

Boundary conditions and linear analysis of finite-cell Rayleigh–Bénard convection

By YIH-YUH CHEN†

Condensed Matter Physics 114-36, California Institute of Technology, Pasadena,
CA 91125, USA

(Received 18 January 1991 and in revised form 24 January 1992)

The linear stability of finite-cell pure-fluid Rayleigh–Bénard convection subject to any homogeneous viscous and/or thermal boundary conditions is investigated via a variational formalism and a perturbative approach. Some general properties of the critical Rayleigh number with respect to change of boundary conditions or system size are derived. It is shown that the chemical reaction–diffusion model of spatial-pattern-forming systems in developmental biology can be thought of as a special case of the convection problem. We also prove that, as a result of the imposed realistic boundary conditions, the nodal surfaces of the temperature of a nonlinear stationary state have a tendency to be parallel or orthogonal to the sidewalls, because the full fluid equations become linear close to the boundary, thus suggesting similar trend for the experimentally observed convective rolls.

1. Introduction

Rayleigh–Bénard convection has received much attention in recent years because of the rich dynamics and patterns it displays in experimentally well-controlled situations (Croquette 1989*a, b*; Niemela, Ahlers & Cannel 1990; Kolodner, Bensimon & Surko 1988; Steinberg, Moses & Feinberg 1987), and because a systematic theoretical treatment in terms of the amplitude equation seems capable of explaining many of the observed features quite successfully (see, for example, Segel 1967; Newell & Whitehead 1969; Cross 1988). Although our understanding of the problem is admittedly complete only if nonlinearity is included, it is also important to have a thorough understanding of the linear stability of the system when the convection cell is finite. This is particularly desirable because recent experiments (Zhong, Ecke & Steinberg 1991; Pfothenauer, Niemela & Donnelly 1987) and numerical calculations (Buell & Catton 1983) have shown that the finite-size effects on the onset stability of a rotating convection cell are very different from those expected for a non-rotating one. It seems reasonable to assume that by isolating the influence of the sidewall on its stability we can gain a better understanding of how rotation affects this system. The major philosophy of this work is, therefore, to study how the linear stability of pure-fluid convection is affected by the presence of the sidewall, and to derive the results on a firmer basis.

In this paper we shall concentrate on the following questions. How do various boundary conditions affect the linear stability of a finite Rayleigh–Bénard cell which has an arbitrary horizontal cross-section? What is the influence of the cell size on the

† Present address: Institute of Physics, Academia Sinica, Nankang, Taipei, Taiwan, ROC.

stability? How do convective rolls behave near the sidewalls? Specifically, we organize the presentation as follows. In §2 we discuss two variational principles: one for the growth rate of the linear stability of the system, and the other for the onset Rayleigh number R_c . Several results concerning the effects of the boundary conditions on the linear stability are then derived as immediate consequences of this variational formalism. Some of them will also serve as preparatory work for §3, which treats the convection problem mainly from the point of view of perturbation theory. There, the size dependence and related issues are studied, and the results derived are compared with the features exhibited by the chemical reaction–diffusion model of pattern formation in developmental biology. We show that the spatial pattern-generation mechanisms in both cases are of a very similar nature, and point out that in the convection problem the incompressibility condition is essential in making the two systems similar. We also show that the critical Rayleigh number differs from its infinite-cell limit by an amount that scales like L^{-2} of the system size L . In §4, we address one aspect of defining the angle between the convective rolls and the sidewalls. In particular, we show that the nodal surfaces of the temperature of both the nonlinear stationary and the linearized solutions tend to be orthogonal to the sidewall, a result that is linear in nature in spite of the fact that the full nonlinear fluid equation is used in the derivation.

2. Variational principles and consequences

Several versions of the variational principle for the linear stability of pure-fluid convection in a laterally infinite cell were derived by Pellew & Southwell (1940) and Chandrasekhar (1954, 1961). A generalized version that applies to a finite cell of arbitrary horizontal geometry which satisfies certain boundary conditions was developed by Sorokin (1953, see also Sorokin & Sushkin 1960), and later rigorously justified independently by Sani (1964) and Pnueli & Iscovici (1967). Joseph (1976) has also discussed several aspects pertinent to the variational formalism in fluid dynamics. In practice, it is desirable to have a variational expression for both the growth rate and the critical Rayleigh number, because it allows one to compute the quantities of interest to high precision with relative ease. And indeed this has already been done for a cylindrical cell (Charlson & Sani 1970, 1971; Shaumeyer, Behringer & Baierlein 1981) and for a rectangular box (Edwards 1988; Davis 1967). On the other hand, one can also employ the variational character of this system to study the analytic properties of the finite-size effects both qualitatively and quantitatively. This approach is favoured over the regular perturbation method because usually it is easier to justify the results one derives whenever a variational formalism does exist. However, one should also keep in mind that there are situations in which a direct application of the variational principle does not provide enough information on what one wishes to learn about the system, in which case perturbation theory will turn out more useful. In this contribution we will not only apply the variational principles discussed below to facilitate the argument, but also supplement it by perturbation theory to illustrate how the combined force of both can yield interesting conclusions. In this section we will first give a quick derivation of two variational principles that apply to a finite cell which satisfies homogeneous boundary conditions, briefly comment on how they can be modified to incorporate a realistic sidewall, then prove several monotonicity results concerning the effects of the sidewall properties on the stability of the system. Some of these results will then be used in §3 when we study the dependence of the stability on the system size.

The equations that describe the motion under gravity g of a layer of pure fluid of thermal diffusivity κ , kinematic viscosity ν , and coefficient of thermal expansion γ heated from below by a temperature difference ΔT across the top and the bottom plates of a convection cell of height d are (i) the Navier–Stokes equation with the Oberbeck–Boussinesq approximation, and (ii) the heat transfer equation. They are usually written in the following non-dimensional form (see, for example, Cross 1980):

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} = \sigma \nabla^2 \mathbf{u} + \sigma \theta \hat{\mathbf{e}}_z - \nabla p, \tag{2.1}$$

$$\frac{\partial \theta}{\partial t} + (\mathbf{u} \cdot \nabla) \theta = \nabla^2 \theta + Ru_z \quad (u_z \equiv \mathbf{u} \cdot \hat{\mathbf{e}}_z), \tag{2.2}$$

$$\nabla \cdot \mathbf{u} = 0 \quad (\text{incompressibility condition}), \tag{2.3}$$

where \mathbf{u} and θ are respectively the deviations of velocity and temperature from the pure-conduction profile, and length, time and temperature are scaled by d , d^2/κ , and $\kappa\nu/\gamma g d^3$, respectively. In the above, σ is the Prandtl number

$$\sigma = \nu/\kappa, \tag{2.4}$$

and R is the Rayleigh number

$$R = \gamma g d^3 \Delta T / (\kappa \nu). \tag{2.5}$$

But to analyse the linear stability of this problem, it turns out to be more convenient to rescale θ by $\theta \rightarrow (R/\sigma)^{\frac{1}{2}} \theta$ so that the linearized equations read

$$\sigma \nabla^2 \mathbf{u} + (\sigma R)^{\frac{1}{2}} \theta \hat{\mathbf{e}}_z - \nabla p = \frac{\partial \mathbf{u}}{\partial t} = \lambda \mathbf{u}, \tag{2.6}$$

$$\nabla^2 \theta + (\sigma R)^{\frac{1}{2}} u_z = \frac{\partial \theta}{\partial t} = \lambda \theta, \tag{2.7}$$

or in operator form

$$\mathcal{L} \psi = \frac{\partial \psi}{\partial t} = \lambda \psi, \tag{2.8}$$

$$\begin{aligned} \mathcal{L} \psi \equiv & \left\{ \begin{pmatrix} \sigma & 0 & 0 & 0 \\ 0 & \sigma & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \nabla^2 + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & (\sigma R)^{\frac{1}{2}} \\ 0 & 0 & (\sigma R)^{\frac{1}{2}} & 0 \end{pmatrix} \right\} \psi - \begin{pmatrix} \partial p / \partial x \\ \partial p / \partial y \\ \partial p / \partial z \\ 0 \end{pmatrix} \\ & \equiv \{ \hat{\mathbf{D}} \nabla^2 + \hat{\mathbf{A}} \} \psi - \begin{pmatrix} \nabla p \\ 0 \end{pmatrix}, \end{aligned} \tag{2.9}$$

where $\hat{\mathbf{D}}$ and $\hat{\mathbf{A}}$ are 4×4 matrices defined in an obvious way, λ is the eigenvalue of this problem, and

$$\psi \equiv \begin{pmatrix} \mathbf{u} \\ \theta \end{pmatrix}. \tag{2.10}$$

If we define the inner product between ψ and ψ' by

$$\langle \psi' | \psi \rangle \equiv \int_{\Omega} \mathbf{u}'^* \cdot \mathbf{u} + \theta'^* \theta \quad (\Omega \equiv \text{entire cell}), \tag{2.11}$$

where $*$ indicates complex conjugation, then it is easy to show that \mathcal{L} is a self-adjoint

operator, and thus its eigenfunctions can be chosen to be real-valued, as will be done from now on, provided that the boundary conditions for \mathbf{u} and θ are homogeneous in the following sense:

$$u_n \equiv \mathbf{u} \cdot \hat{\mathbf{n}} = 0, \quad (2.12)$$

$$\hat{\mathbf{e}}_{\parallel} \cdot \left(\alpha_1 \frac{\partial \mathbf{u}}{\partial n} + \hat{\boldsymbol{\alpha}} \cdot \mathbf{u} \right) = 0, \quad \hat{\boldsymbol{\alpha}} \equiv \boldsymbol{\alpha} + \hat{\mathbf{C}}, \quad (2.13)$$

$$\beta_1 \frac{\partial \theta}{\partial n} + \beta \theta = 0. \quad (2.14)$$

Here $\hat{\mathbf{n}}$ is the outward normal to the boundary $\partial\Omega$, and $\hat{\mathbf{e}}_{\parallel}$ is any vector parallel to $\partial\Omega$, whereas $\hat{\mathbf{C}} \equiv \nabla \hat{\mathbf{n}}$ so that $\hat{\mathbf{e}}_{\parallel} \cdot \hat{\mathbf{C}}$ is the curvature vector ($\hat{\mathbf{e}}_{\parallel} \cdot \nabla$) $\hat{\mathbf{n}}$ of the boundary along the direction of $\hat{\mathbf{e}}_{\parallel}$. In the above, α , α_1 , β and β_1 are scalars that can vary on $\partial\Omega$, though throughout this paper we shall assume that, for the sake of illustrating the underlying idea, they are constants on each of the two horizontal plates and the sidewall. It is worth pointing out that the case $\alpha = 0$, which we shall call the rotation-free condition in this paper, does not correspond to the free-slip condition when the boundary is curved. (The sign of the curvature term is wrong.) This condition is a natural one when we try to use separation of variables to solve the coupled equations, because it guarantees that the vertical vorticity at the sidewall is identically zero, which is essential if the method is to work for the problem at hand. This issue will be taken up in detail later on and in Appendix C. But before we embark on our investigation, it seems reasonable to ask why one would consider a physically unrealizable boundary condition like (2.13), and to what extent the sidewall thermal property can be simulated by (2.14). Actually, (2.13) is only a convenient mathematical construct which will help us understand how the viscous boundary condition affects the system's stability. Eventually we shall relate the no-slip condition to the rotation-free case by way of this construction, and derive results that are true for a rigid cell. However, this is not entirely the case for (2.14). It so happens that this condition can be realized in some limiting cases, and Busse & Riahi (1980) and Sparrow, Goldstein & Jonsson (1964) have applied this boundary condition on the two horizontal plates in their study of convection problem. But more than that, we can actually write down a variational principle for a convection cell with a finite or infinite realistic sidewall, and show that in a well-defined sense the β in (2.14) does represent the effective thermal conductivity of the sidewall. Suffice it to say that the results reported here can be easily generalized to a real convection cell. The issues related to a cell surrounded by a realistic sidewall are reported elsewhere (Chen 1992). For the rest of this paper we will assume α_1 (β_1) is either zero or unity, as the case may be, and the two horizontal plates are held at $\theta = 0$.

We now state the first variational principle.

Variational Principle I

Define a variational functional

$$I[\boldsymbol{\psi}] \equiv \frac{-\left(\int_{\Omega} \sigma |\nabla \mathbf{u}|^2 + |\nabla \theta|^2 + 2(\sigma R)^{\frac{1}{2}} \int_{\Omega} u_z \theta - \left(\int_{\partial\Omega} \sigma \mathbf{u} \cdot \hat{\boldsymbol{\alpha}} \cdot \mathbf{u} + \beta |\theta|^2\right)\right)}{\int_{\Omega} |\mathbf{u}|^2 + |\theta|^2}, \quad (2.15)$$

where $\boldsymbol{\psi}$ must satisfy $\nabla \cdot \mathbf{u} = 0$ in Ω and $u_n = 0$ on $\partial\Omega$. Then $I[\boldsymbol{\psi}]$ takes on a stationary value if and only if the following are satisfied:

(i) On the boundary $\partial\Omega$, \mathbf{u} and θ satisfy

$$\hat{\mathbf{e}}_1 \cdot \left(\frac{\partial \mathbf{u}}{\partial n} + \hat{\mathbf{a}} \cdot \mathbf{u} \right) = 0, \quad (2.16)$$

$$\frac{\partial \theta}{\partial n} + \beta \theta = 0, \quad (2.17)$$

and (ii)

$$\mathcal{L}\psi = \lambda\psi \quad (2.18)$$

for some constant λ . Thus, ψ is an eigenfunction of \mathcal{L} with (2.16) and (2.17) as the boundary conditions. Also, the eigenvalue λ is precisely given by $I[\psi]$.

Indeed, straightforward algebra shows that

$$\begin{aligned} \delta I[\psi] &\equiv I[\psi + \delta\psi] - I[\psi] \\ &= 2 \left\{ \frac{1}{a^2} \int_{\Omega} \delta \mathbf{u} \cdot (a(\sigma \nabla^2 \mathbf{u} + (\sigma R)^{\frac{1}{2}} \theta \hat{\mathbf{e}}_z) - b\mathbf{u}) \right. \\ &\quad \left. - \frac{1}{a} \int_{\partial\Omega} \sigma \delta \mathbf{u} \cdot \left(\frac{\partial \mathbf{u}}{\partial n} + \hat{\mathbf{a}} \cdot \mathbf{u} \right) + \delta \theta \left(\frac{\partial \theta}{\partial n} + \beta \theta \right) \right\}, \end{aligned}$$

where we have performed integration by parts, and a and b are defined by

$$\begin{aligned} a &\equiv \int_{\Omega} |\mathbf{u}|^2 + |\theta|^2, \\ b &\equiv - \int_{\Omega} \sigma |\nabla \mathbf{u}|^2 + |\theta|^2 + 2(\sigma R)^{\frac{1}{2}} \int_{\Omega} u_z \theta - \int_{\partial\Omega} \sigma \mathbf{u} \cdot \hat{\mathbf{a}} \cdot \mathbf{u} + \beta |\theta|^2. \end{aligned}$$

Variation of $\delta \mathbf{u}$ subject to the constraints $\nabla \cdot \delta \mathbf{u} = 0$ and $\delta u_n = 0$ then immediately yields both the linearized Navier-Stokes equation (2.6) and the appropriate boundary condition (2.16). Similarly, the heat transport equation (2.7) can be derived upon variation of $\delta \theta$. We should point out that the trial function ψ we choose to substitute into $I[\psi]$ does not have to satisfy the correct boundary conditions ((2.16) and (2.17)) simply because the variational principle has incorporated within it the appropriate boundary terms to force the 'correct' trial function into satisfying the right boundary condition if it is to make $I[\psi]$ stationary. This gives us more freedom in choosing the trial function when performing numerical computation or doing theoretical analysis. On the other hand, when dealing with the eigenvalue problem with the rigid boundary condition, we can formulate it in terms of a variational principle with a corresponding $I[\psi]$ whose expression contains no explicit boundary term such as $\int_{\partial\Omega} \sigma \mathbf{u} \cdot \hat{\mathbf{a}} \cdot \mathbf{u}$ only if we are willing to consider trial functions that already satisfy $\mathbf{u} = 0$ on $\partial\Omega$. This point was noted by previous workers (see, for instance, Charlson & Sani 1970), but can be seen more transparently later when we investigate how the solutions behave under a continuous change of the parameters α and β .

