J. Fluid Mech. (2005), *vol.* 542, *pp.* 343–363.
doi:10.1017/S0022112005006555 Printed in the United Kingdom

Infinite-Prandtl-number convection. Part 1. Conservative bounds

By S. C. PLASTING[1] AND G. R. IERLEY[2]

[1]Scripps Institution of Oceanography, University of California, San Diego, CA 92093-0230, USA
splasting@ucsd.edu

[2]Cecil H. and Ida M. Green Institute of Geophysics and Planetary Physics, University of California, San Diego, CA 92093-0225, USA

(Received 5 April 2005 and in revised form 26 May 2005)

The methods that have come to be known as the Malkus–Howard–Busse (MHB) and the Constantin–Doering–Hopf (CDH) techniques have, over the past few decades, produced the few rigorous statements available about average properties (e.g. momentum and heat transport) of turbulent flows governed by the Navier–Stokes equation and the heat equation. In this, the first of two papers investigating upper bounds on the heat transport in infinite-Prandtl-number convection, we show that the methods of MHB and CDH yield equivalent optimal bounds: as at a saddle – one from above, and one from below.

We also demonstrate that here, in contrast to earlier applications of the CDH method, the simplest possible, one-parameter, 'test function' does not capture the leading-order scaling associated with the fully optimal solution. We explore the consequences of a two-parameter test function in modifying the scaling of the upper bound. In the case of no-slip, the suggestion is that a hierarchy of test functions of increasing complexity is required to yield the correct limiting behaviour.

1. Introduction

Theories of turbulence are classically based upon dimensional arguments – most notably power spectrum scaling laws of Kolmogorov and Batchelor – whereas results obtained from a rigorous derivation directly from the equations of fluid dynamics – namely the Navier–Stokes equations – are few and far between.

Owing to the formidable difficulties of obtaining exact analytic solutions of the Navier–Stokes equations, an alternative philosophy has arisen that makes use of a subset of the full equations, which subset takes the form of simple, typically integral, constraints (analogous to the use of energy conservation in elementary physics, in lieu of solving the equations of motion). This produces predictions for mean properties, as opposed to details of local structure, of fully developed turbulence. This approach to studying turbulence is complementary to statistical and dimensional analyses. In some cases the same minimalist approach admits a more general use of functional analysis, in which case the selected subset of the governing equations may take various forms beyond the integral possibilities noted above. The latter approach gives the first rigorous proof of certain fundamental results about turbulence that have hitherto been based solely on phenomenological arguments (see for example Doering & Gibbon 1995; Foias *et al.* 2001).

Over the last forty or so years, analytical tools based on variational calculus and on functional estimates have enabled researchers to bound transport properties associated with turbulent flows. Two methods will be discussed below: historically the first is the Malkus–Howard–Busse method (henceforth MHB), which was developed and applied to classical fluid problems; the second is the Constantin–Doering–Hopf method (henceforth CDH), which is similar to MHB but has the added benefit of a solution obtained using simple 'test functions'. In this paper we discuss the application of these methods to bounding heat transport in the classical Boussinesq convection problem with infinite Prandtl number. These methods will be shown to be complementary and we discuss the implications for previous results.

Inspired by a hypothesis on maximizing heat transport (Malkus 1954*a*, *b*), the original paper introducing the formal MHB theory was Howard (1963). Howard established a formal upper bound for turbulent heat transport in a channel, optimizing over a field of trial functions satisfying a subset of the physical constraints imposed on a fluid by the full Navier–Stokes equations. Omitting the incompressibility constraint, Howard obtained an explicit optimal solution with an asymptotic scaling of $Ra^{1/2}$. With incompressibility the Howard problem is more challenging, which led Busse (1969) to develop his theory of multi-α optimal solutions. Busse proved that the asymptotic solution of Howard's problem with incompressibility has a structure which grows in complexity as Ra increases by spawning an increasing hierarchy of horizontal lengthscales. For no-slip boundaries, this boundary layer solution has associated heat transport which scales like $Ra^{1/2}$ while in Vitanov & Busse (1997) a partial numerical treatment of the optimal equations suggests that the same scaling applies to the same problem with free-slip boundaries. The fundamental nature of the variational problems of MHB consists of the estimation of a rigorous upper bound from below; therefore the method is not generally suited to easily extracting elementary, rigorous, bounds for basic asymptotic scalings. For review articles on the MHB method see Howard (1972) and Busse (1978).

