J. math. fluid mech. 6 (2004) 1–20
1422-6928/04/010001-20

DOI 10.1007/s00021-003-0077-3

Journal of Mathematical Fluid Mechanics

A Numerical Verification of Nontrivial Solutions for the Heat Convection Problem

Yoshitaka Watanabe, Nobito Yamamoto, Mitsuhiro T. Nakao and Takaaki Nishida

Communicated by W. Nagata

Abstract. A computer assisted proof of the existence of nontrivial steady-state solutions for the two-dimensional Rayleigh–Bénard convection is described. The method is based on an infinite dimensional fixed-point theorem using a Newton-like operator. This paper also proposes a numerical verification algorithm which generates automatically on a computer a set including the exact nontrivial solution. All discussed numerical examples take into account of the effects of rounding errors in the floating point computations.

Mathematics Subject Classification (2000). 35Q35, 65N15, 76D03.

Keywords. Rayleigh–Bénard convection, computer assisted proof, fixed-point theorem.

1. Introduction

Consider a plane horizontal layer ($0 \le z \le h$) of an incompressible viscous fluid heated from below. At the lower boundary: $z = 0$ the layer of fluid is maintained at temperature $T + \delta T$ and the temperature of the upper boundary ($z = h$) is T (see Fig. 1).

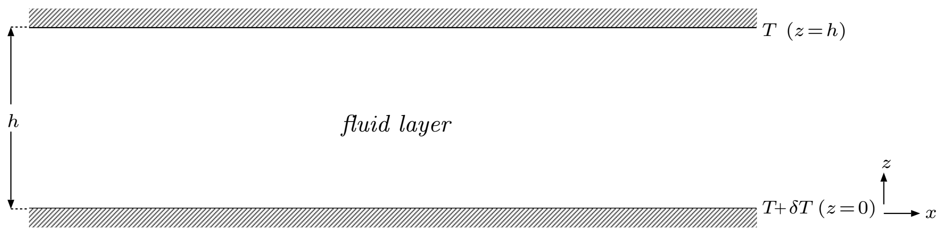


FIG. 1. Geometry of the convection problem

All variations with respect to the y-direction are assumed to vanish, then according to the Oberbeck–Boussinesq approximations [3, 4], the equations governing convection in a layer in the two-dimensional (x-z) are described as follows:

$$\left\{\begin{array}{rcl} u_t + uu_x + wu_z & = & -p_x/\rho_0 + \nu\Delta u, \\ w_t + uw_x + ww_z & = & -(p_z + g\rho)/\rho_0 + \nu\Delta w, \\ u_x + w_z & = & 0, \\ \theta_t + u\theta_x + w\theta_z & = & \kappa\Delta\theta. \end{array}\right. \tag{1}$$

In the above system (1), $(u, 0, w)$ is the velocity vector field in the respective directions (x, y, z); p is the pressure field; θ is the temperature; ρ is the fluid density; ρ_0 is the density at temperature $T + \delta T$; ν is the kinematic viscosity; g is the gravitational acceleration; κ is the coefficient of thermal diffusivity; $*_\xi{:=}\partial/\partial\xi(\xi = x, z, t)$; and $\Delta := \partial^2/\partial x^2 + \partial^2/\partial z^2$. The Oberbeck–Boussinesq approximation also requires that the fluid density is to be independent of pressure and depends linearly on the temperature θ, therefore ρ can be represented by

$$\rho - \rho_0 = -\rho_0\alpha(\theta - T - \delta T),$$

where α is the coefficient of thermal expansion.

The Oberbeck–Boussinesq equations (1) have a stationary solution:

$$u^* = 0, \quad w^* = 0, \quad \theta^* = T + \delta T - \frac{\delta T}{h}z, \quad p^* = p_0 - g\rho_0(z + \frac{\alpha\delta T}{2h}z^2)$$

representing the purely heat conducting state, where p_0 is a constant. By setting

$$\hat{u} := u, \qquad \hat{w} := w, \qquad \hat{\theta} := \theta^* - \theta, \qquad \hat{p} := p^* - p,$$

the perturbed equations

$$\left\{\begin{array}{rcl} \hat{u}_t + \hat{u}\hat{u}_x + \hat{w}\hat{u}_z & = & \hat{p}_x/\rho_0 + \nu\Delta\hat{u}, \\ \hat{w}_t + \hat{u}\hat{w}_x + \hat{w}\hat{w}_z & = & \hat{p}_z/\rho_0 - g\alpha\hat{\theta} + \nu\Delta\hat{w}, \\ \hat{u}_x + \hat{w}_z & = & 0, \\ \hat{\theta}_t + \delta T\hat{w}/h + \hat{u}\hat{\theta}_x + \hat{w}\hat{\theta}_z & = & \kappa\Delta\hat{\theta}, \end{array}\right. \tag{2}$$

are obtained. Moreover, transforming to dimensionless variables

$$t \to \kappa t, \quad u \to \hat{u}/\kappa, \quad w \to \hat{w}/\kappa, \quad \theta \to \hat{\theta}h/\delta T, \quad p \to \hat{p}/(\rho_0\kappa^2)$$

in (2), the dimensionless equations

$$\left\{\begin{array}{rcl} u_t + uu_x + wu_z & = & p_x + \mathcal{P}\Delta u, \\ w_t + uw_x + ww_z & = & p_z - \mathcal{P}R\,\theta + \mathcal{P}\Delta w, \\ u_x + w_z & = & 0, \\ \theta_t + w + u\theta_x + w\theta_z & = & \Delta\theta \end{array}\right. \tag{3}$$

are obtained, where

$$\mathcal{R} := \frac{\delta T\alpha g}{\kappa\nu h}$$

is the Rayleigh number[1] and

$$\mathcal{P} := \frac{\nu}{\kappa}$$

[1] The Rayleigh number is sometimes defined by $\mathcal{R} = (\delta T\alpha g h^3)/(\kappa\nu)$ when the dimensionless equations are reduced to the domain of $0 \le z \le 1$.

is the Prandtl number.

It is well known that for small $\mathcal{R}$ the fluid conducts heat diffusively, and at a critical point $\mathcal{R}_C$, heat is transposed through the fluid by convection. The origin of these rolls is the experiment by Bénard [2] in 1900. He observed the establishment of a regular, steady pattern of flow cells in a thin horizontal layer of molten spermaceti with a free upper surface, then these cells which later came to be known as *Bénard cells*. In 1916, Lord Rayleigh considered the linearized stability of (3) and found $\mathcal{R}_C$ when both the upper and lower boundaries are taken to be stress-free [15]. From the above-mentioned problem (3) (of course including the three dimensional case) is called *Rayleigh–Bénard convection*.

Although a large number of studies have been made on Rayleigh–Bénard convection [4, 10], theoretical results are very few. It has been shown by Joseph [7] that (3) has a unique trivial solution for $\mathcal{R} < \mathcal{R}_C$. Iudovich [6] and Rabinowitz [14] proved that, for each $\mathcal{R}$ slightly exceeding the critical Rayleigh number $\mathcal{R}_C$, the equation (3) has at least two nontrivial steady-state solutions. The stability analysis of the bifurcated solution in a small neighbourhood of the bifurcation points is considered by Kagei and Wahl [8]. However, the global structure of bifurcated solutions after the critical Rayleigh point $\mathcal{R}_C$ has not been known *theoretically* up to now.

In this paper, we propose an approach to prove the existence of the steady-state solutions for given $\mathcal{P}$ and $\mathcal{R}$ by a computer assisted proof, which gives us a tool to study the global bifurcation structure. This method is based on an infinite dimensional fixed-point theorem using a Newton-like operator together with a spectral approximation and constructive error estimates. Another method of computer assisted proof for the Navier–Stokes equations have been presented by Heywood, Nagata and Xie [5]. This method needs a norm bound for the inverse of the linearized problem and, unfortunately, rigorous numerical results cannot be obtained. Our method does not need a norm estimate of the linearized problem and our numerical examples take into account the effects of rounding errors in floating point computations.

The contents of this paper are as follows. The boundary conditions and some function spaces and notations are defined, and the fixed-point formulation is introduced in Section 2. Constructive a priori error estimates for the linearized problems are described in Section 3, which are needed in numerical computations. An existence theorem in certain appropriate function spaces using Newton-like iteration is considered in Section 4. A computable verification condition is given in Section 5. Numerical examples which prove the existence of steady-state solutions are described in Section 6.