Because our later argument depends heavily on the fact that the eigenvalues as expressed by $I[\psi]$ are continuous functions of the adjusting parameters α and β , and can never come into existence out of or disappear into nowhere when we tune them, it is important to show why this is true. The whole idea hinges on the following

inequality, which bounds a surface integral of any function g by a volume integral of ∇g :

$$\frac{\int_{\partial\Omega} g^2}{\int_{\Omega} g^2} < c_1 \left(\frac{\int_{\Omega} |\nabla g|^2}{\int_{\Omega} g^2} \right)^{\frac{1}{2}} + c_2. \tag{2.19}$$

Here c_1 and c_2 are some constants which depend only on the geometry of the domain Ω but not on g . Although (2.19) can be proved by straightforward manipulation of inequalities in the one-dimensional case (Courant & Hilbert 1966) or in spaces of higher dimension (Chen 1991), it is also interesting to note that this inequality is basically the well-known Heisenberg's uncertainty principle in quantum mechanics. This aspect is discussed in Appendix A. Armed with this inequality, we can easily see that each term in the numerator of (2.15), when normalized by the denominator $\int_{\Omega} |\mathbf{u}|^2 + |\theta|^2$, must be bounded for a given $I[\boldsymbol{\psi}]$. This is because if we denote the volume terms involving $|\nabla \mathbf{u}|^2$ and $|\nabla \theta|^2$ by $-\zeta$, then the surface terms are bounded by $c_1 \zeta^{\frac{1}{2}} + c_2$ in virtue of (2.19), whereas the cross-term $2 \int_{\Omega} u_z \theta$ is no bigger than unity by the Cauchy-Schwarz inequality. In short,

$$-\zeta + a \zeta^{\frac{1}{2}} + b \geq I[\boldsymbol{\psi}] \geq -\zeta - a \zeta^{\frac{1}{2}} - b \tag{2.20}$$

for some positive constants a and b . This implies that the volume term ζ must be bounded for the given $I[\boldsymbol{\psi}]$, which then guarantees the boundedness of the surface terms. Incidentally, (2.20) tells us that for any fixed parameters the functional $I[\boldsymbol{\psi}]$ actually is bounded above for all $\boldsymbol{\psi}$. This allows one to maximize $I[\boldsymbol{\psi}]$ to find the 'ground state'. Suppose for a given set of parameters we have found the corresponding ground state $\boldsymbol{\psi}_1$, and assume that the parameter β is changed to a slightly different β' , then by the boundedness of the surface term $\int_{\partial\Omega} |\theta|^2$ of (2.15) we see that $I[\boldsymbol{\psi}_1] - I[\boldsymbol{\psi}_1]$ can be made arbitrarily small as $\beta' \rightarrow \beta$. Therefore, the ground-state eigenvalues associated with β' and β must coalesce as $\beta' \rightarrow \beta$. This not only proves that the ground-state eigenvalue changes continuously as we tune β , but also shows that it cannot suddenly terminate or appear when β is varied. By an elegant inductive procedure, the minimum-maximum principle (Courant & Hilbert 1966), one then concludes that, in fact, all the eigenvalues must obey the same law. In other words, the ordered eigenvalues of \mathcal{L} must change continuously as we tune the parameters.

The previous discussion also gives us the following characterization of the most unstable mode (the 'ground state') λ_1 :

$$\lambda_1 = \max_{\boldsymbol{\psi}} I[\boldsymbol{\psi}]. \tag{2.21}$$

If we control the Rayleigh number so that the system becomes marginally stable at $R = R_c$, then (2.21) becomes $0 = \lambda_1 \geq I[\boldsymbol{\psi}]$ for any trial function $\boldsymbol{\psi}$. Rewriting this inequality then leads us naturally to the second variational principle:

Variational Principle II

The critical Rayleigh number R_c that makes λ_1 vanish is characterized by

$$\frac{1}{R_c^{\frac{1}{2}}} = \max_{\boldsymbol{\psi}} \frac{2 \int_{\Omega} u_z \theta}{\left(\int_{\Omega} |\nabla \mathbf{u}|^2 + \int_{\partial\Omega} \mathbf{u} \cdot \hat{\boldsymbol{\alpha}} \cdot \mathbf{u} \right) + \left(\int_{\Omega} |\nabla \theta|^2 \right) + \left(\int_{\partial\Omega} \beta |\theta|^2 \right)} \equiv \max_{\boldsymbol{\psi}} J[\boldsymbol{\psi}], \tag{2.22}$$

where we have absorbed $\sigma^{\frac{1}{2}}$ into \mathbf{u} .

However, in order for this variational expression for R_c to be meaningful we must make sure that the denominator of (2.22) is positive for all trial functions $\boldsymbol{\psi}$. This

complication can be taken care of quite easily. For instance, we can consider only non-negative values for α and β , and drop the curvature term \mathbf{C} in (2.13). (One does not need to worry about the curvature term if the sidewall is convex.) Once the complication is resolved we can construct a self-adjoint operator \mathcal{R} (defined with respect to an inner product different from the one we use in this paper) such that its eigenfunctions correspond exactly to the functions which make $J[\boldsymbol{\psi}]$ stationary. This permits one to investigate the properties of R_c with rigour more directly (Chen 1992). But we choose not to pursue this line here, because the study of the growth rate λ itself turns out to be intimately related to an interesting biological pattern formation model, the characteristics of which will be addressed and compared with those of the convection problem in this study. We would also like to comment that the following slightly different variational expression for R_c that exists in the literature (see, for example, Busse 1985) can also be derived directly from (2.22):

$$\frac{1}{R_c^2} = \max_{\boldsymbol{\psi}} \frac{2 \int_{\Omega} u_z \theta}{\left[\left(\int_{\Omega} |\nabla \mathbf{u}|^2 + \int_{\partial\Omega} \mathbf{u} \cdot \hat{\mathbf{n}} \cdot \mathbf{u} \right) \cdot \left(\int_{\Omega} |\nabla \theta|^2 + \int_{\partial\Omega} \beta |\theta|^2 \right) \right]^{\frac{1}{2}}} \equiv \max_{\boldsymbol{\psi}} J[\boldsymbol{\psi}]. \quad (2.23)$$

To see this, we can start with (2.22) and simply consider variation of the type $\mathbf{u} \rightarrow \mu \mathbf{u}, \theta \rightarrow \nu \theta$ for arbitrary μ and ν . Clearly, $J[\boldsymbol{\psi}]$ is maximized only if the two parenthesized quantities in the denominator of (2.22) are equal, from which (2.23) immediately follows. It is also easy to check that each of $\delta J'[\boldsymbol{\psi}] = 0$ and $\delta J[\boldsymbol{\psi}] = 0$ alone implies (2.8), with $\lambda = 0$. Another point of interest is that the stationary values of $J[\boldsymbol{\psi}]$ come in pairs of opposite sign, because $(\mathbf{u}, -\theta)$ also makes $J[\boldsymbol{\psi}]$ stationary whenever (\mathbf{u}, θ) does.

In order to bring out the explicit dependence of $I[\boldsymbol{\psi}]$ and $J[\boldsymbol{\psi}]$ on the tuning parameter α (or β and R , as the case may be), for the rest of this paper we will attach an appropriate subscript α (or β and R) to both functionals $I[\boldsymbol{\psi}]$ and $J[\boldsymbol{\psi}]$ whenever such emphasis is necessary. A superscript Ω is sometimes added to remind ourselves of their dependence on the geometry of Ω as well.

As a first application of the variational principles, we observe that the stability of the system is lowered when R is increased. This can be simply shown as follows. Let $\lambda_1(R)$ be the largest eigenvalue for a given Rayleigh number R with the associated eigenfunction $\boldsymbol{\psi}_R$. Then

$$\begin{aligned} \lambda_1(R) &\equiv I_R[\boldsymbol{\psi}_R] = I_R[\boldsymbol{\psi}_R] - (I_{R'}[\boldsymbol{\psi}_R] - I_R[\boldsymbol{\psi}_R]) \\ &= I_R[\boldsymbol{\psi}_R] - 2\sigma^{\frac{1}{2}}(R'^{\frac{1}{2}} - R^{\frac{1}{2}}) \frac{\int_{\Omega} u_z \theta}{\int_{\Omega} |\mathbf{u}|^2 + |\theta|^2}. \end{aligned}$$

But for the ground state it must be true that $\int_{\Omega} u_z \theta \geq 0$, because we could flip the sign of θ to make $I_R[\boldsymbol{\psi}_R]$ larger were $\int_{\Omega} u_z \theta$ less than zero, a contradiction to the assumption that λ_1 is the largest eigenvalue. Therefore,

$$\lambda_1(R) \leq I_R[\boldsymbol{\psi}_R] \leq \max_{\boldsymbol{\psi}} I_R[\boldsymbol{\psi}] \equiv \lambda_1(R'), \quad (2.24)$$

provided $R' > R$. We would like to point out that, contrary to what Sorokin (1953) assumes, (2.24) is not generally true for other eigenvalues λ_j ($j \geq 2$). This is shown in the example studied in Appendix B. We should also mention that the variational principle actually tells us more than what (2.24) promises. This is especially easy to see when the most unstable state is non-degenerate. Indeed, if we begin with

$$\begin{aligned} \lambda_1(R') - \lambda_1(R) &= I_{R'}[\boldsymbol{\psi}_{R'}] - I_R[\boldsymbol{\psi}_R] \\ &= (I_{R'}[\boldsymbol{\psi}_{R'}] - I_R[\boldsymbol{\psi}_{R'}]) + (I_R[\boldsymbol{\psi}_{R'}] - I_R[\boldsymbol{\psi}_R]), \end{aligned}$$

and note that in the limit $R' \rightarrow R$ the second term in parentheses is of second order in $R' - R$ by Variational Principle I, while the first simply reduces to

$$2\sigma^{\frac{1}{2}}\delta R^{\frac{1}{2}} \frac{\int_{\Omega} u_z \theta}{\int_{\Omega} |u|^2 + |\theta|^2},$$

then we immediately obtain

$$\frac{\partial \lambda_1}{\partial R^{\frac{1}{2}}} = 2\sigma^{\frac{1}{2}} \frac{\int_{\Omega} u_z \theta}{\int_{\Omega} |u|^2 + |\theta|^2} \geq 0, \quad (2.25)$$

which result can also be derived directly from a simple regular perturbation calculation. This is the basis of the growth rate derivative calculation done by Shaumeyer *et al.* (1981).

Similarly, if we are interested in how the effective sidewall thermal conductivity β affects the stability, then we can substitute into $I[\psi]$ the ground-state wavefunction ψ_{β} corresponding to the given β , and show that

$$\begin{aligned} \lambda_1(\beta') &\equiv I_{\beta'}[\psi_{\beta'}] = I_{\beta}[\psi_{\beta'}] - (I_{\beta}[\psi_{\beta'}] - I_{\beta'}[\psi_{\beta'}]) \\ &= I_{\beta}[\psi_{\beta'}] - (\beta' - \beta) \frac{\int_{\partial\Omega} |\theta|^2}{\int_{\Omega} |u|^2 + |\theta|^2} \\ &\leq \max_{\psi} I_{\beta}[\psi] - (\beta' - \beta) \frac{\int_{\partial\Omega} |\theta|^2}{\int_{\Omega} |u|^2 + |\theta|^2} \\ &\equiv \lambda_1(\beta) - (\beta' - \beta) \frac{\int_{\partial\Omega} |\theta|^2}{\int_{\Omega} |u|^2 + |\theta|^2}, \end{aligned}$$

which means $\lambda_1(\beta') \leq \lambda_1(\beta)$ when $\beta' > \beta$. Because Courant's minimum-maximum principle applies here, this result is immediately generalizable to all eigenstates, i.e.

$$\lambda_j(\beta') \leq \lambda_j(\beta) \quad \text{for all } j = 1, 2, \dots \quad (2.26)$$

The same method applied to (2.22) yields

$$R_c(\beta') \geq R_c(\beta) \quad \text{if } \beta' > \beta. \quad (2.27)$$

Again, in the non-degenerate case we can show that

$$\frac{\partial \lambda_j}{\partial \beta} = - \frac{\int_{\partial\Omega} |\theta|^2}{\int_{\Omega} |u|^2 + |\theta|^2} \leq 0 \quad (2.28)$$

and

$$\frac{\partial R_c^{\frac{1}{2}}}{\partial \beta} = \frac{\int_{\partial\Omega} |\theta|^2}{2 \int_{\Omega} \theta u_z} \geq 0 \quad (2.29)$$

by the same trick that led to (2.25), provided β does not vary on $\partial\Omega$. The physical interpretation of these results is simply this: the system tends to become more stable if the thermal dissipation on the boundary, as represented here by the effective thermal conductivity β , is stronger. The numerical work of Sparrow *et al.* (1964) has confirmed this prediction. A similar conclusion can be reached with respect to change of the parameter α .

The physical interpretation given above, though necessarily true, nonetheless presents a small puzzle when one encounters it for the first time. The confusion arises because intuitively we expect the two cases $\beta = \pm \infty$ to correspond to the same

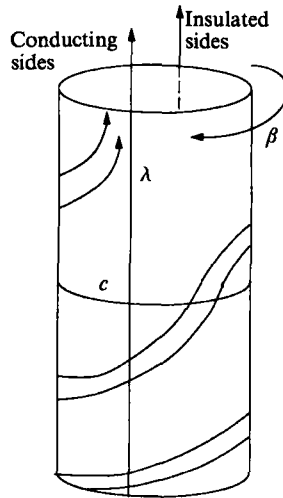


FIGURE 1. Schematic plot of eigenvalues λ_j as functions of β when the two extremes $\beta = -\infty$ and $\beta = +\infty$ are identified on a cylinder. The ring c corresponds to a state that is never perturbed by β . Two runaway solutions for $\beta \rightarrow -\infty$ are shown.

situation in which the sidewall is a perfect conductor, since (2.17) should reduce to $\theta = 0$ in both limits, yet (2.26) seems to indicate otherwise because the eigenvalues are predicted to be monotonically increasing (and thus never come back to their original values) when we tune β from $+\infty$ to $-\infty$. The apparent paradox is easily resolved once we realize that this is just the manifestation of the well-known phenomenon of holonomy in differential geometry (do Carmo 1976, p. 297) or ‘Berry’s phase’ in quantum mechanics (Shapere & Wilczek 1989), i.e. cyclic evolution of a system does not necessarily bring the eigenstates of an operator on the system back to their starting configuration. As will be shown in the following, some eigenvalues actually disappear into infinity when we vary β from $+\infty$ to $-\infty$. Figure 1 shows schematically what happens for this cyclic evolution when β is treated like the angular coordinate on a cylinder and the two extremes $\beta = \pm\infty$ are identified.

The reason why, in the variational formalism, we can only take trial functions whose θ vanish identically on $\partial\Omega$ for a cell with a perfectly conducting sidewall is also clear in this light. As we tune β_1 of $\theta + \beta_1 \partial\theta/\partial n = 0$ from 0^- to 0^+ , several solutions are lost abruptly, and the numbering of the eigenstates is not preserved in this process. Just like in the case when the boundary condition is given by (2.17), we would expect the ordered eigenvalues to be continuous functions of β_1 if it were possible to construct a variational principle that is valid for β_1 in any open interval containing 0, thus allowing the use of trial functions whose θ does not necessarily vanish on $\partial\Omega$.

