to unsteadiness in a cavity of aspect ratio 8. It is shown that these parameters have, indeed, a significant effect on the value of Rayleigh number at which unsteadiness is triggered. We also discuss the nature of the time-periodic solution which is obtained for Ra slightly supercritical values.

1. INTRODUCTION

Natural convection in enclosures is an important problem of heat transfer literature. It is also a popular test problem to test numerical methods to integrate the Navier–Stokes equations for recirculating viscous incompressible fluid flows. In particular, various spectral methods have been applied to the classical thermally differentially heated cavity in the last decade [1–3]. In most cases, these algorithms have treated the Navier–Stokes equations under the Boussinesq approximation, that is, in the limit of infinitely small temperature differences. The limits of validity of the Boussinesq approximation have been investigated in [4, 5], for instance. For a fluid such as air, this assumption is known to be valid for temperature differences smaller than approximately 30K as long as the vertical length scale does not exceed a few meters. On the other hand, for water, the temperature range over which this assumption is valid is much smaller, a few degrees only, essentially due to the rapid variation of the physical properties such as viscosity or thermal diffusivity with temperature.

Many non-isothermal flows found in nature or engineering applications are thus clearly out of the range of validity of the Boussinesq approximation and this has long promoted the development of numerical algorithms to deal with buoyant flows generated by large temperature differences.

Several steps can be taken in that direction, the simplest being to let the viscosity vary with temperature as was done in [6], while retaining the first or higher order expansion of density around the reference temperature in the buoyancy forces. This procedure is particularly suited for the study of non-Boussinesq natural convection in liquids where compressibility effects are generally negligible compared with variations of thermophysical properties with temperature. On the other hand, for gases, compressibility effects may become important and allowance for variable density with temperature and pressure everywhere in the Navier–Stokes equations in addition to the variations of thermophysical properties must be retained. This results in an increased complexity of the governing equations by comparison with the Boussinesq case. The choice of an appropriate formulation and the subsequent numerical solution of the governing equations are then major difficulties.

For most common gases, and for air in particular, it is legitimate to assume that the equation of state is given by the perfect gas law. One straightforward approach is obviously to use the full compressible Navier–Stokes equations in primitive variable formulation in which pressure is eliminated and replaced by the perfect gas law, as was done in [7–9], amongst others. The major difficulty with the resulting system of equations is that it describes acoustic waves. While these acoustic waves carry only a minute fraction of the total energy, they, however, propagate at a finite value, the speed of sound, which is typically two to three orders of magnitude larger than the convective velocities carrying the quasi-totality of energy. Therefore CFL stability criteria are very restrictive in situations where the

0021-9991/92 $5.00

typical time scale for the flow to reach steady state is given by the time scale needed to damp out internal gravity waves as is the case for convective flows in differentially heated cavity [10].

We also note that a compressible vector potential–vorticity formulation was developed by Leonardi and Reizes [11, 12] for solving the steady state equations. This procedure has the additional complexity with respect to the Boussinesq formulation of having to solve an elliptic Poisson equation for pressure in order to obtain the density from the equation of state.

Whichever formulation is used, it is clearly very desirable to work with equations from which these acoustic waves have been eliminated. Such equations were established by Rehm and Baum in the invisid case [13] and Paolucci [16] derived them in the viscous case. The essence of these equations is to split the pressure between a mean value $\bar{P}$ which depends only on time and a hydrodynamic part Π which is responsible for the satisfaction of the continuity equation. These equations are incompressible in nature since the equation for Π is elliptic, but the velocity field is not divergence-free. Note that these equations were intuitively used by Forester and Emery a long time ago [14], and similar types of equations have also long been known to meteorologists (e.g., [15]).

These equations have since been used in many studies to investigate the effects of large temperature differences compared to the Boussinesq limit [17–19]. In particular, some of Chenoweth and Paolucci's results [18] indicate that there might be a rather large influence of this parameter on the transition to unsteady solutions of the 2D equations in vertical differentially heated cavity, for instance. The scheme developed by Chenoweth and Paolucci integrates the time-dependent equations in primitive variable form. The time stepping scheme is explicit and first-order accurate. Spatial discretization relies on second-order accurate finite difference approximation on non-uniform meshes. Even though finite difference discretizations can provide accurate solutions provided enough resolution is used, it seems that there is room for alternative approaches and, in particular, for algorithms based on spectral methods which are known to provide "infinite order accuracy" if the solution is smooth enough. In particular, spectral methods have been shown to perform extremely well for investigating loss of stability of natural convection flows in enclosures under the Boussinesq approximation (see the conclusions of the GAMM workshop [20], for instance). Fröhlich and Peyret have already developed one such spectral algorithm to integrate the non-Boussinesq equations [21, 22], but they have considered a 2D configuration with one direction of periodicity using products of Chebyshev and Fourier polynomials as expansion functions. The aim of this work is to develop an algorithm based on the use of Chebyshev collocation discretization to integrate the 2D Navier–Stokes equations applicable to air in a two-dimensional closed enclosure. This algorithm is then used to investigate the influence of the temperature difference and of total mass contained in the cavity on the transition to unsteadiness in a cavity of aspect ratio eight with adiabatic top and bottom walls.

The paper is organized as follows: in the next section we recall the governing equations in dimensionless form. Section 3 is devoted to the time discretization scheme. In Section 4, we compare several approaches to solve the non-constant coefficient Helmholtz equations which arise from the time discretization and in Section 5 we deal with the treatment of the continuity equation. In Section 6, we consider the determination of auxiliary quantities $\bar{P}$ and $d\bar{P}/dt$ and summarize the different steps of the time-stepping procedure. Section 7 is devoted to a validation of the scheme, mainly by comparison with some of the results by Chenoweth and Paolucci. Applications and results are presented in Section 8.

2. GOVERNING EQUATIONS

Consider the natural convection flow of a newtonian viscous fluid in a two-dimensional rectangular enclosure of width L and height H. The fluid is an ideal gas with constant specific heat coefficients C_p and C_v of ratio $\gamma = C_p/C_v$ equal to 1.4. Its thermal conductivity κ and molecular viscosity μ are allowed to depend on temperature. The coordinate system is defined so that the vertical axis (Oz) points vertically upwards and the x-axis is horizontal. Constant uniform temperatures T_1 and T_2 $(\Delta T = T_2 - T_1 \geqslant 0)$ are imposed at the right and left vertical walls, respectively. The top and bottom walls are taken as thermally insulated. The dimensionless temperature difference is denoted by $\varepsilon = \Delta T/2T_0$, where T_0 is the average temperature $(T_1 + T_2)/2$ which will be used as the reference temperature.

The physical governing equations are those considered by Chenoweth and Paolucci [18]. The equations are made dimensionless by introducing the following reference quantities: L_0 for length, V_0 for velocity, $t_0 = L_0/V_0$ for time, $\bar{P}_0$ for the mean pressure, and ρ_0 for density. The reference quantities for μ, κ, and kinematic viscosity ν, and thermal diffusivity α are taken at the reference temperature T_0. The pressure reference $\bar{P}_0$ and density reference ρ_0 verify the equation of state $\bar{P}_0 = C_v(\gamma - 1)\, T_0 \rho_0$ so that only one of these quantities needs to be specified. Several choices can be made as will be seen later. With this set of reference quantities, the governing equations in dimensionless form read

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \rho \mathbf{V} = 0 \tag{1}$$

$$\rho\left(\frac{\partial \mathbf{V}}{\partial t}+(\mathbf{V}\cdot\nabla)\mathbf{V}\right)=-\nabla\Pi+\frac{1}{\mathrm{Re}}\nabla\cdot\tau - \frac{gL_0}{V_0^2}(\rho-1)\mathbf{z} \tag{2}$$

$$\rho\left(\frac{\partial T}{\partial t}+(\mathbf{V}\cdot\nabla)T\right)=\frac{1}{\mathrm{Pr\,Re}}\nabla\cdot\kappa\nabla T + \left(\frac{\gamma-1}{\gamma}\right)\frac{d\bar{P}}{dt} \tag{3}$$

$$\bar{P}=\rho T, \tag{4}$$

where τ is given by $\tau=\mu(\nabla\mathbf{V}+(\nabla\mathbf{V})^t-\frac{2}{3}(\nabla\cdot\mathbf{V})I)$, $\mathbf{z}$ is the unit vector in the vertical direction, and g is the modulus of gravitational acceleration. Π represents the deviation from the average pressure and the hydrostatic component in the isothermal cavity to the first order in γM^2 ($\Pi=(P-\bar{P}+\rho_0 gz)/\rho_0 V_0^2+O(\gamma M^2)$). Re is a Reynolds number based on the reference scales ($=L_0V_0/\nu_0$) and Pr is the Prandtl number at the reference temperature $\mathrm{Pr}=\nu_0/\alpha_0$.

It was shown by Paolucci [16] that, in the limit $\varepsilon\ll 1$, these equations yield the classical Boussinesq equations. This follows from the fact, that for $\varepsilon\ll 1$, the temperature field for the conduction solution varies almost linearly between $T_1=1-\varepsilon$ and $T_2=1+\varepsilon$. The mean pressure field corresponding to the conservation of initial mass in an isothermal cavity at temperature T_0 and pressure $\bar{P}_0$ is given by $\bar{P}=1/\int_\Omega (1/T)\,dv$, where Ω is the computational domain (see Section 6). Assuming a linear temperature distribution gives $\bar{P}=2\varepsilon/\ln(1+\varepsilon)/(1-\varepsilon)$ which in the limit of small ε gives $\bar{P}=1-\varepsilon^2/3+\mathcal{O}(\varepsilon^4)$. If we define a dimensionless temperature Θ by $T=1-2\varepsilon\Theta$ (Θ is the usual dimensionless temperature of the Boussinesq equations), then $\rho=1-2\varepsilon\Theta+\mathcal{O}(\varepsilon^2)$. Since the isobaric thermal expansion coefficient $\beta=-(1/\rho)(\partial\rho/\partial T)_P$ for a perfect gas at temperature T_0 is equal to $1/T_0$, it follows that $2\varepsilon=\beta\,\Delta T$ and the first-order ε expansion of $\rho=1-2\varepsilon\Theta$ thus reduces to $\rho=1-\beta\,\Delta T\,\Theta$, which is the usual formulation of the buoyancy forces in the Boussinesq approximation. Inserting these expansions for ρ and T into the governing equations first yields $d\bar{P}/dt=0$ from the energy equation (or $\bar{P}=1$) and the Boussinesq equations are then obtained at leading order (zeroth order in the continuity and momentum equations and first order in the energy equation).