2. A fixed-point formulation

We shall find *steady-state solutions*, where u_t, w_t and θ_t are equated to 0 in (3), and assume that all fluid motion is confined to the rectangular region $\Omega := \{0 < x < 2\pi/a,\ 0 < z < \pi\}$ for a given wave number $a > 0$. Let us impose periodic boundary conditions (with period $2\pi/a$) in the horizontal direction, stress-free boundary conditions ($u_z = w = 0$) for the velocity field, and Dirichlet boundary conditions ($\theta = 0$) for the temperature field on the surfaces $z = 0, \pi$. Furthermore, we assume the following evenness and oddness conditions:

$$u(x,z) = -u(-x,z), \quad w(x,z) = w(-x,z), \quad \theta(x,z) = \theta(-x,z).$$

We introduce the stream function Ψ, through the definition

$$u = -\Psi_z, \quad w = \Psi_x$$

so that $u_x + w_z = 0$. Cross-differentiating the equation of motion in (3) in order to eliminate the pressure p and setting $\Theta := \sqrt{\mathcal{PR}}\theta$, we obtain

$$\begin{cases} \mathcal{P}\Delta^2\Psi \;=\; \sqrt{\mathcal{PR}}\,\Theta_x - \Psi_z\Delta\Psi_x + \Psi_x\Delta\Psi_z \ \text{ in } \Omega, \\ \;-\Delta\Theta \;=\; -\sqrt{\mathcal{PR}}\,\Psi_x + \Psi_z\Theta_x - \Psi_x\Theta_z \quad \text{ in } \Omega. \end{cases} \tag{4}$$

From the boundary conditions imposed above, the stream function Ψ and departure of temperature from linear profile Θ can be represented by the following double Fourier series:

$$\Psi = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\sin(amx)\sin(nz), \quad \Theta = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} B_{mn}\cos(amx)\sin(nz). \tag{5}$$

By (5), we introduce following function spaces for $k \geq 0$:

$$X^k := \left\{ \Psi = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\sin(amx)\sin(nz) \;\middle|\; A_{mn} \in \boldsymbol{R}, \right.$$
$$\left. \sum_{m=1}^{\infty}\sum_{n=1}^{\infty}((am)^{2k} + n^{2k})A_{mn}^2 < \infty \right\},$$

$$Y^k := \left\{ \Theta = \sum_{m=0}^{\infty}\sum_{n=1}^{\infty} B_{mn}\cos(amx)\sin(nz) \;\middle|\; B_{mn} \in \boldsymbol{R}, \right.$$
$$\left. \sum_{m=0}^{\infty}\sum_{n=1}^{\infty}((am)^{2k} + n^{2k})B_{mn}^2 < \infty \right\}$$

which are considered as closed subspaces of usual k-th order Sobolev space $H^k(\Omega)$.

For $M_1, N_1, M_2 \geq 1$ and $N_2 \geq 0$, we indicate a relation $N := (M_1, N_1, M_2, N_2)$

and define the finite dimensional approximating subspaces by

$$\begin{aligned}
S_N^{(1)} &= \left\{ \Psi_N = \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} \hat{A}_{mn} \sin(amx)\sin(nz) \;\middle|\; \hat{A}_{mn} \in \boldsymbol{R} \right\}, \\
S_N^{(2)} &= \left\{ \Theta_N = \sum_{m=0}^{M_2} \sum_{n=1}^{N_2} \hat{B}_{mn} \cos(amx)\sin(nz) \;\middle|\; \hat{B}_{mn} \in \boldsymbol{R} \right\}, \\
S_N &= S_N^{(1)} \times S_N^{(2)},
\end{aligned}$$

and denote an approximate solution of (4) by $\hat{u}_N := (\hat{\Psi}_N, \hat{\Theta}_N) \in S_N$ which is obtained by an appropriate method. Then setting

$$\left\{ \begin{aligned}
f_1(\Psi, \Theta) &:= \sqrt{\mathcal{PR}}\,\Theta_x - \Psi_z \Delta\Psi_x + \Psi_x \Delta\Psi_z, \\
f_2(\Psi, \Theta) &:= -\sqrt{\mathcal{PR}}\,\Psi_x + \Psi_z \Theta_x - \Psi_x \Theta_z,
\end{aligned} \right.$$

$$\Psi = \hat{\Psi}_N + w^{(1)}, \qquad \Theta = \hat{\Theta}_N + w^{(2)},$$

(4) is rewritten as the problem to find $(w^{(1)}, w^{(2)}) \in X^4 \times Y^2$ satisfying

$$\left\{ \begin{aligned}
\mathcal{P}\Delta^2 w^{(1)} &= f_1(\hat{\Psi}_N + w^{(1)}, \hat{\Theta}_N + w^{(2)}) - \mathcal{P}\Delta^2 \hat{\Psi}_N \ \text{ in } \Omega, \\
-\Delta w^{(2)} &= f_2(\hat{\Psi}_N + w^{(1)}, \hat{\Theta}_N + w^{(2)}) + \Delta\hat{\Theta}_N \qquad \text{ in } \Omega.
\end{aligned} \right. \tag{6}$$

Note that $(w^{(1)}, w^{(2)})$ is expected to be small if $\hat{u}_N$ is an accurate approximation. Defining

$$\begin{aligned}
w &= (w^{(1)}, w^{(2)}), \\
h_1(w) &= f_1(\hat{\Psi}_N + w^{(1)}, \hat{\Theta}_N + w^{(2)}) - \mathcal{P}\Delta^2 \hat{\Psi}_N, \\
h_2(w) &= f_2(\hat{\Psi}_N + w^{(1)}, \hat{\Theta}_N + w^{(2)}) + \Delta\hat{\Theta}_N, \\
h(w) &= (h_1(w), h_2(w)),
\end{aligned}$$

by virtue of Sobolev embedding theorem and the definition of f_1 and f_2, h is a bounded continuous map from $X^3 \times Y^1$ to $X^0 \times Y^0$. Moreover, it is easily shown that for all $(g_1, g_2) \in X^0 \times Y^0$, the linear problem

$$\left\{ \begin{aligned}
\Delta^2 \bar{\Psi} &= g_1 \ \text{ in } \Omega, \\
-\Delta\bar{\Theta} &= g_2 \ \text{ in } \Omega
\end{aligned} \right. \tag{7}$$

has a unique solution $(\bar{\Psi}, \bar{\Theta}) \in X^4 \times Y^2$. When this mapping is denoted by $\bar{\Psi} = (\Delta^2)^{-1} g_1$ and $\bar{\Theta} = (-\Delta)^{-1} g_2$, an operator:

$$\mathcal{K} := (\mathcal{P}^{-1}(\Delta^2)^{-1}, (-\Delta)^{-1}) : \ X^0 \times Y^0 \to X^3 \times Y^1$$

is a compact map because of the compactness of the imbeddings $H^4(\Omega) \hookrightarrow H^3(\Omega)$, $H^2(\Omega) \hookrightarrow H^1(\Omega)$ and the boundedness of $(\Delta^2)^{-1} : X^0 \to X^4$, $(-\Delta)^{-1} : Y^0 \to Y^2$. Therefore, (6) is rewritten as a fixed-point equation:

$$w = Fw \tag{8}$$

for the compact operator $F := \mathcal{K} \circ h$ on $X^3 \times Y^1$, and Schauder's fixed-point theorem (e.g. Zeidler[17]) asserts that, for a nonempty, closed, bounded and convex set $W \subset X^3 \times Y^1$, if

$$FW \subset W \tag{9}$$

holds, then there exists a fixed-point of (8) in W. In the following sections, we shall propose a computer algorithm to construct a *candidate set* W which satisfies (9).

3. Constructive error estimates for the projections

This section is devoted to the estimation of constructive a priori constants of the finite dimensional approximations.

We denote the L^2-inner product and the L^2-norm on Ω by $(\,\cdot,\,\cdot)_{L^2}$ and $\|\cdot\|_{L^2}$, respectively, and also define the H_0^1-norm: $\|\nabla u\|_{L^2}$ and the H^k-norm: $\|u\|_{H^k}$ on Ω by $\|\nabla u\|_{L^2}^2 = \|u_x\|_{L^2}^2 + \|u_z\|_{L^2}^2$ and $\|u\|_{H^k}^2 = \sum_{i,j\in\mathbf{N},i+j\le k} \|\partial^{i+j}u/\partial^i x\partial^j z\|_{L^2}^2$, respectively.

For $\Psi \in X^3$ and $\Theta \in Y^1$, let us define projections $P_N^{(1)}\Psi \in S_N^{(1)}$ and $P_N^{(2)}\Theta \in S_N^{(2)}$ by

$$\begin{cases} (\Delta(P_N^{(1)}\Psi - \Psi), \Delta v_N^{(1)})_{L^2} \,=\, 0 & \forall v_N^{(1)} \in S_N^{(1)}, \\ (\nabla(P_N^{(2)}\Theta - \Theta), \nabla v_N^{(2)})_{L^2} \,=\, 0 & \forall v_N^{(2)} \in S_N^{(2)}. \end{cases} \tag{10}$$

Then, it is easily shown that for given $(g_1, g_2) \in X^0 \times Y^0$ the projection $(P_N^{(1)}\bar{\Psi}$, $P_N^{(2)}\bar{\Theta})$ of the solution $(\bar{\Psi}, \bar{\Theta}) \in X^4 \times Y^2$ of (7) coincides with (M_1, N_1)-truncation of $\bar{\Psi}$ and (M_2, N_2)-truncation of $\bar{\Theta}$, respectively. From this fact, the following constructive a priori error estimates are derived.