Next, we proceed to investigate more closely the physically absurd case in which β (or α) is tuned negative. This problem is of interest because it actually tells us something about what to expect when β is positive! (But the appreciation of this novel idea must wait until the end of §3.) The first point to note about the negative- β regime is that at least one eigenvalue (λ_1) goes to $+\infty$ when the β of the sidewalls tends to $-\infty$. (A similar result holds when we tune α to $-\infty$.) This is quite obvious if we take a trial function such that its temperature θ does not vanish identically on the sidewall and then allow β to go to $-\infty$. Inspection of (2.15) and (2.21) then shows

$\lambda_1 \geq I_\beta[\psi] \rightarrow +\infty$ as β tends to $-\infty$. The existence of these ‘runaway’ eigenfunctions clearly is symmetry independent, i.e. there is at least one runaway solution associated with each type of symmetry when the convection cell possesses certain symmetries. It is also interesting to note that the above runaway solutions manage to disappear without jeopardizing the completeness of the whole set of orthonormal eigenfunctions by mimicking a zero vector $\psi = 0$ when β approaches $-\infty$. This mimicry is achieved by developing a very thin thermal boundary layer near the sidewalls and suppressing the velocity field in the fluid. In fact, this feature can be derived and the boundary layer width calculated very easily. First, we argue that for a very negative β the term $(\sigma R)^{\frac{1}{2}}u_z$ in (2.7) must be much smaller than $\lambda_1\theta$ in magnitude. Were this not true, the term $(\sigma R)^{\frac{1}{2}}\theta\hat{e}_z$ of (2.6) would become negligible compared to other terms in the same equation, thus implying the decoupling of the temperature and the velocity fields. But then a restricted version of (2.20) applied to the decoupled linearized Navier–Stokes equation, (2.6), alone immediately tells us that λ_1 is bounded above and could never be made arbitrarily large when we tune β to $-\infty$, thus contradicting our assumption. This argument readily implies that the fluid velocity is strongly suppressed and is much smaller than the temperature everywhere in the cell. Next we take the dot product of (2.7) and θ and integrate over the whole cell to yield

$$-\int_{\Omega} |\nabla\theta|^2 + \int_{\partial\Omega} \theta \frac{\partial\theta}{\partial n} + (\sigma R)^{\frac{1}{2}} \int_{\Omega} \theta u_z = \lambda \int_{\Omega} |\theta|^2, \tag{2.30}$$

from which the presence of a thermal boundary layer is obvious because (i) the first term on the left-hand side is always negative, and (ii) the third term is negligible compared to the right-hand side. Hence, near the sidewall (2.7) can be approximated by

$$\frac{d^2\theta}{ds^2} \approx \lambda\theta, \tag{2.31}$$

where s is the distance measured normally inward from the sidewall. Solving (2.31) subject to the boundary condition $-\partial\theta/\partial s + \beta\theta = 0$ yields

$$\theta \propto e^{\beta s} \tag{2.32}$$

and
$$\lambda \approx \beta^2, \tag{2.33}$$

which clearly indicates the existence of a thermal boundary layer of width $\delta = -1/\beta$. The example studied in Appendix B (cf. (B 13) and (B 14)) verifies this prediction. We remark in passing that, because of continuity, the total number of runaway eigenfunctions does not depend on the value of α when we tune β . Combining the previously derived innocent looking results, finally we are able to show something that is less expected:

PROPOSITION. *Fix all the parameters save β . Assume λ_0 is a degenerate eigenvalue of multiplicity k when $\beta = \beta_0$ for some β_0 . Then no matter what value β takes, λ_0 is always an eigenvalue of multiplicity at least $k-1$, provided the total number of runaway eigenfunctions is 1 when $\beta \rightarrow -\infty$.*

In other words, these degenerate states are robust against perturbation by β when we have only one runaway solution. Again, we have a corresponding theorem for α .

The idea behind this statement is very simple and can be appreciated best by looking at figure 2, which plots schematically the eigenvalues as monotonically

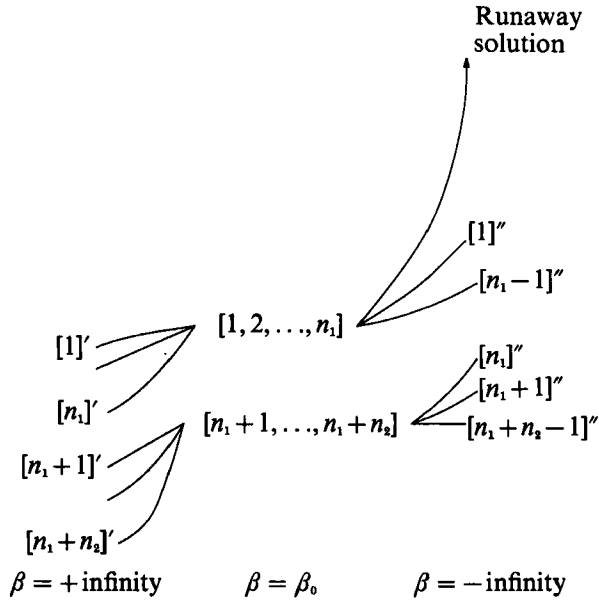


FIGURE 2. Monotonicity of λ_j with respect to β is shown with $[\dots]'$ and $[\dots]''$ denoting, respectively, the eigenvalues for $\beta = +\infty$ and $\beta = -\infty$.

decreasing functions of β . Here we have arranged the eigenvalues in decreasing order, and have assumed very generally that the multiplicities of the eigenvalues are n_1, n_2, n_3, \dots , respectively. The three cases $\beta = \beta_0$ (eigenstates shown in square brackets), $\beta = +\infty$ (eigenstates represented by primed square brackets) and $\beta = -\infty$ (double-primed square brackets) are shown in the same figure for comparison. From this picture we see that

$$[1]'' \geq \dots \geq [n_1 - 1]'' \geq [1]', \tag{2.34}$$

$$[n_1 + 1]'' \geq \dots \geq [n_1 + n_2 - 1]'' \geq [n_1 + 1]'. \tag{2.35}$$

⋮

Since the two cases $\beta = \pm \infty$ correspond to the same physical problem of having a perfectly conducting sidewall, we must have

$$[1]'' = [1]' \tag{2.36}$$

and
$$[n_1 + 1]'' = [n_1 + 1]'. \tag{2.37}$$

Comparing (2.34) with (2.36) immediately shows that there are at least $n_1 - 1$ states in the first group which can never change their eigenvalues when β varies. Similarly, inspection of (2.35) and (2.37) tells us that there is a degeneracy of at least $n_2 - 1$ in the second group, irrespective of what β is. Our claim can then be proved by induction.

We should point out that the previous argument respects the symmetries of the system, too. In other words, the assertion is still valid if we consider only solutions with specific symmetry. As a matter of fact, the proposition is only useful when we can sort out the symmetries of the system first to guarantee the uniqueness of the runaway eigenfunction. An explicit application of this theorem will be given later when we investigate the size dependence of the convection problem. Another point

to be noted is that those eigenvalues which are not perturbed by β are highly exceptional, because (2.28) asserts that their temperature θ must vanish identically on $\partial\Omega$. In view of the boundary condition for θ , (2.17), we see that this also means that $\partial\theta/\partial n$ must vanish identically on $\partial\Omega$, too. Not surprisingly, we will be able to show in the next section that even the eigenfunctions associated with these eigenvalues are not perturbed by β at all.

We close this section with a brief discussion on the behaviour of R_c for a no-slip cell as β is varied. Despite the fact that most of the results and arguments for λ_1 presented above work equally well for R_c , there nevertheless exists a sharp difference between the two when β is negative. This is because the denominator of $J[\psi]$ might vanish when β is negative, in which case Variational Principle II will experience difficulty. In order to prevent this awkward situation from happening, we can only take negative β which satisfies the following constraint:

$$-\beta_{\text{admissible}} \leq \min \frac{\int_{\Omega} |\nabla\theta|^2}{\int_{\partial\Omega} |\theta|^2}. \quad (2.38)$$

We can derive a more useful criterion for the admissible β by dividing the numerator and the denominator of the right-hand side of (2.38) by $\int_{\Omega} |\theta|^2$ and making use of (2.19) and the fact that

$$\frac{\int_{\Omega} |\nabla\theta|^2}{\int_{\Omega} |\theta|^2} \geq \pi^2 + \mu_{\parallel},$$

where μ_{\parallel} is the lowest eigenvalue of the two-dimensional thermal diffusion problem $-\nabla_h^2\theta = \mu_{\parallel}\theta$ subject to the Neumann boundary condition on a horizontal cross-section of the cell. The final result is

$$\max(-\beta_{\text{admissible}}) \geq \frac{\pi^2 + \mu_{\parallel}}{c_1(\pi^2 + \mu_{\parallel})^{\frac{1}{2}} + c_2}. \quad (2.39)$$

(Notice that μ_{\parallel} is non-vanishing only if we restrict ourselves to considering states with specific symmetry.) Equally important is that c_1 and c_2 depend on the geometry of the cell, and c_2 becomes very large when the cell has a very small lateral dimension L . (Our derivation of (2.19), as given in Appendix A, in fact shows $c_2 \propto L^{-1}$ for small L .) This implies that the critical Rayleigh number R_c is very small for a small cell if β can be made negative in a real experiment. This feature is also borne out in the example studied in Appendix B. Although we studied the negative- β regime mainly because it will provide us with an interesting way of deriving results for the physically more relevant case when β is positive, our conclusions nevertheless serve another purpose, namely, they clearly show that a model system can have a very different behaviour when the control parameters of the problem enter the unphysical regime. This caveat may be appropriate if, when applying the amplitude equation to model a system, one has to explore the different areas in the parameter space because the direct relation between the physical boundary conditions and the parameters of the model is not readily available.

3. Perturbation theory, size dependence, and biological pattern formation

In this section we will first develop a formal perturbation theory to deal with the change of stability when the sidewall properties are varied. Several examples are considered in order to illustrate the power of the perturbative approach. Then the

effect of the system size is studied, with special attention directed to comparing the convection problem with the chemical reaction-diffusion model of spatial pattern formation in developmental biology. Finally, the analysis in this section is combined with those proved in §2 to derive the scaling behaviour of the onset Rayleigh number as a function of system size L .

Let $\{|\psi_j\rangle\}$ denote the orthonormal eigenvectors of (2.8) subject to the boundary conditions (2.12)–(2.14), with β_1 of the sidewall being set identically to unity. The corresponding eigenvalue for $|\psi_j\rangle$ is denoted by λ_j . For any given non-degenerate eigenvalue λ_0 we define the Green's function by

$$G_\lambda \equiv \sum_{j \neq 0} \frac{|\psi_j\rangle \langle \psi_j|}{\lambda_j - \lambda}. \tag{3.1}$$

If we use a prime to stand for variables associated with the different thermal boundary condition $\partial\theta'/\partial n + \beta'\theta = 0$, then Lyapunov-Schmidt reduction of $\mathcal{L}\psi' = \lambda'\psi'$ immediately gives

$$(\lambda' - \lambda_0) \int_\Omega \psi_0^\dagger \psi' = \int_{\partial\Omega} \psi_0^\dagger \frac{\partial\psi'}{\partial n} - \frac{\partial\psi_0^\dagger}{\partial n} \psi' = - \int_{\partial\Omega} (\beta' - \beta) \psi_0^\dagger \hat{B}\psi', \tag{3.2}$$

$$\psi'^\perp = \int_{\partial\Omega} (\beta' - \beta) G_\lambda \hat{B}\psi', \tag{3.3}$$

$$\psi' \equiv \psi_0 + \psi'^\perp, \tag{3.4}$$

where \hat{B} is a 4×4 matrix with only one non-vanishing element $B_{44} = 1$, and ψ'^\perp is the projection of ψ' off the one-dimensional subspace spanned by ψ_0 . From the known solutions (ψ_j, λ_j) , one can compute (ψ', λ') by the following iteration scheme whose convergence can be established by standard methods, provided $|\beta' - \beta|$ is small enough:

$$\left. \begin{aligned} \lambda'^{(1)} - \lambda_0 &= - \int_{\partial\Omega} (\beta' - \beta) \psi_0^\dagger \hat{B}\psi_0, \\ \psi'^{\perp(1)} &= \int_{\partial\Omega} (\beta' - \beta) G_{\lambda_0} \hat{B}\psi_0, \\ \lambda'^{(l+1)} - \lambda_0 &= - \int_{\partial\Omega} (\beta' - \beta) \psi_0^\dagger \hat{B}\psi'^{(l)} \quad (l \geq 1), \\ \psi'^{\perp(l+1)} &= \int_{\partial\Omega} (\beta' - \beta) G_{\lambda'^{(l)}} \hat{B}\psi'^{(l)}, \\ \psi'^{(l)} &\equiv \psi_0 + \psi'^{\perp(l)}. \end{aligned} \right\} \tag{3.5}$$

An immediate consequence of this calculation is that the first equation implies

$$\frac{\partial\lambda}{\partial\beta} = - \frac{\int_{\partial\Omega} |\theta|^2}{\int_\Omega |u|^2 + |\theta|^2},$$

a result we derived earlier in (2.28) based on the variational principle.

When two states ψ_1 and ψ_2 are nearly degenerate we can define a Green's function

$$G_\lambda \equiv \sum_{j=1,2} \frac{|\psi_j\rangle \langle \psi_j|}{\lambda_j - \lambda} \tag{3.6}$$

and perform Lyapunov-Schmidt reduction on the two-dimensional subspace spanned by ψ_1 and ψ_2 to obtain

$$\begin{pmatrix} \lambda_1 - \lambda' - b_{11} & -b_{12} \\ -b_{21} & \lambda_2 - \lambda' - b_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \int_{\partial\Omega} (\beta' - \beta) \psi_1^\dagger \hat{\mathbf{B}} \psi'^\perp \\ \int_{\partial\Omega} (\beta' - \beta) \psi_2^\dagger \hat{\mathbf{B}} \psi'^\perp \end{pmatrix}, \tag{3.7}$$

$$\psi'^\perp = \int_{\partial\Omega} (\beta' - \beta) G_{\lambda'} \hat{\mathbf{B}} \psi', \tag{3.8}$$

$$\psi' \equiv x_1 \psi_1 + x_2 \psi_2 + \psi'^\perp, \tag{3.9}$$

$$b_{ij} \equiv \int_{\partial\Omega} (\beta' - \beta) \psi_i^\dagger \hat{\mathbf{B}} \psi_j. \tag{3.10}$$

This suggests the following iteration scheme of solving for (ψ'_1, λ'_1) and (ψ'_2, λ'_2) :

(i) Lowest-order eigenvalue $\lambda^{(1)}$ is to be obtained by solving

$$\begin{pmatrix} \lambda_1 - \lambda^{(1)} - b_{11} & -b_{12} \\ -b_{21} & \lambda_2 - \lambda^{(1)} - b_{22} \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = 0. \tag{3.11}$$

This also determines

$$\frac{x_2^{(1)}}{x_1^{(1)}} = \frac{b_{22} - b_{11} + \lambda_1 - \lambda_2 \pm [(b_{11} - b_{22} + \lambda_2 - \lambda_1)^2 + 4b_{12} b_{21}]^{\frac{1}{2}}}{2b_{12}}, \tag{3.12}$$

where we have assumed $b_{12} \neq 0$ for simplicity. Our normalization convention can then be taken as $x_1 \equiv 1$.

(ii) Iteration:

$$\left. \begin{aligned} &\psi'^\perp{}^{(1)} \equiv 0, \\ &\begin{pmatrix} \lambda_1 - \lambda'^{(l+1)} - b_{11} & -b_{12} \\ -b_{21} & \lambda_2 - \lambda'^{(l+1)} - b_{22} \end{pmatrix} \begin{pmatrix} 1 \\ x_2^{(l+1)} \end{pmatrix} = \begin{pmatrix} \int_{\partial\Omega} (\beta' - \beta) \psi_1^\dagger \hat{\mathbf{B}} \psi'^\perp{}^{(l)} \\ \int_{\partial\Omega} (\beta' - \beta) \psi_2^\dagger \hat{\mathbf{B}} \psi'^\perp{}^{(l)} \end{pmatrix}, \\ &\psi'^\perp{}^{(l+1)} = \int_{\partial\Omega} (\beta' - \beta) G_{\lambda^{(l)}} \hat{\mathbf{B}} \psi'^{(l)}, \\ &\psi'^{(l)} \equiv \psi_1 + x_2^{(l)} \psi_2 + \psi'^\perp{}^{(l)}, \end{aligned} \right\} \tag{3.13}$$

where $x_2^{(l+1)}$ and $\lambda'^{(l+1)}$ are solved by the matrix equation shown above.

To show how this perturbation scheme can produce sharper results than the variational principle alone, we simply note that the eigenfunctions associated with those exceptional eigenvalues studied in §2 which are not perturbed by a varying β are actually not affected by β , either. This is almost trivial from (3.5), because the vanishing of both θ and $\partial\theta/\partial n$ on $\partial\Omega$ guarantees that the eigenfunction remains the same up to any order of iteration. (In case of degeneracy, we can apply (3.13) or resort to a continuity argument to show that the same conclusion still holds.)

