J. Fluid Mech. (1998), *vol.* 376, *pp.* 263–296. Printed in the United Kingdom

Bounds for heat transport in a porous layer

By CHARLES R. DOERING[1] AND PETER CONSTANTIN[2]

[1] Department of Mathematics, University of Michigan, Ann Arbor, MI 48109-1109, USA

[2] Department of Mathematics, University of Chicago, IL 60637, USA

(Received 23 February 1998 and in revised form 23 June 1998)

Bounds on convective heat transport in a porous layer heated from below are derived using the background field variational method (Constantin & Doering 1995*a*, *b*, 1996; Doering & Constantin 1992, 1994, 1996; Nicodemus, Holthaus & Grossmann 1997*a*) based on the technique introduced by Hopf (1941). We consider the infinite Prandtl–Darcy number model in three spatial dimensions, and additionally the finite Prandtl–Darcy number equations in two spatial dimensions, relevant for the related Hele-Shaw problem. The background field method is interpreted as a rigorous implementation of heuristic marginal stability concepts producing rigorous limits on the time-averaged convective heat transport, i.e. the Nusselt number Nu, as a function of the Rayleigh number Ra. The best upper bound derived here, although not uniformly optimal, matches the exact value of Nu up to and immediately above the onset of convection with asymptotic behaviour, $Nu \leqslant \frac{9}{256} Ra$ as $Ra \to \infty$, exhibiting the Howard–Malkus–Kolmogorov–Spiegel scaling anticipated by classical scaling and marginally stable boundary layer arguments. The relationship between these results and previous works of the same title (Busse & Joseph 1972; Gupta & Joseph 1973) is discussed.

1. Introduction

One approach to the problem of fluid turbulence is to explore the bounds imposed by the equations of motion on physical quantities. Although rigorous bounds necessarily apply to all solutions, laminar or turbulent, they provide limits within which the predictions of quantitative approximate theories and numerical simulations must lie. From a practical scientific point of view, experimental violation of such bounds would signal invalidity of the model. And often the derivation of *a priori* bounds is an early step in the analysis of existence, uniqueness and regularity of solutions of partial differential equations. But while estimates corresponding to physical quantities may be sufficient for some technical mathematical purpose, they are often neither explicit enough nor precise enough for quantitative comparison with real data.

Hydrodynamic stability theory, in various guises, plays a number of roles in theoretical fluid dynamics. Linearized stability theory can predict the presence of instabilities while nonlinear energy stability theory can predict their absence. Weakly nonlinear stability theory produces asymptotic approximations to solutions near bifurcations, and heuristic marginal stability arguments are invoked for boundary layer modelling. It is not altogether obvious, however, how stability concepts might be utilized in the derivation of bounds on flow quantities.

In this paper we study a model of convection in a fluid-saturated porous layer which allows a rigorous and precise calculation of heat transport bounds, valid for

'turbulent' as well as laminar convection states, that compare remarkably well with experiment. For the infinite Prandtl–Darcy number problem in two or three spatial dimensions and additionally the arbitrary Prandtl–Darcy number problem in two dimensions, we will derive an upper bound on the Nusselt number (Nu) as a function of the Rayleigh number (Ra) that both (i) exactly captures the bifurcation at onset, and (ii) displays the observed 'turbulent' scaling $Nu \sim Ra$ as $Ra \to \infty$. Moreover, the upper bound analysis in this paper may be naturally interpreted as a rigorous mathematical implementation of heuristic marginal stability ideas. Our aim is to demonstrate the mathematical techniques as well as exhibit the connection with hydrodynamic stability theory.

A previous approach to upper bounds for turbulent convection was initiated by Howard (1963) with his formulation of a variational principle for statistically stationary flows (Howard 1972). This theory was developed by Busse and collaborators (Busse 1978) and it was applied to porous-medium convection problems by Busse & Joseph (1972) and Gupta & Joseph (1973). In this paper we utilize the 'background field' method that has its foundation in a mathematical device introduced by Hopf (1941). Hopf's approach produces estimates of long time averages of bulk transport without statistical hypotheses. It has recently been applied to incompressible turbulence in the Navier–Stokes equations (Doering & Constantin 1992, 1994; Constantin & Doering 1995*a*; Kerswell 1996; Wang 1997) and Boussinesq convection in a fluid layer (Doering & Constantin 1996; Constantin & Doering 1996; Doering & Hyman 1997). When applied to problems with sufficient geometric symmetry that Howard's statistical stationarity hypotheses may be exploited, these two approaches share a closely related mathematical structure (Kerswell 1998).

For the porous-medium convection models considered here, the best bounds derived in this paper coincide with a single-wavenumber formula for the arbitrary Prandtl–Darcy number problem derived by Busse & Joseph (1972). Busse & Joseph determined that that formula produced an upper bound for arbitrary Prandtl–Darcy number in a limited range of Rayleigh numbers ($Ra \leqslant 113$), but here we find that the formula is in fact applicable for all Ra for the infinite Prandtl–Darcy number problem in two or three spatial dimensions and for the arbitrary Prandtl–Darcy number problem in two dimensions.

The rest of this paper is organized as follows. We present the equations of motion, physical set-up and definitions that form the basis of the analysis along with brief phenomenological descriptions of the solutions in §2. In §3 we review fundamental stability concepts for stationary solutions along with a heuristic marginal stability argument and its predictions for high-Ra (ostensibly) turbulent convection. A rigorous implementation of the marginal stability argument via the background field method is developed in §4, where we derive a variational formulation of upper bounds on the heat transport. We construct appropriately constrained test background profiles in §5 to produce explicit bounds on Nu as a function of Ra. We discuss the structure of the variational problem and of the optimal background profiles in §6. The constraints for the minimization problem are strengthened in §7 to allow for optimization in the modified variational problem, improving the estimates derived in §5 (although still leaving room for further improvement). In §8 we show that the results of §7 apply as well to the arbitrary Prandtl–Darcy number system in two dimensions, and in the concluding §9 we compare the best bounds obtained here with experimental data.

2. Model equations, definitions and phenomenology

The flow of fluid through a porous medium is a complex and not thoroughly understood phenomenon (Nield & Bejan 1992). In this paper we study what is perhaps the simplest description of thermal convection in a porous layer, a temperature advection–diffusion equation coupled to a divergence-free velocity vector field obeying Darcy's law via a buoyancy force proportional to the local temperature.

2.1. *Model equations*

Consider a layer of fluid-saturated porous material confined between horizontal planes located at $z = 0$ and $z = h$ with fixed temperatures T_{hot} and T_{cold} imposed on the bottom and top planes respectively. We begin with the Darcy–Oberbeck–Boussinesq equations (Lapwood 1948) for the evolution of the temperature field $T(\boldsymbol{x}, t)$, the Darcy seepage velocity $\boldsymbol{u}(\boldsymbol{x}, t)$, and the pressure field $p(\boldsymbol{x}, t)$:

$$T_t + \boldsymbol{u}\cdot\boldsymbol{\nabla} T = \kappa \Delta T \tag{2.1}$$

and

$$\boldsymbol{u}_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + \frac{\nu}{K}\boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{e}_3\, g\alpha(T - T_0) \tag{2.2}$$

with

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \quad \boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + u_3\boldsymbol{e}_3. \tag{2.3}$$

The vertical boundary conditions are

$$T = T_{hot}, \quad u_3 = 0 \quad \text{at} \quad z = 0 \tag{2.4a}$$

and

$$T = T_{cold}, \quad u_3 = 0 \quad \text{at} \quad z = h. \tag{2.4b}$$

In the above, κ is the thermal conduction coefficient of the fluid-medium mixture, ν is the fluid viscosity, K is the Darcy permeability coefficient, g is the acceleration due to gravity in the $-\boldsymbol{e}_3$ direction, α is the coefficient of thermal expansion, and T_0 is a reference temperature. These equations – in particular the linear velocity damping term with effective friction coefficient ν/K – are relevant for flows that are laminar within the material pores, as expected when length scales in the flow field are much greater than the porous-material microscale. For technical convenience we will consider a layer of extent L_x and L_y in the x- and y-directions with periodic boundary conditions on all variables in those directions, although many of our conclusions would also apply to impermeable insulating sidewall conditions or a horizontal layer of infinite extent.

To obtain a dimensionless description of the problem we measure lengths in units of the layer thickness h, time in units of the thermal diffusion time scale h^2/κ, and the temperature shifted by T_0 in units of $(T_{hot} - T_{cold})$. Then the equations of motion become

$$T_t + \boldsymbol{u}\cdot\boldsymbol{\nabla} T = \Delta T \tag{2.5}$$

and

$$B(\boldsymbol{u}_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}) + \boldsymbol{u} + \boldsymbol{\nabla} p = \boldsymbol{e}_3\, RaT \tag{2.6}$$

with

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0 \quad \boldsymbol{u} = u\boldsymbol{e}_1 + v\boldsymbol{e}_2 + w\boldsymbol{e}_3. \tag{2.7}$$

The vertical boundary conditions, as illustrated in figure 1 are

$$T = 1, \quad \boldsymbol{e}_3\cdot\boldsymbol{u} = w = 0 \quad \text{at} \quad z = 0 \tag{2.8a}$$

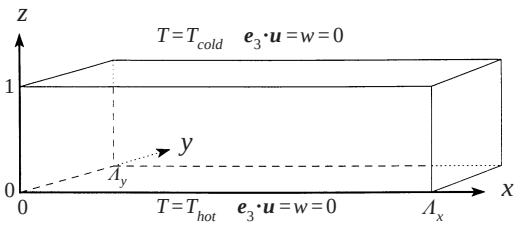

FIGURE 1. The geometry of the set-up and the boundary conditions. Periodic sidewall conditions are imposed.

and

$$T = 0, \quad \boldsymbol{e}_3\cdot\boldsymbol{u} = w = 0 \quad \text{at} \quad z = 1. \tag{2.8b}$$

The remaining parameters, besides the dimensionless horizontal scales $\Lambda_x = L_x/h$ and $\Lambda_y = L_y/h$ of the layer, are the Rayleigh number

$$Ra = \frac{g\alpha(T_{hot} - T_{cold})\,Kh}{\nu\kappa}, \tag{2.9}$$

and the Prandtl–Darcy number B^{-1} defined by

$$B^{-1} = \frac{\nu h^2}{\kappa K}. \tag{2.10}$$

In many cases (Elder 1967; Burreta 1972) the porous-medium microscale $\sim K^{1/2}$ is small enough relative to the gap height h, and the viscosity ν is large enough relative to the thermal diffusion coefficient κ, that the infinite Prandtl–Darcy number limit ($B = 0$) of the equations is an appropriate model. Hence for the three-dimensional analysis we will restrict attention to the infinite Prandtl–Darcy number limit of the equations:

$$T_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}T = \Delta T \tag{2.11}$$

and

$$\boldsymbol{u} + \boldsymbol{\nabla}p = \boldsymbol{e}_3\,RaT \tag{2.12}$$

with

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \quad \boldsymbol{u} = u\boldsymbol{e}_1 + v\boldsymbol{e}_2 + w\boldsymbol{e}_3. \tag{2.13}$$

The two-dimensional ($v = 0$, $\partial_y = 0$) version of the finite Prandtl–Darcy number equations (2.11)–(2.13) is relevant to model Hele-Shaw flow of a viscous fluid between closely spaced no-slip plates parallel to the (x, z)-plane (Koster & Müller 1982). In that case the effective Darcy permeability coefficient is $K = d^2/12$ where d is the spacing between the plates. We will also consider the finite Prandtl–Darcy number problem in two dimensions in this paper.

The dynamical problem is completed by specifying initial conditions in the form of the temperature $T_0(\boldsymbol{x})$ and, in the case $B \neq 0$, the velocity field $\boldsymbol{u}_0(\boldsymbol{x})$, which without loss of physical relevance we may presume to be smooth. For the infinite Prandtl–Darcy number model in equations (2.11)–(2.13), the pressure field is constructed at each instant of time by solving a Poisson equation for p with Neumann boundary conditions at $z = 0$ and $z = 1$ and periodic conditions in the horizontal directions. For the arbitrary Prandtl–Darcy number equations (2.5)–(2.7) in two dimensions, the pressure may be eliminated altogether by going to a stream function representation.

We reiterate that the models under consideration here are not expected to remain valid for any particular 'real' system for arbitrarily large values of Ra. The models break down when length scales in the flows become smaller than the microscale $\sim K^{1/2}$ (or the gap spacing d in the case of Hele-Shaw flow) which ultimately they must become. Darcy's law is not valid in such situations and instead something like Forscheimer's 'quadratic drag law' (Nield & Bejan 1992), in which the square of the local speed is proportional to the local force, should be used. In this paper we do not consider such regimes; what we mean by 'turbulent convection' in the context of the models at hand is the absence of spatial and temporal coherence in the solutions, which may occur without violating the conditions necessary for the models' validity or in direct numerical simulations. These limitations on the physical applicability of the model are clearly illustrated by the experimental data collected by Lister (1990), which will be discussed further in the concluding §9.

2.2. *Heat transport: definitions and identities*

The central goal of theories of convection is to predict the heat transport as a function of the system parameters. The heat flux vector field (modulo a factor of the specific heat) is identified by writing the heat equation in conservation form $T_t + \boldsymbol{\nabla}\cdot\boldsymbol{J} = 0$, so

$$\boldsymbol{J} = \boldsymbol{u}T - \kappa\boldsymbol{\nabla}T. \tag{2.14}$$

In the absence of convection heat is transported by conduction. The corresponding exact stationary solution of the basic equations (2.1)–(2.4) is the pure conduction solution

$$\boldsymbol{u}_{cond} = 0, \tag{2.15a}$$

$$T_{cond} = T_{hot} - \frac{z}{h}(T_{hot} - T_{cold}), \tag{2.15b}$$

$$p_{cond} = g\alpha z\left\{T_{hot} - T_0 - \frac{z}{2h}(T_{hot} - T_{cold})\right\}. \tag{2.15c}$$

In the pure conduction state the current is vertical and proportional to the temperature drop across the layer,

$$\boldsymbol{J}_{cond} = \boldsymbol{e}_3\,\kappa\frac{T_{hot} - T_{cold}}{h}. \tag{2.16}$$

The dimensionless heat transport is given by the Nusselt number Nu, usually defined as the ratio of the long-time and spatially averaged vertical heat flux to its pure conduction value:

$$Nu = \frac{1}{\boldsymbol{e}_3\cdot\boldsymbol{J}_{cond}}\lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}s\,\frac{1}{L_x L_y h}\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\left\{u_3(\boldsymbol{x},s)\,T(\boldsymbol{x},s) - \kappa\frac{\partial T}{\partial z}(\boldsymbol{x},s)\right\}. \tag{2.17}$$

In terms of the dimensionless variables in equations (2.5)–(2.7) or (2.11)–(2.13), the Nusselt number is then simply

$$Nu = \lim_{t\to\infty}\frac{1}{t}\int_0^t \mathrm{d}s\,\frac{1}{\Lambda_x\Lambda_y}\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\left\{w(\boldsymbol{x},s)\,T(\boldsymbol{x},s) - \frac{\partial T}{\partial z}(\boldsymbol{x},s)\right\}. \tag{2.18}$$

The Nusselt number as defined above depends in general on the particular initial conditions; it need not be unique for a specific value of Ra on a given domain.