Theorem 1. *For each* $(g_1, g_2) \in X^0 \times Y^0$*, let* $(\psi, \theta) \in X^4 \times Y^2$ *be the solution of* (7) *and* $(P_N^{(1)}\psi, P_N^{(2)}\theta) \in S_N$ *be finite dimensional approximation defined by* (10)*, then following a priori error estimates hold:*

$$\begin{cases} \|\psi - P_N^{(1)}\psi\|_{L^2} \le C_1\|g_1\|_{L^2}, & \|\nabla(\psi - P_N^{(1)}\psi)\|_{L^2} \le C_2\|g_1\|_{L^2}, \\ \|\Delta(\psi - P_N^{(1)}\psi)\|_{L^2} \le C_3\|g_1\|_{L^2}, & \|\psi - P_N^{(1)}\psi\|_{H^3} \le C_4\|g_1\|_{L^2}, \\ \|(\psi - P_N^{(1)}\psi)_x\|_{L^2} \le C_5\|g_1\|_{L^2}, & \|(\psi - P_N^{(1)}\psi)_z\|_{L^2} \le C_6\|g_1\|_{L^2}, \\ \|\nabla(\psi - P_N^{(1)}\psi)_x\|_{L^2} \le C_7\|g_1\|_{L^2}, & \|\nabla(\psi - P_N^{(1)}\psi)_z\|_{L^2} \le C_8\|g_1\|_{L^2}, \\ \|\Delta(\psi - P_N^{(1)}\psi)_x\|_{L^2} \le C_9\|g_1\|_{L^2}, & \|\Delta(\psi - P_N^{(1)}\psi)_z\|_{L^2} \le C_{10}\|g_1\|_{L^2}, \end{cases} \tag{11}$$

$$\begin{cases} \|\theta - P_N^{(2)}\theta\|_{L^2} \le C_{11}\|g_2\|_{L^2}, & \|\nabla(\theta - P_N^{(2)}\theta)\|_{L^2} \le C_{12}\|g_2\|_{L^2}, \\ \|\theta - P_N^{(2)}\theta\|_{H^1} \le C_{13}\|g_2\|_{L^2}, & \|(\theta - P_N^{(2)}\theta)_x\|_{L^2} \le C_{14}\|g_2\|_{L^2}, \\ \|(\theta - P_N^{(2)}\theta)_z\|_{L^2} \le C_{15}\|g_2\|_{L^2}, & \end{cases} \tag{12}$$

where $C_i (i = 1, \dots, 15)$ are given constructively:

$$
\begin{aligned}
C_1 &= \max\left\{ \frac{1}{(a^2+(N_1+1)^2)^2}, \frac{1}{(a^2(M_1+1)^2+1)^2} \right\},\\
C_2 &= \max\left\{ \frac{1}{(a^2+(N_1+1)^2)^{3/2}}, \frac{1}{(a^2(M_1+1)^2+1)^{3/2}} \right\},\\
C_3 &= \max\left\{ \frac{1}{a^2+(N_1+1)^2}, \frac{1}{a^2(M_1+1)^2+1} \right\},\\
C_4 &= \max\left\{ \left(\sum_{\nu=1}^{4} \frac{1}{(a^2+(N_1+1)^2)^\nu} \right)^{1/2}, \left(\sum_{\nu=1}^{4} \frac{1}{(a^2(M_1+1)^2+1)^\nu} \right)^{1/2} \right\},\\
C_5 &= \max\left\{ \max_{1\le m\le M_1} \frac{am}{(a^2m^2+(N_1+1)^2)^2}, \max_{M_1+1\le m\le\infty} \frac{am}{(a^2m^2+1)^2} \right\},\\
C_6 &= \max\left\{ \max_{1\le n\le N_1} \frac{n}{(a^2(M_1+1)^2+n^2)^2}, \max_{N_1+1\le n\le\infty} \frac{n}{(a^2+n^2)^2} \right\},\\
C_7 &= \max\left\{ \max_{1\le m\le M_1} \frac{am}{(a^2m^2+(N_1+1)^2)^{3/2}}, \max_{M_1+1\le m\le\infty} \frac{am}{(a^2m^2+1)^{3/2}} \right\},\\
C_8 &= \max\left\{ \max_{1\le n\le N_1} \frac{n}{(a^2(M_1+1)^2+n^2)^{3/2}}, \max_{N_1+1\le n\le\infty} \frac{n}{(a^2+n^2)^{3/2}} \right\},\\
C_9 &= \max\left\{ \max_{1\le m\le M_1} \frac{am}{a^2m^2+(N_1+1)^2}, \max_{M_1+1\le m\le\infty} \frac{am}{a^2m^2+1} \right\},\\
C_{10} &= \max\left\{ \max_{1\le n\le N_1} \frac{n}{a^2(M_1+1)^2+n^2}, \max_{N_1+1\le n\le\infty} \frac{n}{a^2+n^2} \right\},\\
C_{11} &= \max\left\{ \frac{1}{(N_2+1)^2}, \frac{1}{a^2(M_2+1)^2+1} \right\},\\
C_{12} &= \max\left\{ \frac{1}{N_2+1}, \frac{1}{(a^2(M_2+1)^2+1)^{1/2}} \right\},\\
C_{13} &= \max\left\{ \frac{(1+(N_2+1)^2)^{1/2}}{(N_2+1)^2}, \frac{(2+a^2(M_2+1)^2)^{1/2}}{a^2(M_2+1)^2+1} \right\},\\
C_{14} &= \max\left\{ \max_{0\le m\le M_2} \frac{am}{a^2m^2+(N_2+1)^2}, \max_{M_2+1\le m\le\infty} \frac{am}{a^2m^2+1} \right\},\\
C_{15} &= \max\left\{ \max_{1\le n\le N_2} \frac{n}{a^2(M_2+1)^2+n^2}, \frac{1}{N_2+1} \right\}.
\end{aligned}
$$

Proof. We show the construction of C_5. The other estimations are quite similar.
Set $\psi = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} A_{mn}\sin(amx)\sin(nz)$ and using

$$\|\sin(amx)\sin(nz)\|_{L^2} = \|\cos(amx)\sin(nz)\|_{L^2} = \pi/\sqrt{2a},$$

$$\begin{aligned}
\|(\psi - P_N^{(1)}\psi)_x\|_{L^2}^2 &= \sum_{m=1}^{M_1}\sum_{n=N_1+1}^{\infty} a^2m^2A_{mn}^2\frac{\pi^2}{2a} + \sum_{m=M_1+1}^{\infty}\sum_{n=1}^{\infty} a^2m^2A_{mn}^2\frac{\pi^2}{2a} \\
&= \sum_{m=1}^{M_1}\sum_{n=N_1+1}^{\infty}\frac{a^2m^2}{(a^2m^2+n^2)^4}(a^2m^2+n^2)^4A_{mn}^2\frac{\pi^2}{2a} \\
&\quad + \sum_{m=M_1+1}^{\infty}\sum_{n=1}^{\infty}\frac{a^2m^2}{(a^2m^2+n^2)^4}(a^2m^2+n^2)^4A_{mn}^2\frac{\pi^2}{2a} \\
&\le C_5^2\sum_{m=1}^{\infty}\sum_{n=1}^{\infty}(a^2m^2+n^2)^4A_{mn}^2\|\sin(amx)\sin(nz)\|_{L^2}^2 \\
&= C_5^2\|\Delta^2\psi\|_{L^2}^2. \qquad \square
\end{aligned}$$

These constructive a priori constants are optimal and have an adequate order, for example, when $\tilde{N} := M_1 = N_1$, we have $C_1 = \mathcal{O}(1/\tilde{N}^4)$ and $C_2 = \mathcal{O}(1/\tilde{N}^3)$.

In actual calculations, L^∞-error estimates are also needed. For this purpose, we use the following estimate by Plum [13].

Lemma 1. (Plum, 1992) *For $u \in X^2$ or $u \in Y^2$, the following assertion holds true:*

$$\|u\|_{L^\infty} \le K_1\|u\|_{L^2} + K_2\|\nabla u\|_{L^2} + K_3\|\Delta u\|_{L^2},$$

where $\|\cdot\|_{L^\infty}$ is the sup-norm and

$$K_1 = \frac{1}{\pi}\sqrt{\frac{a}{2}}, \quad K_2 = 1.1548\sqrt{\frac{4+a^2}{6a}}, \quad K_3 = \frac{0.22361\pi}{3}\sqrt{\frac{144+40a^2+9a^4}{10a^3}}.$$

Theorem 1 and Lemma 1 imply L^∞-error estimates immediately.