Natural convection flows are usually chararacterized by a Rayleigh number Ra, defined in the Boussinesq limit by ($g\beta\,\Delta T L_0^3/\nu_0\alpha_0$). This Rayleigh number can equivalently be written ($2\varepsilon g L_0^3/\nu_0\alpha_0$) which thus constitutes the proper definition of Ra for natural convection flows driven by large temperature differences. This definition of Ra reduces, in the Boussinesq limit ($\varepsilon\ll 1$), to the usual Boussinesq Rayleigh number. An appropriate velocity scale V_0 for highly convective flows in fluids of moderate Prandtl numbers is given by $(\nu_0/L_0)\,\mathrm{Ra}^{1/2}$ with $L_0=L$. With this new set of reference quantities, the governing equations take the final form:

$$\frac{\partial\rho}{\partial t}+\nabla\cdot\rho\mathbf{V}=0 \tag{5}$$

$$\rho\left(\frac{\partial \mathbf{V}}{\partial t}+(\mathbf{V}\cdot\nabla)\mathbf{V}\right)=-\nabla\Pi+\left(\frac{1}{\mathrm{Ra}}\right)^{1/2}\nabla\cdot\tau - \left(\frac{1}{\mathrm{Pr}}\right)\frac{\rho-1}{2\varepsilon}\mathbf{z} \tag{6}$$

$$\rho\left(\frac{\partial T}{\partial t}+(\mathbf{V}\cdot\nabla)T\right)=\frac{1}{\mathrm{Pr\,Ra}^{1/2}}\nabla\cdot\kappa\nabla T + \left(\frac{\gamma-1}{\gamma}\right)\frac{d\bar{P}}{dt} \tag{7}$$

$$\bar{P}=\rho T. \tag{8}$$

In dimensionless coordinates the computational domain extends from $-\frac{1}{2}$ to $\frac{1}{2}$ in the x direction and $-A/2$ to $A/2$ in the vertical direction where A is the vertical aspect ratio of the cavity ($=H/L$). The boundary conditions imposed on the walls on the cavity are the no-slip conditions for the velocity vector. For the dimensionless temperature T, they read

$$\begin{aligned} T(x=-\tfrac{1}{2},z)&=1+\varepsilon;\\ T(x=\tfrac{1}{2},z)&=1-\varepsilon;\\ \frac{\partial T}{\partial z}(x,z=\pm A/2)&=0. \end{aligned} \tag{9}$$

Equations (5)–(8) constitute a system of five equations for six independent quantities (in 2D), ρ, T, u, w, Π, and $\bar{P}$. The missing equation is given by an integral equation governing $d\bar{P}/dt$. Taking into account (7), the continuity equation (5) also reads

$$\nabla\cdot\mathbf{V}=\frac{\nabla\cdot\kappa\nabla T}{(\mathrm{Pr\,Ra}^{1/2})\bar{P}}-\frac{1}{\gamma\bar{P}}\frac{d\bar{P}}{dt}. \tag{10}$$

This equation can be integrated over the fluid domain to give, using the Gauss formula and taking into account boundary conditions on velocity and temperature,

$$\frac{d\bar{P}}{dt}=\frac{\gamma}{A\,\mathrm{Pr\,Ra}^{1/2}}\int_{-A/2}^{A/2}\left(\kappa\left(\frac{1}{2},z\right)\frac{\partial T}{\partial x}\left(\frac{1}{2},z\right)-\kappa\left(-\frac{1}{2},z\right)\frac{\partial T}{\partial x}\left(-\frac{1}{2},z\right)\right)dz. \tag{11}$$

The system of equations is closed by the Sutherland laws for fluid properties,

$$\kappa(T) = T^{3/2}\frac{1+S_k}{T+S_k};\ \mu(T) = T^{3/2}\frac{1+S_\mu}{T+S_\mu}, \tag{12}$$

where $S_k = 0.648$ and $S_\mu = 0.368$. These laws are appropriate for $\varepsilon \leqslant 0.6$ and a large range of T_0 ($210\text{K} \leqslant T_0 \leqslant 673\text{K}$). In all the sequel, the Prandtl number Pr is set to 0.71, corresponding to air.

3. SPATIAL APPROXIMATION AND TIME DISCRETIZATION

3.1. *Spatial Approximation*

The spatial discretization relies on the use of tensor products of Lagrangian interpolants at specified collocation points to approximate the dependent variables. Let (x_i, z_j), $0 \leqslant i \leqslant L$, $0 \leqslant j \leqslant M$, be one set of collocation points. Any 2D function $f(x, z)$ which we consider, for convenience of notation, defined on the domain $[-1, 1] \times [-A, A]$ is thus approximated by a polynomial of $P_L(x) \otimes P_M(z)$, where L and M are the maximum degrees of interpolating polynomials of variables x and z, respectively. One then has

$$f(x, z) = \sum_{i=0}^{L}\sum_{j=0}^{M} f_{ij} L_i(x)\, L_j\left(\frac{z}{A}\right), \tag{13}$$

where f_{ij} are the nodal values at (x_i, z_j) and $L_i(x_k) = \delta_{ik}$, where δ_{ik} is the Kronecker delta function. One classically has $L_i(x) = P_{L+1}(x)/(x - x_i)\, P'_{L+1}(x_i)$, where $P_{L+1}(x)$ is the $(L+1)$th-degree polynomial with roots $(x_i, 0 \leqslant i \leqslant L)$.

For the temperature and velocity components, x_i will be taken as the Chebyshev Gauss Lobatto points which are the roots of $(x^2 - 1) \times T'_L(x)$, where $T_L(x)$ is the Lth-degree Chebyshev polynomial, yielding $x_i = \cos(i\pi/L)$, $0 \leqslant i \leqslant L$, and similarly for z_j. The choice of the polynomial space for the pressure Π will be discussed in Section 4.

Computing the partial derivatives of $f(x, z)$ with respect to x or z thus amounts to differentiating (13) and reevaluating the expansion at the same or at another set of collocation points. This can be done efficiently on vector machines by using fast matrix multiples. Explicit entries for the matrices that correspond to the first- and second-order partial derivations are available in [25].

When the spectral coefficients f_{lm} are needed (i.e., the coefficients such that $f(x, z) = \sum_{l=0}^{L}\sum_{m=0}^{M} f_{lm} T_l(x)\, T_m(z/A)$), they are obtained by using the orthogonality relationships for the discrete cosine series:

$$f_{lm} = \frac{4}{c_l \bar{c}_m LM}\sum_{i=0}^{L}\sum_{j=0}^{M}\frac{1}{\bar{c}_i \bar{c}_j} f(x_i, z_j)\, T_l(x_i)\, T_m\left(\frac{z_j}{A}\right), \tag{14}$$

where $\bar{c}_0 = \bar{c}_L = 2$, $\bar{c}_l = 1$, $0 < l < L$.

3.2. *Time Discretization*

The time discretization scheme is of finite difference type. The use of a Chebyshev collocation method for the spatial resolution makes it necessary to resort to an implicit or semi-implicit discretization of diffusion terms for reasons of stability while the non-linear terms are generally treated explicitly. We have chosen to use the scheme proposed by Vanel *et al.* [3] which combines a second-order backward Euler scheme for the viscous terms with an Adam–Bashforth-type evaluation of convective terms and variable coefficients such as ρ or κ. When applied to a scalar diffusion–advection equation such as

$$\rho\left(\frac{\partial f}{\partial t} + \mathbf{V}\cdot\nabla f\right) = \nabla\cdot\kappa\,\nabla f,$$

this second-order scheme reads

$$\rho^* \frac{3f^{n+1} - 4f^n + f^{n-1}}{2\Delta t} + (\rho\mathbf{V}\cdot\nabla f)^* = \nabla\cdot\kappa^*\,\nabla f^{n+1}, \tag{15}$$

where the starred notation stands for Adams–Bashforth extrapolation $g^* = 2g^n - g^{n-1}$, where n refers to the time $t_n = n\,\Delta t$.

Equation (15) can be cast as a Helmholtz equation for the unknown field f^{n+1},

$$(\nabla\cdot\kappa^*\nabla - \lambda I) f^{n+1} = Sf, \tag{16}$$

where $\lambda(x, z) = 3\rho^*(x, z)/2\Delta t$. The source term Sf is made of known quantities at previous time levels $Sf^{n,n-1} = (\rho\mathbf{V}\cdot\nabla f)^* + \rho^*((-4f^n + f^{n-1})/2\Delta t)$.

When applied to the governing equations (7), (6), (10), this discretization scheme yields three Helmholtz equations for the unknown temperature and velocity fields at time level $n+1$. The velocity components have, furthermore, to satisfy the modified continuity equation (10), which yields a generalized Stokes problem:

$$(\nabla\cdot\kappa^*\nabla - \lambda_T I)\, T^{n+1} = S_T \tag{17}$$

$$(\nabla\cdot\mu^{n+1}\nabla - \lambda_v I)\, u^{n+1} = \mathrm{Ra}^{1/2}\frac{\partial \Pi^{n+1}}{\partial x} + Su \tag{18}$$

$$(\nabla\cdot\mu^{n+1}\nabla - \lambda_v I)\, w^{n+1} = \mathrm{Ra}^{1/2}\frac{\partial \Pi^{n+1}}{\partial z} + Sw \tag{19}$$

$$\left(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z}\right)^{n+1} = S_D, \tag{20}$$

where $\lambda_T = 3\rho^* \text{ Pr Ra}^{1/2}/2\Delta t$ and $\lambda_v = 3\rho^{n+1} \text{ Ra}^{1/2}/2\Delta t$. The source terms are respectively given by

$$S_T = \text{Pr Ra}^{1/2}\left(-\frac{\gamma-1}{\gamma}\left(\frac{d\bar{P}}{dt}\right)^* + (\rho \mathbf{V}\cdot\nabla T)^* + \frac{\rho^*}{2\Delta t}(T^{n-1}-4T^n)\right) \tag{21}$$

$$\mathbf{S_V} = \text{Ra}^{1/2}\left(\rho^{n+1}((\mathbf{V}\cdot\nabla)\mathbf{V})^* + \frac{\rho^{n+1}}{2\Delta t}\times(\mathbf{V}^{n-1}-4\mathbf{V}^n) + \frac{\rho^{n+1}-1}{2\varepsilon \text{ Pr}}\mathbf{z}\right) - \mu^{n+1}\nabla S_D + \frac{2}{3}\nabla(\mu^{n+1}S_D) - (\nabla\mu^{n+1})\cdot(\nabla\mathbf{V})^* \tag{22}$$

$$S_D = \frac{\nabla\cdot\kappa^{n+1}\nabla T^{n+1}}{(\text{Pr Ra}^{1/2})\,\bar{P}^{n+1}} - \frac{1}{\gamma\bar{P}^{n+1}}\left(\frac{d\bar{P}}{dt}\right)^{n+1}. \tag{23}$$

Within each time step the temperature equation is solved first with extrapolated values κ^* and ρ^*. This allows us to compute the new temperature field and, in turn, after determination of $\bar{P}$ and $d\bar{P}/dt$ (see Section 6), the new density and dynamic viscosity fields which are then used in the momentum equations.

We have also examined an alternative discretization scheme for the momentum equations which is obtained by expanding the term $(\nabla\cdot\mu^{n+1}\nabla\mathbf{V}^{n+1})$ which yields

$$(\nabla^2 - \lambda_v I)\,\mathbf{V}^{n+1} = \frac{\text{Ra}^{1/2}}{\mu^{n+1}}\nabla\Pi^{n+1} + \frac{1}{\mu^{n+1}}\times(\mathbf{SV} - (\nabla\mathbf{V})^*\cdot(\nabla\mu^{n+1})), \tag{24}$$

where $\lambda_v = 3\rho^{n+1}\text{ Ra}^{1/2}/2\mu^{n+1}\Delta t$.