Moreover, this definition of the Nusselt number only makes sense when the long-time average exists. The existence of infinite time averages is tacitly assumed in many analyses, but it is one assumption that our approach to upper bounds avoids. Indeed, the analysis in this paper produces upper bounds on the largest possible Nusselt number (for which we shall use the same symbol) for a given domain and specific *Ra* value. That is, we define

$$Nu = \sup_{T_0,\boldsymbol{u}_0} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \frac{1}{\Lambda_x \Lambda_y}\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \left\{ w(\boldsymbol{x},s)\,T(\boldsymbol{x},s) - \frac{\partial T}{\partial z}(\boldsymbol{x},s)\right\} \tag{2.19}$$

which, under our assumptions on the initial data, always exists and is unique for a given geometry and Rayleigh number (Ly & Titi 1996). In the following analysis we will only use the fact that the long-time average of the time derivative of a uniformly bounded function vanishes:

$$0 = \lim_{t\to\infty} \frac{1}{t}\int_0^t f'(s)\,\mathrm{d}s \quad \text{when} \quad \sup_{t>0} |f(t)| < \infty. \tag{2.20}$$

From these considerations, there are a number of identities which may be derived from the equations of motion and the boundary conditions.

First and most directly, the boundary conditions and the incompressibility condition imply

$$Nu = 1 + \sup_{T_0,\boldsymbol{u}_0} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \frac{1}{\Lambda_x \Lambda_y}\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \{ w(\boldsymbol{x},s)\,[T(\boldsymbol{x},s) - f(z)]\} \tag{2.21}$$

for arbitrary functions $f(z)$.

Secondly, consider the derivative with respect to z of a finite-time average of the horizontally integrated heat flux:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}\frac{1}{t}\int_0^t \mathrm{d}s \int \boldsymbol{e}_3\cdot\boldsymbol{J}(x,y,z,s)\,\mathrm{d}x\,\mathrm{d}y &= \frac{1}{t}\int_0^t \mathrm{d}s \int \mathrm{d}x\,\mathrm{d}y \frac{\partial}{\partial z}\left\{ w(\boldsymbol{x},s)\,T(\boldsymbol{x},s) - \frac{\partial T}{\partial z}(\boldsymbol{x},s)\right\} \\ &= \frac{1}{t}\int_0^t \mathrm{d}s \int \mathrm{d}x\,\mathrm{d}y\,\{\boldsymbol{u}\cdot\boldsymbol{\nabla}T - \Delta T\} = \frac{1}{t}\int \mathrm{d}x\,\mathrm{d}y\,\{T(x,y,z,t) - T(x,y,z,0)\}. \end{aligned} \tag{2.22}$$

For reasonable initial data this last expression vanishes as $t\to\infty$ because T remains uniformly bounded (a fact which may be shown to follow from the equations of motion) so this shows that the long-time-averaged heat flux is the same through each horizontal layer. For example, because $\boldsymbol{e}_3\cdot\boldsymbol{J}(x,y,0,t) = -(\partial T/\partial z)(x,y,0,t)$,

$$\limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \int \boldsymbol{e}_3\cdot\boldsymbol{J}(x,y,z,s)\,\mathrm{d}x\,\mathrm{d}y = \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \int \left(-\frac{\partial T}{\partial z}(x,y,0,s)\right)\mathrm{d}x\,\mathrm{d}y. \tag{2.23}$$

Hence Nu is the largest possible long-time average of the horizontally averaged heat flux at any vertical level according to

$$Nu = \sup_{T_0,\boldsymbol{u}_0} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \frac{1}{\Lambda_x \Lambda_y}\int \mathrm{d}x\,\mathrm{d}y \left\{ w(x,y,\cdot,s)\,T(x,y,\cdot,s) - \frac{\partial T}{\partial s}(x,y,\cdot,s)\right\}. \tag{2.24}$$

Finally, then, multiplying the heat equation (2.5) or (2.11) by T and integrating over the entire volume utilizing the boundary conditions we see that

$$Nu = \sup_{T_0, \boldsymbol{u}_0} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \frac{1}{\Lambda_x \Lambda_y}\int \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,|\boldsymbol{\nabla} T|^2. \tag{2.25}$$

The goal of theory and analysis is to estimate the functional relation $Nu(Ra)$.

2.3. *Phenomenology*

A generic summary of the model's behaviour (Kimura, Schubert & Straus 1986, 1989; Graham & Steen 1994) is sketched in figure 2. For low values of Ra, below a critical value $Ra_c \geqslant 4\pi^2$, the pure conduction state is the unique long-time system configuration. In that state the fluid remains at rest in the bulk, there is a linear temperature profile across the layer, and $Nu = 1$.

The conduction state is unstable for $Ra > Ra_c$ where convection sets in. At the onset of convection the velocity field forms stationary convection rolls, the (horizontally averaged) temperature profile across the layer deviates from the linear conduction profile, and the Nusselt number bifurcates away its conduction value of 1. Quantitative details of this behaviour depend on details of the experiment, i.e. the system aspect ratio and the degree of uniformity and homogeneity of the porous medium (Shattuck *et al*. 1995, 1997), but we restrict attention to the ideal case as described above. As the Rayleigh number is further increased, with details depending on the particulars of the domain shape and sidewall conditions, the rolls eventually lose stability and oscillatory and chaotic states may be realized.

For values of Ra far above the transition and presumably into the $Ra\to\infty$ asymptotic regime, spatial coherence of the flow patterns is lost and a 'turbulent' state ensues. The temperature profile establishes a boundary layer structure, and the dominant mode of heat transport is hot (cold) thermal plumes or blobs released from thin thermal boundary layers at the bottom (top) of the domain. In this regime, 'Howard–Malkus–Kolmogorov–Spiegel' scaling $Nu \sim Ra$ is observed in experiments (Elder 1967) and numerical simulations (Graham & Steen 1994). This relation is not improperly referred to as turbulent Kolmogorov–Spiegel scaling because it implies that the macroscopic heat transport is independent (Spiegel 1971) of the microscopic transport coefficient κ:

$$Nu \sim Ra \Rightarrow \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathrm{d}s \int \boldsymbol{e}_3 \cdot \boldsymbol{J}(x, y, \cdot, s)\,\mathrm{d}x\,\mathrm{d}y \sim \Lambda_x \Lambda_y \frac{g\alpha(T_{hot} - T_{cold})^2 K}{\nu}. \tag{2.26}$$

This is in direct analogy to Kolmogorov's scaling theory for incompressible fluid turbulence wherein the rate of macroscopic energy dissipation becomes independent of the microscopic dissipation coefficient, in that case the viscosity. (Note that the heat transport is also independent of the layer thickness h in this situation.) It is also not improperly referred to as Howard–Malkus scaling because it follows from the marginally stable boundary layer hypothesis put forth by Malkus (1954) and quantified by Howard (1964) in the context of convection in a fluid layer. That heuristic scaling argument will be recalled in the next section.

3. Stability, instability and a marginal stability argument

There are several notions of stability, as well as several applications of these notions, that will enter our analysis of heat transport bounds. In this section we briefly review

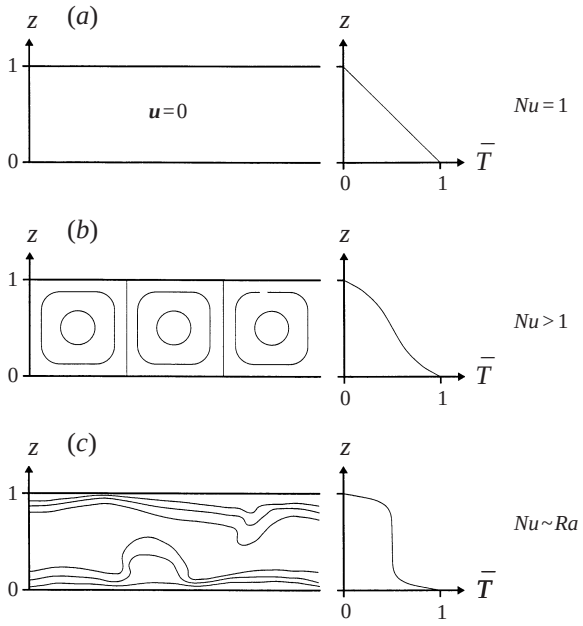

FIGURE 2. Qualitative sketch illustrating the regimes of convection in a porous layer. (*a*) Below onset there is no flow and a linear temperature profile across the layer. (*b*) Immediately above onset, steady rolls appear and the heat flux increases ($Nu > 1$). The rolls may destabilize, oscillate, and eventually break up. (*c*) High-Ra 'turbulent' convection is characterized by thin thermal boundary layers shedding plumes which are transported across the layer by buoyancy forces. (The overbar on T indicates a horizontal and long-time average.)

the applications of linear stability theory, nonlinear energy stability theory, and weakly nonlinear stability theory. Then we recall the heuristic marginal stability argument presumed relevant for the turbulent convection situation and its prediction of Howard–Malkus–Kolmogorov–Spiegel scaling, $Nu \sim Ra$.

3.1. *Linear stability theory*

Linearized stability theory (Chandrasekhar 1963) is capable of providing a sufficient condition for instability. Let $\tau(\boldsymbol{x})$ be a stationary temperature field and $\boldsymbol{U}(\boldsymbol{x})$ and $P(\boldsymbol{x})$ be the corresponding velocity and pressure fields satisfying

$$\boldsymbol{U}\cdot\boldsymbol{\nabla}\tau = \Delta\tau, \tag{3.1}$$

$$\boldsymbol{U} + \boldsymbol{\nabla}P = \boldsymbol{e}_3\, Ra\tau, \tag{3.2}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{U} = 0, \tag{3.3}$$

along with the physical boundary conditions. Infinitesimal perturbations of the temperature, velocity and pressure fields, $\theta(\boldsymbol{x}, t)$, $\boldsymbol{u}(\boldsymbol{x}, t)$ and $p(\boldsymbol{x}, t)$, evolve according to the linearized equations,

$$\theta_t + \boldsymbol{U}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{u}\cdot\boldsymbol{\nabla}\tau = \Delta\theta, \tag{3.4}$$

$$\boldsymbol{u} + \boldsymbol{\nabla}p = \boldsymbol{e}_3 Ra\theta, \tag{3.5}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{3.6}$$

along with homogeneous versions of the boundary conditions.

Presuming an $e^{-\lambda t}$ time dependence for θ, $\boldsymbol{u}$, and p, linear stability is rewritten as a spectral problem for the eigenvalues λ and the eigenfunctions $\theta_\lambda(\boldsymbol{x})$, $\boldsymbol{u}_\lambda(\boldsymbol{x})$, and $p_\lambda(\boldsymbol{x})$ satisfying

$$-\lambda\theta_\lambda + \boldsymbol{U}\cdot\boldsymbol{\nabla}\theta_\lambda + \boldsymbol{u}_\lambda\cdot\boldsymbol{\nabla}\tau = \Delta\theta_\lambda, \tag{3.7}$$

$$\boldsymbol{u}_\lambda + \boldsymbol{\nabla}p_\lambda = \boldsymbol{e}_3\, Ra\theta_\lambda, \tag{3.8}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{u}_\lambda = 0. \tag{3.9}$$

Infinitesimal perturbations grow exponentially if the real part of any eigenvalue is negative. Thus linear stability theory may be summarized, in the context of equations (3.7)–(3.9), as the statement

$$\exists\,\lambda \ni \mathrm{Re}\{\lambda\} < 0 \Rightarrow \text{instability}. \tag{3.10}$$

The application to the pure conduction solution is elementary: for $\tau = 1-z$, $\boldsymbol{U} = 0$ and $P = Ra(z-\frac{1}{2}z^2)$, the eigenvalue problem reduces to

$$\lambda\theta_{\lambda,k^2}(z) = (-\mathrm{D}^2 + k^2)\,\theta_{\lambda,k^2}(z) - w_{\lambda,k^2}(z) \tag{3.11}$$

and

$$0 = (-\mathrm{D}^2 + k^2)\,w_{\lambda,k^2}(z) - Ra\,k^2\theta_{\lambda,k^2}(z) \tag{3.12}$$

with homogeneous Dirichlet boundary conditions at $z = 0$ and 1, and where $k^2 = |\boldsymbol{k}|^2$ is the square of the magnitude of the horizontal wavenumber $\boldsymbol{k}$. Here and henceforth, the symbol D stands for $\mathrm{d}/\mathrm{d}z$. The solutions are

$$\theta_{\lambda,k^2} = A\sin n\pi z, \quad w_{\lambda,k^2} = \frac{Ra\,k^2}{n^2\pi^2 + k^2}A\sin n\pi z, \tag{3.13}$$

and the eigenvalues are

$$\lambda_{n,k^2} = n^2\pi^2 + k^2 - \frac{Ra\,k^2}{n^2\pi^2 + k^2}. \tag{3.14}$$

The lowest eigenvalues occur at $n = 1$, and the critical wavenumber above which the mode of the horizontal wavenumber $\boldsymbol{k}$ is unstable is

$$Ra_c(k^2) = \frac{(\pi^2 + k^2)^2}{k^2}. \tag{3.15}$$

The critical Rayleigh number has the minimum value

$$Ra_c = 4\pi^2 \tag{3.16}$$

at $k_c = \pi$, corresponding to square convection rolls with aspect ratio 1.

3.2. *Nonlinear energy stability theory*

Nonlinear energy stability theory (Joseph 1976) produces a sufficient condition for stability to perturbations of arbitrary amplitude. Consider a stationary solution $\tau(\boldsymbol{x})$, $\boldsymbol{U}(\boldsymbol{x})$ and $P(\boldsymbol{x})$ to the system

$$\boldsymbol{U}\cdot\boldsymbol{\nabla}\tau = \Delta\tau, \tag{3.17}$$

$$\boldsymbol{U} + \boldsymbol{\nabla}\boldsymbol{P} = \boldsymbol{e}_3\, Ra\,\tau, \tag{3.18}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{U} = 0, \tag{3.19}$$

along with the physical boundary conditions. Arbitrary-amplitude perturbations $\theta(\boldsymbol{x}, t)$, $\boldsymbol{u}(\boldsymbol{x}, t)$ and $p(\boldsymbol{x}, t)$ evolve according to

$$\theta_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{U}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{u}\cdot\boldsymbol{\nabla}\tau = \Delta\theta, \tag{3.20}$$

$$\boldsymbol{u} + \boldsymbol{\nabla}p = \boldsymbol{e}_3\, Ra\,\theta, \tag{3.21}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{3.22}$$

along with homogeneous versions of the boundary conditions. Multiplying equation (3.20) by θ and integrating over the volume, we find an evolution equation for the mean-square perturbation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int \theta(\boldsymbol{x}, t)^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = -\int \{|\boldsymbol{\nabla}\theta|^2 + \theta\boldsymbol{u}\cdot\boldsymbol{\nabla}\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \equiv -H_\tau\{\theta(\cdot, t)\}, \tag{3.23}$$

where, noting that $\boldsymbol{u}$ is a linear (albeit non-local) function of θ defined via equations (3.21) and (3.22), we identify the quadratic form H_τ which depends parametrically on the temperature field $\tau(\boldsymbol{x})$.