Corollary 1. *Under the same assumptions of Theorem 1, following a priori error estimates hold:*

$$\begin{aligned}
\|\psi - P_N^{(1)}\psi\|_{L^\infty} &\le C_{16}\|g_1\|_{L^2}, & \|\theta - P_N^{(2)}\theta\|_{L^\infty} &\le C_{17}\|g_2\|_{L^2}, \\
\|(\psi - P_N^{(1)}\psi)_x\|_{L^\infty} &\le C_{18}\|g_1\|_{L^2}, & \|(\psi - P_N^{(1)}\psi)_z\|_{L^\infty} &\le C_{19}\|g_1\|_{L^2},
\end{aligned}$$

where

$$\begin{aligned}
C_{16} &= C_1K_1 + C_2K_2 + C_3K_3, & C_{17} &= C_{11}K_1 + C_{12}K_2 + K_3, \\
C_{18} &= C_5K_1 + C_7K_2 + C_9K_3, & C_{19} &= C_6K_1 + C_8K_2 + C_{10}K_3.
\end{aligned}$$

4. Newton-like iteration

In this section, we apply the Newton-like method for nonlinear elliptic problems proposed by the author [11, 12] to the fixed-point equation (8).

By defining the projection from $X^3 \times Y^1$ to S_N by

$$P_N = (P_N^{(1)}, P_N^{(2)}),$$

the fixed-point problem $w = Fw$ can be uniquely decomposed as the finite dimensional (projection) part and infinite dimensional (error) part as follows:

$$\begin{cases} P_N w = P_N F w, \\ (I - P_N) w = (I - P_N) F w, \end{cases} \tag{13}$$

where I is the identity map on $H^3(\Omega) \times H^1(\Omega)$. We suppose that the restriction of the operator $P_N - P_N \mathcal{K} f'(\hat{u}_N) : X^3 \times Y^1 \longrightarrow S_N$ to S_N has an inverse

$$[I - P_N \mathcal{K} f'(\hat{u}_N)]_N^{-1} : S_N \longrightarrow S_N, \tag{14}$$

where $f'(\hat{u}_N)$ denotes the Fréchet derivative of $f := (f_1, f_2)$ at the approximate solution $\hat{u}_N$ which coincides with $h'(0)$. Note that this assumption is equivalent to the invertibility of a matrix, which is able to be numerically checked in actual verified computations (for example see Rump [16]). Applying the Newton-like method to the first term of (13) we define the operator $\mathcal{N}_N : X^3 \times Y^1 \longrightarrow S_N$ by

$$\mathcal{N}_N w = P_N w - [I - P_N \mathcal{K} f'(\hat{u}_N)]_N^{-1} P_N (I - F) w$$

and the compact map $T : X^3 \times Y^1 \longrightarrow X^3 \times Y^1$ by

$$Tw = \mathcal{N}_N w + (I - P_N) F w.$$

Then under the invertibility assumption of (14), the two fixed-point problems

$$w = Tw \tag{15}$$

and (8) are equivalent. If the approximate solution $\hat{u}_N = (\hat{\Psi}_N, \hat{\Theta}_N)$ is sufficiently good, the finite dimensional part of T will possibly be a contraction. On the other hand, the magnitude of the infinite dimensional part of T is expected to be small when the truncation numbers of S_N are taken to be sufficiently large, because of Theorem 1.

The question which we must consider next is to find a solution of (15) in a set W, referred to as a candidate set. Setting $L_1 := M_1 N_1$, $L_2 := (M_2 + 1) N_2$, $M := L_1 + L_2$ and denoting base functions of $S_N^{(1)}$, $S_N^{(2)}$ by $\{\psi_i\}_{i=1}^{L_1}$, $\{\theta_i\}_{i=1}^{L_2}$, respectively, we write any $(w_N^{(1)}, w_N^{(2)}) \in S_N$ as

$$w_N^{(1)} = \sum_{i=1}^{L_1} a_i \psi_i, \qquad w_N^{(2)} = \sum_{i=1}^{L_2} b_i \theta_i,$$

where a_i and b_i are real coefficients. Then, we define $(w_N^{(1)})_i$ and $(w_N^{(2)})_i$ by

$$(w_N^{(1)})_i := |a_i|, \qquad (w_N^{(2)})_i := |b_i|.$$

Now, for any non-negative real vector $\{W_i\}_{1\le i\le M+2}$, let us define

$$W_N := \left\{ (w_N^{(1)}, w_N^{(2)}) \in S_N \;\middle|\; \begin{array}{ll} (w_N^{(1)})_i \le W_i & (1 \le i \le L_1), \\ (w_N^{(2)})_i \le W_{L_1+i} & (1 \le i \le L_2) \end{array} \right\},$$

$$W_* := \left\{ (w_*^{(1)}, w_*^{(2)}) \in S_N^\perp \;\middle|\; \begin{array}{ll} \|w_*^{(1)}\|_\Psi \le C_i W_{M+1} & (1 \le i \le 10), \\ \|w_*^{(2)}\|_\Theta \le C_i W_{M+2} & (11 \le i \le 15) \end{array} \right\},$$

$$W := W_N \oplus W_*,$$

where $\|\cdot\|_\Psi$ and $\|\cdot\|_\Theta$ are generic symbols for seminorms corresponding to $C_i (1 \le i \le 15)$ in (11) and (12), for example $\|w_*^{(1)}\|_{L^2} \le C_1 W_{M+1}$, $\|(w_*^{(2)})_x\|_{L^2} \le C_{14} W_{M+2}$, and so on. Here, $S_N^\perp$ stands for the image space of $(I - P_N^{(1)}, I - P_N^{(2)}) : X^3 \times Y^1 \to X^3 \times Y^1$. Then, the following verification condition is obtained.

Theorem 2. *Let W_N, W_* and W be sets defined above. If the inclusions*

$$\begin{cases} \mathcal{N}_N W \subset W_N, \\ (I - P_N) F W \subset W_* \end{cases} \tag{16}$$

hold, then there exists a fixed-point w of T in W.

Proof. By definition, W is a non-empty, closed, convex and bounded set in $X^3 \times Y^1$. For any $w \in W$, $\mathcal{N}_N w \in S_N$, $(I - P_N) F w \in S_N^\perp$ and the decomposition $Tw = \mathcal{N}_N w + (I - P_N) F w$ is unique. Hence by (16), we get $\mathcal{N}_N W + (I - P_N) F W \subset W_N + W_*$ in $X^3 \times Y^1$, namely,

$$TW \subset W.$$

Therefore, by the compactness of the operator T and Schauder's fixed-point theorem, the desired result is obtained. □

5. Computable verification condition

In this section, we will show a computable verification algorithm to construct the candidate set in $X^3 \times Y^1$ which is expected to satisfy the verification condition (16).

First, for $k = 0$, take an appropriate initial non-negative real vector $\{W_i^{(0)}\}_{1\le i\le M+2}$ and set $W^{(0)} := W_N^{(0)} \oplus W_*{}^{(0)}$.

For $k \ge 1$, with a given $0 < \delta \ll 1$, set

$$\bar{W}_i^{(k-1)} := W_i^{(k-1)} (1 + \delta) \quad (1 \le i \le M + 2),$$

and for $\{\bar{W}_i^{(k-1)}\}_{1\le i\le M+2}$, define the δ-*inflation* of $W^{(k-1)}$ by

$$\bar{W}^{(k-1)} = \bar{W}_N^{(k-1)} \oplus \bar{W}_*{}^{(k-1)}.$$

For the set $\bar{W}^{(k-1)}$, construct the k-th set $W^{(k)} := W_N^{(k)} \oplus {W_*}^{(k)}$, where

$$\begin{cases} W_N^{(k)} := \mathcal{N}_N \bar{W}^{(k-1)}, \\ W_{M+1}^{(k)} := \mathcal{P}^{-1} \sup_{w \in \bar{W}^{(k-1)}} \|h_1(w)\|_{L^2}, \\ W_{M+2}^{(k)} := \sup_{w \in \bar{W}^{(k-1)}} \|h_2(w)\|_{L^2}. \end{cases} \tag{17}$$

Here, note that it is impossible to calculate each quantity of (17) *exactly* because $\bar{W}^{(k-1)}$ is the infinite dimensional set. However such a set can be obtained by enclosing $W^{(k)}$ in the *over-estimated sense* using a priori norm estimates and interval arithmetic described following subsections. In the actual calculation on a computer, we reset this over-estimated set as the candidate set $W^{(k)}$.

Now, we have the following verification condition in a computer.