Given the dependent variables up to time level n, the determination of the temperature and velocity field at new time level $n+1$ requires the solution of three non-constant coefficient Helmholtz problems and a Stokes problem. The solution of these generalized Helmholtz equation and Stokes problem are the subject of Sections 4 and 5, respectively.

4. SOLUTION OF A HELMHOLTZ PROBLEM

At each time step the temperature field and the velocity components are thus obtained from the solution of a generalized Helmholtz problem which reads (dropping the time index)

$$L(f) = Sf \tag{25}$$

in $\Omega = [-1, 1]\times[-A, A]$ with Dirichlet or Neumann-type boundary conditions on the boundary $\partial\Omega$ of Ω. Depending upon the formulation which is chosen, the operator L is given by

$$L(f) = \nabla\cdot a\nabla f - \lambda f \tag{26}$$

or

$$L(f) = \nabla^2 f - \frac{\lambda}{a}f, \tag{27}$$

where $a(x, z)$ and $\lambda(x, z)$ are scalar fields which depend on space and which satisfy $a(x, z) \geqslant a_0 > 0$ and $\lambda(x, z) \geqslant 0$ on Ω. In the context of time integration of the governing equations, $a(x, z)$ typically represents the molecular viscosity or the thermal conductivity and λ varies like $\frac{3}{2}\rho\sigma$ with $\sigma = \text{Ra}^{1/2}/\Delta t$. Both fields also depend on time through time evolution of the temperature field with the result that the operators L to be solved change at each time step, precluding the use of direct solvers.

Let L_{sp} be the pseudospectral discretization at Chebyshev Gauss Lobatto nodes (x_i, z_j), $0 \leqslant i \leqslant L$, $0 \leqslant j \leqslant M$, of either one of the above second-order elliptic partial operator L. We consider both Dirichlet boundary conditions for the velocity or mixed Dirichlet–Neumann boundary conditions for the temperature. For instance, for (26), L_{sp} reads

$$L_{sp}(f) = \nabla\cdot I_{LM}(a\nabla f) - I_{LM}(\lambda f), \tag{28}$$

where I_{LM} stands for the polynomial interpolation at the collocation nodes (x_i, z_j), $0 \leqslant i \leqslant L$, $0 \leqslant j \leqslant M$. After elimination of the boundary conditions, (28) yields a linear system for the unknown variables at interior nodes (x_i, z_j), $1 \leqslant i \leqslant L-1$, $1 \leqslant j \leqslant M-1$. This system is characterized by a very large bandwidth of order $(LM^2 - LM)$, which prevents direct solution, except for the constant coefficient Helmholtz problem for which an efficient direct solution is available, along the lines proposed in [23]. The inversion of (28) must therefore be carried out iteratively. When solving iteratively a linear system characterized by a matrix $\mathcal{M}$, it is well known that the efficiency of any iterative scheme is linked to the condition number $K(\mathcal{M})$ which is generally defined as the ratio of the extreme moduli of eigenvalues of $\mathcal{M}$. For Chebyshev spatial approximation, $K(L_{sp})$ varies like $\mathcal{O}(L^4)$ which makes it necessary to resort to preconditioning techniques for an efficient iterative inversion of L_{sp}. A good preconditioner H must therefore be such that it be easily invertible and that the condition number $K(H^{-1}L_{sp})$ be as close to one as possible.

In this work, we have considered and tested two possible choices for H. The first one is the pseudo-spectral approximation of the constant coefficient operator derived

from L_{sp} by taking integral averages of both scalar fields a and λ. This preconditioner is noted H_{Δ}. It is given by

$$H_{\Delta} = \Delta - \frac{\bar{\lambda}}{\bar{a}} I \tag{29}$$

or

$$H_{\Delta} = \Delta - \left(\frac{\bar{\lambda}}{a}\right) I, \tag{30}$$

depending on the choice (26) or (27). For a scalar field g, the average $\bar{g}$ is defined by: $\bar{g} = (1/4A) \int_{\Omega} g(x, z)\, dx\, dz$. This version differs slightly from a similar type of spectral preconditioning independently proposed in [24]. H_{Δ} is efficiently inverted by the bi-diagonalization algorithm [25]. The operation count is $4LM(L+M)$. This should be compared to the number of operations required to evaluate the residual which is $4LM(L+M)$ or $2LM(L+M)$ for each type of approximation (26) or (27), respectively.

The second preconditioner we have considered is the classical five-point finite-differences approximation of L_{sp} factorized in incomplete LU-5 form with row sum agreement [25], noted H_{RS5}. With this choice, the operation count per iteration merely reduces to the evaluation of the residual. Depending on the choice of approximation (26) or (27), the operation count per iteration with H_{Δ} preconditioning is between two or three times larger than that for H_{RS5}. The relative performance of either choice of preconditioning thus depends on the condition numbers $K(H_{\Delta}^{-1} L_{sp})$ and $K(H_{RS5}^{-1} L_{sp})$. To solve the Helmholtz equations for the velocity components one thus has four possible choices, depending on the formulation (26) or (27) and on preconditioners H_{Δ} or H_{RS5}. We have compared the condition numbers of these four preconditioned operators on a test problem closely related to the application envisioned, in that the scalar fields a and λ correspond to the definition:

$$T(x, z) = 1 + 2\varepsilon\theta(x, z)$$

$$0(x, z) = -0.5x(1 - 0.1 \times b(x, z)^2)$$

$$\text{with} \quad b(x, z) = \sin(\pi x) \sin(\pi z)$$

$$\bar{P} = 2\varepsilon / \ln((1+\varepsilon)/(1-\varepsilon))$$

$$\rho(x, z) = \bar{P}/T$$

$$\lambda = \tfrac{3}{2}\rho\sigma$$

$$a(x, z) = \mu(T) \qquad \text{given by Sutherland law.}$$

Tables I–III given the condition numbers of $H_{\Delta}^{-1} L_i$ and $H_{RS5}^{-1} L_i$, $i = 1, 2$, where L_1 and L_2 stand for the Helmholtz operators corresponding to (26) and (27), respectively, for three different values of ε.

These tables clearly show that formulations (26) and (27) are equivalent when preconditioned by H_{RS5}. For small values of σ and of ε, H_{Δ} preconditioning is better than H_{RS5} preconditioning even when taking into account the different operation count per iteration. This is, furthermore, confirmed by the fact that H_{Δ} preconditioning is independent of spatial resolution, whereas the condition number of $K(H_{RS5}^{-1} L_i)$ increases with increasing spatial resolution. On the other hand, for large values of σ, which is the case in practice due to the stability constraint resulting from explicit discretization of advection terms, the condition number of $H_{RS5}^{-1} L_i$ becomes comparable or even better than the condition number of $H_{\Delta}^{-1} L_i$, in particular for large values of ε. This shows that H_{RS5} preconditioning should be preferred when computing strongly non-Boussinesq flows.

These tests were made for Dirichlet boundary conditions on $\partial\Omega$, which is the case of the velocity components. For the temperature field, mixed boundary conditions of Dirichlet type on two sides and Neumann type on two other sides were considered. In this case, H_{Δ}-type-preconditioning with

TABLE I

Condition Numbers Corresponding to Dirichlet Boundary Conditions

σ	$H_{\Delta}^{-1} L_1$	$H_{RS5}^{-1} L_1$	$H_{\Delta}^{-1} L_2$	$H_{RS5}^{-1} L_2$
0	1.167	6.262	1.0	6.263
10^2	1.166	3.172	1.239	3.171
10^4	1.195	1.650	1.380	1.650

Note. $\varepsilon = 0.1$; $L = M = 32$.

TABLE II

Condition Numbers Corresponding to Dirichlet Boundary Conditions

σ	$H_{\Delta}^{-1} L_1$	$H_{RS5}^{-1} L_1$	$H_{\Delta}^{-1} L_2$	$H_{RS5}^{-1} L_2$
0	1.367	6.254	1.0	6.263
10^2	1.365	3.265	1.543	3.264
10^4	1.432	1.677	1.919	1.677

Note. $\varepsilon = 0.2$; $L = M = 32$.

TABLE III

Condition Numbers Corresponding to Dirichlet Boundary Conditions

σ	$H_{\Delta}^{-1} L_1$	$H_{RS5}^{-1} L_1$	$H_{\Delta}^{-1} L_2$	$H_{RS5}^{-1} L_2$
0	3.111	6.304	1.0	6.263
10^2	3.078	3.675	4.633	3.671
10^4	3.383	1.777	9.816	1.776

Note. $\varepsilon = 0.6$; $L = M = 32$.

L_1 formulation (formulation of Laplacian as in (26)) was only tested and was found almost insensitive to the type of boundary conditions investigated.

So far, we have discussed only the preconditioning operators which are essential to the performance of iterative schemes. The other important aspect is the nature of the iterative scheme. Several choices are possible, from the simple Richardson iteration with fixed or variable parameter to more sophisticated second-order schemes such as conjugate residual or conjugate gradient schemes. Figure 1 shows the evolution of residuals and errors obtained with Richardson minimum residual (PMR), conjugate residual (PCR), and conjugate gradient (PCG), when used in conjunction with H_{RS5} preconditioning and L_2 formulation (formulation of Laplacian as in (27)). The test function is $f(x, z) = \sin(2\pi x + \pi/4)\sin(2\pi z + \pi/4)$. The scalar fields $a(x, z)$ and $\lambda(x, z)$ are given by Eqs. (26)–(32). In view of the figure, it appears the two second-order schemes perform much better than the first-order scheme PMR at the same cost per iteration. Since the PCR scheme achieves convergence even for non-symmetric definite matrices, it was therefore chosen in most computations.

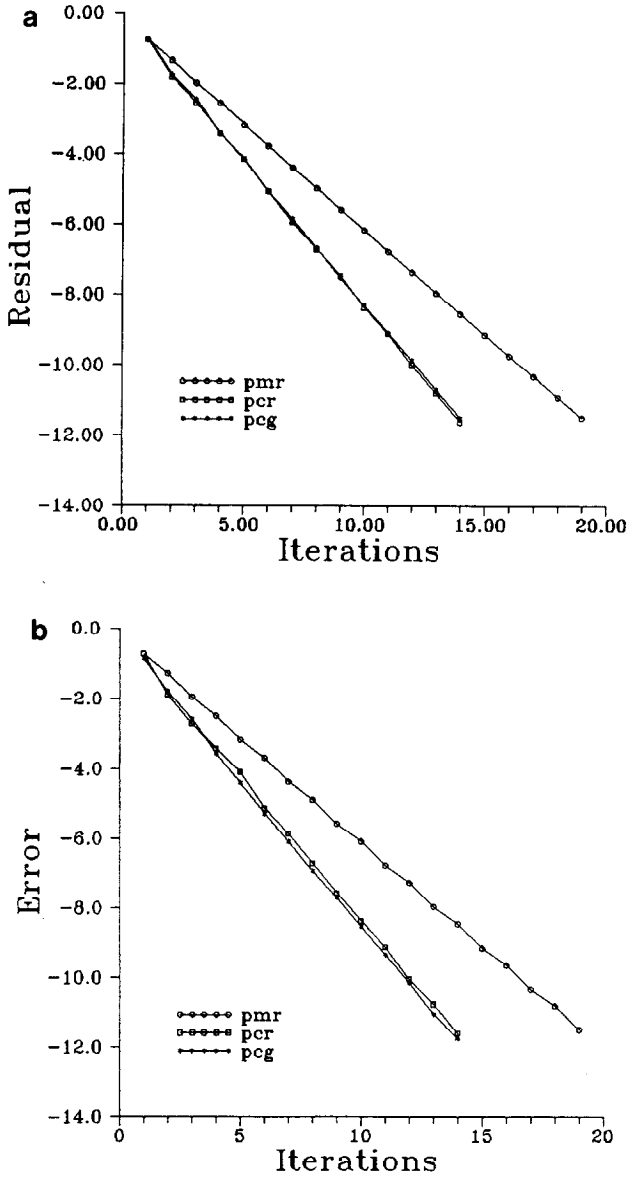


FIG. 1. Evolution of log(*residual*) (a) and of log(*error*) (b) with iteration number; $L = M = 32$; variable coefficient test problem given by Eqs. (26)–(32); test function is $f(x, z) = \sin(2\pi x + \pi/4)\sin(2\pi z + \pi/4)$.