The key observation is that if the temperature field τ happens to be such that the quadratic form H_τ is positive, then the perturbations decay monotonically in time. In fact in this case the L_2 norm of an arbitrary-amplitude perturbation decays exponentially in time at minimal rate

$$\lambda = \inf \frac{H_\tau\{\theta\}}{\int \theta(\boldsymbol{x}, t)^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z}, \tag{3.24}$$

where the infimum is over all square-integrable test functions $\theta(\boldsymbol{x})$ satisfying the (homogeneous) boundary conditions.

The minimum decay rate λ is the solutions of the variational problem in equation (3.24) which may be recast as a spectral problem. Utilizing standard calculus of variations for this constrained optimization problem implies that the minimizing θ satisfies the Euler–Lagrange equations

$$\lambda\theta = -\Delta\theta + \tfrac{1}{2}(\boldsymbol{u}\cdot\boldsymbol{\nabla}\tau - \boldsymbol{e}_3\cdot\boldsymbol{v}), \tag{3.25}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0, \tag{3.26}$$

$$\boldsymbol{u} + \boldsymbol{\nabla}p = Ra\,\boldsymbol{e}_3\,\theta, \tag{3.27}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0, \tag{3.28}$$

$$\boldsymbol{v} + \boldsymbol{\nabla}q = -Ra\,\theta\boldsymbol{\nabla}\tau, \tag{3.29}$$

where $\boldsymbol{u}$ and p are the velocity and pressure fields associated to θ, while $\boldsymbol{v}$ and q are the Lagrange multipliers introduced to enforce the constraints embodied in equations (3.26) and (3.27). The velocity field $\boldsymbol{v}$ satisfies the same boundary conditions as $\boldsymbol{u}$, i.e. $\boldsymbol{e}_3\cdot\boldsymbol{v} = 0$ at $z = 0$ and $z = 1$, and λ is the smallest real number such that the solution may be normalized according to

$$1 = \int \theta^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{3.30}$$

As long as the spectrum of the self-adjoint eigenvalue problem in equations (3.25)–(3.29) is non-negative, arbitrary-amplitude perturbations of the stationary solution decay exponentially in time, indicating absolute stability of the base solution.

This nonlinear energy stability theory may be summarized, in the context of equations (3.25)–(3.30), by the statement

$$\mathrm{Re}\{\lambda\} > 0 \,\forall\, \lambda \Rightarrow \text{stability}. \tag{3.31}$$

Application to the pure conduction solution is particularly straightforward, for when $\tau = 1 - z$ the equations for $\boldsymbol{v}$ and q become identical to those for $\boldsymbol{u}$ and p. Combined with the same boundary conditions, then, $\boldsymbol{v} = \boldsymbol{u}$ so the problem reduces to

$$\lambda\theta_{\lambda,k^2}(z) = (-\mathrm{D}^2 + k^2)\,\theta_{\lambda,k^2}(z) - w_{\lambda,k^2}(z) \tag{3.32}$$

and

$$0 = (-\mathrm{D}^2 + k^2)\, w_{\lambda,k^2}(z) - Ra\, k^2\, \theta_{\lambda,k^2}(z) \tag{3.33}$$

with homogeneous Dirichlet boundary conditions at $z = 0$ and 1, and where k^2 is the square of the horizontal wavenumber. This is precisely the same as the linearized stability problem, so the critical Rayleigh number below which the pure conduction solution is stable to arbitrary perturbations is also

$$Ra_c = 4\pi^2. \tag{3.34}$$

This result establishes the absolute stability of the conduction state when it is not linearly unstable:

$$Ra < 4\pi^2 \Rightarrow Nu = 1. \tag{3.35}$$

Thus the transition from the conduction state is a forward pitchfork bifurcation to the convection state. Linear and nonlinear stability theories provide a complete and precise characterization of the pure conduction state, but we must turn to other methods for the heat transport in the convection state.

3.3. *Weakly nonlinear stability theory*

Although exact analytic expressions are not available for convection solutions above the critical Rayleigh number, it is possible to obtain asymptotic estimates for the temperature and velocity fields in the limit $Ra \downarrow Ra_c$. This analysis via 'amplitude equations' has already been carried out for this problem (Busse & Joseph 1972; Fowler 1997) so here we just describe the approach and its results.

Consider Rayleigh numbers just slightly above the critical value compatible with the cell geometry, i.e.

$$Ra = Ra_c(k_c^2)\,(1 + \varepsilon^2), \tag{3.36}$$

with

$$Ra_c(k_c^2) = \frac{(\pi^2 + k_c^2)^2}{k_c^2}, \tag{3.37}$$

where k_c is the wavenumber of the first horizontal mode to go unstable. When ε is small enough only a single mode ($n = 1, k = k_c$) is unstable and its growth rate is $O(\varepsilon^2)$. Hence it is natural to look for approximate solutions for the perturbations from the conduction solution of the form

$$\theta(\boldsymbol{x}, t) = \varepsilon\theta^{(1)}(\boldsymbol{x}, t, s) + \varepsilon^2\theta^{(2)}(\boldsymbol{x}, t, s) + \ldots, \tag{3.38}$$

where we have introduced the slow time variable $s = \varepsilon^2 t$. Presuming $O(\varepsilon)$ initial data concentrated in the unstable mode, inserting the expansion (3.38) into the equations of motion along with a similar ansatz for the velocity field, and solving order by order with the usual suppression of secular terms (integrability conditions for steady-state solutions), one constructs an asymptotic solution of the form

$$\theta \sim \varepsilon[A(s)\, \mathrm{e}^{\mathrm{i}\boldsymbol{k}_c \cdot (x,y)} \sin \pi z + \text{c.c.}]. \tag{3.39}$$

The velocity field is then also $O(\varepsilon)$ and the slowly evolving amplitude $A(s)$ obeys Landau equation

$$\frac{\mathrm{d}A}{\mathrm{d}s} = 2\pi^2 A - 4\pi^4 |A|^2 A. \tag{3.40}$$

The amplitude is unstable near $A = 0$, so it grows on an $O(s)$ time scale toward its $O(1)$ stable steady state where the steady-state convective heat transport is $O(\varepsilon^2)$.

For the problem at hand, if the cell admits the critical wavenumber $k_c = \pi$ so that the critical Rayleigh number is its minimum possible value $Ra_c = 4\pi^2$, then exact evaluations of the coefficients in the Landau equation yield (Busse & Joseph 1972) the estimate

$$Nu \sim 1 + \tfrac{1}{2}\varepsilon^2 = 1 + \frac{1}{2\pi^2}(Ra - Ra_c). \tag{3.41}$$

Higher-order estimates have also been computed for this problem (Palm, Weber & Oddmund 1972). For a comparison of this prediction with the experimental data, see figure 3. This approach is capable of producing accurate approximations immediately above the onset of convection, along with a reliable indication of the stability of the steady finite-amplitude convection solution. Because of the $\varepsilon \quad 1$ restriction required by this perturbation theory, the large Rayleigh number behaviour must be deduced by another method.

3.4. *Marginal stability for turbulent convection: Howard–Malkus scaling*

There are no exact solutions for large Rayleigh numbers where the state of affairs is generically unsteady 'turbulent' convection. A theoretical indication of the turbulent heat flux may be derived if we invoke further physical assumptions in analogy to Malkus' (1954) 'marginally stable boundary layer' theory for convection in a fluid layer (Howard 1964). Presumably, the temporally and horizontally averaged temperature profile across the layer takes on a boundary layer structure as sketched in figure 2(*c*). On average, half of the entire temperature drop falls across the boundary layers of thickness $\delta \quad h/2$ while the convecting core acts as a thermal short. The vertical component of the velocity vanishes at the boundaries so the heat transport is purely conductive there, controlled by the vertical temperature gradient at the boundary. The dimensional heat flux is approximately the conductive heat flux across the boundary layer,

$$J_3 \approx \kappa \frac{\frac{1}{2}(T_{hot} - T_{cold})}{\delta}, \tag{3.42}$$

in which case the Nusselt number is

$$Nu = \frac{J_3 h}{\kappa(T_{hot} - T_{cold})} = \frac{1}{2}\frac{h}{\delta}. \tag{3.43}$$

Thus an estimate for the boundary layer thickness as a function of the Rayleigh number produces an estimate for the Nusselt number as a function of the Rayleigh number.

The key physical assumption in this theory is that δ adjusts itself precisely so that the boundary layer itself is marginally stable. That is, if the boundary layer were so thick as to be unstable then convection could set in to break it up, perhaps shedding a thermal plume. On the other hand if the boundary layer were so thin as to be stable then it would suppress the vertical flow near it so heat could diffuse further via

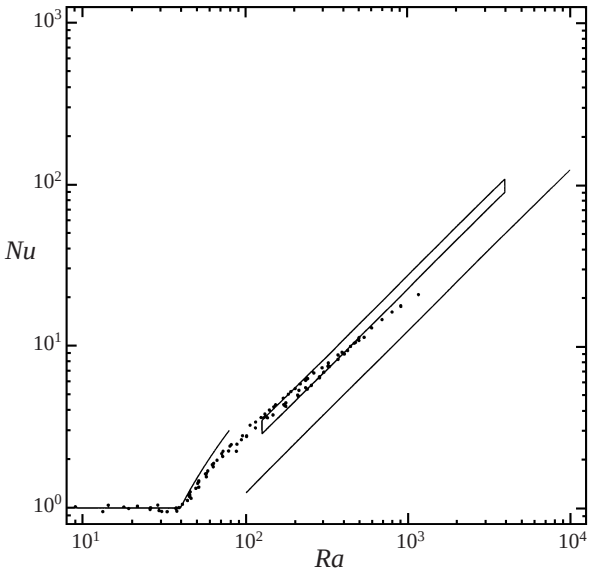

FIGURE 3. Predictions from stability considerations plotted along with experimental data. The discrete data are from Buretta (1972) and the boxed region indicates the range of Elder's (1967) data ($Nu = 0.025\ Ra \pm 10\,\%$ for $100 < Ra < 5000$). The convection state with $Nu = 1$ is absolutely stable for $Ra < \sim 40$, where it bifurcates according to the prediction of weakly nonlinear stability theory. The heuristic marginal stability argument predicts $Nu > 0.013\ Ra$, indicated by the solid line under the high-Ra data.

conduction to thicken the layer. We may then derive a Ra–δ relationship by way of this marginal stability criterion by setting the Rayleigh number Ra_δ based on the boundary layer parameters equal to an appropriate critical Rayleigh number:

$$Ra_c = Ra_\delta = \frac{g a_{\frac{1}{2}}^1 (T_{hot} - T_{cold})\, K\delta}{\nu\kappa} = \frac{1}{2}\frac{\delta}{h} Ra. \tag{3.44}$$

Combining equations (3.43) and (3.44) we deduce

$$Nu \approx \frac{1}{4Ra_c} Ra. \tag{3.45}$$

This marginal stability argument leads to Howard–Malkus–Kolmogorov–Spiegel scaling, $Nu \sim Ra$, and it even provides a prefactor estimate in terms of a critical Rayleigh number. It is not appropriate to simply use $4\pi^2 \approx 40$ for Ra_c because that value was derived for the more constrained problem of a layer confined between two impermeable boundaries; only one of the boundaries is truly impermeable in this scenario. In fact it is not altogether clear just what stability problem *is* appropriate (Howard considered the linearized stability of a thermal front diffusively penetrating into an infinite fluid layer). In any case we may expect that the number is probably less than the common linear and nonlinear critical value for the more constrained problem, so supposing that the relevant $Ra_c < O(20)$ we infer

$$Nu \approx cRa, \quad c > 0.013. \tag{3.46}$$

This prediction is also plotted along with experimental data in figure 3. This scaling appears qualitatively relevant to the large-Ra behaviour. The prefactor guess emerging from this argument is low by (only) about a factor of 2.

4. Variational formulation of heat transport bounds

The physical assumptions entering the marginal stability argument are quite reasonable and it is natural to wonder if these kinds of ideas might be utilized in a rigorous mathematical analysis. But there are some technical points glossed over in the heuristic argument which would have to be clarified before proceeding. For example as alluded to in the previous section, it is not clear precisely what stability problem is the relevant one, i.e. what boundary conditions should be imposed? Moreover, in the previous argument a notion of 'stability' was ambiguous, invoked without discriminating between linearized stability or nonlinear stability; these stability concepts are not generally identical in their predictions. These issues are resolved in this section where we derive a rigorous version of the marginal stability criterion directly from the equations of motion, valid over the full range of Ra.

4.1. *A variational bound on the heat transport*

Let $\tau(\boldsymbol{x})$ be an arbitrary 'background' temperature field that satisfies the physical boundary conditions, $\tau(x, y, 0) = 1$ and $\tau(x, y, 1) = 0$, and a 'non-dissipative' form of the stationary temperature equation, $\boldsymbol{U}\cdot\boldsymbol{\nabla}\tau = 0$, where $\boldsymbol{U}(\boldsymbol{x})$ and $P(\boldsymbol{x})$ are the corresponding velocity and pressure fields defined by

$$\boldsymbol{U} + \boldsymbol{\nabla} P = \boldsymbol{e}_3 \, Ra\, \tau, \tag{4.1}$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{U} = 0, \tag{4.2}$$

along with the boundary conditions $\boldsymbol{e}_3\cdot\boldsymbol{U} = 0$ for $z = 1$ and $z = 1$. (As will be seen, the solution set for this class of background fields is not empty.) Then any solution of the full equations of motion (2.11)–(2.13) may be decomposed according to $T = \tau + \theta$, $\boldsymbol{u} = \boldsymbol{U} + \boldsymbol{v}$ and $p = P + q$ where the 'fluctuations' θ, $\boldsymbol{v}$ and q satisfy

$$\theta_t + \boldsymbol{v}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{U}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{v}\cdot\boldsymbol{\nabla}\tau = \Delta\theta + \Delta\tau, \tag{4.3}$$

$$\boldsymbol{v} + \boldsymbol{\nabla} q = \boldsymbol{e}_3 \, Ra\, \theta, \tag{4.4}$$

and

$$\boldsymbol{\nabla}\cdot\boldsymbol{v} = 0, \tag{4.5}$$

along with homogeneous boundary conditions on θ and $\boldsymbol{e}_3\cdot\boldsymbol{v}$ at the top and bottom.