Theorem 3. *If, for a step K,*

$$W_i^{(K)} < \bar{W}_i^{(K-1)} \qquad (1 \le i \le M+2)$$

hold, then in the set $\bar{W}^{(K-1)} = \bar{W}_N^{(K-1)} \oplus \bar{W}_*^{(K-1)}$ defined by $\{\bar{W}_i^{(K-1)}\}_{1 \le i \le M+2}$, there exists an element w satisfying $w = Tw$.

Proof. Applying Theorem 2, it is sufficient to check (16) holds for $\bar{W}^{(K-1)}$. By assumption and the definition of the set $W^{(K)}$, we have

$$\mathcal{N}_N \bar{W}^{(K-1)} \subset \bar{W}_N^{(K-1)}.$$

Besides, for any $v_* \in (I - P_N) F \bar{W}^{(K-1)}$, we can take $w \in \bar{W}^{(K-1)}$ such that

$$v_* = ((I - P_N^{(1)}) \mathcal{P}^{-1} (\Delta^2)^{-1} h_1(w), (I - P_N^{(2)})(-\Delta)^{-1} h_2(w)),$$

where we have used the same notation I both for the identity map on X^3 and Y^1. By Theorem 1 and (17), we get

$$\begin{aligned} \|(I - P_N^{(1)}) \mathcal{P}^{-1} (\Delta^2)^{-1} h_1(w)\|_\Psi &\le C_i \mathcal{P}^{-1} \|h_1(w)\|_{L^2} \\ &\le C_i \bar{W}_{M+1}^{(K-1)} \qquad (1 \le i \le 10), \\ \|(I - P_N^{(2)})(-\Delta)^{-1} h_2(w)\|_\Theta &\le C_i \|h_2(w)\|_{L^2} \\ &\le C_i \bar{W}_{M+2}^{(K-1)} \qquad (11 \le i \le 15). \end{aligned}$$

Hence, we obtain $v_* \in \bar{W}_*{}^{(K-1)}$, and thus

$$(I - P_N) F \bar{W}^{(K-1)} \subset \bar{W}_*{}^{(K-1)}$$

holds. $\square$

By virtue of the Newton-like operator $\mathcal{N}_N$ and the constructive a priori error estimate of Theorem 1, the above iteration process should be successful as the truncation numbers: M_1, M_2, N_1 and N_2 become large.

5.1. Computation of $W_N^{(k)}$

In order to decide the finite dimensional set $W_N^{(k)}$ in (17), it is needed to compute $\mathcal{N}_N W$, for a candidate set $W = W_N \oplus W_*$ which is generated from a positive vector $\{W_i\}_{1\le i\le M+2}$. We consider some detailed computational procedures below.

For any fixed $w \in W$ we have

$$\begin{aligned}\mathcal{N}_N w &= [I - P_N\mathcal{K}f'(\hat{u}_N)]_N^{-1}\,(\,[I - P_N\mathcal{K}f'(\hat{u}_N)]_N P_N w - P_N(I-F)w)\\ &= [I - P_N\mathcal{K}f'(\hat{u}_N)]_N^{-1} P_N\mathcal{K}\,(\,h(w) - f'(\hat{u}_N)P_N w\,)\,.\end{aligned}$$

Then we need to get an enclosure of $[I - P_N\mathcal{K}f'(\hat{u}_N)]_N^{-1} P_N\mathcal{K}(h(w) - f'(\hat{u}_N)P_N w)$ for any $w \in W$. By setting

$$(v_N^{(1)}, v_N^{(2)}) := \mathcal{N}_N w = (\sum_{i=1}^{L_1} a_i\psi_i, \sum_{i=1}^{L_2} b_i\theta_i),$$

taking account of

$$P_N(I - P_N\mathcal{K}f'(\hat{u}_N))\mathcal{N}_N w = P_N\mathcal{K}(\,h(w) - f'(\hat{u}_N)P_N w\,)$$

and the definition of the projection P_N, we have

$$\begin{cases}(\mathcal{P}\Delta^2 v_N^{(1)} - f_1{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)}), \psi_i)_{L^2} = (h_1(w) - f_1{}'(\hat{u}_N)P_N w, \psi_i)_{L^2} \\ \qquad\qquad (1 \le i \le L_1), \\ (-\Delta v_N^{(2)} - f_2{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)}), \theta_i)_{L^2} = (h_2(w) - f_2{}'(\hat{u}_N)P_N w, \theta_i)_{L^2} \\ \qquad\qquad (1 \le i \le L_2).\end{cases} \tag{18}$$

Here, let us introduce a Jacobian operator by

$$J(u,v) := u_x v_z - v_x u_z = -J(v,u).$$

Since

$$\begin{aligned}f_1{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)}) &= J(v_N^{(1)}, \Delta\hat{\Psi}_N) + J(\hat{\Psi}_N, \Delta v_N^{(1)}) + \sqrt{\mathcal{P}R}(v_N^{(2)})_x,\\ f_2{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)}) &= -\sqrt{\mathcal{P}R}(v_N^{(1)})_x + J(\hat{\Theta}_N, v_N^{(1)}) + J(v_N^{(2)}, \hat{\Psi}_N),\end{aligned}$$

we get

$$\begin{aligned}&\mathcal{P}\Delta^2 v_N^{(1)} - f_1{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)})\\ &\quad= \sum_{j=1}^{L_1}\Big(\mathcal{P}\Delta^2\psi_j + J(\Delta\hat{\Psi}_N, \psi_j) + J(\Delta\psi_j, \hat{\Psi}_N)\Big)\, a_j - \sqrt{\mathcal{P}R}\sum_{j=1}^{L_2} b_j(\theta_j)_x,\\ &-\Delta v_N^{(2)} - f_2{}'(\hat{u}_N)(v_N^{(1)}, v_N^{(2)})\\ &\quad= \sqrt{\mathcal{P}R}\sum_{j=1}^{L_1}\Big((\psi_j)_x + J(\psi_j, \hat{\Theta}_N)\Big)\, a_j + \sum_{j=1}^{L_2}\Big(-\Delta\theta_j + J(\hat{\Psi}_N, \theta_j)\Big)\, b_j.\end{aligned}$$

Therefore, define the $L_i \times L_j$ matrix G_{ij}, for $1 \le i, j \le 2$, and the $M \times M$ matrix G by

$$\begin{aligned}
(G_{11})_{ij} &= (\mathcal{P}\Delta^2\psi_j + J(\Delta\hat{\Psi}_N, \psi_j) + J(\Delta\psi_j, \hat{\Psi}_N), \psi_i)_{L^2},\\
(G_{12})_{ij} &= -\sqrt{\mathcal{P}R}((\theta_j)_x, \psi_i)_{L^2},\\
(G_{21})_{ij} &= \sqrt{\mathcal{P}R}((\psi_j)_x + J(\psi_j, \hat{\Theta}_N), \theta_i)_{L^2},\\
(G_{22})_{ij} &= (-\Delta\theta_j + J(\hat{\Psi}_N, \theta_j), \theta_i)_{L^2},\\
G &= \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix},
\end{aligned}$$

also define the $M \times 1$ vector $\boldsymbol{d}$ by

$$d_i = (h_1(w) - f_1{}'(\hat{u}_N)P_N w, \psi_i)_{L^2} \quad (1 \le i \le L_1), \tag{19}$$

$$d_{L_1+i} = (h_2(w) - f_2{}'(\hat{u}_N)P_N w, \theta_i)_{L^2} \quad (1 \le i \le L_2), \tag{20}$$

and let $\boldsymbol{f} := [a_1, \ldots, a_{L_1}, b_1, \ldots, b_{L_2}]$. Then $\mathcal{N}_N w$ is obtained by solving the linear equation:

$$G\boldsymbol{f} = \boldsymbol{d}.$$

On the verified solution of the above problem, which assures also the invertibility of the matrix G, i.e., the the existence of the inverse (14), we used the method in Rump [16]. In order to take account of the rounding error in computer, the matrix G is generated as an *interval matrix* based upon an approximate solution $\hat{u}_N = (\hat{\Psi}_N, \hat{\Theta}_N)$, given numbers $\mathcal{P}$ and $\mathcal{R}$, and the base functions $\{\psi_i\}_{i=1}^{L_1}$ and $\{\theta_i\}_{i=1}^{L_2}$ using *interval arithmetic* [1]. On the other hand, the element of the vector $\boldsymbol{d}$ can also be evaluated as an *interval vector* which is the upper and lower bound of L^2-inner products of (19) and (20) for all $w \in W$ described in the next subsection.