5. SOLUTION OF STOKES PROBLEM

Equations (18)–(20) constitute a Stokes problem coupling velocity and pressure fields at time level $(n+1)$. Eliminating the velocity components between these equations gives an equation for the pressure field which reads

$$(\nabla \cdot L_1^{-1} \nabla) \Pi = S_\Pi \tag{31}$$

or

$$\left(\nabla \cdot L_2^{-1} \frac{1}{\mu} \nabla\right) \Pi = S_\Pi, \tag{32}$$

depending on the formulation used for the Helmholtz equation for the velocity components. The left-hand sides of Eq. (31) or (32) define generalized Uzawa operators which are noted $\mathcal{U}_1$ and $\mathcal{U}_2$.

The discrete system of equations depends on the choice of the collocation grids for the velocity components u and w and for the pressure field Π. The choice of these grids corresponds to a choice of polynomial space for each corresponding variable, noted $\mathcal{V}_U$, $\mathcal{V}_W$, $\mathcal{V}_\Pi$, respectively. The gradient operator for pressure, noted $\mathcal{G}$, is thus defined as a linear operator from $\mathcal{V}_\Pi$ into $\mathcal{V}_U \times \mathcal{V}_W$, whereas the divergence operator for the velocity, noted $\mathcal{D}$, takes $\mathcal{V}_U \times \mathcal{V}_W$ into $\mathcal{V}_\Pi$. The Uzawa operators $\mathcal{U}_1$ and $\mathcal{U}_2$ thus read

$$\mathcal{U}_1 = \mathcal{D} L_1^{-1} \mathcal{G} \tag{33}$$

or

$$\mathcal{U}_2 = \mathcal{D} L_2^{-1} \mathcal{G}, \tag{34}$$

respectively. It is well known that if $\mathcal{V}_U = \mathcal{V}_W = \mathcal{V}_\Pi = P_L(x) \otimes P_M(z)$, where L and M are the maximum degree of interpolating polynomials in directions x and z, respectively, the Uzawa operator suffers from spurious pressure modes, which were first identified by Morchoisne [26]. The remedy to this problem is to choose a polynomial space $\mathcal{V}_\Pi$ smaller than those for the velocity components,

which amounts to the use of a staggered grid algorithm. The first staggered grid algorithm proposed in [27] still suffers from one spurious mode for the pressure and a three-grid algorithm was then proposed by Bernardi and Maday [28] which is free of spurious modes except for the constant pressure mode. In this algorithm, the two velocity components are not defined at the same collocation points, which thus requires the inversion of two different Helmholtz operators. This was done in [29] for the Boussinesq equations, but the extension of the algorithm to the non-Boussinesq case requires a lot of programming.

In order to simplify the programming in the non-Boussinesq context we follow the method proposed in [30], in which the two velocity components belong to $P_L(x) \otimes P_M(z)$ and are defined at the corresponding Gauss–Lobatto points (x_i, z_j), $0 \leqslant i \leqslant L$, $0 \leqslant j \leqslant M$. Choosing $\mathscr{V}_H = P_{L-2}(x) \otimes P_{M-2}(z)$ allows one to eliminate all spurious pressure modes. There remains the choice of the collocation points to define the pressure field and to collocate the continuity equation. The choice that was made in [30] is the $(L-1) \times (M-1)$ Gauss points which are the roots of $T_{L-1}(x)\, T_{M-1}(z/A)$. This choice results in four matrix multiples to evaluate the gradient or divergence operators. This number can be reduced to two by choosing to define the pressure field and to collocate the continuity equation at the internal Gauss–Lobatto points (x_i, z_j), $1 \leqslant i \leqslant L-1$, $1 \leqslant j \leqslant M-1$.

For obvious reasons, the inversion of (31) or (32) must be carried out iteratively and the same questions as those raised in the previous paragraph arise, i.e., choice of an iterative scheme and choice of a good preconditioner. It must be kept in mind that in practical computations, the coefficient λ of the Helmholtz equations which varies like $1/\Delta t$ is very large since, as we have already said, the time-step Δt is limited by stability criteria due to the explicit discretization of the convective terms. The rather severe stability criteria met in actual computations result in values of λ, such that the spectrum of the Uzawa operator has the same magnitude as the spectrum of the second-order differentiation operator and the iterative inversion of the Uzawa operator thus necessitates a very efficient preconditioner.

We have chosen to use the Uzawa operator corresponding to the constant coefficient case as preconditioner. For either choice of Helmholtz approximation, these preconditioners read $U_1 = \mathscr{D} H_\Delta^{-1} \mathscr{G}$ and $U_2 = \mathscr{D} H_\Delta^{-1}(1/\mu(x, z)) \mathscr{G}$. These operators are explicitly built by letting the pressure field span the basis associated with the internal Gauss points (x_i, z_j), $1 \leqslant i \leqslant L-1$, $1 \leqslant j \leqslant M-1$, thus generating a matrix of order $(L-1) \times (M-1)$. The rank of this matrix is $(L-1) \times (M-1) - 1$ and its inversion is carried out by direct Gauss elimination after regularization, which amounts to fixing the pressure increment in one point. The inverse matrix is used to calculate the pressure increment during the iterations. The cost of the preconditioning is thus $2((L-1) \times (M-1))^2$ per iteration. Each iteration also requires the evaluation of the residual which is done in $4LM(L+M)(1 + i \times nit)$ operations with $i = 1$ or 2, depending on the choice L_i and nit is the number of iterations for the Helmholtz problems. In practical computations, where $nit \simeq 7$, the preconditioning cost is thus one to two times larger than the cost of computing the residual, depending on each formulation, which makes $U_1^{-1}\mathscr{U}_1$ approximately one and one-half times more expensive per iteration than $U_2^{-1}\mathscr{U}_2$.

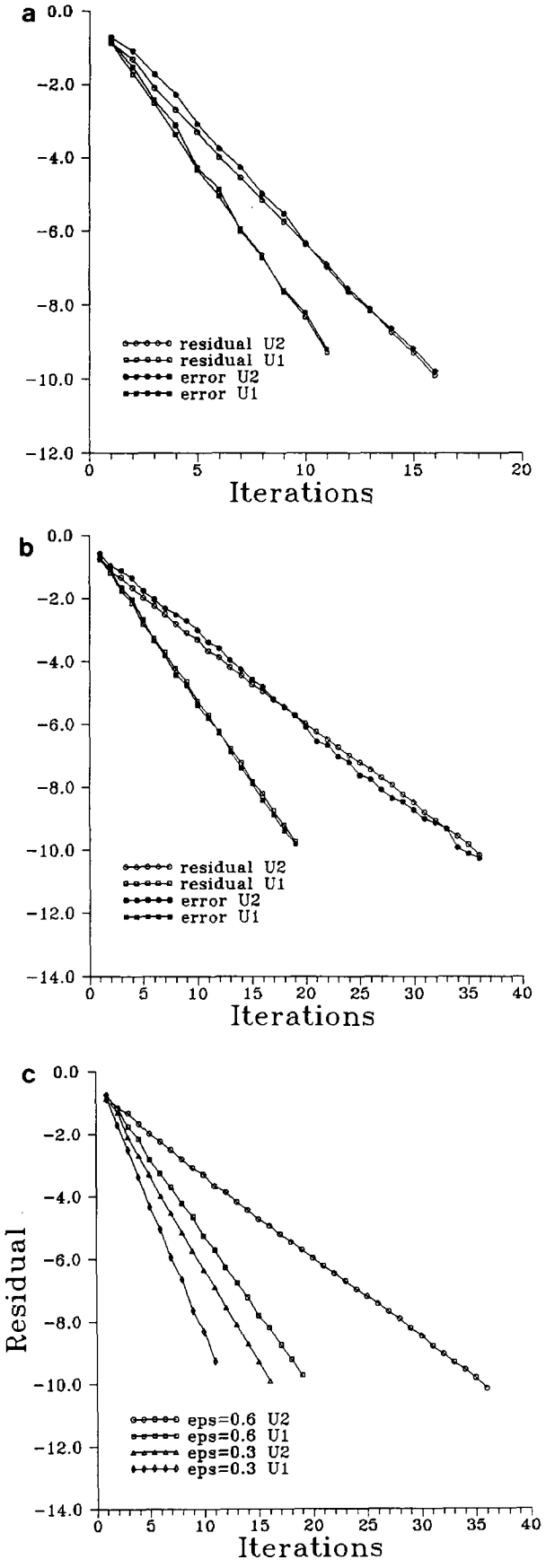
The efficiency of this preconditioning is shown in Table IV, which presents the condition number of $U_1^{-1}\mathscr{U}_1$ and $U_2^{-1}\mathscr{U}_2$ for the test problem corresponding to Eqs. (26)–(32). This table shows that for small values of σ, the condition number of $U_2^{-1}\mathscr{U}_2$ is much better than that of $U_1^{-1}\mathscr{U}_1$, the inverse being true for large values of σ. This is so because the condition number of $U_2^{-1}\mathscr{U}_2$ strongly depends on the spatial inhomogeneity of $\lambda(x, z)/a(x, z)$ which varies more than $\lambda(x, z)$ or $a(x, z)$ alone, due to the variations of physical properties of air with temperature. These conclusions might therefore not be true for other types of variable coefficient problems. On the other hand, the condition number of $U_1^{-1}\mathscr{U}_1$ depends very little on σ and this formulation of the Uzawa operator would likely apply to a large class of problems.