Consider the evolution of the mean-square temperature fluctuation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\int \theta(\boldsymbol{x}, t)^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = -\int \{\boldsymbol{\nabla}\tau\cdot\boldsymbol{\nabla}\theta + |\boldsymbol{\nabla}\theta|^2 + \theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.6}$$

The τ, θ cross-term can be replaced in favour of T and θ,

$$-\int \boldsymbol{\nabla}\tau\cdot\boldsymbol{\nabla}\theta\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{1}{2}\int \{|\boldsymbol{\nabla}\theta|^2 + |\boldsymbol{\nabla}\tau|^2 - |\boldsymbol{\nabla}T|^2\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.7}$$

so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \theta(\boldsymbol{x}, t)^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \int |\boldsymbol{\nabla}T|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z - \int \{|\boldsymbol{\nabla}\theta|^2 + 2\theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.8}$$

Taking the long-time average and recalling the expression for Nu in equation (2.25) we find

$$Nu = \frac{1}{\Lambda_x\Lambda_y}\int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \sup_{T_0, \boldsymbol{u}_0}\limsup_{t\to\infty}\frac{1}{t}\int_0^t (-2H_\tau^{(1/2)}\{\theta(\cdot, s)\})\,\mathrm{d}s. \tag{4.9}$$

In the above we have identified the quadratic form

$$H_\tau^{(1/2)}\{\theta\} \equiv \frac{1}{\Lambda_x \Lambda_y}\int \{\tfrac{1}{2}|\boldsymbol{\nabla}\theta|^2 + \theta \boldsymbol{v}\cdot\boldsymbol{\nabla}\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{4.10}$$

defined for functions θ satisfying the fluctuations' boundary conditions, where $\boldsymbol{v}\{\theta\}$ is the linear non-local functional of θ specified by equations (4.4) and (4.5) and the velocity boundary conditions. Note that modulo the factor $\frac{1}{2}$ in the $|\boldsymbol{\nabla}\theta|^2$ term, the quadratic form $H_\tau^{(1/2)}$ is the same as H_τ defined in equation (3.23) which determines the nonlinear energy stability of a stationary temperature field.

If the background profile τ happens to be such that $H_\tau^{(1/2)}$ is a non-negative quadratic form, i.e. $H_\tau^{(1/2)}\{\theta\} \geqslant 0$ for all relevant argument functions θ, then the background temperature profile produces an upper bound on Nu:

$$H_\tau^{(1/2)}\{\theta\} \geqslant 0 \Rightarrow Nu \leqslant \frac{1}{\Lambda_x \Lambda_y}\int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1 + \frac{1}{\Lambda_x \Lambda_y}\int |\boldsymbol{\nabla}\tau + \boldsymbol{e}_3|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.11}$$

This volume average of $|\boldsymbol{\nabla}\tau + \boldsymbol{e}_3|^2$ is precisely what the actual convective heat transport would be if τ were a steady solution of the equations of motion (it already satisfies the physical boundary conditions). By analogy with the mathematical criterion for nonlinear energy stability we naturally refer to arbitrary temperature functions $\tau(\boldsymbol{x})$ as marginally '$\frac{1}{2}$-stable' if $H_\tau^{(1/2)}\{\theta\} \geqslant 0$. Strictly speaking, a temperature field $\tau(\boldsymbol{x})$ is $\frac{1}{2}$-stable *if it would be* energy stable at Rayleigh number $2Ra$ *if it were* a steady solution. For example, the linear conduction profile $\tau = 1 - z$ is $\frac{1}{2}$-stable for $Ra < \frac{1}{2}Ra_c$, as are many other profiles. Hence we have a rigorous marginal stability criterion: the actual heat transport is bounded from above by that in a temperature field which is $\frac{1}{2}$-stable. In practice we will produce explicit bounds on Nu as a function of Ra by constructing appropriately stable τ (this will be carried out in detail in §§5 and 7).

These bounds may be optimized by minimizing over acceptable background temperature profiles:

$$Nu \leqslant \inf_{\tau(x,y,0)=1;\,\tau(x,y,1)=0}\left\{\frac{1}{\Lambda_x \Lambda_y}\int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \,|\, H_\tau^{(1/2)} \geqslant 0\right\}. \tag{4.12}$$

The $\frac{1}{2}$-stability constraint has also been referred to as the 'spectral' constraint on acceptable background temperature profiles (Constantin & Doering 1995*a*, *b*; Doering & Constantin 1994, 1996).

This variational problem is simplified by considering horizontally translation-invariant background temperature profiles, i.e. taking $\tau = \tau(z)$. Then the associated flow field $\boldsymbol{U} = 0$ and the bound in equation (4.12) becomes

$$Nu - 1 \leqslant \inf_{\tau(0)=1;\,\tau(1)=0}\left(\int_0^1 (\tau'(z)+1)^2\,\mathrm{d}z \,|\, H_\tau^{(1/2)} \geqslant 0\right\}, \tag{4.13}$$

where

$$H_\tau^{(1/2)}\{\theta\} \equiv \frac{1}{\Lambda_x \Lambda_y}\int \{\tfrac{1}{2}|\boldsymbol{\nabla}\theta|^2 + \tau'(z)\, w\theta\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{4.14}$$

is defined for functions $\theta(\boldsymbol{x})$ satisfying $\theta(x,y,0) = 0 = \theta(x,y,1)$, where

$$-\Delta w = Ra(\mathrm{D}^2 - \Delta)\,\theta \tag{4.15}$$

with $w(x,y,0) = 0 = w(x,y,1)$.

4.2. *Improved variational bound on the heat transport*

Following Nicodemus, Grossmann & Holthaus (1997*a*) we may introduce another variational parameter to reduce the upper bound. We return to equation (4.6) and use equation (4.7) to eliminate only part of the cross-term. Let $c > 1$. Adding $c \times$ (4.6) and $2 \times$ (4.7) we find

$$\frac{\mathrm{d}t}{\mathrm{d}t}\frac{c}{2}\int \theta(\boldsymbol{x},t)^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \int |\boldsymbol{\nabla} T|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$
$$= \int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z - \int \{(c-2)\,\boldsymbol{\nabla}\tau\cdot\boldsymbol{\nabla}\theta + (c-1)\,|\boldsymbol{\nabla}\theta|^2 + c\theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \quad (4.16)$$

Taking the long-time average, then,

$$Nu = \frac{1}{\Lambda_x \Lambda_y}\int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$
$$+ \sup_{T_0}\limsup_{t\to\infty}\frac{1}{t}\int_0^t \left(\frac{-1}{\Lambda_x\lambda_y}\int \{(c-1)\,|\boldsymbol{\nabla}\theta|^2 + c\theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau + (2-c)\,\theta\Delta\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\right)\mathrm{d}s. \quad (4.17)$$

An upper bound on Nu is obtained by replacing the second term above by its absolute minimum over all relevant θ and $\boldsymbol{v}$:

$$Nu \leqslant \frac{1}{\Lambda_x\Lambda_y}\int |\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}_z - \inf_{\theta} F_\tau(\theta), \quad (4.18)$$

where

$$F_\tau(\theta\} = \frac{1}{\Lambda_x\Lambda_y}\int \{(c-1)\,|\boldsymbol{\nabla}\theta|^2 + c\theta\boldsymbol{v}\cdot\boldsymbol{\nabla}\tau + (2-c)\,\theta\Delta\tau\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \quad (4.19)$$

Restricting attention to horizontally translation-invariant background temperature profiles, the Euler–Lagrange equations for the fields minimizing F_τ are

$$0 = -2(c-1)\,\Delta\theta + c\tau'(z)\,w + c(\mathrm{D}^2-\Delta)\,W + (2-c)\,\tau''(z), \quad (4.20)$$
$$0 = \Delta w + Ra(\mathrm{D}^2-\Delta)\,\theta, \quad (4.21)$$
$$0 = \Delta W + Ra\tau'(z)\,\theta, \quad (4.22)$$

where the Lagrange multiplier field $W(\boldsymbol{x})$ satisfies periodic horizontal conditions and homogeneous Dirichlet conditions at $z = 0$ and $z = 1$. The relevant solutions are independent of the horizontal coordinates: $\theta = \Theta(z)$, $w = 0$ and $W = W(z)$ so equation (4.20) reduces to

$$0 = -2(c-1)\,\Theta'' + (2-c)\,\tau''. \quad (4.23)$$

Integrating and using the boundary conditions $\Theta(0) = 0 = \Theta(1)$ we find

$$\Theta(z) = -\frac{c-2}{2(c-1)}(\tau(z) - 1 + z). \quad (4.24)$$

Inserting this minimizer into F_τ in equation (4.18), we deduce

$$Nu \leqslant \int_0^1 \tau'(z)^2\,\mathrm{d}z + \frac{(c-2)^2}{4(c-1)}\int_0^1 (\tau'(z)+1)^2\,\mathrm{d}z = 1 + \frac{c^2}{4(c-1)}\int_0^1 (\tau'(z)+1)^2\,\mathrm{d}z. \quad (4.25)$$

It is ensured that $\Theta(z)$ is indeed a minimizer so long as the quadratic part of F_τ is positive definite – or equivalently, the linear operator in the quadratic part of the last term in equation (4.18) is positive definite. In fact slightly less is required: it is enough that the linear operator in the quadratic part of F_τ be non-negative with a null space orthogonal to the inhomogeneous term proportional to $\tau''(z)$. What this means is that it is necessary that the eigenvalues of

$$\lambda\theta = -2(c-1)\,\Delta\theta + c\tau'(z)\,w + c(\mathrm{D}^2-\Delta)\,W, \tag{4.26}$$

$$0 = \Delta w + Ra(\mathrm{D}^2-\Delta)\,\theta, \tag{4.27}$$

$$0 = \Delta W + Ra\tau'(z)\,\theta, \tag{4.28}$$

with periodic horizontal conditions and homogeneous Dirichlet boundary conditions at $z = 0$ and $z = 1$, satisfy $\lambda \geqslant 0$ and the eigenfunction(s) corresponding to $\lambda = 0$ must be orthogonal to τ''. The orthogonality of the null-space condition turns out to be automatically satisfied. Indeed, with $\lambda = 0$ the horizontally Fourier transformed version of equations (4.26)–(4.28) are

$$0 = -2(c-1)\,(\mathrm{D}^2-k^2)\,\theta_k + c\tau'(z)\,w_k + ck^2\,W_k, \tag{4.29}$$

$$0 = (\mathrm{D}^2-k^2)\,w_k + Ra\,k^2\theta_k, \tag{4.30}$$

$$0 = (\mathrm{D}^2-k^2)\,W_k + Ra\,\tau'(z)\,\theta_k. \tag{4.31}$$

The eigenfunctions will be orthogonal to $\tau''(z)$ (actually orthogonal to any function of z alone) if $k^2 \neq 0$, and it is easy to see that any such null eigenfunction necessarily has $k^2 > 0$. This is because if $k^2 = 0$, then equation (4.30) and the boundary conditions force $w = 0$ so that equation (4.29) becomes

$$0 = -2(c-1)\,\mathrm{D}^2\theta, \tag{4.32}$$

where the homogeneous Dirichlet boundary conditions force θ to vanish as well. The upshot of these considerations is that non-negativity of the spectrum of the operator and boundary conditions in equations (4.26)–(4.28) (or equivalently, of the quadratic part of F_τ) is sufficient for the validity of the bound in (4.25). This non-negative definite requirement is directly analogous to a 'marginal stability' criterion.

From this point forward we switch the new parameter from $c(1 < c < \infty)$ to

$$a = \frac{c-1}{c} \quad (0 < a < 1). \tag{4.33}$$

Then the variational bound is

$$Nu - 1 \leqslant \inf_{0<a<1} \inf_{\tau(0)=1;\,\tau(1)=0} \left\{ \frac{1}{4a(1-a)} \int_0^1 (\tau'(z)+1)^2\,\mathrm{d}z \,|\, H_\tau^{(a)} \geqslant 0 \right\}, \tag{4.34}$$

where

$$H_\tau^{(a)}\{\theta\} \equiv \frac{1}{A_x A_y} \int \{a|\nabla\theta|^2 + \tau'(z)\,w\theta\}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \tag{4.35}$$

is defined for functions $\theta(\boldsymbol{x})$ satisfying $\theta(x, y, 0) = 0 = \theta(x, y, 1)$, and where

$$-\Delta w = Ra(\mathrm{D}^2-\Delta)\,\theta \tag{4.36}$$

with $w(x, y, 0) = 0 = w(x, y, 1)$. We say that a temperature profile $\tau(z)$ is 'a-stable' if $H_\tau^{(a)}$ is a positive quadratic form. We say that a temperature profile $\tau(z)$ is 'marginally a-stable' if $H_\tau^{(a)}$ is a non-negative quadratic form.

This is the generalized rigorous marginal stability criterion: For every $a \in (0,1)$ the convective heat transport is bounded from above by $(4a(1-a))^{-1}$ times the convective heat transport in a marginally a-stable background temperature profile. The linear conduction profile $\tau = 1-z$ is marginally a-stable for $Ra \leqslant a \times Ra_c$, so that we recover the energy stability result $Nu = 1$ for all $Ra < Ra_c$. In the next section we construct marginally stable background profiles to derive bounds on Nu for $Ra > Ra_c$.

5. Bounding the bounds

The best upper bound that the rigorous marginal stability criterion can produce is given by the solution of the variational problem in equation (4.34). We may derive an upper bound on the best bound directly by producing an a-stable test background temperature profile. Referring to equation (4.35), it is evident that a-stable profiles should have gradients confined close to the boundaries at $z = 0$ and $z = 1$ where w and θ vanish. This suggests that we use test profiles of the form

$$\tau_\delta(z) = \begin{cases} 1 - z/(2\delta); & 0 \leqslant z \leqslant \delta \\ \frac{1}{2}; & \delta \leqslant z \leqslant 1-\delta \\ (1-z)/(2\delta); & 1-\delta \leqslant z \leqslant 1 \end{cases} \tag{5.1}$$

as illustrated in figure 4. The parameter $\delta < \frac{1}{2}$ will be referred to as the boundary layer thickness for this family of test profiles. As will be shown explicitly in this section, for a given value of Ra and of a we will be able to choose δ small enough so that τ_δ is a-stable. To establish this we must consider the sign of

$$H^{(a)}_{\tau_\delta}\{\theta\} = \sum_k \int_0^1 \{a|\mathrm{D}\theta_k|^2 + ak^2|\theta_k|^2 + \tau'_\delta(z)\,\mathrm{Re}[w_k^*\,\theta_k]\}\,\mathrm{d}z \tag{5.2}$$

for functions $\theta_k(z)$ satisfying $\theta_k(0) = 0 = \theta_k(1)$ and where

$$(-\mathrm{D}^2 + k^2)\,w_k = Ra\,k^2\theta_k \tag{5.3}$$

with $w_k(0) = 0 = w_k(1)$. Here we have gone to Fourier transformed variables, but this is the same a-stability criterion as in equations (4.35) and (4.36). $H^{(a)}_{\tau_\delta}$ will be non-negative if and only if for each Fourier mode,

$$h_k = \int_0^1 \{a|\mathrm{D}\theta_k|^2 + ak^2|\theta_k|^2 + \tau'_\delta(z)\,\mathrm{Re}[w_k^*\,\theta_k]\}\,\mathrm{d}z \geqslant 0. \tag{5.4}$$

Hence in order to establish marginal a-stability it is sufficient to establish marginal a-stability wavenumber by wavenumber.

What we must do in order to show that $h_k \geqslant 0$ when δ is small enough, is to show that the indefinite term $\int \tau'_\delta(z)\,\mathrm{Re}[w_k^*\,\theta_k]\,\mathrm{d}z$ is smaller in magnitude than the positive definite terms $\int \{a|\mathrm{D}\theta_k|^2 + ak^2|\theta_k|^2\}\,\mathrm{d}z$. We do this in several steps.