To give another example, let us study how the stability of a convection cell with a smooth, rigid, and perfectly conducting sidewall is affected by the size of the system. Assume Ω' is a cell that is completely inside a slightly bigger cell Ω . A Lyapunov-Schmidt reduction yields

$$(\lambda' - \lambda_0) = \frac{\int_{\partial\Omega'} \psi_0^\dagger \hat{\mathbf{D}} \frac{\partial \psi'}{\partial n}}{\int_{\Omega'} \psi_0^\dagger \psi'}, \tag{3.14}$$

$$\psi'^\perp = - \int_{\partial\Omega'} G_{\lambda_0} \hat{\mathbf{D}} \frac{\partial \psi'}{\partial n}, \tag{3.15}$$

where we have extended the definition of ψ' by setting $\psi' \equiv 0$ outside Ω' , and $\psi' \equiv \psi_0 + \psi'^{\perp}$. Here $\partial\Omega'$ refers to the sidewall alone, and notice the volume integral is over Ω' . These equations suggest the following iteration scheme for solving the solutions for Ω' :

$$\left. \begin{aligned} \lambda^{(1)} - \lambda_0 &= \frac{\int_{\partial\Omega'} \psi_0^\dagger \mathbf{D} \frac{\partial \psi_0}{\partial n}}{\int_{\Omega'} \psi_0^\dagger \psi_0}, \\ \psi'^{\perp(1)} &= - \int_{\partial\Omega'} G_{\lambda_0} \mathbf{D} \frac{\partial \psi_0}{\partial n}, \\ \lambda^{(l+1)} - \lambda_0 &= \frac{\int_{\partial\Omega'} \psi_0^\dagger \mathbf{D} \frac{\partial \psi'^{(l)}}{\partial n}}{\int_{\Omega'} \psi_0^\dagger \psi'^{(l)}}, \\ \psi'^{\perp(l+1)} &= - \int_{\partial\Omega'} G_{\lambda_0} \mathbf{D} \frac{\partial \psi'^{(l)}}{\partial n}. \end{aligned} \right\} \quad (3.16)$$

With the formalism at hand, we can show that the eigenvalues of the primed and the unprimed systems satisfy $\lambda_j > \lambda'_j$ for $j = 1, 2, \dots$. In other words, Ω' is strictly more stable than Ω . (Notice this result can be trivially translated in terms of R_c by virtue of (2.25).) To make our argument simple, we will assume without loss of generality that the two cells almost coincide, and that the state under consideration is non-degenerate.

First, if $\partial\psi_0/\partial n$ is not identically zero on $\partial\Omega$, then the lowest-order iteration of (3.16) yields

$$\lambda^{(1)} - \lambda_0 = \frac{\int_{\partial\Omega'} \psi_0^\dagger \mathbf{D} \frac{\partial \psi_0}{\partial n}}{\int_{\Omega'} \psi_0^\dagger \psi_0} \approx - \int_{\partial\Omega} \frac{\partial \psi_0^\dagger}{\partial n} \mathbf{D} \frac{\partial \psi_0}{\partial n} s < 0,$$

where s is the inward normal distance from a point on $\partial\Omega$ to $\partial\Omega'$. This clearly is what we set out to prove. So we only need to consider the remaining possibility that $\partial\psi_0/\partial n = 0$ on $\partial\Omega$. If this does happen, then the next order iteration of (3.16) yields

$$\lambda^{(2)} - \lambda_0 = \frac{\int_{\partial\Omega'} \psi_0^\dagger \mathbf{D} \frac{\partial \psi'^{(1)}}{\partial n}}{\int_{\Omega'} \psi_0^\dagger \psi'^{(1)}} \approx - \int_{\partial\Omega} \frac{1}{2} \frac{\partial^2 \psi_0^\dagger}{\partial n^2} \mathbf{D} \frac{\partial^2 \psi_0}{\partial n^2} s^3 \leq 0.$$

But a closer look at this case shows that '=' is never attained in the above expression. In fact, the equality can hold only if $\psi_0 = \partial\psi_0/\partial n = \partial^2 \psi_0/\partial n^2 = 0$ on the sidewall. Yet upon rewriting the original fluid equations ((2.8) and (2.3)) in local coordinates near the sidewall, one immediately arrives at the following initial-value problem for fluid variables as functions of s :

$$\frac{\partial^2 \psi}{\partial s^2} = f_1, \quad \frac{\partial p}{\partial s} = f_2,$$

where f_1 and f_2 are 'driving terms' which (i) only involve lower-order derivatives in s and derivatives along the sidewall, and (ii) vanish identically on the sidewall if

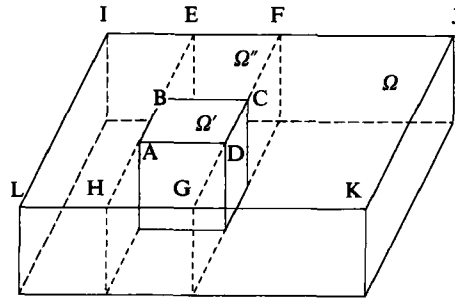


FIGURE 3. An intermediate cell $\Omega'' = EFGH$ is introduced to allow one to compare the stability of $\Omega' = ABCD$ and $\Omega = IJKL$.

$\psi_0 = \partial\psi_0/\partial n = \partial^2\psi_0/\partial n^2 = 0$ on it. This implies that $\psi \equiv 0$ inside Ω after integrating along s , thus contradicting our assumption that ψ is an eigenvector. Therefore, ' $=$ ' can never be attained, and the claim is verified. We emphasize again that the variational principle itself does not offer a simpler proof for this case, because the trial function for the smaller domain must be carefully chosen to take care of the $O(s^3)$ possibility we treated above if strict inequality is to be proved. We should also point out that although the previous argument assumes that the sidewall is smooth, thus excluding rectangular cells that are commonly used in experiments, we can easily modify the proof to cover the case when the relative stability of two rectangular cells is the subject of interest. The idea is to compare them with an intermediate rectangular cell Ω'' whose two pairs of parallel sidewalls are each in contact with the sidewalls of Ω and Ω' , respectively; see figure 3.

The previous result already gives us a hint that somehow the presence of the sidewall has the effect of stabilizing the convection system. However, it still is not general enough in the sense that it is only proved under the assumption that the sidewall is perfectly conducting. What can we say if the sidewall is not a perfect conductor? As it turns out, the sidewall still has a stabilizing effect on the system, but we cannot expect to see the same monotonic behaviour, because another competing factor due to the convective rolls trying to fit themselves snugly into the cell is acting. Roughly speaking, the system is more stable if its dimension is such that an integral number of convective rolls can be accommodated completely into the cell. In other words, we would expect to see an oscillatory behaviour if we plot the growth rate (or the critical Rayleigh number) as a function of the system size. The stabilizing effect coming from the thermally dissipative nature of the sidewall, though always at work, may not be strong enough to wipe out this background oscillation. A look at figure 11 (Appendix B) immediately convinces one that this is indeed the case. In the following, we plan to work out the precise meaning of this physical interpretation by analysing a rectangular box as an example, and then relate this observation to one mechanism that has been proposed to explain several pattern formation features observed in a developing embryo.

Let Ω be a rectangular box whose x -dimension is L . Stretch Ω by a factor $l > 1$ in the x -direction to obtain a larger box Ω' . Assume the boundary conditions for both boxes are (i) the two horizontal plates are rigid or free slip, and perfectly conducting, (ii) $\partial\theta/\partial n + \beta\theta = 0$ ($\beta > 0$) on the sidewalls, and (iii) $\mathbf{u} = 0$ on the sidewalls. Then the claim is that there exists an l_1 such that $R_c^{\Omega'} < R_c^{\Omega}$ for all $l > l_1$. To see how this finite-size stabilizing mechanism is caused by the thermal property of the sidewall, we first observe that we will lose no generality in assuming that θ does not vanish identically

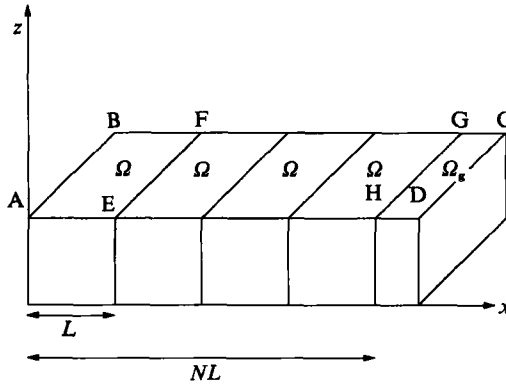


FIGURE 4. N copies of identical box $\Omega = ABFE$ are joined lengthwise to fill up cell $\Omega' = ABCD$, with an unfilled gap $\Omega_g = GCDH$.

on the two sides that are orthogonal to the x -axis, because that case is already covered by our previous discussion. We now try to fill up Ω' by placing side by side as many copies of Ω as possible. In general there will be a gap Ω_g left that is too small for any copy of Ω to fit in; see figure 4. Assume it takes N copies $\Omega_1, \Omega_2, \dots, \Omega_N$ to almost fill Ω' . Next, we construct a trial function ψ' for Ω' in the following manner:

(i) Inside each of $\Omega_1, \Omega_3, \Omega_5, \dots$ we simply take the ground-state wavefunction ψ of Ω as ψ' .

(ii) Inside each of $\Omega_2, \Omega_4, \Omega_6, \dots$ we define ψ' to be the reflection of ψ with respect to the wall common to its neighbouring cell.

(iii) In the gap Ω_g we define ψ' as $u' \equiv 0$ and $\theta'(x, y, z) \equiv \theta(N, L, y, z)$.

Note that the ‘kinks’ of ψ' across the cell walls can be smoothed out and present no real difficulty to our argument. Then

$$\begin{aligned} \frac{1}{J^{\Omega}[\psi]} &= \frac{N(\int_{\Omega} |\nabla u|^2 + |\nabla \theta|^2) + N \int_{\partial\Omega} \beta |\theta|^2}{2N \int_{\Omega} u_z \theta} \\ &= \frac{\int_{\Omega'} |\nabla u'|^2 + |\nabla \theta'|^2 + \int_{\partial\Omega'} \beta |\theta'|^2 + N \int_{\partial\Omega} \beta |\theta|^2}{2 \int_{\Omega'} u_z \theta'} + O\left(\frac{1}{N}\right) \\ &= \frac{1}{J^{\Omega'}[\psi']} + \frac{\int_{\partial\Omega} \beta |\theta|^2}{2 \int_{\Omega} u_z \theta} + O\left(\frac{1}{N}\right). \end{aligned} \tag{3.17}$$

By Variational Principle II we see this implies that $R_c^{\Omega} > R_c^{\Omega'}$ for all large enough N , and hence the claim is verified. Clearly, this also is true if Ω' has a much larger dimension than Ω in both the x - and y -directions. The main point to be made here lies, however, in (3.17), which says that *the finite system is more stable than an infinite system by an amount that is about what the boundary contributes in the variational expression for $R_c^{\frac{1}{2}}$.*

As was remarked before, our inability to prove strict monotonicity for all Ω' bigger than Ω in the previous argument stems from the fact that oscillatory behaviour does indeed exist when Ω' is not much bigger than Ω . What is interesting is that this oscillation can be viewed as the remnant of a very special spatial pattern selection mechanism, the Turing instability (Turing 1952), that one encounters in developmental biology. To see the connection between the two, we recall in the study

of biological pattern formation that one usually analyses the stability of the following model (see, for example, Murray 1989, chaps. 14, 15)

$$\hat{D}\nabla^2\psi + \hat{A}\psi = \lambda\psi \quad \text{in } \Omega, \tag{3.18}$$

$$\hat{D}\frac{\partial\psi}{\partial n} + \hat{B}\psi = 0 \quad \text{on } \partial\Omega, \tag{3.19}$$

where $\psi \in \mathcal{C}^m$ is an m -component chemical concentration field in a ‘cell’ Ω , and \hat{D} , \hat{A} , and \hat{B} are $m \times m$ matrices. Usually \hat{D} is a (positive definite) diagonal matrix which is comprised of the diffusion constants of the chemical species, and \hat{B} is identically zero when no-flux boundary condition is assumed.

This problem can be solved analytically for the no-flux boundary condition by taking the solution to be of the form

$$\psi = \varphi\bar{\psi}, \tag{3.20}$$

where $\bar{\psi}$ is some $m \times 1$ constant vector, and φ is an eigenfunction of a Laplacian subject to the Neumann boundary condition:

$$\nabla^2\varphi = -\mu\varphi \quad \text{in } \Omega, \tag{3.21}$$

$$\frac{\partial\varphi}{\partial n} = 0 \quad \text{on } \partial\Omega. \tag{3.22}$$

Substituting (3.20) into (3.18) we get the dispersion relation

$$\det(-\mu\hat{D} + \hat{A} - I\lambda) = 0. \tag{3.23}$$

When the control parameters in the matrix \hat{A} are suitably adjusted, there will only be one branch of solution that becomes marginally stable when μ equals some value μ_c . For all other μ the associated growth rate λ will have a negative real part, and that corresponds to a stable configuration. Because μ scales in size L like L^{-2} , a cell that expands indefinitely will lower the μ of each mode, thus allowing different modes to become unstable in succession when their respective μ decreases and passes through μ_c , and only to recover their stability again after further expansion. Thus, a plot of the real part of the growth rate λ versus L for this scenario looks very much like figure 8(a) (Appendix B). The most prominent feature of this figure is the many intersections between different curves. This signifies the constant switching of unstable modes as one increases L . In fact, this characteristic is also shared by the convection system if the sidewall boundary conditions are properly (but artificially) assumed. Our next task is, therefore, to bring out the connection between the two systems, and then make very good use of this peculiarity later when we try to deduce the scaling behaviour for R_c .

The best way to see the similarity between the biological model and the convection problem is to note that for a convection cell whose α is identically zero on the boundaries and whose sidewalls are perfectly insulating, we can adopt an ansatz that is closest to the above method in spirit by writing the solution ψ to the problem as

$$\psi \equiv \begin{pmatrix} u_x \\ u_y \\ u_z \\ \theta \end{pmatrix} = \begin{pmatrix} \bar{u} \partial\varphi/\partial x \\ \bar{u} \partial\varphi/\partial y \\ \bar{u}_z \partial\varphi/\partial z \\ \bar{\theta} \partial\varphi/\partial z \end{pmatrix}, \tag{3.24}$$

$p = \bar{p}\varphi$

where \bar{u} , \bar{u}_z , $\bar{\theta}$ and \bar{p} are constants to be determined by the boundary conditions

imposed on the two horizontal plates, and φ satisfies (3.21) and (3.22). Upon substitution of (3.24) into (2.6), (2.7) and (2.3) one obtains

$$\begin{pmatrix} -1 & -\sigma\mu - \lambda & 0 & 0 \\ -1 & 0 & -\sigma\mu - \lambda & (\sigma R)^{\frac{1}{2}} \\ 0 & 0 & (\sigma R)^{\frac{1}{2}} & -\mu - \lambda \\ 0 & -\mu + \mu_z & -\mu_z & 0 \end{pmatrix} \begin{pmatrix} \bar{p} \\ \bar{u} \\ \bar{u}_z \\ \bar{\theta} \end{pmatrix} = 0, \quad (3.25)$$

where we have separated out z -dependence by introducing $\partial^2\varphi/\partial z^2 \equiv -\mu_z\varphi$ for some appropriate constant μ_z . Solving (3.25) for λ we get

$$\frac{\lambda}{\mu} = \frac{-1 - \sigma \pm \{(1 + \sigma)^2 - 4\sigma[1 - (R/\mu^2)(1 - \mu_z/\mu)]\}^{\frac{1}{2}}}{2}, \quad (3.26)$$

from which we see that the pattern selection mechanism described before for biological systems works equally well here if we tune R to $R_{c0} \equiv \min \mu^3/(\mu - \mu_z)$.

This ansatz is particularly suitable for a one-dimensional problem, and in figure 8(a) (Appendix B) we show the λ versus L plot for this case when $R = R_{c0} = \frac{27}{4}\pi^4$. Again, readers are reminded to pay special attention to the many intersections between different modes. When we change the viscous and/or thermal properties on the sidewalls the intersections generally will break up and form many oscillations, as shown in figure 8(b). This clearly justifies our calling the wavy curves remnants of a special biological pattern selection mechanism. The implication is that Rayleigh-Bénard convection can be thought of as a canonical pattern formation problem that bears features that some more specialized reaction-diffusion models possess.

At this point it is only fair for us to stress that the applicability of the previous ansatz, (3.24), to a three-dimensional cell is not as simple as it sounds. The problem is that this trick is manifestly equivalent to solving the coupled equations by the method of separation of variables, which requires the vertical vorticity of the eigenfunctions we are interested in to vanish, a point that is addressed by Joseph (1976) and discussed more fully in Appendix C. Yet this is not guaranteed for the problem at hand. Fortunately, this difficulty does not exist for a large three-dimensional cell whose two horizontal plates satisfy either the more general boundary condition (2.16) with a *positive* α or the physically realizable rigid condition, provided that we adopt the following more general ansatz when the sidewall is rotation-free and insulating (the subtlety involved and the validity of this method are discussed in Appendix C):

$$\psi = \begin{pmatrix} \bar{u} \partial\varphi/\partial x \\ \bar{u} \partial\varphi/\partial y \\ \bar{u}_z \varphi \\ \bar{\theta} \varphi \end{pmatrix}, \quad (3.27)$$

$$p = \bar{p}\varphi.$$

Here, φ is the eigenfunction of the two-dimensional Laplacian on the horizontal cross-section of the cell subject to the Neumann boundary condition ($\nabla_{\mathbf{h}}^2\varphi = -\mu\varphi$), and \bar{u} , \bar{u}_z , $\bar{\theta}$ and \bar{p} are functions of z properly chosen so that the governing equations and the boundary conditions at $z = 0, 1$ are satisfied. Once again, this ansatz will give us a (slightly more complicated) determinantal equation which allows us to write the growth rate λ for the most unstable branch as a function of μ and R : $\lambda = \lambda(\mu, R)$. When R equals $R_{c\infty}$, the critical Rayleigh number for an infinite cell, there is only one

value of μ (denoted by μ_c from now on) which can make λ vanish. For other choices of μ the corresponding λ is always less than zero. It is in this respect that we clearly see how the model biological system is intimately related to the convection problem. But more interestingly, later in this study we will even take advantage of the simplicity of this ansatz and derive from it a scaling property for the critical Rayleigh number R_c that actually holds for almost all boundary conditions (the physically realizable cases included)!