As for the Helmholtz equations, three types of iteratives schemes were tried and a PCR iterative scheme was finally retained for the same reasons as described in the previous paragraph. Figure 2 shows the evolution of the residuals and errors of the continuity equation for two values of ε and each type of approximation. It is seen that, for the test problem considered (which corresponds to a large value of σ), the convergence rate of formulation $U_1^{-1}\mathscr{U}_1$ is always smaller than that of formulation $U_2^{-1}\mathscr{U}_2$, the difference increasing with increasing ε. The conclusion is that for small values of ε, formulation $U_2^{-1}\mathscr{U}_2$ should be preferred despite its larger condition number since the cost per iteration is about one and one-half times smaller, whereas for larger values of ε, both formulations are almost equivalent when the cost per iteration is taken into account. These tests were

TABLE IV

Condition Numbers; $L = M = 32$

	$\varepsilon = 0.3$		$\varepsilon = 0.6$	
σ	$U_1^{-1}\mathscr{U}_1$	$U_2^{-1}\mathscr{U}_2$	$U_1^{-1}\mathscr{U}_1$	$U_2^{-1}\mathscr{U}_2$
10^{-2}	1.616	1.001	3.093	1.002
10	1.615	1.385	3.090	2.165
10^2	1.672	1.933	3.067	4.568
10^4	1.728	2.711	3.375	9.782
10^6	1.829	2.941	3.852	11.88



performed for $L = M = 32$. For large values of ε and larger spatial resolutions we have retained the $U_1^{-1}\mathscr{U}_1$ formulation because of the fact that its condition number is almost independent of the time-step and spatial resolution.

In the computations of time-periodic solutions which are reported below, the pressure at a previous time level was used to provide the initial guess at the current time level. Thus, for values of ε up to 0.3, at most one to two iterations of $U_1^{-1}\mathscr{U}_1$ or two to three iterations of $U_2^{-1}\mathscr{U}_2$ were required per time step to bring the error of the continuity equation down to 10^{-7} in absolute value. For the larger values of ε that were considered ($\simeq 0.6$) and formulation $U_1^{-1}\mathscr{U}_1$ the number of outer iterations for the Uzawa operator was at most three, the number of inner iterations (iterations for the Helmholtz equations for the velocity components) being less than seven. Note also that, for transient solutions, viscosity depends on time through dependence of temperature on time. U_2-type of preconditioning was then also periodically reconstructed, typically every 1000 to 2000 timesteps. The frequency of the update of the preconditioner only influences the rate of convergence of the iterative inversion of the Uzawa operator and not the solution which is produced, provided of course that the residual is decreased enough.

6. TREATMENT OF $d\bar{P}/dt$ AND $\bar{P}$

As we have already said, within each time step, temperature equation is solved first, giving the new temperature field at time level $n+1$. The thermodynamic properties μ^{n+1} and κ^{n+1} are then obtained from Sutherland laws. The computation of ρ^{n+1} and of the source term S_D^{n+1} of the continuity equation require the evaluation of $d\bar{P}/dt$ and $\bar{P}$, which are discussed below.

6.1. *Computation of* $d\bar{P}/dt$

Evaluation of $d\bar{P}/dt$ is done through the use of Eq. (11), which simply states that, in a cavity with impervious walls, the variation in time of $\bar{P}$ is only governed by the energy flux at the boundaries. The integrals in (11) can be evaluated exactly after obtaining the spectral coefficients of $\partial T/\partial x$ and computation of $d\bar{P}/dt$ should not in principle pose any difficulty. In particular, when computing a steady state solution, $d\bar{P}/dt$ should tend to zero as time tends to infinity. This follows from the fact that, at steady state, energy conservation should result in an exact cancellation of the energy fluxes at the two vertical boundaries. In the non-Boussinesq

FIG. 2. Evolution of log(*residual*) and log(*error*) of continuity equation for $\varepsilon = 0.3$ (a) and $\varepsilon = 0.6$ (b) for U_1 and U_2-types of preconditioning; in (c) the evolutions of log(*residual*) for $\varepsilon = 0.6$ and 0.3 and for U_1- and U_2-types of preconditioning are compared.

case exact cancellation is not warranted in contrast to the Boussinesq case, where it is trivially obtained by symmetry. Due to truncation errors, $d\bar{P}/dt$ is thus not really zero in actual computations even though steady state is attained. In fact, $d\bar{P}/dt$ is a measure of the truncation error of the temperature field and should tend to zero as spatial discretization is increased. The absolute value of $d\bar{P}/dt$ is, however, not very meaningful by itself and it is clear from (11) that a relative measure of the error is obtained by scaling the difference of heat fluxes at the boundaries with the mean value of Nusselt number. This quantity is reported in Table V for one test case and, as expected, tends to zero as spatial resolution increases.

6.2. *Computation of* $\bar{P}$

Since $d\bar{P}/dt$ is in principle known at each time step, this could be used to compute $\bar{P}$ at time $n+1$, by integration of this equation starting from the initial value of $\bar{P}$ at time $n=0$. For several reasons it seems preferable not to do so. One reason is linked to the accuracy of this time integration, but the major reason is that, as just stated above, $d\bar{P}/dt$ is obtained from a discretized form of the temperature equation which involves truncation errors. In particular, since $d\bar{P}/dt$ does not tend to zero for steady state unless sufficient resolution is used, $\bar{P}$ would slowly drift in time and possibly be prevented from reaching any steady state even though one exists. Thus, whenever this is possible, it seems preferable to obtain $\bar{P}$ from other constraints imposed on the solution directly. One such constraint for our problem, since the mass flux at the boundary is zero, is the conservation of the mass initially contained within the cavity. The equation of state can be written $\rho = \bar{P}/T$ and integrated over the fluid domain to give $\bar{P} = \int_\Omega \rho / \int_\Omega 1/T$ or

$$\bar{P} = M_0 \Big/ \frac{1}{\Omega} \int_\Omega \frac{dv}{T(x,z)}. \tag{35}$$

where M_0 is the initial mass of gas within the cavity. Here again the integral in (35) can be exactly evaluated after computation of the spectral coefficients of $1/T(x,z)$.

Several choices are possible for M_0, depending on the way the heating was applied. Suppose that the cavity is initially at T_0, $\bar{P}_0$ and that the walls are heated and cooled symmetrically, i.e., the temperature of the hot wall is raised to $T=1+\varepsilon$ and that of the cold wall is decreased to $T=1-\varepsilon$. In that case, with our choice of reference quantities, $M_0=1$. Another possibility that is often used in the experiments is that, starting at room temperature T_1 and pressure $\bar{P}_0$, only one wall is heated up to T_2, the other being kept at T_1. Then it is immediate to show that $M_0 = 1/(1-\varepsilon)$. Symmetrically, if the cavity is initially at $(T_2, \bar{P}_0)$ and the cold wall temperature is decreased to T_1, then $M_0 = 1/(1+\varepsilon)$.

A fourth possibility is to drop the requirement of mass conservation and to consider instead $\bar{P}=1$, in which case $\rho = 1/T$. This corresponds to a leaky cavity in which the mean pressure inside the cavity remains equal to the mean pressure outside the cavity. This possibility is likely to be particularly true for natural convection experiments with gases, where the time scales are often very long, giving much time for the inner and outer pressure to equalize if one does not make particular effort to built a leakproof cavity. Numerical experiments show that, in the first case considered above (when $M_0=1$), $\bar{P}$ is always smaller than 1. So the "leaky" cavity, in fact, corresponds to a cavity with a total mass larger than 1, which means that the cavity has filled up, compared to the reference state $M_0=1$.

We will show on an example in Section 8.2 that the critical Ra value corresponding to transition to unsteadiness in the cavity of aspect ratio 8 that was studied is in fact very sensitive to this value of M_0 even for small ε.

TABLE V

$d\bar{P}/dt$ Scaled by mean Nusselt Number at Steady State for ($\mathrm{Ra}=10^3$, $\varepsilon=0.6$)

$L=M=$	6	8	12	16	20	24	32
$d\bar{P}/dt$	9×10^{-2}	8×10^{-3}	3×10^{-3}	4×10^{-4}	5×10^{-5}	8×10^{-6}	2×10^{-7}

6.3. *Overall Time-Stepping Scheme*

Given the solution up to time level $n\,\Delta t$, the solution at time level $(n+1)\,\Delta t$ is obtained in the following way:

- Determination of temperature field T^{n+1} after computation of source term S_T (Eq. (21)) and inversion of (17).
- Determination of thermophysical properties μ^{n+1} and κ^{n+1} from Sutherland laws.
- Computation of $(d\bar{P}/dt)^{n+1}$ from (11) and of $\bar{P}$ from (35).
- Determination of ρ^{n+1} from equation of state (8).
- Computation of the divergence of velocity field S_D^{n+1} (Eq. (23))
- Resolution of generalized Stokes problem.
  - Guess initial pressure field Π_e. In general $\Pi_e = \Pi^n$.
  - Computation of corresponding divergence $D_e = \mathscr{D} L_{sp}^{-1}(\mathscr{G}\Pi_e + \mathbf{S_V})$.
  - Compute the correcting pressure by iterative inversion of the Uzawa operator: $\Pi_c = \mathscr{U}^{-1}(S_D - D_e)$.
  - Obtain new pressure field $\Pi^{n+1} = \Pi_e + \Pi_c$.
  - Obtain new velocity field $\mathbf{V}^{n+1} = L_{sp}^{-1}(\mathscr{G}\Pi^{n+1} + \mathbf{S_V})$.
- Proceed to the next time step.

7. VALIDATION

The validation of the code has been carried out in several ways. First, we have tested that it was immaterial to choose the internal Gauss–Lobatto points to define the pressure field Π, instead of the more classical Gauss points. This was done on a Boussinesq code by comparing the results with those previously obtained with a three-grid algorithm. The test case was a cavity of vertical aspect ratio 8 with adiabatic top and bottom walls and with both codes it was possible to bound the critical value of the Rayleigh number corresponding to transition to unsteadiness between 3×10^5 and 3.2×10^5. This confirms the critical value previously obtained in [31], in which the assumption of Hopf bifurcation was used to produce an accurate value of Ra_c equal to 3.1×10^5.

Concerning the non-Boussinesq code, it was first tested in the limit of small ε to compare the results with those produced by a Boussinesq code ($\varepsilon = 0$). In this case, the preconditioner used for the Helmholtz equations is H_Δ (see Section 3) which allows us to obtain the convergence in one iteration only. Such a comparison is shown in Fig. 3, where two solutions corresponding to the same Rayleigh number but different values of ε ($\varepsilon = 0$ and $\varepsilon = 0.001$) are displayed in the form of temperature, stream function, and pressure iso-contours. (In the non-Boussinesq case, since at steady state $\partial\rho/\partial t = 0$, it is possible to introduce a stream-function ψ such that $\rho \mathbf{V} = \nabla \times \psi \mathbf{k}$, where $\mathbf{k}$ is the unit vector normal to the xOz plane. ψ is then obtained from the solution of $\nabla^2\psi = \partial\rho w/\partial x - \partial\rho u/\partial z$). As can be seen, these two solutions are in extremely good quantitative agreement.