First consider the indefinite term which may be bounded as follows:

$$\left|\int_0^1 \tau'(z)\,\mathrm{Re}[w_k^*\,\theta_k]\,\mathrm{d}z\right| \leqslant \int_0^1 |\tau_\delta^*(z)|\,|w_k^*|\,|\theta_k|\,\mathrm{d}z = \frac{1}{2\delta}\int_0^\delta |w_k^*|\,|\theta_k|\,\mathrm{d}z + \frac{1}{2\delta}\int_{1-\delta}^1 |w_k^*|\,|\theta_k|\,\mathrm{d}z. \tag{5.5}$$

Now consider the growth of $\theta_k(z)$ away from the boundary at $z = 0$. For $0 \leqslant z \leqslant \frac{1}{2}$, the fundamental theorem of calculus and the Schwarz inequality imply

$$|\theta_k(z)| = \left|\int_0^z \mathrm{D}\theta_k(z')\,\mathrm{d}z'\right| \leqslant z^{1/2}\left(\int_0^{1/2} |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z'\right)^{1/2}. \tag{5.6}$$

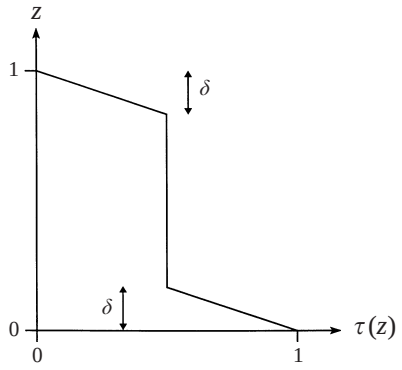

FIGURE 4. Test background temperature profile with the boundary layer structure.

Hence we see that θ_k is controlled pointwise by the integral of the square of its derivative.

The same kind of estimate may be used for w_k because we may use the $k^2|\theta_k|^2$ term to control the integral of the square of its derivative. Indeed, using equation (5.3) and integration by parts,

$$\int_0^1 k^2|\theta_k|^2\,\mathrm{d}z = \frac{1}{Ra^2}\int_0^1 \left\{\frac{1}{k^2}|\mathrm{D}^2 w_k|^2 + 2|\mathrm{D}w_k|^2 + k^2|w_k|^2\right\}\mathrm{d}z. \tag{5.7}$$

More integration by parts, the Schwarz inequality, and $2AB \leqslant (1/k^2)\,A^2 + k^2B^2$ imply

$$\begin{aligned}\int_0^1 |\mathrm{D}w_k|^2\,\mathrm{d}z = -\int_0^1 w_k^*\,\mathrm{D}^2 w_k\,\mathrm{d}z &\leqslant \left(\int_0^1 |w_k(z')|^2\,\mathrm{d}z'\right)^{1/2}\left(\int_0^1 |\mathrm{D}^2 w_k(z')|^2\,\mathrm{d}z'\right)^{1/2}\\ &\leqslant \frac{1}{2}\int_0^1 \left\{\frac{1}{k^2}|\mathrm{D}^2 w_k|^2 + k^2|w_k|^2\right\}\mathrm{d}z.\end{aligned} \tag{5.8}$$

Thus equation (5.7) may be rewritten

$$\int_0^1 k^2|\theta_k|^2\,\mathrm{d}z \geqslant \frac{4}{Ra^2}\int_0^1 |\mathrm{D}w_k|^2\,\mathrm{d}z. \tag{5.9}$$

Combining these elements we see that

$$\begin{aligned}\frac{1}{2\delta}\int_0^\delta |w_k^*|\,|\theta_k|\,\mathrm{d}z &\leqslant \frac{1}{2\delta}\int_0^\delta z\,\mathrm{d}z\left(\int_0^{1/2} |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z'\right)^{1/2}\left(\int_0^{1/2} |\mathrm{D}w_k(z'')|^2\,\mathrm{d}z''\right)^{1/2}\\ &\leqslant \frac{\delta}{4}\left\{\frac{c}{2}\int_0^{1/2} |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z' + \frac{1}{2c}\int_0^{1/2} |\mathrm{D}w_k(z'')|^2\,\mathrm{d}z''\right\}\end{aligned} \tag{5.10}$$

for any $c > 0$. Precisely analogous considerations near $z = 1$ yield the estimate

$$\begin{aligned}\frac{1}{2\delta}\int_{1-\delta}^1 |w_k^*|\,|\theta_k|\,\mathrm{d}z &\leqslant \frac{1}{2\delta}\int_{1-\delta}^1 (1-z)\,\mathrm{d}z\left(\int_{1/2}^1 |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z'\right)^{1/2}\left(\int_{1/2}^1 |\mathrm{D}w_k(z'')|^2\,\mathrm{d}z''\right)^{1/2}\\ &\leqslant \frac{\delta}{4}\left\{\frac{c}{2}\int_{1/2}^1 |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z' + \frac{1}{2c}\int_{1/2}^1 |\mathrm{D}w_k(z'')|^2\,\mathrm{d}z''\right\}.\end{aligned} \tag{5.11}$$

Adding these together and recalling equation (5.5),

$$\left|\int_0^1 \tau_\delta'(z)\,\mathrm{Re}[w_k^*\,\theta_k]\,\mathrm{d}z\right| \leqslant \frac{\delta}{8}\left\{c\int_0^1 |\mathrm{D}\theta_k(z')|^2\,\mathrm{d}z' + \frac{1}{c}\int_0^1 |\mathrm{D}w_k(z'')|^2\,\mathrm{d}z''\right\}. \tag{5.12}$$

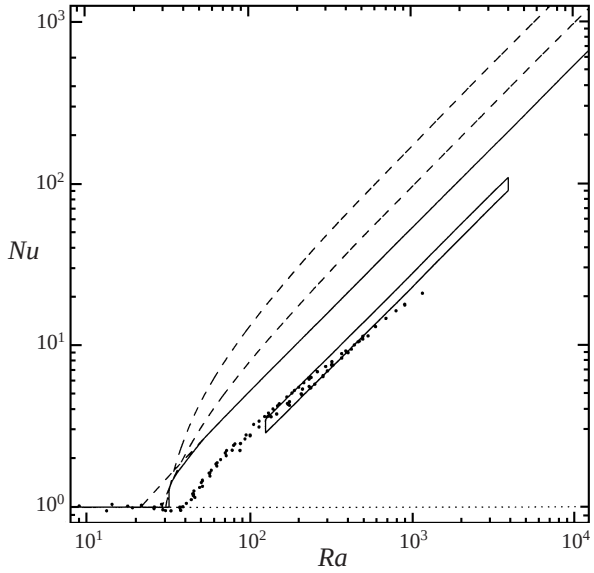

FIGURE 5. Rigorous upper bounds plotted along with experimental data (see caption to figure 3). Dashed lines: bounds for $a = 0.95$, 0.9 and $\frac{2}{3}$. Solid line: lower envelope for all $a \in (0, 1)$.

Now let us choose c and δ so that $\frac{1}{8}\delta c = a$ and $\delta/(8c) = 4a/Ra^2$, i.e.,

$$\delta = 16a/Ra \tag{5.13}$$

and

$$c = \tfrac{1}{2}Ra. \tag{5.14}$$

Then

$$\left| \int_0^1 \tau_\delta'(z)\, \mathrm{Re}[w_k^* \theta_k]\, \mathrm{d}z \right| \leqslant \int_0^1 a|\mathrm{D}\theta_k|^2\, \mathrm{d}z + \frac{4a}{Ra^2}\int_0^1 |\mathrm{D}w_k|^2\, \mathrm{d}z$$

$$\leqslant \int_0^1 a|\mathrm{D}\theta_k|^2\, \mathrm{d}z + \int_0^1 ak^2|\theta_k|^2\, \mathrm{d}z, \tag{5.15}$$

where in the end we have recalled equation (5.9).

What we have shown is that

$$\delta = \frac{16a}{Ra} \Rightarrow h_k \geqslant 0. \tag{5.16}$$

This means that τ_δ is a-stable when $\frac{1}{2} \geqslant \delta = 16a/Ra$ and we may use it to write down an upper bound according to equation (4.33):

$$Nu \leqslant 1 + \frac{1}{4a(1-a)}\int_0^1 (\tau_\delta'(z)+1)^2\, \mathrm{d}z = \frac{Ra}{128a^2(1-a)} - \frac{(1-2a)^2}{4a(1-a)} \tag{5.17}$$

for each a, $0 < a < 1$, and any $Ra \geqslant 32a$. Several of these bounding curves are plotted in figure 5. The prefactor of Ra is minimized with the choice $a = \frac{2}{3}$ yielding the asymptotic estimate

$$Nu \leqslant \tfrac{27}{512}\, Ra\left(1 + O\left(\frac{1}{Ra}\right)\right) \quad \text{as} \quad Ra \to \infty. \tag{5.18}$$

For each value of $Ra > 32$, the bound in equation (5.17) may be optimized by a choice of $a(Ra) \in (0, 1)$ but we need to solve a quartic polynomial to do so. It is just as easy to perform the minimization numerically, and in figure 5 we also show the lower envelope of this family of bounds. The optimized bound qualitatively captures some aspects of behaviour of the heat transport data. The bound bifurcates from the conduction state ($Nu = 1$) at a critical value of the Rayleigh number ('Ra_c' $= 32$) which is about 80 % of the true critical value, $4\pi^2$. The high-Rayleigh-number behaviour of the bound is Howard–Malkus–Kolmogorov–Spiegel scaling $Nu \sim Ra$ with a prefactor $\frac{27}{512} \approx 0.053$, about a factor 2 above Elder's (1967) data. We stress that the results in equations (5.17) and (5.18) and illustrated in figure 5, although clearly not optimal, are completely rigorous; they follow from the equations of motion alone without any statistical or other assumptions on the solutions. We can improve on these results by heading toward the ultimate optimization, and that is the subject of the next two sections.

6. The optimization problem: variations on a variation

In this section we describe the mathematical structure of the optimal problem for the best bound this approach has to offer. Although we cannot analytically solve the Euler–Lagrange equations for the optimal background temperature profile, we can establish a number of basic facts about the problem, notably the existence and uniqueness of optimal marginally a-stable profiles along with a glimpse of the functional geometry of the solution.

Consider, for each $0 < a < 1$, the variational problem at hand:

$$Nu - 1 \leqslant \inf_{\tau(0)=1;\,\tau(1)=0} \left\{ \frac{1}{4a(1-a)} \int_0^1 (\tau'(z)+1)^2 \,\mathrm{d}z \,|\, H_\tau^{(a)} \geqslant 0 \right\}, \tag{6.1}$$

where

$$H_\tau^{(a)}\{\theta\} \equiv \frac{1}{\Lambda_x \Lambda_y} \int \{a|\boldsymbol{\nabla}\theta|^2 + \tau'(z)\, w\theta\} \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \tag{6.2}$$

is defined for functions $\theta(\boldsymbol{x})$ satisfying $\theta(x, y, 0) = 0 = \theta(x, y, 1)$, and where

$$-\Delta w = Ra\,(\mathrm{D}^2 - \Delta)\,\theta \tag{6.3}$$

with $w(x, y, 0) = 0 = w(x, y, 1)$. Both θ and w are periodic in the horizontal directions.

A change of variables from $\tau(z)$ to

$$\psi(z) = \frac{1}{2[a(1-a)]^{1/2}} (\tau'(z)+1) \tag{6.4}$$

recasts the problem as

$$Nu - 1 \leqslant \inf_{\psi \in \boldsymbol{H}} \left\{ \int_0^1 \psi(z)^2 \,\mathrm{d}z \,|\, H_\psi \geqslant 0 \right\} \tag{6.5}$$

with

$$H_\psi\{\theta\} \equiv \frac{1}{\Lambda_x \Lambda_y} \int \left\{ a|\boldsymbol{\nabla}\theta|^2 + (2[a(1-a)]^{1/2}\,\psi(z) - 1)\, w\theta \right\} \mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z \tag{6.6}$$

and where the Hilbert space

$$\boldsymbol{H} = \left\{ \psi \in L^2[0,1], \int_0^1 \psi(z) \,\mathrm{d}z = 0 \right\}.$$

In equations (6.5) and (6.6) we have suppressed the a-dependence of everything; the parameter a maintains a fixed value in the following considerations. The problem then becomes that of finding the function with the smallest norm in the Hilbert space $\boldsymbol{H}$ that satisfies the constraint $H_\psi \geqslant 0$. Equivalently, the constraint can be expressed as the non-negativity of the spectrum of the self-adjoint linear operator L_ψ defined by $H_\psi\{\theta\} = \langle \theta, L_\psi \theta \rangle$, where $\langle \cdot, \cdot \rangle$ signifies the inner product in $L^2[0,1]$. Explicitly,

$$L_\psi \theta = -a\Delta\theta + \tfrac{1}{2}(2[a(1-a)]^{1/2}\psi(z) - 1)\, w + \tfrac{1}{2}(\mathrm{D}^2 - \Delta)\, v, \tag{6.7}$$

$$-\Delta w = Ra(\mathrm{D}^2 - \Delta)\,\theta, \tag{6.8}$$

$$-\Delta v = Ra(2[a(1-a)]^{1/2}\psi(z) - 1)\,\theta, \tag{6.9}$$

where v is the Lagrange multiplier enforcing the slaving of w to θ; it satisfies the same boundary conditions as w, i.e. v vanishes for $z = 0$ and $z = 1$ and is periodic in the horizontal directions. Let $\mu = \mu(\psi\}$ denote the lowest eigenvalue of L_ψ.

Let S be the set of functions in $\boldsymbol{H}$ that satisfies the constraint $H_\psi \geqslant 0$. Then S is the set of all mean zero square-integrable functions ψ such that $\mu\{\psi\} \geqslant 0$. S is not empty because $0 \in S$ when $Ra \leqslant 4\pi^2 a$ (recall that the linear conduction profile, corresponding to $\psi = 0$, is marginally a-stable for $Ra \leqslant a \times Ra_c$) and we explicitly constructed acceptable profiles for $Ra > 4\pi^2 a > 32a$ in §5.

Moreover, S is convex. This means that if ψ_1 and ψ_2 are in S, then convex combinations $t\psi_1 + (1-t)\,\psi_2 \in S$ for $0 \leqslant t \leqslant 1$. This fact follows from the observation that ψ appears linearly in H_ψ,

$$H_{t\psi_1 + (1-t)\psi_2}\{\theta\} = tH_{\psi_1}\{\theta\} + (1-t)\,H_{\psi_2}\{\theta\}, \tag{6.10}$$

so when H_{ψ_1} and $H_{\psi_2}\{\theta\}$ are non-negative, so is $H_{t\psi_1 + (1-t)\psi_2}\{\theta\}$. The boundary of S is the 'level set' of ψ satisfying $\mu\{\psi\} = 0$. When $Ra \leqslant 4\pi^2 a$, $0 \in S$, so the optimal element is precisely $\psi \equiv 0$ and $Nu = 1$ in that case. On the other hand, $0 \notin S$ when $Ra > 4\pi^2 a$ because then the linear conduction profile is not (marginally) a-stable. The existence of a minimizing solution when $Ra > 4\pi^2 a$ follows from the geometry of convex sets in the Hilbert space $\boldsymbol{H}$: given the closed convex set S and the point $0 \in \boldsymbol{H}$ not in S, there is an element of S with minimal distance to 0 (Reed & Simon 1980). Such a minimizer is unique, for if there were two distinct elements ψ_1 and ψ_2 realizing the minimum then $(\psi_1 + \psi_2)/2$ would also be in S with an even smaller norm. Finally, the unique minimizer ψ_{opt} is on the boundary of S, i.e. $\mu\{\psi_{opt}\} = 0$.