5.2. Computation of $\boldsymbol{d}$

We define the candidate set W of the form

$$W = (w_N^{(1)} + \alpha^{(1)}, w_N^{(2)} + \alpha^{(2)}),$$

where $(w_N^{(1)}, w_N^{(2)}) \subset S_N$, $(\alpha^{(1)}, \alpha^{(2)}) \subset S_N^{\perp}$. Moreover, $w_N^{(1)}$ and $w_N^{(2)}$ are supposed to be represented by

$$w_N^{(1)} = \sum_{i=1}^{L_1} A_i\psi_i, \qquad w_N^{(2)} = \sum_{i=1}^{L_2} B_i\theta_i,$$

where $A_i := [\,-W_i, W_i\,]$ and $B_i := [\,-W_{L_1+i}, W_{L_1+i}\,]$ are real intervals. The explicit forms of $\alpha^{(1)}$ and $\alpha^{(2)}$ are unknown, however, these norms and seminorms can be estimated as below. We get

$$h_1(w) - {f_1}'(\hat{u}_N)P_N w = f_1(\hat{\Psi}_N, \hat{\Theta}_N) - \mathcal{P}\Delta^2\hat{\Psi}_N + J(w_N^{(1)}, \Delta w_N^{(1)}) + \\ J(\hat{\Psi}_N + w_N^{(1)}, \Delta\alpha^{(1)}) + J(\alpha^{(1)}, \Delta(\hat{\Psi}_N + w_N^{(1)})) + J(\alpha^{(1)}, \Delta\alpha^{(1)}) + \sqrt{\mathcal{P}R}(\alpha^{(2)})_x,$$

$$h_2(w) - {f_2}'(\hat{u}_N)P_N w = f_2(\hat{\Psi}_N, \hat{\Theta}_N) + \Delta\hat{\Theta}_N + J(w_N^{(2)}, w_N^{(1)}) + \\ J(\alpha^{(2)}, \hat{\Psi}_N + w_N^{(1)}) + J(\hat{\Theta}_N + w_N^{(2)}, \alpha^{(1)}) + J(\alpha^{(2)}, \alpha^{(1)}) - \sqrt{\mathcal{P}R}(\alpha^{(1)})_x.$$

In (19) and (20), the L^2-inner products

$$(f_1(\hat{\Psi}_N, \hat{\Theta}_N) - \mathcal{P}\Delta^2\hat{\Psi}_N + J(w_N^{(1)}, \Delta w_N^{(1)}), \psi_i)_{L^2}$$

and

$$(f_2(\hat{\Psi}_N, \hat{\Theta}_N) + \Delta\hat{\Theta}_N + J(w_N^{(2)}, w_N^{(1)}), \theta_i)_{L^2}$$

include no infinite dimensional quantities, thus these values can be estimated by direct computations using usual interval arithmetic. The L^2-inner products

$$(J(\hat{\Psi}_N + w_N^{(1)}, \Delta\alpha^{(1)}), \psi_i)_{L^2}, \quad (J(\alpha^{(1)}, \Delta(\hat{\Psi}_N + w_N^{(1)}), \psi_i)_{L^2}, \quad ((\alpha^{(2)})_x, \psi_i)_{L^2},$$
$$(J(\alpha^{(2)}, \hat{\Psi}_N + w_N^{(1)}), \theta_i)_{L^2}, \quad (J(\hat{\Theta}_N + w_N^{(2)}, \alpha^{(1)}), \theta_i)_{L^2}, \quad ((\alpha^{(1)})_x, \theta_i)_{L^2}$$

are evaluated using the a priori constants of Theorem 1. Let us show one example. In what follows, we use the canonical basis of trigonometric functions for S_N. The inner product

$$((\hat{\Psi}_N + w_N^{(1)})_x \Delta(\alpha^{(1)})_z, \psi_i)_{L^2},$$

can be rewritten as

$$(\hat{\Psi}_N + w_N^{(1)})_x \psi_i = \sum_{m=1}^{2M_1}\sum_{n=0}^{2N_1} \hat{Z}_{mn} \sin(amx)\cos(nz)$$

with some interval coefficients $\hat{Z}_{mn}$. Also note that the coefficients of $\Delta(\alpha^{(1)})_z$ are zero when $1 \le m \le M_1$ and $1 \le n \le N_1$. Hence, we have

$$\begin{aligned}
&((\hat{\Psi}_N + w_N^{(1)})_x \Delta(\alpha^{(1)})_z, \psi_i)_{L^2} \\
&\subset [-1,1]\ \|\Delta(\alpha^{(1)})_z\|_{L^2} \times \\
&\quad \|\sum_{m=1}^{2M_1}\sum_{n=N_1+1}^{2N_1} \hat{Z}_{mn}\sin(amx)\cos(nz) + \sum_{m=M_1+1}^{2M_1}\sum_{n=0}^{2N_1} \hat{Z}_{mn}\sin(amx)\cos(nz)\|_{L^2} \\
&\subset [-1,1]\ C_{10} W_{M+1} \frac{\pi}{\sqrt{2a}} \left(\sum_{m=1}^{2M_1}\sum_{n=N_1+1}^{2N_1} \hat{Z}_{mn}^2 + \sum_{m=M_1+1}^{2M_1}\sum_{n=0}^{2N_1} \hat{Z}_{mn}^2 \right)^{1/2}.
\end{aligned}$$

The remaining L^2-inner products

$$(J(\alpha^{(1)}, \Delta\alpha^{(1)}), \psi_i)_{L^2}, \quad (J(\alpha^{(2)}, \alpha^{(1)}), \theta_i)_{L^2}$$

are similarly evaluated, for example,

$$\begin{aligned}
((\alpha^{(1)})_z \Delta(\alpha^{(1)})_x, \psi_i)_{L^2} &\subset [-1,1]\ \|(\alpha^{(1)})_z\|_{L^2} \|\Delta(\alpha^{(1)})_x\|_{L^2} \\
&\subset [-1,1]\ C_6 C_9 W_{M+1}^2.
\end{aligned}$$

5.3. Computation of $W_{M+1}^{(k)}$ and $W_{M+2}^{(k)}$

In order to describe the computation of $W_{M+1}^{(k)}$ and $W_{M+2}^{(k)}$ in (17), we will show how to estimate

$$\sup_{w\in W}\|h_1(w)\|_{L^2}, \qquad \sup_{w\in W}\|h_2(w)\|_{L^2}$$

for a candidate set $W = (w_N^{(1)} + \alpha^{(1)}, w_N^{(2)} + \alpha^{(2)})$. Since

$$\begin{aligned} h_1(W) \;=&\; f_1(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) - \mathcal{P}\Delta^2\hat{\Psi}_N + J(\hat{\Psi}_N + w_N^{(1)}, \Delta\alpha^{(1)}) \\ &+\; J(\alpha^{(1)}, \Delta(\hat{\Psi}_N + w_N^{(1)})) + J(\alpha^{(1)}, \Delta\alpha^{(1)}) + \sqrt{\mathcal{P}R}(\alpha^{(2)})_x, \end{aligned}$$

we have

$$\begin{aligned} \|h_1(W)\|_{L^2} \;\le&\; \|f_1(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) - \mathcal{P}\Delta^2\hat{\Psi}_N\|_{L^2} \\ &+\|(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty\|\Delta(\alpha^{(1)})_x\|_{L^2} \\ &+\|\Delta(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty\|(\alpha^{(1)})_z\|_{L^2} \\ &+\|(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty\|\Delta(\alpha^{(1)})_z\|_{L^2} \\ &+\|\Delta(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty\|(\alpha^{(1)})_x\|_{L^2} + \|(\alpha^{(1)})_z\|_\infty\|\Delta(\alpha^{(1)})_x\|_{L^2} \\ &+\|(\alpha^{(1)})_x\|_\infty\|\Delta(\alpha^{(1)})_z\|_{L^2} + \sqrt{\mathcal{P}R}\|(\alpha^{(2)})_x\|_{L^2} \\ \le&\; \|f_1(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) - \mathcal{P}\Delta^2\hat{\Psi}_N\|_{L^2} \\ &+(C_9\|(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty + C_6\|\Delta(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty \\ &+C_{10}\|(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty + C_5\|\Delta(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty)W_{M+1} \\ &+(C_{19}C_9 + C_{18}C_{10})W_{M+1}^2 + \sqrt{\mathcal{P}R}C_{14}W_{M+2}. \end{aligned}$$