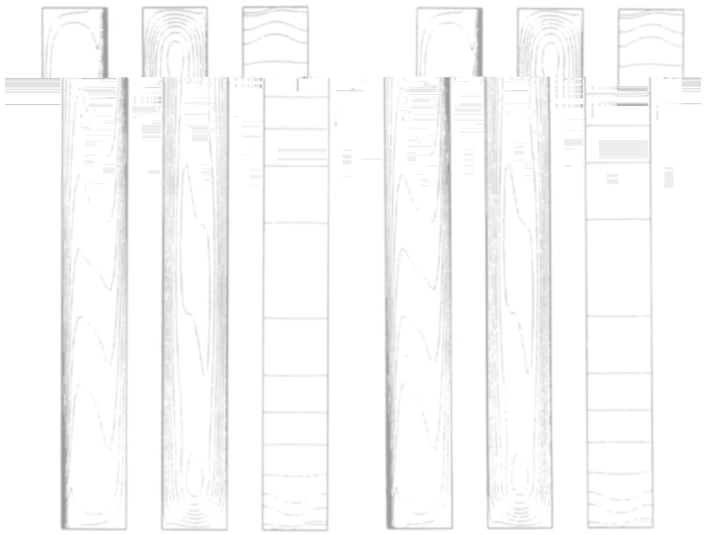


FIG. 3. Isocontours of temperature (left), stream-function (center), and pressure (right) for the Boussinesq case (a) and for $\varepsilon = 0.001$ (b); $\mathrm{Ra} = 3 \times 10^5$, $A = 8$; temperature isovalues are: -0.5 (0.1) 0.5; stream function isovalues are: 0.0002, 0.001, 0.002, 0.003, 0.004, 0.005, 0.0057; pressure isovalues are: -1400, -1200, -1000, -800, -600, -400, -200, 0, 40.

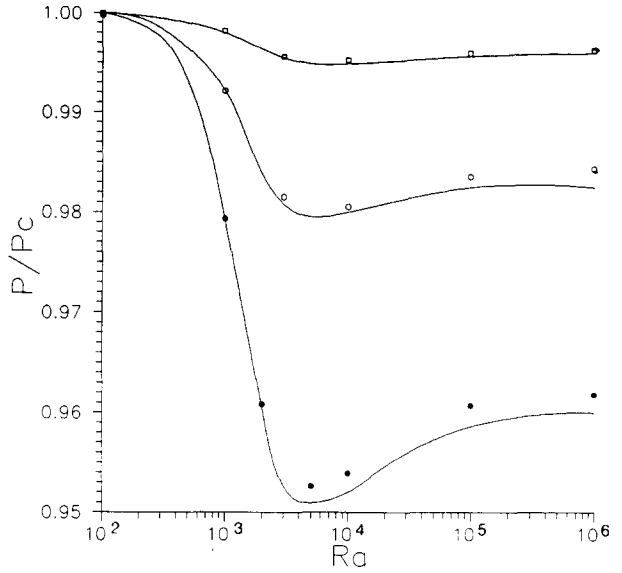


FIG. 4. Evolution of $\bar{P}/\bar{P}_c$ with Ra for different values of ε; (□), $\varepsilon = 0.2$; (○), $\varepsilon = 0.4$; (●), $\varepsilon = 0.6$. Full lines are from [18].

For larger values of ε, the tests were performed against the results obtained by Chenoweth and Paolucci [18] concerning the evolution of $\bar{P}/P_c$, where $\bar{P}_c$ is the mean pressure of the variable property conduction solution found for $u = w = 0$. This comparison is carried out in Fig. 4 and it shows that both sets of results are in very good agreement. Likewise, comparison of global Nusselt numbers at steady state (not shown) also shows very good agreement.

8. TRANSITION TO UNSTEADINESS IN A TALL CAVITY

It was thus decided to investigate the influence of variable properties on the critical value of Rayleigh number corresponding to the onset of unsteadiness in a vertical differentially heated cavity of aspect ratio 8. The methodology that was classically used was to integrate long enough the time-dependent equations with given values of Ra and ε to be able to determine the nature (steady or unsteady) of the asymptotic solution. Let us recall here that, in the Boussinesq case and for an aspect ratio of 8, stable steady state 2D solutions are found up to a critical Ra value close to 3.1×10^5. For smaller values of Ra the flow develops smoothly from conduction solution into the boundary layer regime. In particular, in the Boussinesq case, the instability of the conduction regime can only be found in cavities of aspect ratio larger than a critical value between 11 and 12.

Some of Chenoweth and Paolucci's results suggest that there might be pronounced effects of the temperature dif-

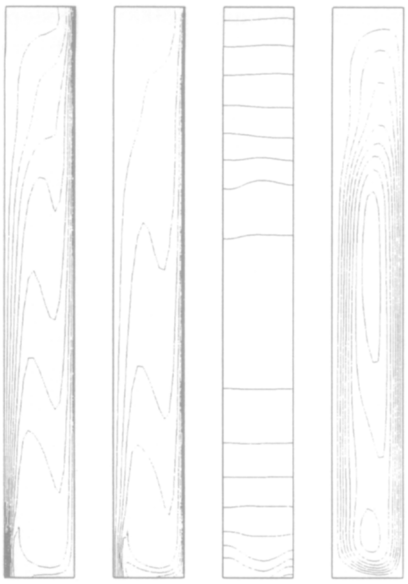


FIG. 5. Solution for $Ra = 10^5$; $A = 8$; $\varepsilon = 0.6$. From left to right: Temperature 0.4 (0.1) 1.6, density 0.7 (0.2) 2.3, pressure Π and streamfunction 0.001 (0.001) 0.008.

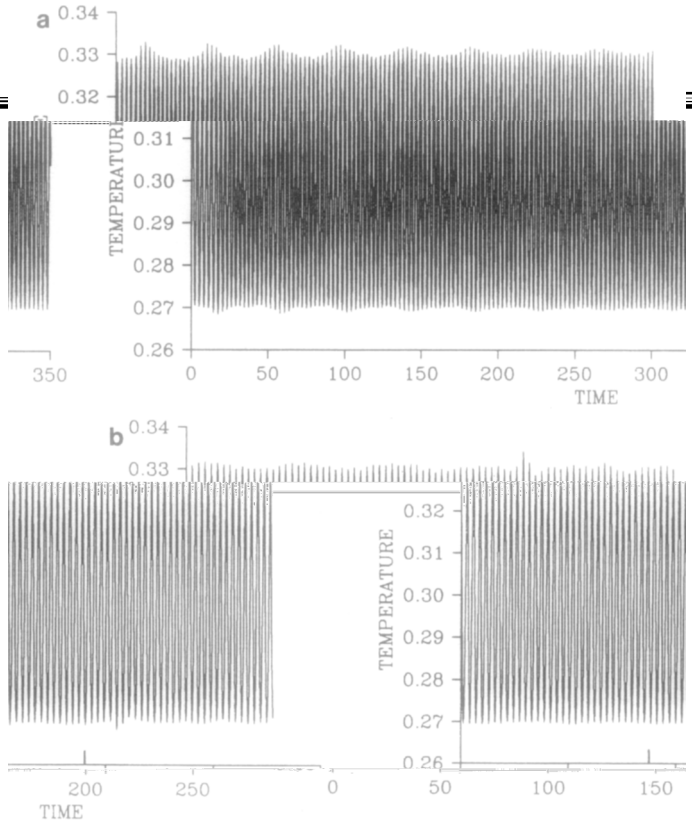


FIG. 6. (a) Time trace of temperature at location $(x, z/A) = (-0.71, 0.71)$ for $Ra = 3 \times 10^5$; $A = 8$; $\varepsilon = 0.3$; $(L, M) = (24, 72)$. Time integration performed with $\Delta t = 1.75 \times 10^{-2}$. (b) Concatenation of time traces of temperature at monitoring location for three different spatial resolutions: up to $t \simeq 88$, $(L, M) = (24, 72)$; from 88 to $\simeq 140$, $(L, M) = (32, 96)$; after 140, $(L, M) = (20, 64)$. Times at which spatial resolution was changed are indicated by thick marks on the inner side of the time axis.

ference on the flow regimes. In particular, the temperature fields corresponding to values of $Ra = 1 \times 10^5$, $\varepsilon = 0.6$ [18, Figs. 10b and c] show very irregular isotherms that the authors attribute to the cross-roll multicellular instability. They also claim that the corresponding solutions are steady. The solution for an aspect ratio of 8 was computed and its spatial structure is displayed in Fig. 5 in the form of isotemperature and streamfunction contours. The accuracy of that solution was checked with two different spatial resolutions ($L = 20$, $M = 64$) and ($L = 24$, $M = 72$). These plots do not obviously show the wiggles that are found in the corresponding solution in Chenoweth and Paolucci [18]. A possible explanation would be the existence of multiple solutions depending on initial condition.

8.1. *Influence of* ε

For the same configuration (cavity of $A = 8$ and adiabatic top and bottom walls) we then have determined the limits between steady and unsteady solutions for different values of ε. As an example, Fig. 6a shows that the asymptotic solution corresponding to a Ra value of 3×10^5 and an ε-value of 0.3 is time periodic. This was obtained with spatial resolution $(L, M) = (24, 72)$. To show that spatial resolution is adequate the following test was performed. The fields of the

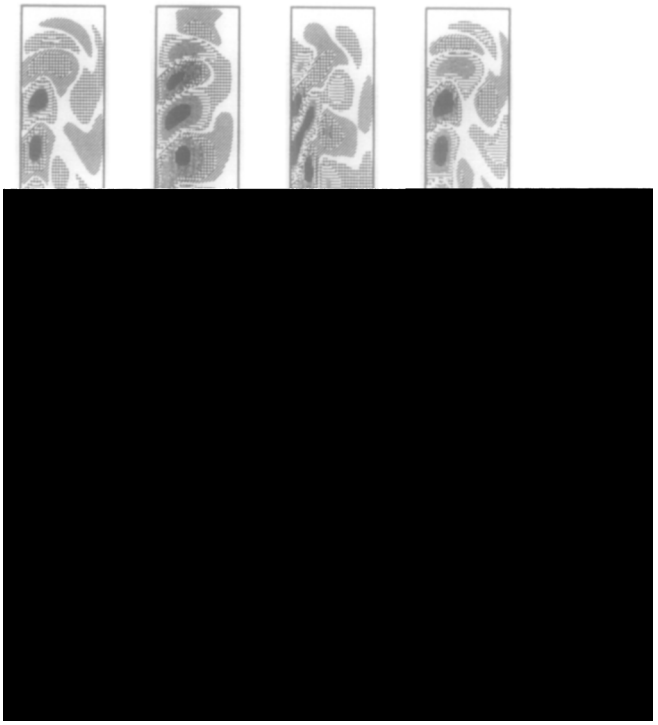


FIG. 7. Instantaneous fluctuating fields: from left to right: Temperature, u-velocity, w-velocity, and density; $Ra = 3 \times 10^5$; $A = 8$; $\varepsilon = 0.6$. Isovalue ranges correspond to the greyscale shown; smallest positive and negative values around zero are in white; the other isovalue ranges are symmetrical with respect to zero.