The Euler–Lagrange equations for the optimal profile may be derived from similar geometric considerations (Doering & Constantin 1996). The condition for the optimal profile is that the ray along ψ_{opt} should be parallel to the gradient to the level set $\mu\{\psi\} = 0$ evaluated at ψ_{opt} and projected into $\boldsymbol{H}$ (i.e. with its mean subtracted). So long as there is a unique lowest eigenstate ('ground state') for ψ_{opt}, the functional gradient $\delta\mu/\delta\psi$ projected onto $\boldsymbol{H}$, i.e. $P(\delta\mu/\delta\psi)$, where

$$P(f)(z) = f(z) - \int_0^1 f(z')\,\mathrm{d}z',$$

is given explicitly in terms of the associated ground-state eigenfunction for L_ψ via regular non-degenerate spectral perturbation theory. Then the Euler–Lagrange equation satisfied by the optimal profile ψ_{opt} is

$$\psi = \alpha P\left(\frac{\delta\mu}{\delta\psi}\right), \tag{6.11}$$

where the proportionality factor α is a Lagrange multiplier. Note that from the geometry (see, for example, the illustrations in Doering & Constantin 1996) we may conclude that $\alpha > 0$ when $Ra > 4\pi^2 a$ and that α approaches 0 from above as $Ra \downarrow 4\pi^2 a$. If the ground state for ψ_{opt} is not unique, say it is n-fold degenerate, then $P(\delta\mu/\delta\psi)$ has n different directions associated with it. Equation (6.11) still holds in this case, but with the Lagrange multiplier α an n-tuple of real numbers so that ψ is a linear combination of the different directions. In fact it is this latter case of a degenerate ground state which is apparently the generic situation in problems such as the one at hand (Howard 1963, 1972; Busse & Joseph 1972; Gupta & Joseph 1973; Joseph 1974; Busse 1978; Doering & Constantin 1996).

This structure may be seen in the Euler–Lagrange equations. With the eigenfunction normalization

$$\int |\theta|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1, \tag{6.12}$$

the variation of the eigenvalue μ with respect to ψ is

$$\frac{\delta\mu}{\delta\psi(z)} = \int 2[a(1-a)]^{1/2}\, w(x,y,z)\,\mathrm{d}x\,\mathrm{d}y, \tag{6.13}$$

so the Euler–Lagrange equation is

$$\psi_{opt}(z) = \alpha 2[a(1-a)]^{1/2} \int \left\{ w(x,y,z)\,\theta(x,y,z) - \int_0^1 w(x,y,z')\,\theta(x,y,z')\,\mathrm{d}z' \right\} \mathrm{d}x\,\mathrm{d}y. \tag{6.14}$$

In the above θ is the ground-state eigenfunction of $L_{\psi_{opt}}$ when the ground-state eigenvalue $\mu = 0$ *if the ground state of* $L_{\psi_{opt}}$ *is unique*. If that is the case, inserting equation (6.13) into $0 = L_{\psi_{opt}}\theta$ yields a closed nonlinear boundary value problem for θ (and the associated w and v) in which α is to be adjusted until the normalization condition in equation (6.12) is satisfied. In general, if the ground state of $L_{\psi_{opt}}$ is N-fold degenerate, $1 \leqslant N < \infty$, with eigenfunctions $\theta_1, \ldots, \theta_N$, then the Euler–Lagrange equation is

$$\psi_{opt}(z) = \sum_{n=1}^{N} \alpha_n\, 2[a(1-a)]^{1/2} \int \left\{ w_n(x,y,z)\,\theta_n(x,y,z) - \int_0^1 w_n(x,y,z')\,\theta_n(x,y,z')\,\mathrm{d}z' \right\} \mathrm{d}x\,\mathrm{d}y. \tag{6.15}$$

Then this is to be inserted into $0 = L_{\psi_{opt}}\theta$ and all the α_n must be adjusted until the normalization condition in equation (6.12) is satisfied. By the uniqueness argument, for a given value of Ra and of a there will be one and only one profile $\psi_{opt}(z)$ and Lagrange multiplier (the n-tuple $\{\alpha_n\}$) satisfying the spectral 'stability' constraint $\mu(\psi_{opt}) \geqslant 0$, the normalization in equation (6.12), and the closure in equation (6.15), i.e. $0 = L_{\psi_{opt}}\theta_n$ for $n = 1, \ldots, N$. For a given value of Ra and of a that solution will yield a bound on Nu which should then be minimized over the auxiliary parameter a to deduce the best bound this approach has to offer.

We cannot solve these equations analytically, although numerical solution is possible (Gupta & Joseph 1973; Doering & Hyman 1997; Nicodemus, Grossmann & Holthaus 1997*b*; Vitanov & Busse 1997). We will return to this problem in the

conclusion of this paper. But now we turn to modification of the variational problem for which we can analytically carry out the optimization and compute an explicit upper limit on $Nu(Ra)$.

7. A modified variational problem and its optimized bounds

We can make practical progress to derive improved explicit bounds by modifying the variational problem. Strengthening the constraints on the profiles over which we minimize in a certain way transforms the problem into a new one which may be optimized. The optimal bound for the new variational problem is necessarily higher than the ultimate optimal bound for the original problem described in the previous section, but nevertheless the net result is an improved upper bound compared to that established in §5.

We start with the variational problem as derived in §5 and restated in equations (6.1)–(6.4):

$$Nu-1 \leqslant \inf_{\psi\in\boldsymbol{H}}\left\{\int_0^1 \psi(z)^2\,\mathrm{d}z \,|\, H_\psi \geqslant 0\right\} \tag{7.1}$$

with $H_\psi(\theta\} = \sum_{\boldsymbol{k}} h_\psi^{(k)}\{\theta_{\boldsymbol{k}}\}$, where

$$h_\psi^{(k)}\{\theta\} \equiv \int_0^1 \{a|\mathrm{D}\theta_{\boldsymbol{k}}|^2 + ak^2|\theta_{\boldsymbol{k}}|^2 + (2[a(1-a)]^{1/2}\,\psi(z)-1)\,\mathrm{Re}[w_{\boldsymbol{k}}^*\,\theta_{\boldsymbol{k}}]\}\,\mathrm{d}z). \tag{7.2}$$

As in §5, we have gone to the horizontal Fourier transform representation for θ and w where, mode by mode, the slaving of w to θ is

$$(-\mathrm{D}^2+k^2)\,w_{\boldsymbol{k}} = Ra\,k^2\,\theta_{\boldsymbol{k}}. \tag{7.3}$$

Both $\theta_{\boldsymbol{k}}(z)$ and $w_{\boldsymbol{k}}(z)$ satisfy homogeneous Dirichlet conditions at $z=0$ and $z=1$.

Note now that equation (7.3) and the boundary conditions on $w_{\boldsymbol{k}}$ imply

$$\begin{aligned}\int_0^1 ak^2|\theta_{\boldsymbol{k}}|^2\,\mathrm{d}z &= \int_0^1 \frac{a}{Ra^2\,k^2}|(-\mathrm{D}^2+k^2)\,w_{\boldsymbol{k}}|^2\,\mathrm{d}z \\ &= \frac{a}{Ra^2}\int_0^1\left(\frac{1}{k^2}|\mathrm{D}^2 w_{\boldsymbol{k}}|^2 + 2|\mathrm{D}w_{\boldsymbol{k}}|^2 + k^2|w_{\boldsymbol{k}}|^2\right)\mathrm{d}z.\end{aligned} \tag{7.4}$$

Integrations by parts, Schwarz's inequality and the fact that $(1/k^2)\,A^2 + k^2\,B^2 \geqslant 2AB$ ensure that

$$\int_0^1\left(\frac{1}{k^2}|\mathrm{D}^2 w_{\boldsymbol{k}}|^2 + k^2|w_{\boldsymbol{k}}|^2\right)\mathrm{d}z \geqslant 2\int_0^1(-\mathrm{D}^2 w_{\boldsymbol{k}}^*)\,w_{\boldsymbol{k}}\,\mathrm{d}z = 2\int_0^1|\mathrm{D}w_{\boldsymbol{k}}|^2\mathrm{d}z, \tag{7.5}$$

Hence

$$\int_0^1 ak^2|\theta_{\boldsymbol{k}}|^2\,\mathrm{d}z \geqslant \frac{4a}{Ra^2}\int_0^1|\mathrm{D}w_{\boldsymbol{k}}|^2\,\mathrm{d}z \tag{7.6}$$

and $H_\psi\{\theta\} \geqslant \sum_{\boldsymbol{k}} J_\psi\{\theta_{\boldsymbol{k}}, w_{\boldsymbol{k}}\}$ where

$$J_\psi\{\theta_{\boldsymbol{k}}, w_{\boldsymbol{k}}\} = \int_0^1\left(a|\mathrm{D}\theta_{\boldsymbol{k}}|^2 + \frac{4a}{Ra^2}|\mathrm{D}w_{\boldsymbol{k}}|^2 + [(2a(1-a)]^{1/2}\,\psi(z)-1)\,\mathrm{Re}[w_{\boldsymbol{k}}^*\,\theta_{\boldsymbol{k}}]\right)\mathrm{d}z. \tag{7.7}$$

Observe that if the quadratic form $J_\psi\{\theta, w\} \geqslant 0$ for all $\theta(z)$ and $w(z)$ satisfying homogeneous Dirichlet conditions at $z=0$ and $z=1$ (even if θ and w are not otherwise

related), then $H_\psi\{\theta\} \geqslant 0$. Changing variables to $f(z) = \theta(z) + (2/Ra)\, w(z)$ and $g(z) = \theta(z) - (2/Ra)\, w(z)$, we see that J_ψ is non-negative when $K_\psi\{f, g\} = K_\psi^{(+)}\{f\} + K_\psi^{(-)}\{g\}$ is non-negative, where

$$K_\psi^{(\pm)}\{h\} = \int_0^1 \left\{ |\mathrm{D}h|^2 \pm \frac{Ra}{4a}(2[a(1-a)]^{1/2}\,\psi(z) - 1)\,|h|^2 \right\} \mathrm{d}z \tag{7.8}$$

for functions vanishing at $z = 0$ and $z = 1$. In particular, J_ψ is non-negative when both $K_\psi^{(\pm)}$ are non-negative. That is, the set $S = \{\psi \in \boldsymbol{H} \,|\, H_\psi \geqslant 0\}$ of appropriately constrained profiles contains the (convex) set S' defined by

$$S' = \{\psi \in \boldsymbol{H} \,|\, \text{both } K_\psi^{(\pm)} \geqslant 0\}. \tag{7.9}$$

Thus the minimization of the norm of ψ over S is necessarily not greater than the minimization of the norm of ψ over S' so we have an upper bound on the optimal upper bound,

$$Nu - 1 \leqslant \inf_{\psi(z) \in S} \int_0^1 \psi(z)^2 \,\mathrm{d}z \leqslant \inf_{\psi(z) \in S'} \int_0^1 \psi(z)^2 \,\mathrm{d}z. \tag{7.10}$$

The point is that the constraints $K_\psi^{(\pm)} \geqslant 0$ are considerably simpler than the full constraint $H_\psi \geqslant 0$ (in particular, the dependence on the horizontal wavenumber $\boldsymbol{k}$ has been eliminated) and the optimization process may now be carried out exactly. In fact what we shall do is find the optimal ψ constrained only by $K_\psi^{(+)} \geqslant 0$ and afterwards confirm that indeed $K_\psi^{(-)} \geqslant 0$ for that choice of ψ. This is now essentially the same 'optimal marginal stability' problem that arose in the context of deriving variational bounds on heat transport in a fluid layer (Doering & Constantin 1996).

The sign of $K_\psi^{(+)}$ is the sign of the lowest (ground-state) eigenvalue $\lambda\{\psi\}$ of the linear operator in $K_\psi^{(+)}$, i.e. the eigenvalue problem

$$\lambda f = -\mathrm{D}^2 f + \frac{Ra}{4a}(2[a(1-a)]^{1/2}\,\psi(z) - 1) f. \tag{7.11}$$

For $\psi = 0$, corresponding to the pure conduction state, the lowest eigenvalue is exactly

$$\lambda\{0\} = \pi^2 - \frac{Ra}{4a} \tag{7.12}$$

which is non-negative if and only if $Ra \leqslant 4\pi a^2$. Note also that $K_0^{(-)} > 0$, so we observe that $0 \in S'$ precisely when $Ra \leqslant 4\pi a^2$, the sharpest possible value. This indicates that the modification of the problem from the original constraint (minimization over S) to these strengthened constraints (minimization over S') does not change the bifurcation point of the bound from the conduction state $Nu = 1$.

For $Ra > 4\pi a^2$, the considerations of the previous section carry through so the optimal ψ will be the unique point on the isospectral surface $\lambda\{\psi\} = 0$ where the ray $\sim \psi$ is parallel to the projection onto $\boldsymbol{H}$ of the functional gradient of $\lambda\{\psi\}$. That is,

$$\psi = \alpha P\left(\frac{\delta\lambda}{\delta\psi}\right), \tag{7.13}$$

where α is a Lagrange multiplier. As before, we have $\alpha > 0$ for $Ra > 4\pi^2 a$ and α approaches 0 from above as $Ra \downarrow 4\pi^2 a$. The ground state is unique for the

one-dimensional Schrödinger equation (7.11) so the functional gradient may be evaluated by standard non-degenerate regular perturbation theory. With eigenfunction normalization

$$\int_0^1 |f|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1, \tag{7.14}$$

the variation of the eigenvalue λ with respect to ψ is

$$\frac{\delta\lambda}{\delta\psi(z)} = \frac{Ra}{2a}[a(1-a)]^{1/2}\,|f(z)|^2 \tag{7.15}$$

and the Euler–Lagrange equation is

$$\psi_{opt}(z) = \frac{\alpha}{2[a(1-a)]^{1/2}}\{f(z)^2-1\}, \tag{7.16}$$

where $f(z)$ (without loss of generality real) is the ground-state eigenfunction for ψ_{opt} and for convenience we have scaled the Lagrange multiplier. Inserting equation (7.15) into (7.11) with $\lambda = 0$ yields a closed nonlinear boundary value problem for f in which α is to be adjusted until the normalization condition in equation (7.14) is satisfied. That is,

$$0 = -\mathrm{D}^2 f + \frac{Ra}{4a}(\alpha\{f(z)^2-1\}-1)f, \tag{7.17}$$

with $f(0) = 0 = f(1)$ and α fixed by

$$\int_0^1 |f|^2 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 1. \tag{7.18}$$

Equation (7.17) is a nonlinear Schrödinger equation, a.k.a. Duffing's equation, whose exact solutions are Jacobi elliptic functions (Abramowitz & Stegun 1965). After considerable calculation, it may be verified that $2[a(1-a)]^{1/2}\psi_{opt}(z) \leqslant 1$ so that $K^{(-)}_{\psi_{opt}} \geqslant 0$ and it is ensured that the α-stability constraint is satisfied (see Doering & Constantin 1996).