The CDH method or 'background method' was introduced in the letter Doering & Constantin (1992) where, in the spirit of MHB, upper bounds were sought for the energy dissipation rate in plane Couette flow. The method uses a decomposition due to Hopf (1941) of the velocity field into a 'background' field, which assumes the flow boundary conditions, and a 'fluctuation' field with homogeneous boundary conditions. In contrast to the MHB method, however, estimation of the rigorous upper bound from above can be made using simple test functions in conjunction with elementary functional estimates. For instance the original Howard problem of heat transport in Boussinesq convection with incompressibility was addressed in Doering & Constantin (1996) and simple estimates were used to infer the same $Ra^{1/2}$ scaling of the heat transport, for either no-slip or free-slip boundaries, but without the need for the elaborate structure of Busse's multi-α solutions. The exact relation between the MHB and CDH method was not apparent until Kerswell (1997, 1998) proved that the methods constitute dual variational problems for both Couette flow and Boussinesq convection at finite Prandtl number. Given this duality the methods could, for the aforementioned problems, be viewed as a single theory expressed in different languages.

In recent years a full numerical calculation of the optimal solution to the CDH problem in Couette flow – Plasting & Kerswell (2003) – has validated Busse's asymptotic proposal of a multi-α solution. Additionally a number of numerical procedures for calculating semi-optimal test function bounds in the CDH formulation have been developed, notably Nicodemus, Grossmann & Holthaus (1998), Otero (2002) and Otero *et al.* (2004). In this paper we follow Otero's solution technique.

The focus of this paper is the application of MHB and CDH methods to Boussinesq convection at infinite Prandtl number. In that case the momentum equation is a linear, diagnostic, equation for $\boldsymbol{u}$ in terms of T and can thus be imposed as a stronger, pointwise, constraint instead of a purely integral constraint. We discuss both no-slip and free-slip boundary conditions.

Following the programme established by Kerswell, we first prove that the methods are also dual for this particular problem, and then calculate semi-optimal bounds using the CDH method utilizing a two-parameter family of test functions. We begin by rederiving each method as in the original literature and then proceed to demonstrate their duality.

On the basis of this duality, there emerges an apparent anomaly on comparing the results for (i) the no-slip case in which the MHB method was applied by Chan (1971) and a solution found similar to Busse's multi-k solutions with asymptotic $Ra^{1/3}$ dependence of the heat flux; and (ii) the same problem addressed by means of the CDH method by Doering & Constantin (2001) resulting in a rigorous bound on the optimal asymptotic scaling of the form $Ra^{2/5}$.† The issue at hand is how the gap between exponents of 1/3 and 2/5 is closed; either or both might need refinement in order to saturate the optimal scaling behaviour. For previous work in the case of free-slip boundaries see Vitanov (1998) and later papers, a further examination of which appears in Part 2 of this paper (Ierley, Kerswell & Plasting 2005).

1.1. *Derivation*

We consider Boussinesq convection between plates at fixed temperature, with the temperature drop between plates $\Delta T = T_{bot} - T_{top}$ positive so that the fluid layer is convectively unstable. All material properties are assumed constant: kinematic viscosity ν, thermometric conductivity κ, thermal expansion α and mean density ρ. Without loss of generality all fields are taken to be periodic in the horizontal direction, as the geometry is plane-parallel. The non-dimensionalized Rayleigh–Bénard equations assume the form

$$\left.\begin{aligned} \frac{1}{\sigma}\left(\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\nabla\boldsymbol{u}\right) + \nabla p &= \nabla^2\boldsymbol{u} + Ra\, T\widehat{\boldsymbol{z}}, \\ \frac{\partial T}{\partial t} + \boldsymbol{u}\cdot\nabla T &= \nabla^2 T, \\ \nabla\cdot\boldsymbol{u} &= 0, \end{aligned}\right\} \tag{1.1}$$

where velocity is measured in units of κ/d; lengths by d; time by d^2/κ; pressure by $\rho(\nu\kappa/d^2)$; and temperatures by ΔT. The dimensionless parameters σ and Ra in (1.1) are, respectively, the Prandtl number and the Rayleigh number:

$$\sigma = \frac{\nu}{\kappa}, \qquad Ra = \frac{\alpha g \Delta T d^3}{\kappa\nu}.$$