Despite our showing the similarity of pattern generation between convection and biological systems, there exists a difference in which the incompressibility of fluid plays some role. Specifically, if we consider the eigenvalue problem of the biological model ((3.18) and (3.19)) with a self-adjoint \hat{A} and a positive semidefinite \hat{B} , so that the governing equation will look very much like that for the convection problem (cf. (2.9)) except that the pressure term is absent to account for the fact that we do not have an incompressibility constraint for this system, then we can show that the growth rate, which must be real-valued because of our assumptions, displays no oscillatory behaviour as we stretch the system in some direction. As a matter of fact, the system simply becomes less stable when we stretch it. Because of its intrinsic interest, and because this is not the same as a similar statement which holds true for domain expansion of any kind when the Dirichlet boundary condition is imposed,† we shall present a sketch of the argument in what follows.

Suppose Ω' is a cell obtained by linearly stretching Ω in the x -direction by a factor $l > 1$, then we claim that $\lambda'_j \geq \lambda_j$ for all j . To facilitate the proof, we observe that this eigenvalue problem is equivalent to the variational principle of the following functional:

$$K^{\Omega}[\psi] \equiv \frac{-\int_{\Omega} \nabla \psi^{\dagger} \hat{D} \nabla \psi + \int_{\Omega} \psi^{\dagger} \hat{A} \psi - \int_{\partial \Omega} \psi^{\dagger} \hat{B} \psi}{\int_{\Omega} \psi^{\dagger} \psi}.$$

Now take the ‘ground state’ $\psi(x, y, z)$ for the cell Ω (so that $\lambda_1 = K^{\Omega}[\psi]$) and define a trial function ψ' for the stretched domain Ω' as $\psi'(x', y', z') \equiv \psi(x'/l, y', z')$. Then for any function $f(x, y, z)$ we have

$$\begin{aligned} \int_{\Omega} f(x, y, z) \, dx \, dy \, dz &= \frac{1}{l} \int_{\Omega'} f\left(\frac{x'}{l}, y', z'\right) \, dx' \, dy' \, dz' \\ &\equiv \frac{1}{l} \int_{\Omega'} f', \end{aligned}$$

where $f'(x', y', z') \equiv f(x'/l, y', z')$, and similarly

$$\int_{\Omega} \frac{\partial \psi^{\dagger}}{\partial x} \hat{D} \frac{\partial \psi}{\partial x} = l^2 \frac{1}{l} \int_{\Omega'} \frac{\partial \psi'^{\dagger}}{\partial x'} \hat{D} \frac{\partial \psi'}{\partial x'}.$$

We also have

$$\int_{\partial \Omega} f = \iint \frac{f}{\cos \theta} \, dy \, dz = \int_{\partial \Omega'} f' \frac{\cos \theta'}{\cos \theta},$$

where $\theta(\theta')$ is the angle between the surface area element of $\partial \Omega$ ($\partial \Omega'$) and the (y, z) -plane. Hence

$$l \int_{\partial \Omega} f = \int_{\partial \Omega'} f' l \frac{\cos \theta'}{\cos \theta}.$$

† One should recall that a smaller membrane subject to the Neumann boundary condition is not necessarily more stable than a larger one; see, for example, Courant & Hilbert (1966).

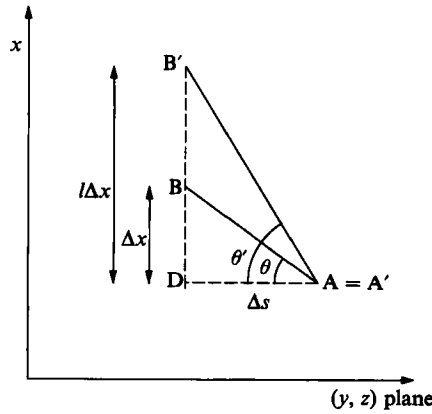


FIGURE 5. Side view of a small segment AB of boundary $\partial\Omega$ and its stretched version $A'B'$ on $\partial\Omega'$. A and A' are identified for convenience.

But from figure 5 we readily see that

$$l \frac{\cos \theta'}{\cos \theta} = l \frac{\Delta s / [(\Delta s)^2 + (l\Delta x)^2]^{\frac{1}{2}}}{\Delta s / [(\Delta s)^2 + (\Delta x)^2]^{\frac{1}{2}}} > 1,$$

thus, $l \int_{\partial\Omega} f \geq \int_{\partial\Omega'} f'$, if f is non-negative on $\partial\Omega$. Combining all the previous results, one obtains

$$\begin{aligned} \lambda_1 &= K^Q[\psi] \\ &\leq \frac{-\int_{\Omega'} \nabla' \psi'^{\dagger} \hat{D} \nabla' \psi' + \int_{\Omega'} \psi'^{\dagger} \hat{A} \psi' - \int_{\partial\Omega'} \psi'^{\dagger} \hat{B} \psi'}{\int_{\Omega'} \psi'^{\dagger} \psi'} \\ &\leq \lambda'_1, \end{aligned}$$

by the variational principle. Finally, to show that this is actually true for all the excited states we can simply invoke the minimum-maximum principle.

To summarize: the model described above cannot be a candidate for Turing type biological pattern formation, because the incompressibility condition is not present in this problem to prevent us from ‘dilating’ the original solution for Ω to produce a state of lower stability for the larger cell Ω' . Besides, this new less stable state is not a result of switching of modes that is sought in the standard model. In contrast, the corresponding fluid system, which must satisfy the incompressibility condition, can remain in the most unstable state only if it constantly switches to a new mode with an extra convective roll to account for the extra room gained as a result of our stretching of the cell. This also partially explains why in our earlier analysis of the finite-size effect we had to put identical copies of smaller cell side by side in order to construct a good trial function for the stretched system Ω' .

We now concentrate on the scaling behaviour of R_c when the cell dimension L varies from 0 to $+\infty$.

The small- L behaviour can be easily extracted from the fact that the critical Rayleigh number corresponding to other boundary conditions is no smaller than $\mu^3/(\mu - \mu_z)$, the critical Rayleigh number for the same cell with $\alpha = 0$ everywhere on the boundaries and $\beta = 0$ on the sidewalls, thanks to the analysis of §2. Thus

$$R_c \geq \frac{\mu^3}{\mu - \mu_z} \propto \frac{1}{L^4} \tag{3.28}$$

for small L , because μ scales like L^{-2} . We can also see this from (2.23) because each of ∇u and $\nabla \theta$ contributes a L^{-1} factor when L is small. Since this small-aspect-ratio

problem is ‘trivial’ in this respect, and is already well documented in the literature (Verhoeven 1969; Yih 1959; Hales 1937), we will proceed to study the large- L limit for the rest of this section.

We have not been able to prove with complete rigour that for a large cell of arbitrary geometry the difference between R_c and its infinite-cell limit $R_{c\infty}$ is of order L^{-2} , though a very plausible argument is presented below to show that this is indeed the case. The idea is to study the scaling property of a cylindrical cell first, and then compare cells of other shapes with this fiducial one to establish that in general the geometry will not alter the scaling law. The examination of the scaling behaviour of the cylindrical cell itself will be broken down into two steps. (i) We will consider a cell of radius L which has an insulating sidewall and an α that is identically zero on $\partial\Omega$. This simple example will provide us with enough motivation to ask if the features we learned from it carry over to the more realistic problem in which both horizontal plates are rigid. (ii) We then show that, indeed, the basic features we observed in the previous example do not change appreciably for the more realistic cell, and, as a result, $R_c - R_{c\infty}$ for any cell having rigid bounding walls scales like L^{-2} , with a prefactor that is practically independent of the azimuthal quantum number m .

Consider, therefore, the zero vertical vorticity states of a side-insulated cylindrical cell of radius L . Assume α is identically zero all over $\partial\Omega$. Then we have, in cylindrical coordinates,

$$\varphi = \sin n\pi(z + \frac{1}{2}) e^{im\phi} J_m\left(k_{m,j} \frac{r}{L}\right), \tag{3.29}$$

where m, n are integers, J_m is the Bessel function of order m , and $k_{m,j}$ is the j th root of $dJ_m(\rho)/d\rho = 0$. The corresponding λ versus L curves for fixed m and n look very much like that shown in figure 8 (Appendix B) when $R = R_{c0}$. It is clear that at the intersection of any two curves we can find a state φ_1 in the two-dimensional degenerate subspace such that $\varphi_1 = \partial\varphi_1/\partial r = 0$ at $r = L$. This means that the state satisfies $\theta = \partial\theta/\partial r = 0$ on the sidewall, thus implying it will never be perturbed by β from our previous discussions. We will call it an ‘anchored state’ in what follows. In addition, this implies that there is only one runaway solution when $\beta \rightarrow -\infty$. Hence the most unstable eigenstate for a cell with $\alpha = 0$ and a conducting sidewall is bounded below by the second unstable state of a cell with an insulating sidewall, a result that can be seen by the same reasoning used in proving the Proposition in §2. But it is easy to verify that the second unstable curve λ_2 in the λ vs. L plot is monotonic between the two anchored states that are intersections of curves corresponding to $k_{m,j}, k_{m,j+1}$ and $k_{m,j}, k_{m,j+2}$, respectively, because each of these curves is locally quadratic in L . By the asymptotic expression (Morse & Feshbach 1968)

$$J_m(z) \approx \left(\frac{2}{\pi m \tan \gamma}\right)^{\frac{1}{2}} \cos(m(\tan \gamma - \gamma) - \frac{1}{4}\pi) \quad (1 \ll z \equiv m \sec \gamma, 0 < \gamma < \frac{1}{2}\pi) \tag{3.30}$$

one can derive

$$k_{m,j} \equiv m \sec \gamma_{m,j} \approx \begin{cases} m + \frac{1}{2}(3\pi(j - \frac{3}{4}))^{\frac{2}{3}} m^{\frac{1}{3}} & (m \gg 1, m \gg j); \\ \left(j + \frac{m}{2} - \frac{3}{4}\right)\pi - \frac{m^2}{2\pi(j + \frac{1}{2}m - \frac{3}{4})} & (j \gg 1, j \gg m), \end{cases} \tag{3.31}$$

and from this formula and (3.26) we conclude that $\lambda_2 = O(L^{-2})$ for large L . Hence, the effect of perturbation on the sidewall thermal property is always of $O(L^{-2})$. Translating this result in terms of R_c , we immediately have $R_c - R_{c0} = O(L^{-2})$.

Similar analysis shows that perturbing the sidewall viscous boundary condition for the $m = 0$ modes has the same effect. However, it is to be noted that simultaneous perturbation by α and β immediately destroys the anchored states. Also, we do not have any control on $m \neq 0$ modes when we vary α . This apparent difficulty can be easily handled once we realize that there are states that behave virtually like anchored states under any perturbations, an interesting assertion that will be made precise below for a rigid-rigid cell as demonstration.

For a cell with two rigid horizontal plates we have to use (3.27) as the unperturbed system and tune α and β of the sidewall to study their effects. In contrast to the previous case, we do not have any anchored states this time, because j , the quantum number that characterizes the number of rolls in the radial direction, is coupled to the z -dependence of the eigenstates, thus preventing us from constructing a linearly superposed state out of the two intersecting solutions to satisfy $\theta = 0$ or $u = 0$ at $r = L$, as the case may be. However, we do have states that are practically anchored for all positive α and β ! To understand why this is so, consider states that have the same number of vertical rolls. (In reality, only the states of one vertical roll are of interest to us.) Fix R to be $R_{c\infty}$. Again, for a given m , the intersecting solutions at $L = L_{\text{int}}$ correspond to those that have different j values. In particular, the most unstable intersections consist of two solutions whose j values differ only by one. By (3.30) and (3.31) we see $|J_m(k_{m,j+1}) + J_m(k_{m,j})| = O(1/L_{\text{int}})$ at the very worst. Therefore, up to a phase, the intersecting solutions differ at most only by $O(1/L_{\text{int}})$ on the sidewall. Equation (3.12) then tells us that one of the perturbed eigenstates will have θ almost zero on the sidewall, because it is almost proportional to $J_m(k_{m,j+1}) + J_m(k_{m,j})$. By (3.3) and (3.4) we see that this state, in turn, must remain practically fixed during the process of perturbation. More specifically, its perturbation will be $O(L^{-1})$ smaller than that experienced by other states or states of different L . We would like to emphasize that our argument clearly is very general and applies to simultaneous perturbation of α and β as well. For example, we can choose $\alpha \equiv \tan(\frac{1}{2}\pi\beta/\beta_0)$ and let $\beta \rightarrow \beta_0$. The final configuration then corresponds to a cell with a rigid sidewall whose thermal property is characterized by $\partial\theta/\partial n + \beta_0\theta = 0$.

We now examine what happens when maximal perturbation is achieved, i.e. $\alpha = \beta = +\infty$. The previous argument plus an earlier assertion about the strict decrease of stability when we expand this system implies $\lambda = -O(L^{-2})$ monotonically in L . In fact, the proportionality constant can also be computed in principle, because the locations of the practically anchored states can be well approximated by the intersections of the unperturbed curves, which themselves can be computed analytically. To compute this prefactor, we note that, for a given m , the system size L_j which renders $\lambda = 0$ must be solved from

$$k_{m,j}^2/L_j^2 = \mu_c. \tag{3.32}$$

(Notice that the left-hand side is just the eigenvalue of the two-dimensional Laplacian $-\nabla_h^2$.) For a large cell the difference ΔL between L_{j+1} and L_j is well approximated by

$$\Delta L \approx \frac{1}{2\mu_c L_j} (k_{m,j+1}^2 - k_{m,j}^2) \approx \frac{k_{m,j+1} - k_{m,j}}{\mu_c^{1/2}}, \tag{3.33}$$

and so the intersection of the j and the $j + 1$ curves occurs at $L = L_{\text{int}} \equiv L_j + (\frac{1}{2}\Delta L)$, i.e. midway between L_j and L_{j+1} , as can be checked by solving

$$\frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \left(k_{m,j}^2 - \mu_c \right)^2 = \frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \left(k_{m,j+1}^2 - \mu_c \right)^2, \tag{3.34}$$

where we have expanded the j curve about L_j and the $j+1$ curve about L_{j+1} , respectively. Likewise, the growth rate at the intersection can be computed, and is given by

$$\lambda_{\text{int}} = \frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \mu_c^2 \left(\frac{k_{m,j+1}^2 - k_{m,j}^2}{k_{m,j+1}^2 + k_{m,j}^2} \right)^2 \approx \frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \mu_c \frac{(k_{m,j+1} - k_{m,j})^2}{L_j^2}. \quad (3.35)$$

Since $k_{m,j+1} - k_{m,j} \approx \pi / \sin \gamma_{m,j}$, we conclude that in the large-cell limit we must have

$$\left. \begin{aligned} \Delta L &= \frac{\pi}{\mu_c^2 \sin \gamma_{m,j}}, \\ \lambda_{\text{int}}(L_j) &= \frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \frac{\mu_c \pi^2}{\sin^2 \gamma_{m,j} L_j^2}. \end{aligned} \right\} \quad (3.36)$$

In fact, we even expect the $\lambda(L)$ curve to be well approximated by $\lambda_{\text{int}}(L)$, i.e. L_j is replaced by a general value of L in (3.36), because of the monotonicity of $\lambda(L)$ appropriate for the particular boundary conditions imposed. This clearly shows that the most unstable solution for a cylindrical cell with a rigid conducting side is given by

$$\lambda(L) = \frac{1}{2} \frac{\partial^2 \lambda}{\partial \mu_c^2} \frac{\mu_c \pi^2}{L^2}, \quad (3.37)$$

which corresponds to a very small m , because then $\gamma_{m,j} \rightarrow \frac{1}{2}\pi$. But we must quickly add that this does not mean that the $m = 0$ mode is necessarily the most unstable mode! Our argument cannot resolve the stability of the small- m modes. The best we can say is that m/j for the most unstable configuration tends to zero in the large-aspect-ratio limit. To our knowledge, numerical confirmation of this prediction is still lacking at this time, though interesting results have already been reported by Charlson & Sani (1971).