Also, since

$$\begin{aligned} h_2(W) \;=&\; f_2(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) + \Delta\hat{\Theta}_N + J(\alpha^{(2)}, \hat{\Psi}_N + w_N^{(1)}) \\ &+\; J(\hat{\Theta}_N + w_N^{(2)}, \alpha^{(1)}) + J(\alpha^{(2)}, \alpha^{(1)}) - \sqrt{\mathcal{P}R}(\alpha^{(1)})_x, \end{aligned}$$

we get

$$\begin{aligned} \|h_2(W)\|_{L^2} \;\le&\; \|f_2(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) + \Delta\hat{\Theta}_N\|_{L^2} \\ &+\|(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty\|(\alpha^{(2)})_x\|_{L^2} \\ &+\|(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty\|(\alpha^{(2)})_z\|_{L^2} + \|(\hat{\Theta}_N + w_N^{(2)})_x\|_\infty\|(\alpha^{(1)})_z\|_{L^2} \\ &+\|(\hat{\Theta}_N + w_N^{(2)})_z\|_\infty\|(\alpha^{(1)})_x\|_{L^2} + \|(\alpha^{(1)})_z\|_\infty\|(\alpha^{(2)})_x\|_{L^2} \\ &+\|(\alpha^{(1)})_x\|_\infty\|(\alpha^{(2)})_z\|_{L^2} + \sqrt{\mathcal{P}R}\|(\alpha^{(1)})_x\|_{L^2} \\ \le&\; \|f_2(w_N^{(1)} + \hat{\Psi}_N, w_N^{(2)} + \hat{\Theta}_N) + \Delta\hat{\Theta}_N\|_{L^2} \\ &+(\|(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty C_{14} + \|(\hat{\Psi}_N + w_N^{(1)})_x\|_\infty C_{15})W_{M+2} \end{aligned}$$

$$
\begin{aligned}
&+(\|(\hat{\Theta}_N + w_N^{(2)})_x\|_\infty C_6 + \|(\hat{\Theta}_N + w_N^{(2)})_z\|_\infty C_5 + \sqrt{\mathcal{P}\mathcal{R}} C_5) W_{M+1} \\
&+(C_{19}C_{14} + C_{18}C_{15}) W_{M+1} W_{M+2}.
\end{aligned}
$$

In the above estimations, upper bounds of each L^2-norm and L^∞-norm can be computed by interval arithmetical approaches. For example, if $(\hat{\Psi}_N + w_N^{(1)})_z$ is expressed as

$$
(\hat{\Psi}_N + w_N^{(1)})_z = \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} n Z_{mn} \sin(amx)\cos(nz),
$$

by some interval coefficients Z_{mn}, then we use the following estimates:

$$
\|(\hat{\Psi}_N + w_N^{(1)})_z\|_\infty \leq \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} |n Z_{mn}|.
$$

Note that these seem to be overestimates, but usually, by virtue of the small multipliers C_i, there are no serious enlargements of computed quantities.

6. Numerical examples

In the verification step, interval arithmetic is used to take account of the effects of rounding errors in the floating point computations. We use Fortran 90 library `INTLIB_90` coded by Kearfott [9] with DIGITAL Fortran V5.2-705 on an Alpha Server XP1000 (CPU:Alpha 2126 500MHz, OS: Tru64 UNIX 4.0E).

6.1. First bifurcated solutions from the trivial solution

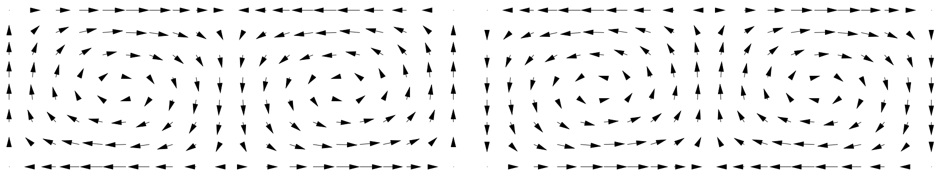
In 1916, Rayleigh [15] considered the linearized stability and found the critical Rayleigh number as follows

$$
\mathcal{R}_C = \inf_{m,n} \frac{(a^2m^2 + n^2)^3}{a^2m^2} = 6.75 \quad (m = 1, n = 1, a = 1/\sqrt{2}).
$$

The usual bifurcation theory implies that the stationary bifurcation occurs from the above critical point. We select $a = 1/\sqrt{2}$ and $\mathcal{P} = 10$ in the following numerical experiments. After the critical Rayleigh number $\mathcal{R}_C = 6.75$, we obtain two non-trivial approximate solutions for various Rayleigh numbers $\mathcal{R}$ of the form:

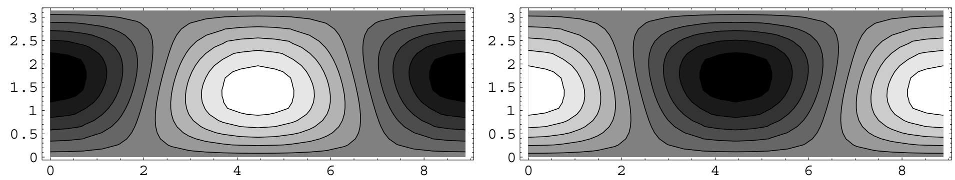
$$
\hat{\Psi}_N = \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} \hat{A}_{mn} \sin(amx)\sin(nz), \qquad \hat{\Theta}_N = \sum_{m=0}^{M_2} \sum_{n=1}^{N_2} \hat{B}_{mn} \cos(amx)\sin(nz)
$$

for some M_1, M_2, N_1 and N_2 by Fourier-Galerkin method combined with Newton-Raphson iteration. Fig. 2 and Fig. 3 show the velocity field $(-(\hat{\Psi}_N)_z, (\hat{\Psi}_N)_x)$ and the isotherm of $\hat{\Theta}_N$ which means the deviation of the temperature from the linear profile at $\mathcal{R} = 7, M_1 = N_1 = M_2 = N_2 = 10$, respectively. We indicate the particular values of coefficients, under the figures, which have the maximum absolute values of $\{\hat{A}_{mn}\}$ and $\{\hat{B}_{mn}\}$, respectively.



$\hat{A}_{11} \approx 0.9456$ $\hat{A}_{11} \approx -0.9456$

FIG. 2. The velocity field of the first bifurcated solutions ($\mathcal{R} = 7$)



$\hat{B}_{11} \approx -3.5966$ $\hat{B}_{11} \approx 3.5966$

FIG. 3. Isotherms of the deviation of the temperature from the linear profile ($\mathcal{R} = 7$)

We verify two exact solutions of (4) corresponding to several Rayleigh numbers in the set

$$(\hat{\Psi}_N + w_N^{(1)} + w_*^{(1)}, \hat{\Theta}_N + w_N^{(2)} + w_*^{(2)}) \subset X^3 \times Y^1,$$

where $(w_N^{(1)}, w_N^{(2)}) \subset S_N$ and $(w_*^{(1)}, w_*^{(2)}) \subset S_N^\perp$ whose norms can be numerically estimated, for example, as in Table 1. 'Step' means the iteration number of Theorem 3, and each values of norms rounded-up at last mantissa unit.

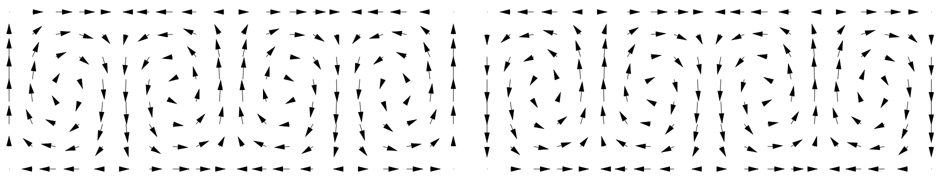
$\mathcal{R}$	step	$\Vert\hat{\Psi}_N\Vert_{L^2}$	$\Vert\hat{\Theta}_N\Vert_{L^2}$	$\Vert W_N^{(1)}\Vert_{L^\infty}$	$\Vert W_N^{(2)}\Vert_{L^\infty}$	$\Vert W_*^{(1)}\Vert_{H^3}$	$\Vert W_*^{(2)}\Vert_{H^1}$
8	3	2.14	7.75	1.24×10^{-11}	1.43×10^{-12}	2.09×10^{-12}	4.28×10^{-12}
10	3	3.53	11.70	3.10×10^{-11}	2.40×10^{-12}	6.86×10^{-12}	1.76×10^{-11}
15	3	5.86	16.83	1.01×10^{-10}	5.02×10^{-12}	2.88×10^{-11}	9.57×10^{-11}
20	3	7.67	20.08	1.97×10^{-10}	7.80×10^{-12}	6.69×10^{-11}	2.55×10^{-10}
25	5	9.23	22.67	1.82×10^{-10}	1.15×10^{-11}	5.17×10^{-11}	2.22×10^{-10}
30	5	10.65	24.89	2.93×10^{-10}	1.08×10^{-11}	1.12×10^{-10}	4.78×10^{-10}
35	5	11.96	26.88	4.89×10^{-10}	1.06×10^{-11}	2.15×10^{-10}	9.51×10^{-10}
40	5	13.18	28.70	9.05×10^{-10}	1.80×10^{-11}	4.21×10^{-10}	2.09×10^{-9}

TABLE 1. Verification results ($M_1 = N_1 = M_2 = N_2 = 35$)

6.2. Second bifurcated solutions from the trivial solution

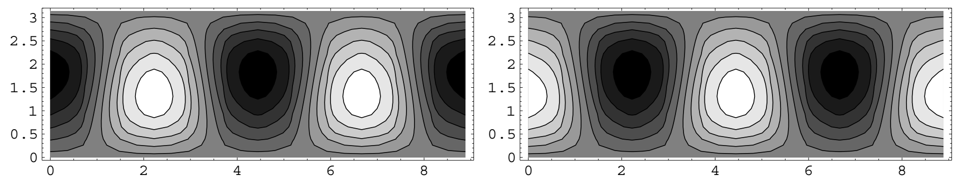
In the case $a = 1/\sqrt{2}$, it seems that at Rayleigh number $\mathcal{R} = 13.5$, a second bifurcation from the trivial solution occurs. Fig. 4 and Fig. 5 show the velocity field

$(-(\hat{\Psi}_N)_z, (\hat{\Psi}_N)_x)$ and the isotherms of $\hat{\Theta}_N$ at $\mathcal{R} = 14, M_1 = N_1 = M_2 = N_2 = 10$, respectively.