We study the case where the Prandtl number is infinite ($\sigma = \infty$), a rigorous justification of which limit has recently been given in Wang (2004). The velocity equation in this

† This seeming disparity was heightened still further on Constantin & Doering (1999) finding that the incorporation of extra information into CDH method yields an upper bound of the form $Ra^{1/3}(\log Ra)^{2/3}$. Subsequently Yan (2004) followed with a proof that $Nu < c\, Ra^{4/11}$. Both of these results are not to be directly compared to the present communication because they use additional estimates on derivatives of the fields, which cannot be derived within the standard CDH method.

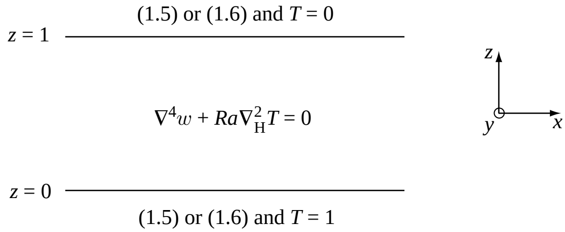

FIGURE 1. Momentum equation and boundary conditions for infinite-Prandtl-number convection.

limit is linear and has no explicit time-dependence:

$$\nabla^2 \boldsymbol{u} + Ra\, T\hat{\boldsymbol{z}} = \nabla p. \tag{1.2}$$

In the variational calculation to follow, equation (1.2) will be employed as a pointwise constraint in space that couples $\boldsymbol{u}$ to T. Inclusion of this constraint means the variational problem enriches the original Howard problem for arbitrary-Prandtl-number convection. In consequence, a scaling exponent less than $1/2$ is achievable for Nu, the Nusselt number.

Two boundary conditions on the velocity field will be used: (i) no-slip conditions to describe a solid boundary

$$\boldsymbol{u} = 0 \quad \text{at } z = 0 \text{ and } 1, \tag{1.3}$$

which, owing to incompressibility, can be equivalently written as $w = w_z = 0$ at $z = 0$ and 1; (ii) free-slip (or slip) conditions, which describe an undisplaced free surface

$$w = 0; \quad u_z = v_z = 0 \quad \text{at } z = 0 \text{ and } 1,$$

and incompressibility in this case then implies that one can write $w = w_{zz} = 0$ at $z = 0$ and 1. For a full discussion of the derivation of the Rayleigh–Bénard equations and of the boundary conditions see Chandrasekhar (1961).

In order to eliminate pressure, the momentum equation (1.2) is written as a fourth-order equation for w in terms of T, as follows:

$$\nabla^4 w + Ra \nabla_H^2 T = 0. \tag{1.4}$$

This can be seen by taking $\nabla \cdot (1.2)$ and $\nabla^2[(1.2) \cdot \hat{\boldsymbol{z}}]$, and noticing that u and v then depend only on the pressure. The horizontal Laplacian is defined as $\nabla_H^2 = \partial_x^2 + \partial_y^2$ and the two types of boundary conditions are, again: no-slip

$$w = w_z = 0 \quad \text{at} \quad z = 0 \text{ and } 1; \tag{1.5}$$

and free-slip

$$w = w_{zz} = 0 \quad \text{at} \quad z = 0 \text{ and } 1. \tag{1.6}$$

The associated dimensionless boundary conditions on T are: $T(0) = 1$ and $T(1) = 0$. See figure 1 for a summary of the geometry and boundary conditions.