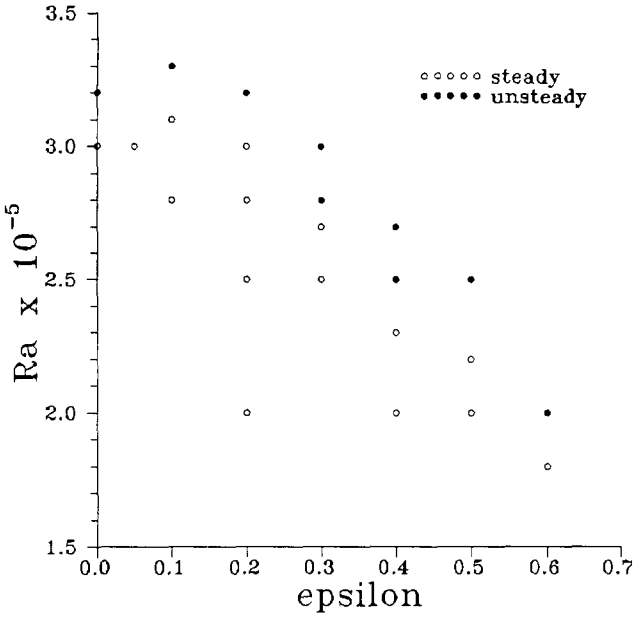


FIG. 8. Stability region: $A = 8$; adiabatic top and bottom walls; (●) unsteady solutions; (○) steady solutions.

extended by polynomial interpolation to $(L, M) = (32, 96)$. Using these interpolated fields as the initial condition, time integration was then resumed for 3000 time steps with the same value of the time step. The location of the monitoring point was the same as in the previous time sequence. Spatial resolution was then decreased to $(L, M) = (20, 64)$ and time integration was carried out for 5000 time steps with again the same value of the time step. Figure 6b shows the concatenation of the temperature traces at the monitoring point for the three different spatial resolutions. It can be seen that the increase of spatial resolution from (24, 72) to (32, 96) can hardly be noticed on the time trace and that even the smaller resolution, except for the startup which shows up on the trace, has a negligible effect on the amplitude of the fluctuation. We thus feel confident that the solutions computed with $(L, M) = (24, 72)$ are very accurate and most of the solutions were therefore obtained with that resolution.

Once the asymptotic time-periodic solution is reached it is possible to compute the time-averaged fields from a simple arithmetic mean and to obtain the fields of fluctuating quantities such as temperature, velocity components and density. Typical instantaneous fields for the corresponding solution are shown in Fig. 7. These fluctuating fields are very similar to those found in the Boussinesq case for small supercritical values of the Rayleigh number. One notable difference though is that they break the centro-symmetry property, the fluctuations being of larger amplitude in the ascending boundary layer along the hot wall than along the cold wall. What is also particularly noteworthy is the close resemblance between the temperature fluctuations $\tilde{T}$ and the density fluctuations $\tilde{\rho}$ since these fields are obviously linked by a relationship of the form $\bar{\rho}\tilde{T} + \tilde{T}\tilde{\rho} = 0$ for small supercritical values of the Rayleigh number, where $\bar{T}$ and $\bar{\rho}$ stand for the time-averaged fields.

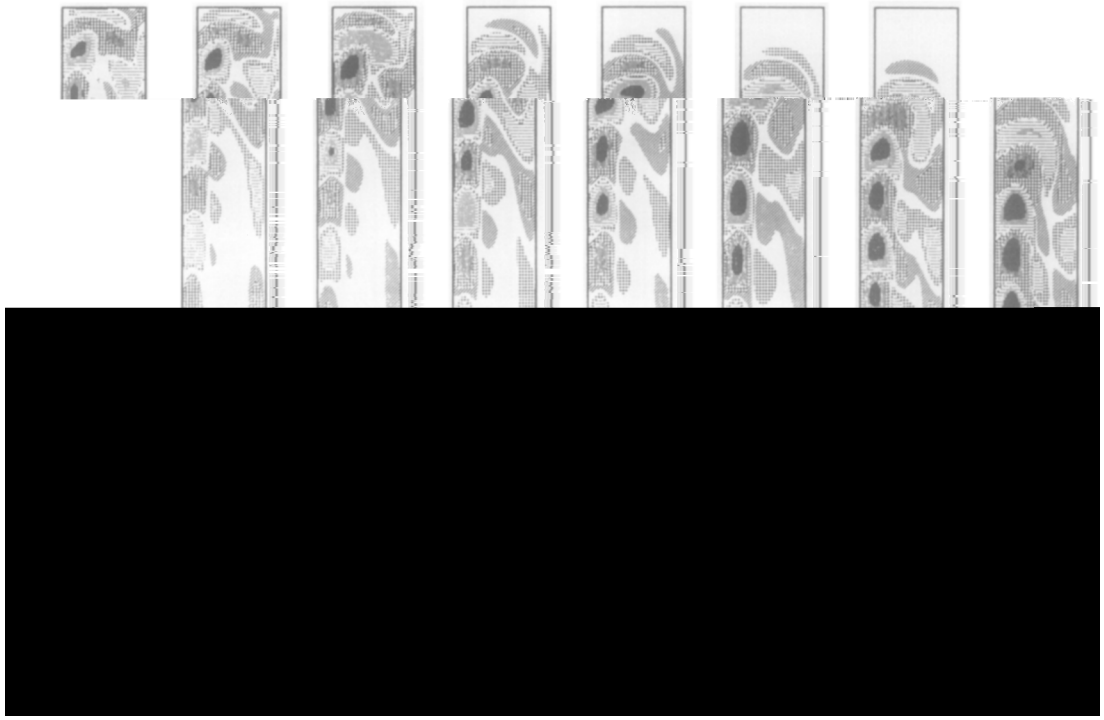


FIG. 9. Instantaneous fluctuating temperature fields. From left to right: $\mathrm{Ra} = 3.2 \times 10^5$, $\varepsilon = 0$; $\mathrm{Ra} = 3.3 \times 10^5$, $\varepsilon = 0.1$; $\mathrm{Ra} = 3.2 \times 10^5$, $\varepsilon = 0.2$; $\mathrm{Ra} = 2.8 \times 10^5$, $\varepsilon = 0.3$; $\mathrm{Ra} = 2.5 \times 10^5$, $\varepsilon = 0.4$; $\mathrm{Ra} = 2.5 \times 10^5$, $\varepsilon = 0.5$; $\mathrm{Ra} = 2.0 \times 10^5$, $\varepsilon = 0.6$. Isovalue ranges correspond to the greyscale shown; smallest positive and negative values around zero are in white; the other isovalue ranges are symmetrical with respect to zero.

Similar determinations of upper and lower bounds of the critical Rayleigh value were carried out for other values of ε. The results are shown in Fig. 8 in the $\mathrm{Ra}-\varepsilon$ plane for values of ε up to 0.6. Increasing ε from 0 to 0.6 decreases the critical value of the Rayleigh number from 3.1×10^5 to less than 2×10^5. It should also be noted that the solution corresponding to a value of ε of 0.1 is found slightly more stable than the Boussinesq solution. Computing all these solutions has required about 50 h of VP-200[1] and we have not undertaken to determine accurately the critical values for obvious reasons. Increasing the value of ε also influences the structure of the unsteady solution. First, as already noted, it breaks the centro-symmetry property and solutions become increasingly dissymmetrical with increasing ε. The boundary layer on the heated side becomes thicker than that on the cold side, which agrees with Chenoweth and Paolucci's results. Figure 9 shows several instantaneous fluctuating temperature fields corresponding to the time-periodic solutions found for increasing values of ε. As was previously shown in the Boussinesq case, the fluctuating temperature field is made of several wavelengths of variable length which circulate around the cavity undergoing successive amplification and damping. As can be seen from Fig. 9, these sequences of amplification and damping become increasingly dissymmetrical with increasing ε, and at the larger values of ε, one even reaches a configuration in which fluctuations are almost constantly amplified in the upward boundary layer and damped as they travel downstream along the cold wall. On the other hand, Fig. 9a shows that the time-periodic solution which appears at criticality in the Boussinesq case is made of 11 wavelengths. (We define one wavelength as being made of two consecutive structures of alternate sign. Note also that an odd number of structures is still compatible with the so-called centro-symmetry property of the solution.) Figures 9b and c show that the solution obtained close to criticality for small ε is also made of 11 wavelengths, although this solution breaks the centro-symmetry property. The fact that the time-periodic solution found at onset of unsteadiness in a cavity of aspect ratio 8 is made of 11 circulating structures thus seems to be a robust property with respect to imperfections which break the centro-symmetry of the solutions.

8.2. *Influence of Initial Mass*

As was already discussed in Section 6, the same final configuration can be reached from several initial states and only differs through the mass M_0 of fluid initially contained in the cavity. Even though any value of M_0 could be in principle considered, we have seen that the most logical assumptions are to consider either $M_0=1$, $M_0=1/(1-\varepsilon)$ or $M_0=1/(1+\varepsilon)$ corresponding to conservation of mass for an isothermal initial state at $(T_0,\bar{P}_0)$, $(T_1,\bar{P}_0)$, or $(T_2,\bar{P}_0)$, respectively. The fourth possibility is simply to set $\bar{P}=1$, in which case the mass M in the asymptotic state is not known a priori and is given by $\int_\Omega 1/T$.

For the purpose of comparison with experiments and in view of the uncertainty in the actual operating conditions, one should therefore investigate the influence of this parameter on the onset of unsteady convection. From preceding considerations, it seems logical to limit the range of values of M_0 from $1/(1+\varepsilon)$ to $1/(1-\varepsilon)$. We have only considered the case of a small ε value of 0.1 and an aspect ratio 8 cavity, and one should keep in mind that larger values of ε should have an increasing effect. For $M_0=1$, it

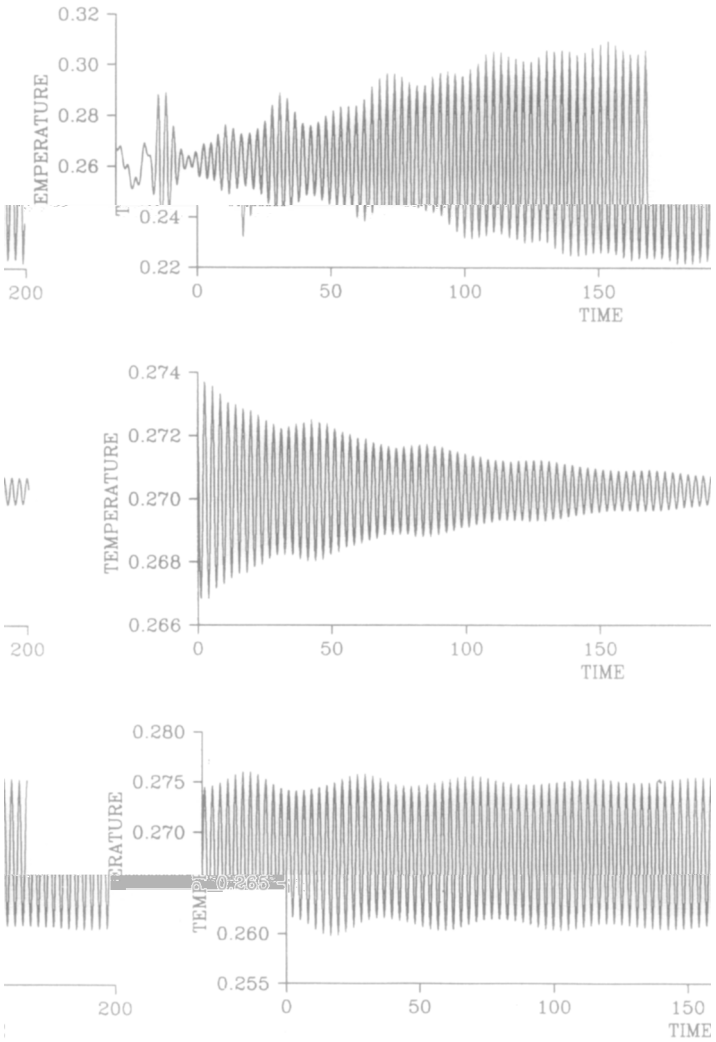


FIG. 10. Time trace of temperature at location $(x, z/A)=(-0.71, 0.71)$ for $(A=8, \varepsilon=0.1, M_0=1/(1-\varepsilon))$: (a, top) $\mathrm{Ra}=3.1\times10^5$, development of unsteadiness; (b, middle) $\mathrm{Ra}=2.4\times10^5$, damped oscillations that will utimately yield a steady solution; (c, bottom) $\mathrm{Ra}=2.6\times10^5$, asymptotic solution in time is unsteady. Time integrations performed with $\Delta t=2.5\times10^{-2}$ and $(L, M)=(24, 72)$.