To understand why in general we still expect to see the same scaling behaviour when the sidewall properties are different from the case we have just studied, we can start with this known solution and perturb the boundary conditions backward. This time the unperturbed curve $\lambda(L)$ is smooth, as opposed to the analytically solvable case in which the most unstable mode exhibits a strong oscillatory behaviour and has a kink at every intersection, and we conclude that the effect of the perturbation is at most of order $O(L^{-3})$, because all the practically anchored states are perturbed by $O(L^{-3})$ only. This is true for perturbations of the type given in (2.13) and (2.14) when α and β are tuned from $+\infty$ to any finite positive values.

The implication of this argument is: if we turn on the perturbation from the analytically solvable case ($\alpha = \beta = 0$ on the sidewall) by a small amount, say $\partial\theta/\partial n + \beta\theta = 0$ for some very small but fixed $\beta > 0$, then there must be a cross-over size L_{c-0} above which $\lambda(L) = \lambda_{\text{int}}(L) + O(L^{-3})$ is valid. In other words, no matter how small the perturbation is, not only will the large- L section of the $\lambda(L)$ curve never touch the L -axis again, but it will also be replaced by a smooth curve that is practically independent of the strength of the perturbation. That smooth curve is well represented by (3.37). Of course, L_{c-0} will depend sensitively on how strong the perturbation is – the smaller the perturbation, the larger L_{c-0} becomes.

All the previous predictions concerning the behaviour of λ for a cylindrical cell can be easily translated in terms of R_c . What is worth emphasizing is that our argument is essentially one-dimensional since m does not introduce any uncontrollable complications, and, thus, not surprisingly, the system bears all the features of the

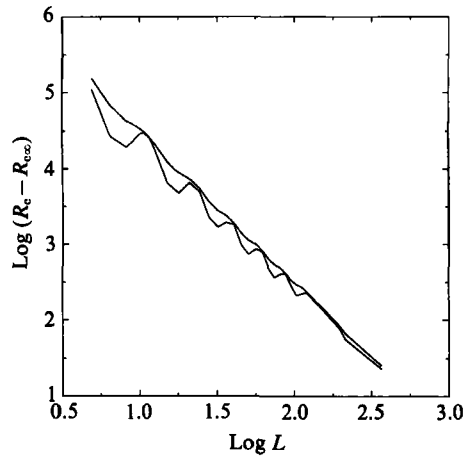


FIGURE 6. $\text{Log}(R_c - R_{c\infty})$ vs. $\text{log}(L)$ for a rigid cylindrical cell with conducting sides (top curve) and insulating sides (wavy curve). Data taken from Charlson & Sani (1970).

example we study in detail in Appendix B. Interestingly, this prediction is verified in Charlson & Sani's (1970) numerical calculation, though they did not address this scaling behaviour. In figure 6 we show the log-log plot of $R_c - R_{c\infty}$ vs. L for the axisymmetric state, with data taken from their tables 1–3. As can be seen, the slope is 2, an unmistakable signature of the L^{-2} scaling. In fact, analysis based on the amplitude equation also gives the same prediction (Ahlers *et al.* 1981; Cross *et al.* 1983), a result that should be contrasted with that derived by Chana & Daniels (1989) for a rigid rectangular channel.

For a cell Ω of arbitrary geometry we can find two concentric cylindrical cells Ω_1 and Ω_2 such that $\Omega_1 \subset \Omega \subset \Omega_2$. Suppose the three cells all have a rigid and perfectly conducting sidewall, then the above analysis for cylindrical cells and the fact that a smaller cell with this sidewall property necessarily is more stable than a bigger one imply that $R_c - R_{c\infty}$ for Ω must obey the same L^{-2} scaling law. (As an aside, we note that Ostrach & Pnueli 1963 actually applied the same kind of comparison theorem to compute bounds for convection cells of more practical configurations.) To derive the same conclusion for a rigid cell with an imperfect conducting side, we can perturb the sidewall parameter β from $\beta = +\infty$. If Ω does not possess any symmetry that causes degeneracy in this (unperturbed) most stable configuration, then the effect of boundary perturbation is expected to give rise to an $O(L^{-3})$ effect, and our basic conclusion will not change. The same idea can also be applied to a convection cell with symmetries once we realize that the annoying degeneracy due to the intrinsic symmetry of the problem can be avoided if we treat each class of symmetric solutions separately. Therefore, it is expected that the L^{-2} scaling and the immediate destruction of the large- L wavy 'tail' of the $\lambda(L)$ curve under perturbation from the analytically solvable case in which the sidewall is rotation-free and insulating are generally true.

As a final remark, we observe that the L^{-2} scaling we derived above together with (3.17) suggest that the temperature (and velocity as well, if we include an α term in it to make it completely general) must drop to at least $O(L^{-1})$ near the sidewall for general boundary conditions. Although we only derived (3.17) for a rectangular cell, it clearly holds for cells of other shapes as long as the sidewalls do not become too rugged. This can be seen by perturbing the boundary of a rectangular cell: one can

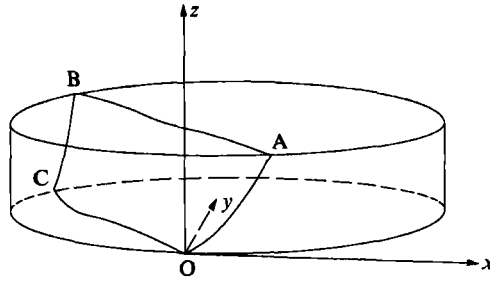


FIGURE 7. Local Cartesian coordinates x, y, z are chosen such that the origin O is at the intersection of temperature nodal surface $OABC$ and the lower edge of the convection cell, and the x -axis is tangent to the sidewall.

derive formulae similar to (3.13) to show that the perturbation to the eigenstates is always of $O(L^{-2})$ near the sidewall and cannot alter our observation. Again, the amplitude equation for a one-dimensional problem gives the same prediction (Cross *et al.* 1983). Our analysis, therefore, not only gives an independent verification to that approach, but also predicts that this phenomenon is generally true for a three-dimensional convection cell.

4. Convective rolls near the sidewalls

It is observed that in the visualization of stationary convection patterns using the shadowgraphic technique the convective rolls tend to align themselves normal to the sidewalls (Croquette 1989*a, b*; Croquette, Mory & Schosseler 1983). However, the interpretation of this observation is not immediately clear because the light rays projected on the screen above the cell to form the observed patterns necessarily have to go through a series of focusing and defocusing processes caused by density variation when they pass through the entire fluid layer. A sensible definition of a roll that is theoretically simple and experimentally justifiable does not seem to exist at this moment. With this precaution in mind, and in view of the fact that intuitively we know the density variations on the two sides of a convective roll, whatever its definition may be, must be opposite in sign to account for the rising and sinking fluid motion about the roll axis, we have decided to take a simplistic approach of identifying the nodal surfaces of the density variation with typical convective rolls. This necessarily makes this section somewhat controversial, though we believe it still offers some insight into what is observed in the laboratories. By the assumed linearity between temperature and density, this means the nodal surfaces of θ will represent convective rolls in our convention.

For a time-independent fully nonlinear solution, (2.2) reduces to $\nabla^2\theta = 0$ on the boundary, provided the sidewall is rigid, a condition we shall assume throughout this section. For the given nodal surface $OABC$ of θ in figure 7 we choose a local Cartesian coordinate such that the origin O is on the lower edge of the container and x is parallel to the sidewall. Because $\theta = 0$ at $z = 0$, we can expand θ in powers of z and write

$$\theta = fz + gz^3 + O(z^4), \quad (4.1)$$

where f and g are functions of x and y , and we have used the fact $\nabla^2\theta = 0$ at $z = 0$ to eliminate the $O(z^2)$ term. Therefore, we see that to the lowest order the nodal surfaces of θ are given by $z = 0$ (the bottom plate) and $f(x, y) = 0$. Notice that the

latter implies that OABC is orthogonal to the bottom plate. Similarly it must be orthogonal to the top plate.

Next, we investigate the behaviour of OABC near the point O. Expanding $f \equiv ax + by + (\text{higher-order terms})$ and imposing the thermal boundary condition $\partial\theta/\partial n + \beta\theta|_{y=0} = 0$ one obtains $b = 0$. Thus to the lowest order $\theta = axz$ near the origin. The case $a \neq 0$ implies OABC is also orthogonal to the sidewall near point O. Parallelism of nodal surfaces to the sidewall corresponds to a non-generic case in our analysis.

It is clear that the nodal surfaces still behave the same if the system is evolving according to linear stability analysis. The argument is also valid for binary fluid convection when the Dufour effect can be neglected, because then the thermal equation is the same as (2.2).

The implication of the previous argument seems to be: if the nodal surfaces of θ more or less correspond to the convective rolls one observes in experiments, then, so long as the boundary condition (2.14) is a good approximation to the real sidewall, the orthogonality of rolls to the sidewalls is a simple consequence of the imposed realistic boundary conditions which render the governing equations linear close to the boundary. We should also point out that our orthogonality argument is compatible with the analysis done by Cross (1982) that is based on the amplitude equation approach. Although his conclusion is that it is possible to have rolls that end at any angle to the sidewall (when viewed a distance away from the wall), it is also suggested that the excitation of the conjugate rolls at the wall is a necessary consequence of the imposed boundary conditions. In fact, the interference at the sidewall between the slowly decaying conjugate rolls $\sin(k_1 x - k_2 y)$ and the incident rolls $\sin(k_1 x + k_2 y)$ simply generates a cellular pattern $\sin k_1 x \cos k_2 y$ whose nodal surfaces are orthogonal to the sidewall. The agreement of both approaches indicates that the range of validity of the orthogonality is at least of order $O(1)$. However, we must admit that our treatment is limited because it cannot explain rigorously why experiments using Doppler anemometry or other visualization techniques still observe rolls ending perpendicular to the sidewalls (Gollub & Steinman 1981; Bergé & Dubois 1980; Somerscales & Dropkin 1966), though intuitively we feel that the correlation between u_z and θ should imply the orthogonality of the former to the sidewalls if the latter exhibits this feature. As pointed out by one referee, the situation is even more complicated in binary fluid convection since the concentration can also change the index of refraction of light. Clearly, more theoretical investigation is awaiting us before this simple-looking phenomenon can be fully explained and understood.

5. Conclusion

We have demonstrated that the combination of variational principles and perturbation theory can help us understand the linear stability of pure-fluid convection in a very general manner. When the geometry of the convection cell is fixed, sidewall properties affect the stability in a monotonic way. We also show that the effects of system size on the stability can be satisfactorily analysed via the existence of virtually anchored states when the cell has high symmetries. In particular, we have applied this idea to a cylindrical cell to demonstrate why the critical Rayleigh number must approach its infinite-cell limit through an L^{-2} scaling law. The same conclusion holds for a cell of other geometry so long as the curvature of its side scales like L^{-1} . This analysis is facilitated by two facts: the first is that the

variational principles allow us to compare the stability of two cells of different shapes for certain physically realizable boundary conditions, and the second is that the equally powerful boundary perturbation technique bridges the gap left untreated by the former. The method we introduce, though unconventional amongst the fluid dynamics community, is reminiscent of the geometrical concept of 'Berry's phase' whose theoretical and experimental importance in quantum mechanics is under intensive investigation. One merit of our method is that it links together the convection problem and another canonical pattern formation model, the Turing mechanism, and shows that both systems exhibit similar features. We also suggest the possibility of identifying convective rolls with the nodal surfaces of the temperature field θ , and offer a simple explanation for why convective rolls tend to bend normally toward the sidewall. It is our hope that, through this detailed study of the boundary effects on the onset convective instability, the role of the bounding walls is further clarified so that future work on other more complex problems, such as rotating convection, will reveal the mutual interaction between the boundary and the extra complications introduced in a more comprehensible way.

I would like to thank Professor M. C. Cross for the many stimulating and helpful discussions.

Appendix A

In this Appendix we will show how (2.19) can be derived from the well-known commutation relation between the momentum operator $\mathbf{P} \equiv (1/i)\nabla$ and the position operator \mathbf{r} in quantum mechanics, thus establishing a link between it and the famous Heisenberg's uncertainty principle. The Hilbert space consists of 'wavefunctions' g defined in a finite domain Ω which satisfy the Neumann boundary condition. (This restriction will be lifted after the main idea of the proof is introduced.) First, we note that for any vector-valued potential energy operator $V(\mathbf{r}) \equiv \mathbf{n}V(\mathbf{r})$, where \mathbf{n} is a unit vector field and V is a scalar function, the commutation relation becomes

$$\mathbf{V} \cdot \mathbf{P} - \mathbf{P} \cdot \mathbf{V} = i\nabla \cdot \mathbf{V}. \quad (\text{A } 1)$$

If we now operate (A 1) on g and take the inner product of it with g , then we obtain

$$i\langle g | \nabla \cdot \mathbf{V} | g \rangle = \langle g | \mathbf{V} \cdot \mathbf{P} | g \rangle - \langle g | \mathbf{P} \cdot \mathbf{V} | g \rangle = \langle \mathbf{V}g | \mathbf{P}g \rangle - \langle \mathbf{P}g | \mathbf{V}g \rangle.$$

We also have

$$i\langle g | \nabla \cdot \mathbf{V} | g \rangle = i\langle g | \mathbf{n} \cdot \nabla V | g \rangle + i\langle g | \mathbf{V} \nabla \cdot \mathbf{n} | g \rangle$$

by definition. Together, they imply

$$\begin{aligned} |i\langle g | \mathbf{n} \cdot \nabla V | g \rangle| &= |-i\langle g | \mathbf{V} \nabla \cdot \mathbf{n} | g \rangle + \langle \mathbf{V}g | \mathbf{P}g \rangle - \langle \mathbf{P}g | \mathbf{V}g \rangle| \\ &\leq |\langle g | \mathbf{V} \nabla \cdot \mathbf{n} | g \rangle| + |\langle \mathbf{V}g | \mathbf{P}g \rangle| + |\langle \mathbf{P}g | \mathbf{V}g \rangle| \\ &\leq |\langle g | \mathbf{V} \nabla \cdot \mathbf{n} | g \rangle| + 2\langle \mathbf{V}g | \mathbf{V}g \rangle^{\frac{1}{2}} \cdot \langle \mathbf{P}g | \mathbf{P}g \rangle^{\frac{1}{2}}, \end{aligned}$$

where the last step is by Cauchy-Schwarz inequality. Written out explicitly, it becomes

$$\left| \int_{\Omega} g^2 \mathbf{n} \cdot \nabla V \right| \leq \left| \int_{\Omega} g^2 \mathbf{V} \nabla \cdot \mathbf{n} \right| + 2 \left(\int_{\Omega} V^2 g^2 \right)^{\frac{1}{2}} \left(\int_{\Omega} |\nabla g|^2 \right)^{\frac{1}{2}}. \quad (\text{A } 2)$$

Suppose S is a surface that divides Ω into an interior Ω_1 and an exterior region Ω_2 . Choose \mathbf{n} to be a vector field such that the outward normal on S coincides with it, and define V to be identically zero in Ω_1 and unity otherwise; then after dividing (A 2) through by $\int_{\Omega} g^2$ we arrive at

$$\frac{\int_S g^2}{\int_{\Omega} g^2} \leq 2 \langle V^2 \rangle^{\frac{1}{2}} \left(\frac{\int_{\Omega} |\nabla g|^2}{\int_{\Omega} g^2} \right)^{\frac{1}{2}} + |\langle V \nabla \cdot \mathbf{n} \rangle|, \tag{A 3}$$

which is essentially (2.19), because the average values $\langle V^2 \rangle$ and $|\langle V \nabla \cdot \mathbf{n} \rangle|$ evidently are bounded above by some constants which are independent of g . Note that imposing the Neumann boundary condition on g is unessential for (A 3) to hold. This is because under the surgery of flattening it out near $\partial\Omega$ any function can be easily made to satisfy the Neumann boundary condition, and the inequality expressed in (A 3) is not modified by this procedure. Equation (2.19) is just a limiting case of (A 3) when $S \rightarrow \partial\Omega$. We should also point out that the bounding constant c_2 can become very large when the domain Ω is small. This is especially clear from (A 3), because c_2 is proportional to the integral of $\nabla \cdot \mathbf{n}$, which itself is proportional to the mean curvature of the surface $\partial\Omega$. Also notice that if we take \mathbf{n} as \hat{e}_x , the unit vector in the x -direction, and V as x , then (A 2) gives us a weaker version of the uncertainty principle which can be cast into the more familiar form if P and V are replaced by $P - \langle P \rangle$ and $V - \langle V \rangle$, respectively.