Now let us turn to the heat equation, denoted by $\mathscr{H}$:

$$\mathscr{H} := \frac{\partial T}{\partial t} + \boldsymbol{u} \cdot \nabla T - \nabla^2 T. \tag{1.7}$$

This can be written in terms of the conductive heat flux ($\boldsymbol{j} := -\nabla T$) and the convective heat flux ($\boldsymbol{J} := \boldsymbol{u}T$) as

$$\frac{\partial T}{\partial t} = -\nabla \cdot (\boldsymbol{j} + \boldsymbol{J}).$$

In a purely conductive state (when $\boldsymbol{u} = \boldsymbol{0}$) the heat flux between the parallel plates is simply $\langle \widehat{\boldsymbol{z}} \cdot (-\nabla T) \rangle = -T(1) + T(0) = 1$, whereas in general the heat flux between the plates is $\langle \widehat{\boldsymbol{z}} \cdot (\boldsymbol{j} + \boldsymbol{J}) \rangle = 1 + \langle wT \rangle$. Regarding notation, we use $\overline{(\cdot)}$, $\langle \cdot \rangle$ and $\|\cdot\|^2$ as horizontal, volume and time, and L_2 averages respectively, that is

$$\overline{(\cdot)} = \lim_{L_{x,y}\to\infty} \frac{1}{4L_x L_y} \int_{-L_x}^{L_x} \mathrm{d}x \int_{-L_y}^{L_y} \mathrm{d}y(\cdot),$$

$$\langle \cdot \rangle = \lim_{\tau\to\infty} \frac{1}{\tau} \int_0^{\tau} \mathrm{d}t \int_{-1/2}^{1/2} \overline{(\cdot)}\, \mathrm{d}z, \qquad \|f\|^2 = \langle |f|^2 \rangle.$$

The dimensionless heat flux, or Nusselt number (hereafter denoted Nu), is defined as the ratio of the long-time average of the total heat flux to the conductive heat flux between the plates

$$Nu = 1 + \langle wT \rangle. \tag{1.8}$$

A second expression for Nu can be deduced from the global entropy flux balance $\langle T\mathscr{H} \rangle = 0$

$$\|\nabla T\|^2 = 1 + \langle wT \rangle, \tag{1.9}$$

where, on appeal to the temperature maximum principle for the advection–diffusion equation (which bounds the temperature field by its value on the boundary: $0 \leqslant T \leqslant 1$; e.g. Protter & Weinberger 1984), the long-time average eliminates the contribution from $\mathrm{d}(\int_{-1/2}^{1/2} \overline{T^2}\, \mathrm{d}z)/\mathrm{d}t$. The net result is

$$Nu = \|\nabla T\|^2. \tag{1.10}$$

The MHB and CDH bounding methods will be based entirely on the information in (1.8), (1.10) and the momentum constraint (1.4). In the next section we derive the MHB and CDH variational problems in accord with the derivations in Chan (1971) and Doering & Constantin (2001) respectively. Note that there is no preferred horizontal direction for the variational solutions since horizontal derivatives only appear in the combination ∇_H^2 in equation (1.4), (1.8), and (1.10).

2. Two bounding methods

2.1. *The MHB method*

The MHB method is based upon the assumption of statistical stationarity for all horizontal averages. So in particular we have $\mathrm{d}\bar{T}/\mathrm{d}t = 0$ and $\langle wT \rangle$ is time-independent. We use the Reynolds decomposition of the temperature field into a mean and fluctuating part, namely $T(\boldsymbol{x}, t) = \bar{T}(z) + \widehat{\theta}(\boldsymbol{x}, t)$. Periodicity and incompressibility imply that $\bar{w} = 0$.

Under the mean-fluctuation decomposition the heat equation (1.7) becomes

$$\mathscr{H} := \frac{\partial \theta}{\partial t} + w\bar{T}_z + \boldsymbol{u} \cdot \nabla \widehat{\theta} - \bar{T}_{zz} - \nabla^2 \widehat{\theta} = 0. \tag{2.1}$$

Two pieces of information are used to derive the variational functional: namely $\overline{\mathscr{H}}=0$ and $\langle T\mathscr{H}\rangle=0$, which are respectively

$$\bar{T}_z = w\widehat{\theta} - \langle w\widehat{\theta}\rangle - 1, \qquad \|\bar{T}_z\|^2 + \bar{T}_z|_{z=0} = -\|\nabla\widehat{\theta}\|^2; \tag{2.2}$$

and can be combined to deduce the so-called second power integral

$$\|\nabla\widehat{\theta}\|^2 + \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 = \langle w\widehat{\theta}\rangle. \tag{2.3}$$