After considerably more calculation the upper bound may be written in closed form by reparameterizing $Ra \in (4\pi^2 a, \infty)$ in terms of $m \in (0, 1)$ according to

$$Ra = 4aK(m)\,[8E(m) + 4(m-1)\,K(m)], \tag{7.19}$$

where the complete elliptic integrals $K(m)$ and $E(m)$ are

$$K(m) = \int_0^1 \frac{\mathrm{d}t}{[(1-t^2)\,(1-mt^2)]^{1/2}} \quad \text{and} \quad E(m) = \int_0^1 \left(\frac{1-mt^2}{1-t^2}\right)^{1/2} \mathrm{d}t. \tag{7.20}$$

Defining the auxiliary variables

$$\eta(m) = 8K(m)\,[K(m) - E(m)] \quad \text{and} \quad \sigma(m) = \frac{2mK(m)^3}{K(m)-E(m)}, \tag{7.21}$$

the rigorous upper bound on the Nusselt number is then

$$Nu \leqslant 1 + \frac{4\eta(m)\,[\alpha\eta(m) + Ra - 4a\sigma(m)]}{3Ra^2(1-a)}. \tag{7.22}$$

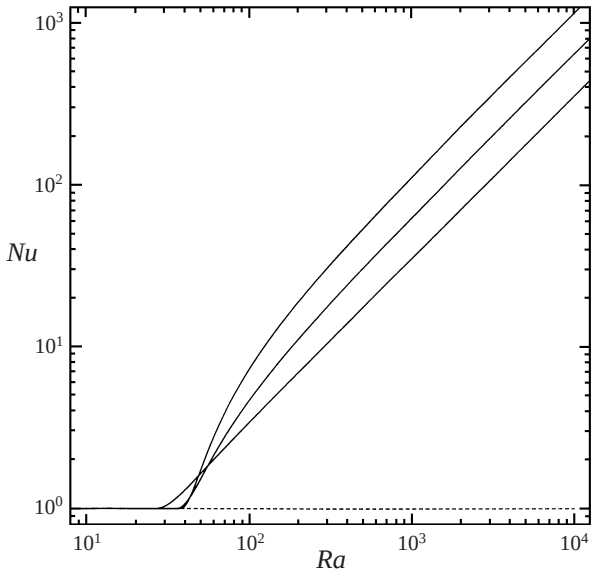

FIGURE 6. Rigorous 'semi-optimal' upper bounds for $a = 0.95$, 0.9 and $\frac{2}{3}$.

These formulae are easily evaluated numerically and several bounding curves are shown in figure 6. Let us examine some limits. For $m \to 0$, both K and $E \to \frac{1}{2}\pi$, so that $Ra \to 4\pi^2 a$ and the bound decreases down to the conduction value. In the other limit as $m \to 1$, $K \sim \frac{1}{2}\ln(16/(1-m)) \to \infty$ while $E \to 0$. Hence $Ra \sim 32aK$, $\eta \sim 8K^2 - 8K$, and $\sigma \sim 2K^2 + 2K$ so the asymptotic bound is

$$Nu - 1 \leqslant \frac{4\eta(m)\,[\alpha\eta(m) + Ra - 4a\sigma(m)]}{3Ra^2(1-a)} \sim \frac{1}{192a^2(a-a)}\,Ra. \tag{7.23}$$

This asymptotic bound is optimized over $a \in (0,1)$ with the choice $a = \frac{2}{3}$, yielding the large-Ra result

$$Nu \leqslant \tfrac{9}{256}\,Ra\left(1 + O\left(\frac{1}{Ra}\right)\right) \quad \text{as} \quad Ra \to \infty. \tag{7.24}$$

This bound also exhibits the Howard–Malkus–Kolmogorov–Spiegel scaling prefactor $\frac{9}{256} \approx 0.035$, representing a one-third decrease compared to the bound derived in §5.

8. Arbitrary Prandtl–Darcy numbers in two spatial dimensions

The analysis up to this point applies as well to the arbitrary Prandtl–Darcy number problem in two spatial dimensions. To see this, consider the evolution of the vorticity $\omega = u_z - w_x$,

$$B\left(\frac{\partial\omega}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\omega\right) + \omega = -Ra\frac{\partial\theta}{\partial x}. \tag{8.1}$$

What will ultimately make this two-dimensional calculation work are the facts that (i) there is no vortex stretching term, and (ii) the dissipative term in the vorticity evolution equation does not utilize any boundary conditions on ω. The system of dynamical

equations is now supplemented with an initial vorticity field $\omega_0(x,z)$. At each instant of time the velocity vector field $\boldsymbol{u}=(u,w)$ is determined by solving the Poisson equation $-\Delta w=\omega_x$ (with periodic boundary conditions in x and homogeneous Dirichlet boundary conditions at $z=0$ and $z=1$) for w, and subsequently $u_x=-w_z$ (with periodic boundary conditions in x) for u.

Decomposing the temperature field, as before, into the background and fluctuation fields, $T(x,z,t)=\tau(z)+\theta(x,z,t)$ with $\tau(0)=1$ and $\tau(1)=0$, we find the following set of equations for bulk integral quantities:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\int\theta^2\,\mathrm{d}x\,\mathrm{d}z=-\int|\boldsymbol{\nabla}\theta|^2\,\mathrm{d}x\,\mathrm{d}z-\int\tau' w\theta\,\mathrm{d}x\,\mathrm{d}z+\int\tau''\theta\,\mathrm{d}x\,\mathrm{d}z, \tag{8.2}$$

$$\frac{1}{2}\int|\boldsymbol{\nabla}T|^2\,\mathrm{d}x\,\mathrm{d}z=\frac{1}{2}\int|\boldsymbol{\nabla}\theta|^2\,\mathrm{d}x\,\mathrm{d}z+\frac{1}{2}\int|\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}z-\int\tau''\theta\,\mathrm{d}x\,\mathrm{d}z, \tag{8.3}$$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}B\int\omega^2\,\mathrm{d}x\,\mathrm{d}z=-\int\omega^2\,\mathrm{d}x\,\mathrm{d}z-Ra\int\omega\theta_x\,\mathrm{d}x\,\mathrm{d}z. \tag{8.4}$$

Choose parameters $b>0$ and $c>1$, add together $c\times(8.2)+2\times(8.3)+(b/Ra^2)\times(8.4)$, and take the long-time average (supremum) to find

$$Nu=\int_0^1\tau'(z)^2\,\mathrm{d}z+\sup_{T_0,\omega_0}\limsup_{t\to\infty}\frac{1}{t}\int_0^t\left(\frac{-1}{\Lambda_x}\int\left\{(c-1)\,|\boldsymbol{\nabla}\theta|^2+\frac{b}{Ra^2}\omega^2+\frac{b}{Ra}\omega\theta_x+c\tau' w\theta+(2-c)\,\theta\tau''\right\}\mathrm{d}x\,\mathrm{d}z\right)\mathrm{d}s. \tag{8.5}$$

An upper bound on Nu is obtained by replacing the second term above by its absolute minimum over all relevant θ and ω:

$$Nu\leqslant\frac{1}{\lambda_x}\int|\boldsymbol{\nabla}\tau|^2\,\mathrm{d}x\,\mathrm{d}z-\inf_{\theta,\omega}G_\tau\{\theta,\omega\}, \tag{8.6}$$

where

$$G_\tau\{\theta,\omega\}=\frac{1}{\Lambda_x}\int\left\{(c-1)|\boldsymbol{\nabla}\theta|^2+\frac{b}{Ra^2}\omega^2+\frac{b}{Ra}\omega\theta_x+c\tau' w\theta+(2-c)\,\theta\tau''\right\}\mathrm{d}x\,\mathrm{d}z. \tag{8.7}$$

The Euler–Lagrange equations for the fields minimizing G_τ are

$$0=-2(c-1)\,\Delta\theta-\frac{b}{Ra}\omega_x+c\tau'(z)\,w+(2-c)\,\tau''(z), \tag{8.8}$$

$$0=2\frac{b}{Ra^2}\omega+\frac{b}{Ra}\theta_x-W_x, \tag{8.9}$$

$$0=\Delta W+c\tau'(z)\,\theta, \tag{8.10}$$

$$0=\Delta w+\omega_x, \tag{8.11}$$

where the Lagrange multiplier field $W(x,z)$, introduced to enforce the $0=\omega_x+\Delta w$ constraint, satisfies homogeneous Dirichlet conditions at $z=0$ and $z=1$ and periodic horizontal conditions. The relevant solution is horizontally translation invariant, i.e. $\theta=\Theta(z)$, $w=0$ and $\omega=0$ and $W=W(z)$, so equation (8.8) reduces to

$$0=-2(c-1)\,\Theta''+(2-c)\,\tau''. \tag{8.12}$$

Integrating and using the boundary conditions $\Theta(0) = 0 = \Theta(1)$ we find

$$\Theta(z) = -\frac{c-2}{2(c-1)}(\tau(z)-1+z). \tag{8.13}$$

Inserting the minimizer into G_τ in equation (8.6), we deduce

$$Nu \leqslant 1+\frac{c^2}{4(c-1)}\int_0^1 (\tau'(z)+1)^2\,\mathrm{d}z. \tag{8.14}$$

It is ensured that $\Theta(z)$ is indeed the minimizer and that this is a rigorous upper bound so long as the quadratic part of G_τ is positive definite.

As in the infinite Prandtl–Darcy number case, slightly less is required: it is enough that the linear operator in the quadratic part of G_τ be non-negative with a null space orthogonal to the inhomogeneous term proportional to $\tau''(z)$. This means that it is necessary that the eigenvalues of

$$\lambda\theta = -2(c-1)\,\Delta\theta - \frac{b}{Ra}\omega_x + c\tau'(z)\,w, \tag{8.15}$$

$$\lambda\omega = 2\frac{b}{Ra^2}\omega + \frac{b}{Ra}\theta_x - W_x, \tag{8.16}$$

$$0 = \Delta W + c\tau'(z)\,\theta, \tag{8.17}$$

$$0 = \Delta\omega + \omega_x \tag{8.18}$$

(with periodic horizontal conditions and homogeneous Dirichlet boundary conditions at $z = 0$ and $z = 1$ for θ, w and W) satisfy $\lambda \geqslant 0$ and the eigenfunction(s) corresponding to $\lambda = 0$ must be orthogonal to τ''. The orthogonality of the null-space condition turns out to be automatically satisfied: with $\lambda = 0$ the horizontally Fourier transformed version of equations (8.15)–(8.18) are

$$0 = -2(c-1)\,(\mathrm{D}^2-k^2)\,\theta_k - \mathrm{i}k\frac{b}{Ra}\omega_k + c\tau'(z)\,w_k, \tag{8.19}$$

$$0 = 2\frac{b}{Ra^2}\omega_k + \mathrm{i}k\frac{b}{Ra}\theta - \mathrm{i}kW_k, \tag{8.20}$$

$$0 = (\mathrm{D}^2-k^2)\,W_k + c\tau'(z)\,\theta_k, \tag{8.21}$$

$$0 = (\mathrm{D}^2-k^2)\,w_k + \mathrm{i}k\omega_k. \tag{8.22}$$

The eigenfunctions will be orthogonal to $\tau''(z)$ (actually orthogonal to any function of z alone) if $k \neq 0$, and it is easy to see that this is necessarily so for any null eigenfunction. This is because if $k = 0$, then equation (8.22) and the w_k boundary conditions force $w_k = 0$, and then equation (8.19) and the θ_k boundary conditions force $\theta_k = 0$ and equation (8.20) forces $\omega_k = 0$.

Hence the non-negativity of the spectrum of the operator and boundary conditions in equations (8.15)–(8.18) (or equivalently, of the quadratic part of G_τ) is sufficient for the validity of the bound in (8.14). Again, this non-negative definite requirement is analogous to a 'marginal stability' criterion.

Now we switch the new parameter from $c\,(1 < c < \infty)$ to

$$a = \frac{c-1}{c} \quad (0 < a < 1). \tag{8.23}$$

The variational bound for arbitrary Prandtl–Darcy number in two dimensions is

$$Nu-1 \leqslant \inf_{0<a<1} \inf_{\tau(0)=1;\,\tau(1)=0} \left\{ \frac{1}{4a(1-a)} \int_0^1 (\tau'(z)+1)^2 \,\mathrm{d}z \,|\, Q_\tau^{(a,b)} \geqslant 0 \right\}, \tag{8.24}$$

where for any $0 < a < 1$ and $b > 0$,

$$Q_\tau^{(a,b)}\{\theta,\omega\} \equiv \frac{1}{A_x} \int \left\{ a|\boldsymbol{\nabla}\theta|^2 + \frac{b}{cRa^2}\omega^2 + \frac{b}{cRa}\omega\theta_x + \tau' w\theta \right\} \mathrm{d}x\,\mathrm{d}z \tag{8.25}$$

is defined for functions $\theta(\boldsymbol{x})$ satisfying $\theta(x,y,0) = 0 = \theta(x,y,1)$, with

$$-\Delta w = \omega_x \tag{8.26}$$

and boundary conditions $w(x,0) = 0 = w(x,1)$.

In order to recover the variational bounds derived in § 7, it is sufficient to strengthen the constraint $Q_\tau^{(a,b)} \geqslant 0$ to reproduce the same variational problem as that solved in § 7.