Appendix B

In this Appendix we will study a one-dimensional problem in detail to illustrate the ideas proposed in the main text. Sophistication of the model will increase as we proceed.

We start with a cell Ω of dimension L that is completely free-slip and has perfectly insulated sides at $x = \pm \frac{1}{2}L$. The two horizontal plates are located at $z = \pm \frac{1}{2}$. The solution for this problem is

$$\lambda = \frac{-1 - \sigma \pm \{(1 + \sigma)^2 - 4\sigma[1 - (R/\mu^2)(1 - \mu_z/\mu)]\}^{\frac{1}{2}}}{2} \cdot \mu, \tag{B 1}$$

$$\left. \begin{aligned} \theta &= \sin n\pi(z + \frac{1}{2}) \cos j\pi\left(\frac{x}{L} + \frac{1}{2}\right), \\ u_z &\propto -\frac{j}{nL} \sin n\pi(z + \frac{1}{2}) \cos j\pi\left(\frac{x}{L} + \frac{1}{2}\right), \\ u_x &\propto \cos n\pi(z + \frac{1}{2}) \sin j\pi\left(\frac{x}{L} + \frac{1}{2}\right), \\ p &\propto \cos n\pi(z + \frac{1}{2}) \cos j\pi\left(\frac{x}{L} + \frac{1}{2}\right), \end{aligned} \right\} \tag{B 2}$$

where $\mu = ((n\pi)^2 + (j\pi/L)^2)$ and $\mu_z = (n\pi)^2$ for integers n and j . We see from (B 1) that it is possible to have $\partial\lambda/\partial R < 0$, although the most unstable state clearly satisfies $\partial\lambda/\partial R > 0$.

For illustration, we will fix $n = 1$ and $R = R_{c0} \equiv \frac{27}{4}\pi^4$. The resulting λ vs. L curves with different j are shown in figure 8. For any given j , λ vanishes when $L = \sqrt{2}j$. The

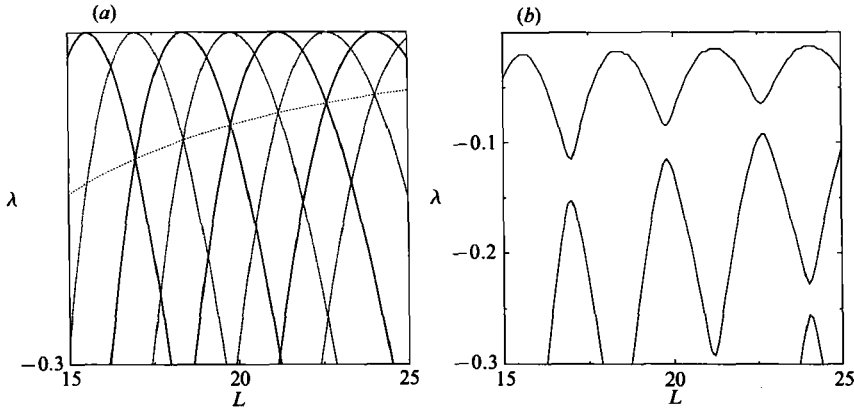


FIGURE 8. (a) Growth rates λ vs. size L for a free-slip insulated one-dimensional cell are shown as solid or dashed, according to whether their symmetry under $x \rightarrow -x$ is even or odd, when $R = R_{c0}$. The dotted curve, given by (B 3), connects intersections of curves j and $j+2$ ($j = \text{integer}$) that have the same symmetry. Biological pattern formation models show the same features. (b) Perturbing the boundary condition in general removes the degeneracies of (a) and forms many wavy curves.

intersection of curves j and $j+2$ which have the same symmetry under $x \rightarrow -x$ can be easily computed from (B 1), and is given by

$$\lambda_{\text{int}}(L_{\text{int}}) = -\frac{4\pi^2\sigma}{1 + \sigma L_{\text{int}}^2} \tag{B 3}$$

for large L_{int} , the size of the cell when degeneracy occurs. (The curve $\lambda_{\text{int}}(L)$ is shown dotted in figure 8.) The difference in cell sizes between two intersections (j and $j+2$, and $j+2$ and $j+4$) is $\sqrt{2}$.

Also notice that in each degenerate two-dimensional subspace spanned by ψ_j and ψ_{j+2} we can find a vector $\psi \equiv \psi_j - \psi_{j+2}$ such that the associated

$$\theta \equiv \theta_j - \theta_{j+2} \propto (\cos j\pi(x/L + \frac{1}{2}) - \cos (j+2)\pi(x/L + \frac{1}{2}))$$

vanishes identically at $x = \pm \frac{1}{2}L$. Hence $\theta = \partial\theta/\partial n = 0$, and the solution satisfies $\partial\theta/\partial n + \beta\theta = 0$ on $\partial\Omega$ for any β , i.e. perturbation of β has no effect on ψ at all! These ‘anchored’ solutions are clearly manifested in figure 9(c) when translated in terms of an R_c vs. L plot. Similarly, one can find anchored solutions when the viscous boundary condition is changed, though they will be different from the ones constructed above. The critical Rayleigh number at $L = L_{\text{int}}$ can also be calculated to give

$$R_c \approx R_{c0} + 18 \frac{\pi^4}{L^2}. \tag{B 4}$$

Next, we consider boundary perturbation of the form $\partial\theta/\partial n + \beta\theta = 0$. To be specific, we will consider θ that is odd in x . (Even solutions can be obtained by the replacement rule $k_j l \rightarrow k_j l + \frac{1}{2}\pi$.) The problem is reduced to the following determinantal equation:

$$\det \begin{pmatrix} \beta \sin k_1 l + k_1 \cos k_1 l & \beta \sin k_2 l + k_2 \cos k_2 l & \beta \sin k_3 l + k_3 \cos k_3 l \\ (\lambda + \pi^2 + k_1^2) k_1 \cos k_1 l & (\lambda + \pi^2 + k_2^2) k_2 \cos k_2 l & (\lambda + \pi^2 + k_3^2) k_3 \cos k_3 l \\ \frac{(\lambda + \pi^2 + k_1^2)}{k_1} \cos k_1 l & \frac{(\lambda + \pi^2 + k_2^2)}{k_2} \cos k_2 l & \frac{(\lambda + \pi^2 + k_3^2)}{k_3} \cos k_3 l \end{pmatrix} = 0 \tag{B 5}$$

where $l \equiv \frac{1}{2}L$ and k_m^2 ($m = 1, 2, 3$) are solutions to the dispersion relation

$$(\lambda + \pi^2 + k_m^2)(\lambda + \sigma(\pi^2 + k_m^2))(\pi^2 + k_m^2) - R\sigma k_m^2 = 0, \quad (\text{B } 6)$$

and θ is some suitable linear combinations of $\sin \pi(z + \frac{1}{2}) \sin k_j x$. To simplify the algebra, we will consider the problem of determining R_c as a function of l . We can easily show that for small deviation of $\Delta R \equiv R_c - R_{c0}$, the k_m^2 are given by

$$k_{1,2} \approx \frac{\pi}{\sqrt{2}} \left(1 \pm \frac{(\Delta R)^{\frac{1}{2}}}{3\pi^2} + \frac{7}{162} \frac{\Delta R}{\pi^4} \right), \quad (\text{B } 7)$$

$$k_3 \approx 2i\pi \left(1 + \frac{2}{81} \frac{\Delta R}{\pi^4} \right). \quad (\text{B } 8)$$

For $\Delta R > 0$, the lowest-order solution to (B 5) can be obtained by (i) dividing through the third column of the determinant by $\cos k_3 l$ and approximating $\tan k_3 l \approx i$ for large l , (ii) allowing all the k_m that do not appear as arguments of sinusoidal functions to assume their unperturbed values. The result is $\beta \sin(k_2 - k_1)l = 0$, or $\Delta R = 18\pi^4/L^2$, provided β is not too small. (We are extracting the leading order of (B 5). Hence, the case when β is very small must be handled separately. This is the first indication of the existence of the cross-over size L_{c-0} we discuss in §3.) Notice that this lowest-order solution coincides with that given by (B 4)! As a matter of fact, this result can also be predicted using a two-scale amplitude equation (Cross *et al.* 1983). The agreement between the two methods is not too surprising once we realize that both approaches employ the same global phase drift $(k_2 - k_1)l$ to ensure that the correct boundary conditions are satisfied to the lowest order. To get the next-order correction, we can rewrite (B 5) as $\beta \sin(k_2 - k_1)l = f$ for some function f that also involves sinusoidal functions itself. Then one replaces the k_m contained in f by their first-order (correct to $O(\Delta R)^{\frac{1}{2}}$) values so that the equation is solved by $(k_2 - k_1)l = \pi - \phi_0 \equiv \pi - \sin^{-1}f/\beta$, or

$$\Delta R = \frac{18\pi^2(\pi - \phi_0)^2}{L^2} \quad (\text{B } 9)$$

for some complicated function ϕ_0 that also involves L . Correct to $O(L^{-1})$ we find that ϕ_0 is given by

$$\phi_0 = \sin^{-1} \left(- \left(\frac{\sqrt{2}}{9} \sin k_1 l + \left(\frac{4}{9} + \frac{2\pi}{\beta} \right) \cos k_1 l \right) \frac{\cos k_2 l}{l} \right). \quad (\text{B } 10)$$

The point to be made is that for any fixed β , the oscillatory behaviour as carried by ϕ_0 is always $O(L^{-1})$ smaller than the 'background' L^{-2} decay law given by (B 4), thus confirming our claim in §3. Figure 9 shows the comparison between (B 9) and numerical calculations.

For any given small positive β , the cross-over size can be estimated by requiring ϕ_0 to be well-defined for all $L > L_{c-0}$. Thus,

$$L_{c-0} \approx \frac{4\pi}{\beta}. \quad (\text{B } 11)$$

This agrees with our observation in the main text that, for a given positive β , the stability curve $R_c(L)$ must be monotonic and independent of the perturbation strength for large enough L , as is predicted by (B 4), though the concept of largeness

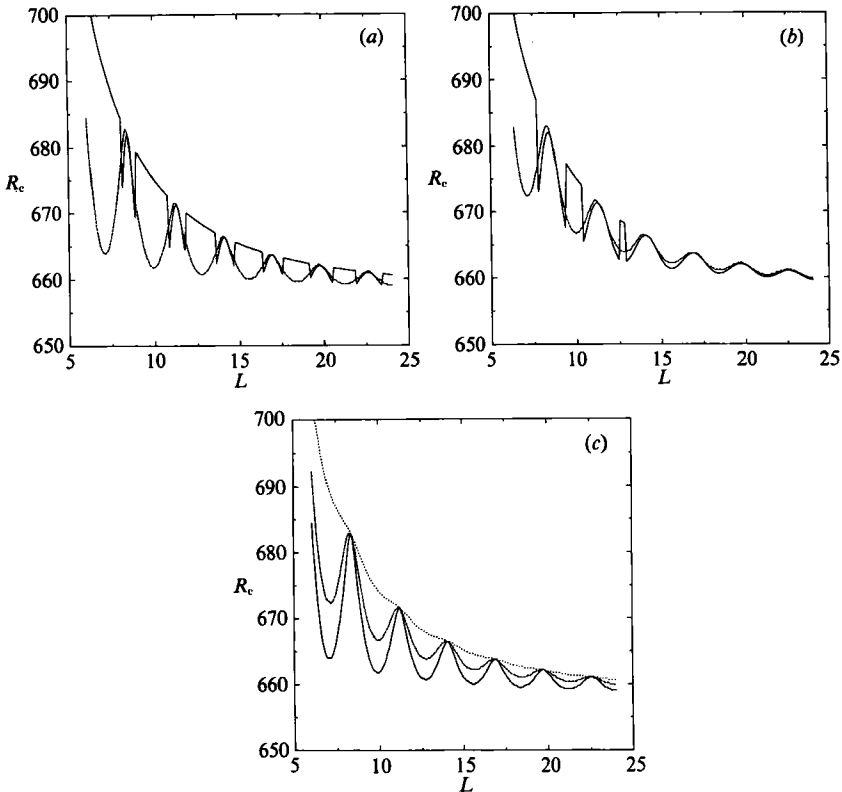


FIGURE 9. (a), (b) The prediction of (B 9) (solid lines) and numerical results (dashed) for $\beta = 0.3$ and 1, respectively. ϕ_0 is set to zero whenever it is not defined for a given L . Solutions for $\beta = 0.3$ (solid), 1 (dashed) and 50 (dotted) are displayed in (c) to show the existence of ‘anchored’ states.

for L must be measured relative to the cross-over size given by (B 11), which predicts its sensitive dependence on β when β is small.

The R_c vs. L plot for solutions whose θ is even under $x \rightarrow -x$ looks very similar to that for odd θ when $\beta > 0$. However, their respective runaway solutions look entirely different once β is tuned negative. This is shown in figure 10. The detachment of an even solution from the R_c axis immediately after β turns negative agrees with our observation in §2 that the variational principle associated with $J[\psi]$ might encounter difficulty when β is negative, especially if the system is small. The odd branch behaves differently simply because its symmetry requires excitations in the x -direction, thus contributing a non-zero μ_{\parallel} in (2.39) to compensate the size effect coming from c_2 . (Notice that μ_{\parallel} scales like L^{-2} .) Specifically, the intersection of the runaway curve with the L -axis for the current problem is given by

$$L_{\text{odd}} = -\frac{2}{\pi} \tanh^{-1} \left(\frac{\pi}{\beta} \right), \quad L_{\text{even}} = -\frac{2}{\pi} \tanh^{-1} \left(\frac{\beta}{\pi} \right). \quad (\text{B } 12)$$

Notice that L_{even} is proportional to $-\beta$ for small negative β , a result that is expected from the discussion at the end of §2. Likewise, the largest possible $-\beta$ allowed for the intersection between the runaway curve and the L -axis not to happen is in inverse proportion to L_{odd} when L_{odd} is small, which is also correctly predicted by (2.39).

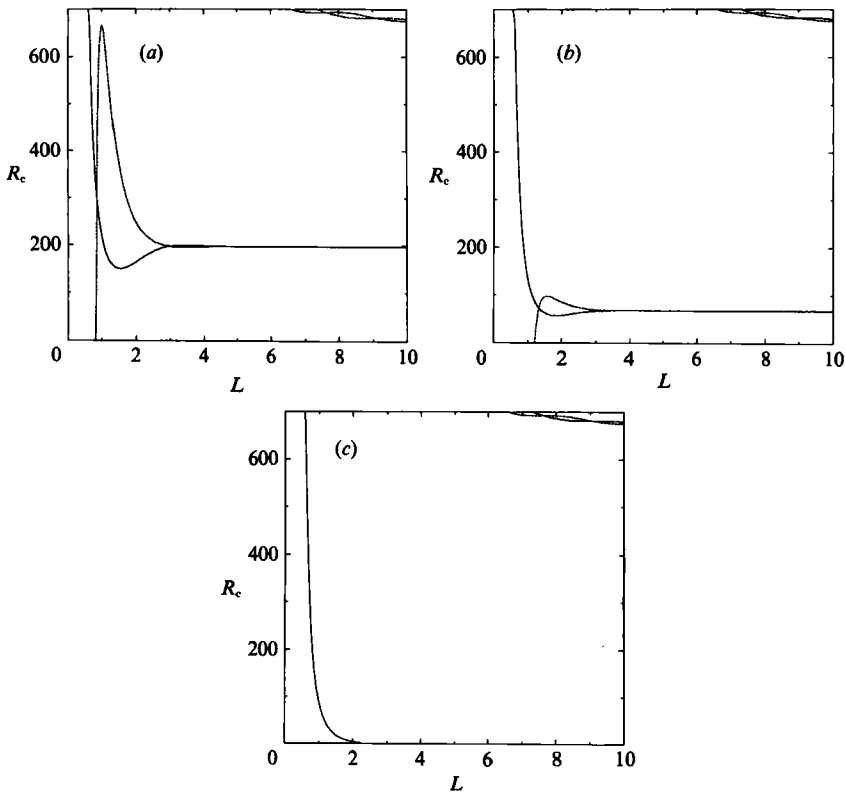


FIGURE 10. Snapshots of the runaway solutions when (a) $\beta = -2.7$, (b) -3 , and (c) -3.145 . Solid curves show odd modes; dashed curves, even modes.