Taking the ratio of terms in the previous balance we arrive at the identity

$$1 = \frac{\langle w\widehat{\theta}\rangle - \|\nabla\widehat{\theta}\|^2}{\|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2},$$

which, when multiplied by $\langle w\widehat{\theta}\rangle$, produces a homogeneous functional

$$F = \frac{\langle w\widehat{\theta}\rangle^2 - \langle w\widehat{\theta}\rangle\|\nabla\widehat{\theta}\|^2}{\|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2}, \tag{2.4}$$

the supremum of which provides an upper bound on the Nusselt number, $Nu - 1 \leqslant \sup F$. The maximization of F is performed over steady fields that satisfy the momentum constraint (1.4), the specific boundary conditions, and the power constraint in equation (2.3), which is imposed after the fact by normalizing $\langle w\widehat{\theta}\rangle = F$. This is essentially the homogeneous functional that Chan (1971) maximizes.

In fact, Chan solved the Euler–Lagrange equations of the following Lagrangian functional:

$$G = F - \big\langle q(\boldsymbol{x})\big(\nabla^4 w + Ra\nabla_H^2\widehat{\theta}\big)\big\rangle, \tag{2.5}$$

where $q(\boldsymbol{x})$ is a Lagrange multiplier imposing the pointwise momentum constraint in equation (1.4). This is the functional in equation (25) of Chan (1971) (save for the normalization $\langle w\widehat{\theta}\rangle = 1$). Taking variations of this functional with respect to w and $\widehat{\theta}$ and substituting in $\|\nabla\widehat{\theta}\|^2 = \langle w\widehat{\theta}\rangle - \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2$ we deduce the following Euler–Lagrange equations:

$$\begin{aligned} &\widehat{\theta}[\langle w\widehat{\theta}\rangle + \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2] - 2F\widehat{\theta}[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] - \nabla^4 q\ \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 = 0,\\ &2\nabla^2\widehat{\theta}\ \langle w\widehat{\theta}\rangle + w[\langle w\widehat{\theta}\rangle + \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2] - 2Fw[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] - Ra\nabla_H^2 q\ \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 = 0. \end{aligned} \tag{2.6}$$

If we normalize w and $\widehat{\theta}$ as Chan does, namely $w \to \langle w\widehat{\theta}\rangle^{-1/2}Ra^{-1/2}w$ and $\widehat{\theta} \to \langle w\widehat{\theta}\rangle^{-1/2}Ra^{1/2}\widehat{\theta}$ so that $\langle w\widehat{\theta}\rangle \to 1$ then equations (2.6) become exactly the Euler–Lagrange equations (27) in Chan (1971).

2.2. *CDH method*

By contrast, the CDH method relies on a decomposition of the temperature field originally due to Hopf (1941), into background and fluctuation parts

$$T(\boldsymbol{x}, t) = \tau(z) + \theta(\boldsymbol{x}, t), \tag{2.7}$$

where the background field takes on the temperature boundary conditions: $\tau(0)=1$ and $\tau(1)=0$ but is otherwise arbitrary, and θ satisfies Dirichlet boundary conditions. The fluctuation field θ is not necessarily a zero-mean variable, hence to distinguish these variables from the equivalent MHB set, we do not use hatted variables at this stage. Substituting (2.7) into (1.7) gives

$$\frac{\partial\theta}{\partial t} + \boldsymbol{u}\boldsymbol{\cdot}\nabla\theta = \nabla^2\theta + \tau'' - w\tau'. \tag{2.8}$$

Multiplying this equation by θ produces

$$0 = -\|\nabla\theta\|^2 + \langle\theta\tau'' - w\theta\tau'\rangle \tag{2.9}$$

(as above the time derivative drops out). Adding $b \times$ (2.9) to the identity $\|\nabla T\|^2 = \|\nabla\theta\|^2 + \|\tau'\|^2 - 2\langle\theta\tau''\rangle$ then yields

$$\|\nabla T\|^2 = \|\tau'\|^2 - \mathscr{G}_{\tau,b}(w,\theta) \tag{2.10}$$

where $\mathscr{G} = \langle(b-1)\|\nabla\theta\|^2 - (b-2)\theta\tau'' + bw\theta\tau'\rangle$.†

Using expression (1.10) the following upper bound on Nu emerges:

$$Nu \leqslant \|\tau'\|^2 - \inf_{w,\theta}\mathscr{G}_{\tau,b}(w,\theta) \tag{2.11}$$

where the infimum is attained by steady fields.