In terms of Fourier transformed variables, $(-\mathrm{D}^2 + k^2)\, w_k = \mathrm{i}k\omega_k$, and

$$Q_\tau^{(a,b)} = \sum_k \int_0^1 \left\{ a|\mathrm{D}\theta_k|^2 + ak^2|\theta_k|^2 + \frac{b}{cRa^2}\frac{1}{k^2}|(-\mathrm{D}^2+k^2)\, w_k|^2 \right.$$
$$\left. -\frac{b}{cRa}[(-\mathrm{D}^2+k^2)\, w_k]\,\theta_k^* + \tau' w_k^*\,\theta_k \right\} \mathrm{d}z. \tag{8.27}$$

Note then that

$$\left| \frac{b}{cRa}[(-\mathrm{D}^2+k^2)\, w_k]\,\theta_k^* \right| \leqslant \frac{1}{4}\frac{b^2}{ac^2\, Ra^2\, k^2}|(-\mathrm{D}^2+k^2)\, w_k|^2 + ak^2|\theta_k|^2, \tag{8.28}$$

so

$$Q_\tau^{(a,b)} \geqslant \sum_k \int_0^1 \left\{ a|\mathrm{D}\theta_k|^2 + \frac{1}{Ra^2k^2}\left(\frac{b}{c} - \frac{b^2}{4ac^2}\right)|(-\mathrm{D}^2+k^2)\, w_k|^2 + \tau' w_k^*\,\theta_k \right\} \mathrm{d}z. \tag{8.29}$$

Now choose $b = 2ac = 2(c-1) > 0$ to find

$$Q_\tau^{(a,2ac)} \geqslant \sum_k \int_0^1 \left\{ a|\mathrm{D}\theta_k|^2 + \frac{a}{Ra^2k^2}|(-\mathrm{D}^2+k^2)\, w_k|^2 + \tau' w_k^*\,\theta_k \right\} \mathrm{d}z. \tag{8.30}$$

Recall equations (7.4) and (7.5):

$$\int_0^1 \frac{a}{Ra^2k^2}|(-\mathrm{D}^2+k^2)\, w_{\boldsymbol{k}}|^2 \mathrm{d}z = \frac{a}{Ra^2}\int_0^1 \left(\frac{1}{k^2}|\mathrm{D}^2 w_{\boldsymbol{k}}|^2 + 2|\mathrm{D}w_{\boldsymbol{k}}|^2 + k^2|w_{\boldsymbol{k}}|^2 \right) \mathrm{d}z$$
$$\leqslant \frac{4a}{Ra^2}\int_0^1 |\mathrm{D}w_{\boldsymbol{k}}|^2 \,\mathrm{d}z. \tag{8.31}$$

Hence, using $\psi(z) = [a(1-a)]^{-1/2}\ (\tau'(z)+1)$, we deduce that $Q_\tau^{(a,2ac)} \geqslant \sum_{\boldsymbol{k}} J_\psi\{\theta_{\boldsymbol{k}}, \omega_{\boldsymbol{k}}\}$, where

$$J_\psi\{\theta_{\boldsymbol{k}}, w_{\boldsymbol{k}}\} = \int_0^1 a|\mathrm{D}\theta_k|^2 + \frac{4a}{Ra^2}|\mathrm{D}w_{\boldsymbol{k}}|^2 + (2[a(1-a)]^{1/2}\,\psi(z) - 1)\,\mathrm{Re}[w_{\boldsymbol{k}}^*\,\theta_{\boldsymbol{k}}]\}\,\mathrm{d}z. \tag{8.32}$$

Observe that *if* the quadratic form $J_\psi\{\theta, w\} \geqslant 0$ for all $\theta(z)$ and $w(z)$ satisfying homogeneous Dirichlet conditions at $z = 0$ and $z = 1$, *then* $Q_\tau^{(a,2ac)} \geqslant 0$ and the calculation proceeds from this point precisely as from equation (7.7) onward. We

conclude that the 'semi-optimal' bounds in equations (7.19)–(7.24) are valid for the arbitrary Prandtl–Darcy number problem in two spatial dimensions. We remark that these bounds are not truly optimal in the sense that we have optimized an over-constrained problem. Better bounds would follow from optimization with the constraint $Q_\tau^{(a,b)} \geqslant 0$ followed by optimization over a and b.

9. Conclusion

In figure 7 we plot the lower envelope of the 'semi-optimal' bounds in equations (7.19)–(7.22) along with experimental data of Buretta (1972) and Elder (1967). As is evident, this upper bound is in striking accord with these experimental measurements of the heat transport. We have not analytically minimized the semi-optimal a-stable bounds over the parameter $a \in (0, 1)$ for all values of Ra, but we have done so for limiting cases. The bounds are minimized for $Ra < 4\pi^2$ by choosing $\tau(z) = 1 - z$ and $a = 1$. The best upper bound in figure 7 bifurcates from the conduction value $Nu = 1$ exactly at the critical value $Ra_c = 4\pi^2$. Immediately above onset the bound is optimized by $a \sim 4\pi^2/Ra$ yielding $Nu \leqslant 1 + (1/2\pi^2)\,(Ra - 4\pi^2)$ which agrees precisely with the asymptotic weakly nonlinear stability prediction. The optimal value of a decreases monotonically from 1 to $a = \frac{2}{3}$ for further increasing Ra, yielding $Nu \leqslant \frac{9}{256}\,Ra \approx 0.035\,Ra$ as $Ra \to \infty$. Elder fit his highest Ra Howard–Malkus–Kolmogorov–Spiegel scaling data to $Nu \approx 0.025\,Ra$ so the bound is only about 40 % above the experimental data.

Also plotted in figure 7 – and falling directly on top of the envelope of the curves from our equations (7.19)–(7.22) – is the curve for Busse & Joseph's (1972) 'single wavenumber' bound for the arbitrary Prandtl–Darcy number problem from equations (4.11) and (4.12) in their paper, a bound they proved valid for $Ra < 113$ only. The Euler–Lagrange equations for their more general problem, derived via Howard's method for statistically stationary flows, are for a limited range of Ra of the same form as those for the modified variational problem in §7 of this paper. Their bound is only valid for $Ra < 113$ because, in the language of this paper, their optimal profile has a non-degenerate ground state only for $Ra < 113$; above that its ground state is multiply degenerate and the optimal solution solves different Euler–Lagrange equations equations.

The results in this paper have now established that the formula in equations (4.11) and (4.12) of Busse & Joseph's (1972) paper is indeed a rigorous upper bound for all Ra for the infinite Prandtl–Darcy number problem in three dimensions, and for arbitrary Prandtl–Darcy numbers in two dimensions. Conveniently, this formula was plotted along with an even larger collection of experimental data by Lister (1990) in figure 2 of his paper, further showing the consistency and precision of the bound for a variety of experimental situations. (Lister's plot also neatly displays the inevitable breakdown of the model at high Ra depending on the pore length scale in the various experiments.)

Gupta & Joseph (1972) considered the infinite Prandtl–Darcy number problem in the context of Howard's method for statistically stationary flows. They derived Euler–Lagrange equations of the same mathematical form as the optimal profile problem derived in §6 of this paper (there is no parameter a in that theory). They solved the equations numerically for a limited range of Ra finding a bifurcation from 'single wavenumber' (non-degenerate ground state, in our language) solutions to 'multiple wavenumber' (degenerate ground state, to us) solutions. For $Ra < 500$, their numerically computed optimal bounds are in exceptionally sharp agreement with Buretta's (1972) data. Although we have not solved the truly optimal background

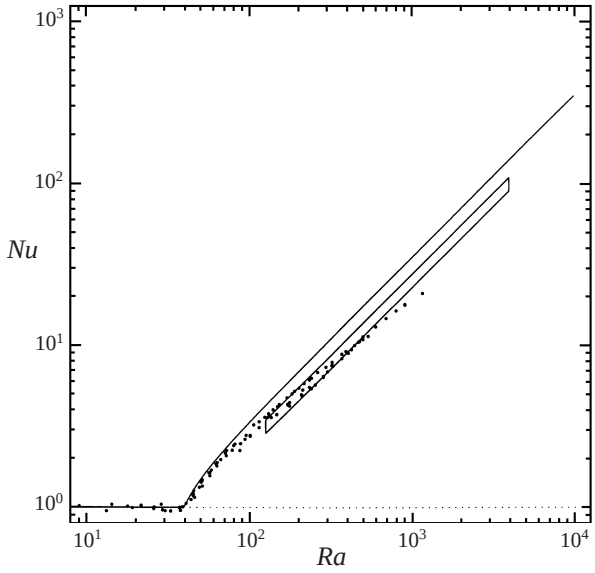

FIGURE 7. Rigorous upper bound plotted along with experimental data (see caption to figure 3). The solid curve is the lower envelope of the curves in figure 6 for all $a \in (0, 1)$.

profile equations from §6, we might expect similar results from this approach. Gupta & Joseph could not access the $Ra \to \infty$ asymptotic regime rigorously, but an approximate analysis using Busse's (1978) multiple boundary layer theory yielded $Nu \sim Ra$ scaling with a prefactor (≈ 0.016) which falls below Elder's (1967) highest-Ra 'granular material' experimental data that displayed Howard–Malkus–Kolmogorov–Spiegel scaling. (For porous media made up of larger glass beads – and hence larger pore scales – Elder did not observe the $Nu \sim Ra^1$ scaling at the highest Ra, but rather $Nu \sim Ra^\beta$ with a smaller exponent β as expected from the breakdown of Darcy's Law in the model analysed here.) It remains to be seen if a truly optional rigorous analysis of the problem at hand via the background method will produce such a smaller prefactor.

We have established a number of new results in this paper. First, we have produced a uniformly valid rigorous upper bound on the heat transport which agrees quantitatively with the predictions of linear, nonlinear, and weakly nonlinear stability theory. Secondly, the 'turbulent' $Ra \to \infty$ limit is rigorously described by the scaling expected from the physically appealing albeit heuristic marginally stable boundary layer argument, and even the prefactor is in reasonable agreement with experimental data. Thirdly, we have developed a fundamental mathematical connection between (nonlinear) hydrodynamic stability theory and the analysis of turbulent dynamics, the antithesis of stability. In this way we have rigorously realized some of the intuition embodied in the marginally stable boundary layer argument, producing relatively sharp and experimentally relevant estimates of the nonlinear heat transport.

We thank P. Steen for suggesting this problem to us, and D. Joseph for providing us with a copy of Buretta's (1972) thesis. We are grateful to S. Grossmann, M. Holthaus, R. Kerswell, R. Nicodemus, and E. Titi for providing access to their works prior to publication. For interesting, helpful, and/or challenging comments, discussions and remarks we acknowledge A. Barvinok, R. Behringer, F. Busse, L. Howard, C. Miller,

E. Spiegel and R. Worthing. The authors began this project while in residence at the Isaac Newton Institute, Cambridge. This research was supported in part by awards from the US National Science Foundation and US Department of Energy.

REFERENCES

Abramowitz, M. & Stegun, I. A. 1965 *Handbook of Mathematical Functions*. Dover, 1046 pp.

Buretta, R. J. 1972 Thermal convection in a fluid filled porous layer with uniform internal heat sources. PhD Thesis, University of Minnesota.

Busse, F. H. 1978 The optimal theory of turbulence. *Adv. Appl. Mech.* **18**, 77–121.

Busse, F. H. & Joseph, D. D. 1972 Bounds for heat transport in a porous layer. *J. Fluid Mech.* **54**, 521–543.

Chandrasekhar, S. 1961 *Hydrodynamic and Hydromagnetic Stability*, 1st Edn. Oxford University Press, 652 pp.

Constantin, P. & Doering, C. R. 1995*a* Variational bounds on energy dissipation in incompressible flows. II. Channel flow. *Phys. Rev.* E **51**, 3192–3198.

Constantin, P. & Doering, C. R. 1995*b* Variational bounds in dissipative systems. *Physica* D **82**, 221–228.

Constantin, P. & Doering, C. R. 1996 Heat transfer in convective turbulence. *Nonlinearity* **9**, 1049–1060.

Doering, C. R. & Constantin, P. 1992 Energy dissipation in shear-driven turbulence. *Phys. Rev. Lett.* **69**, 1648–1651.

Doering, C. R. & Constantin, P. 1994 Variational bounds on energy dissipation in incompressible flows: Shear flow. *Phys. Rev.* E **49**, 4087–4099.

Doering, C. R. & Constantin, P. 1996 Variational bounds on energy dissipation in incompressible flows. III. Convection. *Phys. Rev.* E **53**, 5957–5981.

Doering, C. R. & Hyman, J. M. 1997 Energy stability bounds on convective heat transport: Numerical study. *Phys. Rev.* E **55**, 7775–7778.

Elder, J. W. 1967 Steady free convection in a porous medium heated from below. *J. Fluid Mech.* **27**, 29–48.

Fowler, A. C. 1997 *Mathematical Models in the Applied Science*. Cambridge University Press, 402 pp.

Graham, M. D. & Steen, P. H. 1994 Plume formation and resonant bifurcations in porous-media convection. *J. Fluid Mech.* **272**, 67–89.

Gupta, V. P. & Joseph, D. D. 1973 Bounds for heat transport in a porous layer. *J. Fluid Mech.* **57**, 491–514.

Hopf, E. 1941 Ein allgemeiner Endlichkeitssatz der Hydrodynamik. *Mathematische Annalen* **117**, 764–775.

Howard, L. N. 1963 Heat transport in turbulent convection. *J. Fluid Mech.* **17**, 405–432.

Howard, L. N. 1964 Convection at high Rayleigh number. *Applied Mechanics*, *Proc. 11th Cong. Applied Mech.* (ed. H. Görtler), pp. 1109–1115.

Howard, L. N. 1972 Bounds on flow quantities. *Ann. Rev. Fluid Mech.* **17**, 473–494.

Joseph, D. D. 1974 Repeated supercritical branching of solutions arising in the variational theory of turbulence. *Arch. Rat. Mech. Anal.* **53**, 101–130.

Joseph, D. D. 1976 *Stability of Fluid Motions I*, 1st Edn. Springer.

Kerswell, R. R. 1996 Upper bounds on the energy dissipation in turbulent precession. *J. Fluid Mech.* **321**, 335–370.

Kerswell, R. R. 1998 Unification of variational principles for turbulent flows: the background method of Doering–Constantin and Howard–Busse's mean-fluctuation formulation. *Physica* D (in press).

Kimura, S., Schubert, G. & Straus, J. M. 1986 Route to chaos in porous medium thermal convection. *J. Fluid Mech.* **166**, 305–324.

Kimura, S., Schubert, G. & Straus, J. M. 1989 Time dependent convection in a fluid saturated porous cube heated from below. *J. Fluid Mech.* **207**, 153–189.

Koster, J. N. & Müller, U. 1982 Free convection in vertical gaps. *J. Fluid Mech.* **125**, 429–451.

Lapwood, E. R. 1948 Convection of a fluid in a porous medium. *Proc. Camb. Phil. Soc.* **44**, 508–521.

Lister, C. R. B. 1990 An explanation for the multivalued heat transport found experimentally for convection in a porous medium. *J. Fluid Mech.* **214**, 287–320.

Ly, H. V. & Titi, E. S. 1998 Global Gevrey regularity for 3-D Bénard convection in a porous medium with zero Darcy-Prandtl number. *J. Nonlinear Sci.* (in press).

Malkus, W. V. R. 1954 The heat transport and spectrum of thermal turbulence. *Proc. R. Soc. Lond.* A **225**, 196–212.

Nicodemus, R., Grossmann, S. & Holthaus, M. 1997*a* Improved variational principle for bounds on energy dissipation in turbulent shear flow. *Physica* D **101**, 178–190.

Nicodemus, R., Grossmann, S. & Holthaus, M. 1997*b* Variational bound on energy dissipation in turbulent shear flow. *Phys. Rev. Lett.* **79**, 4170–4173.

Nield, D. A. & Bejan, A. 1992 *Convection in Porous Media*. Springer, 408 pp.

Palm, E., Weber, J. E. & Oddmund, K. 1972 On steady convection in a porous medium. *J. Fluid Mech.* **54**, 153–161.

Reed, M. & Simon, B. 1980 *Functional Analysis*, Revised Edn. Academic, 400 pp.

Shattuck, M. D., Behringer, R. P., Johnson, G. A. & Georgiadis, J. G. 1995 Onset and stability of convection in porous media: Visualization by magnetic resonance imaging. *Phys. Rev. Lett.* **75**, 1934–1937.

Shattuck, M. D., Behringer, R. P., Johnson, G. A. & Georgiadis, J. G. 1997 Convection and flow in a porous media. Part 1. Visualization by magnetic resonance imaging. *J. Fluid Mech.* **332**, 215–245.

Speigel, E. A. 1971 Convection in Stars I. Basic Boussinesq Convection. *Ann. Rev. Astron. Astrophys.* **9**, 323–352.

Vitanov, N. K. & Busse, F. H. 1997 Bounds on heat transport in a horizontal fluid layer with stress-free boundaries. *Z. Angew. Math. Phys.* **48**, 310–324.

Wang, X. 1997 Time averaged energy dissipation rates for shear driven flows in R^n. *Physica* D **99**, 555–563.