It is also easy to compute the solutions to (B 5) and (B 6) when R is fixed while β is negatively very large. We find

$$\left. \begin{aligned} k_1^2 \approx -\lambda, \quad k_2^2 \approx -\frac{\lambda}{\sigma}, \quad k_3^2 \approx -\pi^2, \\ i\beta + k_1 \approx 0. \end{aligned} \right\} \quad (\text{B } 13)$$

To lowest order the associated eigenvector is given by

$$u_x \approx u_z \approx 0, \quad \theta \approx \sin \pi z \frac{\sin k_1 x}{\sin k_1 l}. \quad (\text{B } 14)$$

This shows clearly that the runaway solution $\lambda^{\frac{1}{2}} = -\beta$ develops a thermal boundary layer of thickness $-\beta^{-1}$ and strongly suppresses the velocity field.

It is obvious that all the previous results we have obtained through detailed analysis of a cell with free-slip sidewalls must similarly hold for a cell with rigid sidewalls because the methods we used are equally valid for the latter. Instead of repeating the same analysis, we simply show in figure 11 the numerical calculations as further evidence supporting our claims.

As a final example, we present the numerical results of our study of a cell whose α on the two horizontal plates is also tunable. ($\alpha = \infty$ corresponds to a rigid-rigid cell.) Again, sidewall properties are tuned by (a different) α and β . The critical Rayleigh number is computed by maximizing $J[\psi]$ for ψ that is in a Hilbert subspace of finite dimension whose basis vectors θ and u are independently proportional to

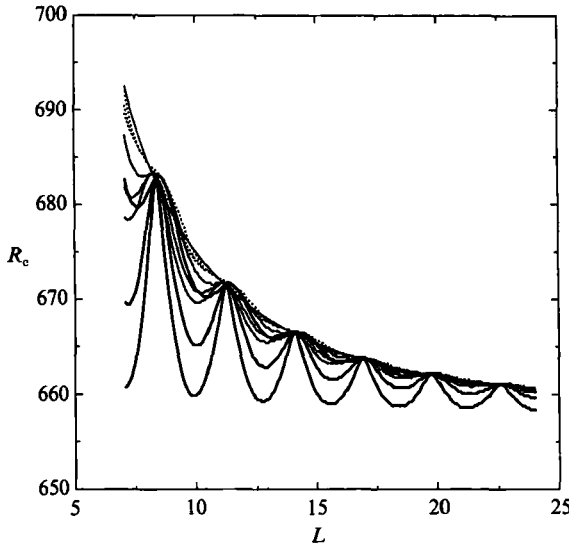


FIGURE 11. Solutions for a free-free cell with (α, β) taking all possible combinations from three values: 0.1, 2, and 50. The existence of the practically anchored states is obvious. The light solid monotonic curve is calculated from (B 4).

their corresponding variables given in (B 2), with 15 modes of odd j and five modes of odd n . It is easy to show that this approach is actually just a variant of the Galerkin expansion method one would use for solving a general linear problem in the sense that now we are simultaneously projecting both the governing equations and the boundary conditions onto the same conveniently chosen basis vectors. This particular projection scheme is favoured over others because it is manifestly equivalent to the variational principle, and the resulting generalized matrix eigenvalue problem is Hermitian. The results are shown in figure 12. One is quickly convinced of the existence of the practically anchored states from this figure. Again, the basic features are not different from those for a free-slip cell.

Appendix C

A simple sufficient condition for the method of separation of variables to hold for the convection problem is derived. This criterion is then shown to be satisfied for the ansatz described by (3.27), thus justifying the approach used in §3. The basic ingredient of our argument is contained in the following estimate for a Rayleigh-Ritz ratio:

LEMMA. Let $f(z)$ be a trial function defined on $[0, 1]$. Let Λ_1 denote the ground-state eigenvalue for $-\text{d}^2g/\text{d}z^2 = \Lambda g$, where g must satisfy $\pm \text{d}g/\text{d}z + \alpha g = 0$ at $z = 0, 1$ for some non-negative constant α . Then for any f that is not identically zero we have

$$\frac{\alpha(f^2(0) + f^2(1)) + \int_0^1 \left(\frac{\text{d}f}{\text{d}z}\right)^2 \text{d}z}{\int_0^1 f^2 \text{d}z} \geq \Lambda_1 \geq \left(\frac{-\alpha + (\alpha^2 + 8\pi^2\alpha)^{1/2}}{4\pi}\right)^2. \tag{C 1}$$

Before proving it, we note that the first inequality is the standard variational characterization for the eigenvalue Λ_1 . Therefore, we only need to verify the second

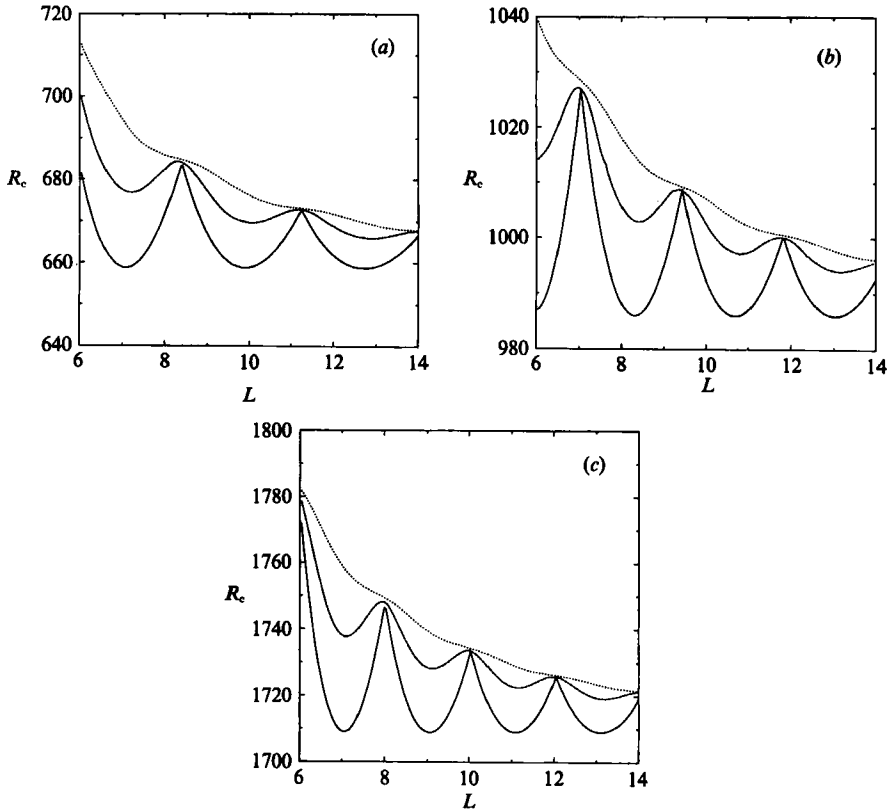


FIGURE 12. Solutions for a cell with sidewall $\alpha = \beta$ taking three values: 0 (solid), 1 (dashed) and 15 (dotted). The two horizontal plates have (a) $\alpha = 0$, (b) 4 and (c) 50, respectively. The practically anchored states are clearly exhibited in this figure.

inequality. Let $k \equiv A_1 \in [0, \pi]$, then the boundary condition implies $\alpha = k \tan \frac{1}{2}k$. by convexity of cosine function in the interval $[0, \frac{1}{2}\pi]$, we know $\cos y \geq (1 - 2y/\pi)$ for all $y \in [0, \frac{1}{2}\pi]$. Hence,

$$\int_0^{\frac{1}{2}k} \frac{1}{(1 - 2y/\pi)^2} dy \geq \int_0^{\frac{1}{2}k} \frac{1}{\cos^2 y} dy,$$

which implies, upon evaluating the integrals, that

$$\frac{(k/\pi)^2}{1 - (k/\pi)} \geq \frac{2}{\pi^2} k \tan \frac{1}{2}k \equiv \frac{2}{\pi^2} \alpha.$$

Equation (C 1) is then obtained by simply inverting this inequality and noting that $A_1 \equiv k^2$.

Notice that (C 1) yields the following well-known assertion when applied to the case $\alpha = \infty$:

COROLLARY. Any non-zero trial function f which satisfies the boundary condition $f(0) = f(1) = 0$ always satisfies

$$\frac{\int_0^1 \left(\frac{df}{dz}\right)^2 dz}{\int_0^1 f^2 dz} \geq \pi^2. \tag{C 2}$$

More generally, the previous two assertions are still true if f is a function defined on a three-dimensional cell Ω of constant horizontal cross-section, provided we substitute df/dz by ∇f , $f^2(0) + f^2(1)$ by $\int_{\partial\Omega} f^2$, and $\int_0^1 dz$ by \int_{Ω} , respectively.

Suppose the sidewall is perfectly insulating and rotation-free, i.e., $\alpha = \beta = 0$ on the sidewall S . Then we can expand both θ and u_z in an orthonormal basis φ_j that are eigenfunctions of the two-dimensional Neumann Laplacian on the horizontal cross section: $-\nabla_h^2 \varphi_j = \mu_j \varphi_j$. Of course, the 'Fourier coefficients' will then be functions of z . Now if we look at (2.9) it is clear that in order to decouple all the φ_j modes (so that we do have separation of variables) we must be able to write the horizontal component \mathbf{u}_h of the fluid velocity as linear combinations of $\nabla_h \varphi$. (The pressure p is automatically expandable in φ_j .) This apparently implies that the vertical vorticity ω must vanish. Thus, we are led to investigate when we can expect ω to vanish identically.

Taking the curl of the Navier–Stokes equation and projecting it on the $\hat{\mathbf{e}}_z$ direction we obtain

$$\lambda\omega = \sigma\nabla^2\omega. \quad (\text{C } 3)$$

Assume ω is not identically zero, then upon multiplying (C 3) by ω and integrating over the fluid cell Ω we get

$$\lambda = -\sigma \frac{\int_{\partial\Omega \setminus S} \alpha \omega^2 + \int_{\Omega} |\nabla\omega|^2}{\int_{\Omega} \omega^2} \leq \sigma A_1 \leq -\sigma \left(\frac{-\alpha + (\alpha^2 + 8\pi^2\alpha)^{1/2}}{4\pi} \right)^2. \quad (\text{C } 4)$$

In arriving at this result we have performed integration by parts, imposed the boundary conditions of (2.12) and (2.16) on $\partial\Omega$, and made use of (C 1) just derived. (Notice that the boundary term involves no contribution from the sidewall S , because ω vanishes on it by the rotation-free condition.) Hence, ω must vanish identically, thus permitting separation of variables, provided the growth rate of interest lies above the bound set up by (C 4). If we recall the derivation of the large-cell scaling behaviour of λ or R_c for positive α in §3, we immediately see that the argument presented there is well justified, because for all large enough size L the relevant growth rate λ considered there actually goes to zero, and therefore does lie above the bound given by (C 4).

We would also like to point out that had we started studying R_c directly without invoking λ , as opposed to what is done in §3, then all the orthonormal eigenfunctions would have zero vertical vorticity, and separation of variables certainly is valid if the sidewall is insulating and rotation free.

REFERENCES

- AHLERS, G., CROSS, M. C., HOHENBERG, P. C. & SAFRAN, S. 1981 *J. Fluid Mech.* **110**, 297.
 BERGÉ, P. & DUBOIS, M. 1980 In *Systems Far from Equilibrium* (ed. L. Garrido). Springer.
 BUELL, J. C. & CATTON, I. 1983 *Phys. Fluids* **26**, 892.
 BUSSE, F. H. 1985 In *Hydrodynamic Instabilities and the Transition to Turbulence* (ed. H. L. Swinney & J. P. Gollub). Springer.
 BUSSE, F. H. & RIAHI, N. 1980 *J. Fluid Mech.* **96**, 243.
 CARMO, M. P. DO 1976 *Differential Geometry of Curves and Surfaces*. Prentice-Hall.
 CHANA, M. S. & DANIELS, P. G. 1989 *J. Fluid Mech.* **199**, 257.
 CHANDRASEKHAR, S. 1954 *Am. Math. Mon.* **61**, 170.
 CHANDRASEKHAR, S. 1961 *Hydrodynamic and Hydromagnetic Stability*, pp. 27ff. Oxford University Press.

- CHARLSON, G. S. & SANI, R. L. 1970 *Intl J. Heat Mass Transfer* **13**, 1469.
- CHARLSON, G. S. & SANI, R. L. 1971 *Intl J. Heat Mass Transfer* **14**, 2157.
- CHEN, Y.-Y. 1991 Effects of boundaries on Rayleigh-Bénard convection. PhD thesis, California Institute of Technology.
- CHEN, Y.-Y. 1992 *Phys. Rev. A* **45**, 3727.
- COURANT, R. & HILBERT, D. 1966 *Methods of Mathematical Physics*, vol. I, chapter VI. Interscience.
- CROQUETTE, V. 1989a *Contemp. Phys.* **30**, 113.
- CROQUETTE, V. 1989b *Contemp. Phys.* **30**, 153.
- CROQUETTE, V., MORY, M. & SCHOSSELER, F. 1983 *J. Phys. Paris* **44**, 293.
- CROSS, M. C. 1980 *Phys. Fluids* **23**, 1727.
- CROSS, M. C. 1982 *Phys. Fluids* **25**, 936.
- CROSS, M. C. 1988 *Phys. Rev. A* **38**, 3593.
- CROSS, M. C., DANIELS, P. G., HOHENBERG, P. C. & SIGGIA, E. D. 1983 *J. Fluid Mech.* **127**, 155.
- DAVIS, S. H. 1967 *J. Fluid Mech.* **30**, 465.
- EDWARDS, B. F. 1988 *J. Fluid Mech.* **191**, 583.
- GOLLUB, J. P. & STEINMAN, J. F. 1981 *Phys. Rev. Lett.* **47**, 505.
- HALES, A. L. 1937 *Geophys. Suppl. Mon. Not. R. Astr. Soc.* **4**, 122.
- JOSEPH, D. D. 1976 *Stability of Fluid Motions*, vols. I & II. Springer.
- KOLODNER, P., BENSIMON, D. & SURKO, C. M. 1988 *Phys. Rev. Lett.* **60**, 1723.
- MORSE, P. M. & FESHBACH, H. 1968 *Methods of Theoretical Physics*. vol. I & II. McGraw-Hill.
- MURRAY, J. D. 1989 *Mathematical Biology*. Springer.
- NEWELL, A. C. & WHITEHEAD, J. A. 1969 *J. Fluid Mech.* **38**, 279.
- NIEMELA, J. J., AHLERS, G. & CANNEL, D. S. 1990 *Phys. Rev. Lett.* **64**, 1365.
- OSTRACH, S. & PNUELI, D. 1963 *Trans. ASME C: J. Heat Transfer* **85**, 346.
- PELLEW, A. & SOUTHWELL, R. V. 1940 *Proc. R. Soc. Lond. A* **176**, 312.
- PFOTENHAUER, J. M., NIEMELA, J. J. & DONNELLY, R. J. 1987 *J. Fluid Mech.* **175**, 85.
- PNUELI, D. & ISCOVICI 1967 *Israel J. Tech.* **5**, 243.
- SANI, R. L. 1964 *Z. Angew. Math. Phys.* **15**, 381.
- SEGEL, L. A. 1967 *J. Fluid Mech.* **30**, 625.
- SHAPER, A. & WILCZEK, F. (ed.) 1989 *Geometric Phases in Physics*. World Scientific.
- SHAUMEYER, J. N., BEHRINGER, R. P. & BAIERLEIN, R. 1981 *J. Fluid Mech.* **109**, 339.
- SOMERSCALES, E. & DROPKIN, D. 1966 *Intl J. Heat Mass Transfer* **9**, 1189.
- SOROKIN, V. S. 1953 *Prikl. Mat. Mekh.* **17**, 39.
- SOROKIN, V. S. & SUSHKIN, I. V. 1960 *Sov. Phys. JETP* **11**, 440.
- SPARROW, E. M., GOLDSTEIN, R. J. & JONSSON, V. H. 1964 *J. Fluid Mech.* **18**, 513.
- STEINBERG, V., MOSES, E. & FEINBERG, J. 1987 In *Chaos '87: Proc. Intl Conf. on the Physics of Chaos and Systems Far from Equilibrium, Monterey, CA, January 1987* (ed. M. Duong-Van). North-Holland.
- TURING, A. M. 1952 *Phil. Trans. R. Soc. Lond.* **B237**, 37.
- VERHOEVEN, J. D. 1969 *Phys. Fluids* **12**, 1733.
- YIH, C.-S. 1959 *Q. Appl. Maths* **17**, 25.
- ZHONG, F., ECKE, R. & STEINBERG, V. 1991 Asymmetric modes and the transition to vortex structures in rotating Rayleigh-Bénard convection. Preprint.