[1] The Fujitsu VP-200 is a 500 Mflops-class machine.

was shown in the previous section that the critical value is bound between 3.1×10^5 and 3.3×10^5. Starting from the steady state solution found for 3.1×10^5 and changing the value of M_0 to $1/(1-\varepsilon)$ yields an unsteady solution as can be seen from Fig. 10a, showing that increasing the density has a destabilizing effect. For that value of M_0, unsteady solutions are found for values of Ra down to 2.6×10^5 (Fig. 10c), whereas the solution for $\mathrm{Ra} = 2.4 \times 10^5$ remains steady (Fig. 10b). On the other hand, values of M_0 smaller than one have a stabilizing influence on the solution. For $M_0 = 1/(1+\varepsilon)$ and a Ra value of 3.3×10^5, integration using the corresponding solution for $M_0 = 1$ as initial condition results in a steady solution after sufficiently long time integration can be seen from Fig. 11a. For this value of M_0 one has to increase the Ra value to 4×10^5 (Fig. 11b) to see an unsteady solution appear, the solution for 3.8×10^5 being steady (Fig. 11c).

The third possibility we have examined, $\bar{P} = 1$, has a very minor effect on the solution. This is so because the mean pressure $\bar{P}$ corresponding to $M_0 = 1$ is about 0.992 with small fluctuations in time. Setting $\bar{P}$ to 1 thus only changes the density by less than 1%, which is much smaller that the approximate 10% change for the two other possibilities.

Increasing the value of M_0 from $1/(1+\varepsilon)$ to $1/(1-\varepsilon)$ thus decreases the critical value of Ra from 3.9×10^5 to 2.5×10^5 approximately. This large difference is, however, fully consistent with the definition of the Rayleigh number as $(\rho_0^2 2\varepsilon g L_0^3/\mu_0\kappa_0)$ which shows that the effective Rayleigh numbers corresponding to values of M_0 equal $1/(1+\varepsilon)$ or $1/(1-\varepsilon)$ are respectively $(1/1+\varepsilon))^2$ or $(1/(1-\varepsilon))^2$ times the nominal value of Ra.

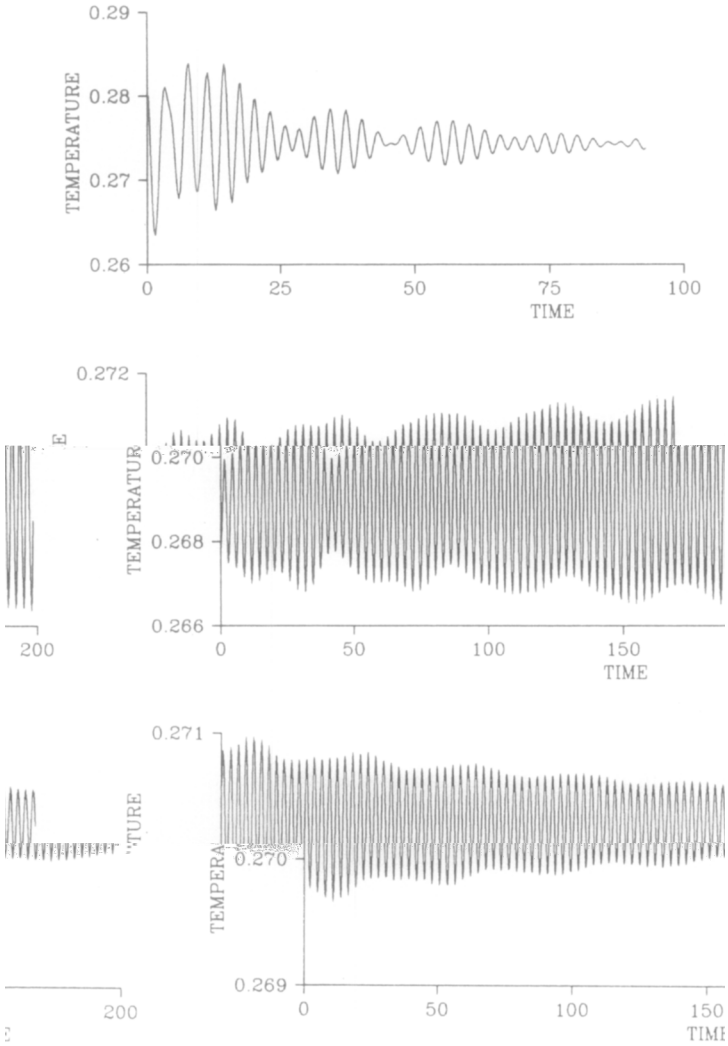


FIG. 11. Time trace of temperature at location $(x, z/A) = (-0.71, 0.71)$ for $(A = 8, \varepsilon = 0.1, M_0 = 1/(1+\varepsilon))$: (a, top) $\mathrm{Ra} = 3.3 \times 10^5$, return to steady state; (b, middle) $\mathrm{Ra} = 4.0 \times 10^5$, unsteady solution developing in time; (c, bottom) $\mathrm{Ra} = 3.8 \times 10^5$, damped oscillations that will utimately yield a steady solution. Time integrations performed with $\Delta t = 2.5 \times 10^{-2}$ and $(L, M) = (24, 72)$.

9. CONCLUSION

We have proposed a pseudo-spectral algorithm to compute non-Boussinesq natural convection in a cavity under large temperature differences. The results we have obtained shows that the algorithm is effective since we have been able to reliably and accurately bound the critical value of the Rayleigh number corresponding to the onset of unsteadiness for a large range of temperature differences. In the case considered, the temperature difference has a significant effect on the value of the Rayleigh number at which unsteadiness sets in and on the structure of the unsteady solution.

ACKNOWLEDGMENTS

This work was supported by DRET under Contract 88/169. Computing time on the VP200 at CIRCE was provided by "Direction Scientifique du SPI."

REFERENCES

1. P. Haldenwang, Thèse d'Etat, University of Provence, Dec. 1984 (unpublished).
2. P. Le Quéré and T. Alziary de Roquefort, *J. Comput. Phys.* **57**, 210 (1985).
3. J. M. Vanel, R. Peyret, and P. Bontoux, in *Numer. Methods Fluid Dyn. II*, edited by Moreton and Baines (Clarendon Press, Oxford, 1986), p. 477.
4. D. D. Gray and A. Giorgini, *Int. J. Heat Mass Transf.* **19**, 545 (1976).
5. E. A. Spiegel and G. Veronis, *Astrophys. J.* **131**, 442 (1960).
6. R. K. McGregor and A. F. Emery, *J. Heat Transf.* **391** (1969).
7. V. I. Polezhaev, *Fluid Dyn. (USSR)* **2**, 70 (1967).
8. L. W. Spradley and S. W. Churchill, *J. Fluid Mech.* **70**, 705 (1975).
9. E. Graham, *J. Fluid Mech.* **70**, 689 (1975).
10. J. Patterson and J. Imberger, *J. Fluid Mech.* **100**, 65 (1980).

11. E. Leonardi and J. A. Reizes, in *Num. Methods in Thermal Problems*, edited by Lewis and Morgan (Pineridge Press, Swansea, UK, 1979), p. 297.
12. E. Leonardi and J. A. Reizes, in *Num. Methods in Heat Transf.*, edited by Lewis, Morgan, and Zienkiewicz (Wiley, New York, 1981), p. 387.
13. R. G. Rehm and H. R. Baum, *J. Res. Nat. Bur. Stand.* **83**, 297 (1978).
14. C. K. Forester and A. F. Emery, *J. Comput. Phys.* **10**, 487 (1972).
15. D. O. Gough, *J. Atmos. Sci.* **26**, 448 (1969).
16. S. Paolucci, Sandia National Lab. Report SAND 82-8257, 1982 (unpublished).
17. P. Le Quéré, J. A. C. Humphrey, and F S. Sherman, *Num. Heat Transf.* **4**, 249 (1981).
18. D. R. Chenoweth and S. Paolucci, *J. Fluid Mech.* **169**, 173 (1986).
19. S. Paolucci and D. R. Chenoweth, *J. Fluid Mech.* **201**, 379 (1989).
20. Roux (Ed.), *Notes on Numerical Fluid Mechanics*, Vol. 27 (Vieweg, Munich, 1990), p. 227.
21. J. Fröhlich and R. Peyret, *Computer Methods in Applied Mechanics and Engineering*, Vol. 80 (North-Holland, Amsterdam, 1990), p. 425.
22. J. Fröhlich, Thesis, Université de Nice, June 1990 (unpublished).
23. D. Haidvogel and T. A. Zang, *J. Comput. Phys.* **30**, 167 (1979).
24. H. Guillard and J. A. Désidéri, *Computer Methods in Applied Mechanics and Engineering*, Vol. 80 (North-Holland, Amsterdam, 1990), p. 305.
25. C. Canuto, M. Y. Hussaini, A. Quarteroni, T. A. Zang, *Spectral Methods in Fluid Dynamics* (Springer-Verlag, New York, 1988).
26. Y. Morchoisne, in *3ème Cong. Int. Meth. Num. Ing., Paris, 1983*, edited by P. Lascaux, p. 275.
27. F. Montigny-Rannou and Y. Morchoisne, *Int. J. Num. Meth. Fluids* **7**, 175 (1987).
28. C. Bernardi and Y. Maday, *Int. J. Num. Meth. Fluids* **8**, 537 (1988).
29. P. Le Quéré, in *Notes on Numerical Fluid Mechanics*, Vol. 27, edited by Roux (Vieweg, Munich, 1990), p. 227.
30. E. M. Rønquist, Ph.D. thesis, MIT, June 1988 (unpublished).
31. P. Le Quéré and T. Alziary de Roquefort, in *Significant Questions in Buoyancy Affected Enclosure or Cavity Flows*, HTD Vol. 60, (ASME WAM, New York, 1986), p. 29.