$\hat{A}_{21} \approx 0.6705$ $\hat{A}_{21} \approx -0.6705$

FIG. 4. The velocity field of the second bifurcated solutions ($\mathcal{R} = 14$)



$\hat{B}_{21} \approx -1.1405$ $\hat{B}_{21} \approx 3.6068$

FIG. 5. Isotherms of the deviation of the temperature from the linear profile

We verify the two exact solutions of (4) for these second bifurcated solutions as in Table 2.

$\mathcal{R}$	step	$\Vert\hat{\Psi}_N\Vert_{L^2}$	$\Vert\hat{\Theta}_N\Vert_{L^2}$	$\Vert W_N^{(1)}\Vert_{L^\infty}$	$\Vert W_N^{(2)}\Vert_{L^\infty}$	$\Vert W_*^{(1)}\Vert_{H^3}$	$\Vert W_*^{(2)}\Vert_{H^1}$
13.75	3	0.48	2.58	4.86×10^{-12}	4.99×10^{-13}	9.21×10^{-13}	1.26×10^{-12}
15	3	1.18	6.23	1.77×10^{-11}	8.45×10^{-13}	3.98×10^{-12}	4.42×10^{-12}
20	3	2.57	12.52	8.88×10^{-11}	2.03×10^{-12}	2.43×10^{-11}	3.55×10^{-11}
25	3	3.55	16.27	2.03×10^{-10}	3.87×10^{-12}	5.85×10^{-11}	1.04×10^{-10}
30	3	4.37	19.17	3.32×10^{-10}	4.29×10^{-12}	1.08×10^{-10}	2.06×10^{-10}
35	5	5.12	21.59	1.58×10^{-10}	5.77×10^{-12}	4.14×10^{-11}	1.02×10^{-10}
40	5	5.80	23.71	2.63×10^{-10}	5.56×10^{-12}	8.96×10^{-11}	2.85×10^{-10}

TABLE 2. Verification results ($M_1 = N_1 = M_2 = N_2 = 35$)

From these results, for example, at $\mathcal{R} = 20$, by using some L^∞-norm estimates in Corollary 1, we can be assured that the four nontrivial solutions are indeed different. Fig. 6 illustrates the bifurcation curves obtained by our scheme. The vertical axis shows the value of the coefficient of the approximate solution: $\hat{\Psi}_N = \sum_{m=1}^{M_1} \sum_{n=1}^{N_1} \hat{A}_{mn} \sin(amx)\sin(nz)$ where the absolute value is equal to $\max\{|\hat{A}_{mn}|;\ 1 \le m \le M_1,\ 1 \le n \le N_1\}$. Each dot implies that the verification procedure succeeded.

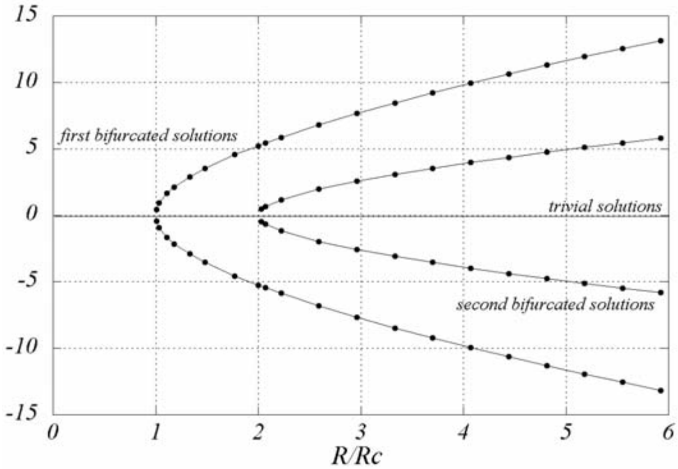


FIG. 6. The bifurcation curves

We actually verified numerically all solutions on both the upper and lower branches indicated in Fig. 6. However, by consideration of the symmetries of problem (4), the existence of solutions corresponding to the lower branches would be naturally concluded from the verification of the upper branch solutions. We cannot say for certain whether the verified solutions are *bifurcated* solutions, or depend continuously on the Rayleigh number, or are locally unique in the candidate sets. These questions must be solved in our future work.

Acknowledgements. We thank the referees for their valuable comments to refine our manuscript.

References

[1] G. ALEFELD AND J. HERZBERGER (translated by J. Rokne), *Introduction to Interval Computations*, Academic Press, New York, 1983.

[2] H. BÉNARD, Les tourbillons cellulaires dans une nappe liquide, *Revue Gén. Sci. Pure Appl.* **11** (1900), 1261–1271, 1309–1328.

[3] S. CHANDRASEKHAR, *Hydrodynamic and Hydromagnetic Stability*, Oxford University Press, 1961.

[4] A. V. GETLING, *Rayleigh–Bénard Convection: structures and dynamics*, World Scientific, Singapore, 1998.

[5] J. G. HEYWOOD, W. NAGATA AND W. XIE, A numerically based existence theorem for the Navier–Stokes equations, *J. Math. Fluid Mech.* **1** (1999), 5–23.

[6] V. I. IUDOVICH, On the origin of convection, *J. Appl. Math. Mech.* **30** (1966), 1193–1199.

[7] D. D. JOSEPH, On the stability of the Boussinesq equations, *Arch. Rational Mech. Anal.* **20** (1965), 59–71.

[8] Y. KAGEI AND W. VON WAHL, The Eckhaus criterion for convection roll solutions of the Oberbeck–Boussinesq equations, *Int. J. Non-Linear Mechanics* **32** (1997), 563–620.

[9] R. B. Kearfott and V. Kreinovich (eds.), *Applications of Interval Computations*, Kluwer Academic Publishers, Netherlands, 1996. (`http://interval.usl.edu/kearfott.html`)
[10] R. Krishnamurti, Some further studies on the transition to turbulent convection, *J. Fluid Mech.* **60** (1973), 285–303.
[11] M. T. Nakao, A numerical verification method for the existence of weak solutions for nonlinear boundary value problems, *J. Math. Anal. Appl.* **164** (1992), 489–507.
[12] M. T. Nakao, Solving nonlinear elliptic problems with result verification using an H^{-1} type residual iteration, *Computing, Suppl.* **9** (1993), 161–173.
[13] M. Plum, Explicit H_2-estimates and pointwise bounds for solutions of second-order elliptic boundary value problems, *J. Math. Anal. Appl.* **165** (1992), 36–61.
[14] P. H. Rabinowitz, Existence and nonuniqueness of rectangular solutions of the Bénard problem, *Arch. Rational Mech. Anal.* **29** (1968), 32–57.
[15] J. W. S. Rayleigh, On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side, *Phil. Mag.* **32** (1916), 529–546; and *Sci. Papers* **6**, 432–446.
[16] S. M. Rump, On the solution of interval linear systems, *Computing* **47** (1992), 337–353.
[17] E. Zeidler (translated by Q. Juergen), *Nonlinear Functional Analysis and It's Applications IV: Application to Mathematical Physics*, Springer-Verlag, New York, 1988.

Y. Watanabe
Kyushu University
Computing and Communications Center
Fukuoka 812-8581
Japan
e-mail: watanabe@cc.kyushu-u.ac.jp

N. Yamamoto
The University of Electro-Communications
Department of Computer Science
and Information Mathematics
Chofu 182-8585
Japan
e-mail: yamamoto@masuo.im.uec.ac.jp

M. T. Nakao
Kyushu University
Graduate School of Mathematics
Fukuoka 812-8581
Japan
e-mail: mtnakao@math.kyushu-u.ac.jp

T. Nishida
Kyoto University
Graduate School of Science
Kyoto 606-8502
Japan
e-mail: a50258@sakura.kudpc.kyoto-u.ac.jp

(accepted: June 20, 2002)