The task of calculating the upper bound on Nu in equation (2.11) is greatly simplified by reducing the problem to one just involving w and the zero-mean field $\widehat{\theta} \equiv \theta - \bar{\theta}$. In this way we can make connection to the variables $(w, \widehat{\theta})$ in the MHB problem. Consider the functional

$$L = \mathscr{G} - \left\langle q(\boldsymbol{x})\left(\nabla^4 w + Ra\nabla_H^2\theta\right)\right\rangle$$

where the Lagrange multiplier $q(\boldsymbol{x})$ is used to impose equation (1.4) and satisfies the natural boundary conditions for the problem, namely those satisfied by w. The infimum over θ is attained by setting the θ variation to zero:

$$\frac{\delta L}{\delta\theta} = -2(b-1)\nabla^2\theta - (b-2)\tau'' + bw\tau' - Ra\nabla_H^2 q = 0.$$

Taking a horizontal average of this equation, and noting periodicity and the boundary conditions, we find that the mean part of the optimal fluctuation field is given by

$$2(b-1)\bar{\theta}'' + (b-2)\tau'' = 0, \tag{2.12}$$

which can be integrated twice to obtain

$$\bar{\theta} = -\frac{(b-2)}{2(b-1)}[\tau + z - 1]. \tag{2.13}$$

Therefore by setting $\widehat{\theta} = \theta - \bar{\theta}$ we can restate problem (2.11) as an optimization over zero-mean fields, namely

$$Nu - 1 \leqslant \frac{b^2}{4(b-1)}(\|\tau'\|^2 - 1), \tag{2.14}$$

subject to the spectral constraint $\widetilde{\mathscr{G}}_{\tau,b}(w,\widehat{\theta}) = (b-1)\|\nabla\widehat{\theta}\|^2 + b\langle w\widehat{\theta}\tau'\rangle \geqslant 0$. Here $(w, \widehat{\theta})$ satisfy equation (1.4) and the specific boundary conditions. It is clear that if $\widetilde{\mathscr{G}}_{\tau,b}$ is

† The parameter b is known in the field as the *balance parameter* and was originally introduced into the formulation of the CDH upper-bound problem for plane Couette flow in Nicodemus, Grossmann & Holthaus (1997); prior to this the special case $b=2$ had only been considered. The extra leverage provided by the balance parameter is what enabled Kerswell (1997, 1998) to prove that the CDH and MHB theories applied to plane Couette flow constituted dual variational problems. b also plays a crucial role in proving the duality for the problem studied here.

to have a minimum value then $b > 1$. The term 'spectral constraint' is used purely for historical reasons and is related to the fact that non-negativity of the quadratic form $\tilde{\mathscr{G}}_{\tau,b}$ can be interpreted as an eigenvalue problem. The Lagrangian associated with this simplified variational problem is

$$\tilde{L} = \frac{b^2}{4(b-1)}(\|\tau'\|^2 - 1) - \tilde{\mathscr{G}}_{\tau,b}(w,\widehat{\theta}) - \left\langle q(\boldsymbol{x})\left(\nabla^4 w + Ra\nabla_H^2\widehat{\theta}\right)\right\rangle. \tag{2.15}$$

The form of this variational problem is similar to that in the Couette flow variational problem solved in Plasting & Kerswell (2003), save that a pointwise momentum constraint is here imposed.

2.3. *Unifying the CDH and MHB methods*

We have seen that the MHB variational method seeks the supremum of a homogeneous functional – equation (2.4) – in order to bound Nu and therefore any specific point $(w,\widehat{\theta})$ yields an underestimate of the upper bound, whereas the CDH method seeks to minimize a functional – namely $\|\tau'^2\|$ – subject to a spectral constraint and therefore any arbitrary τ found to satisfy the spectral constraint yields an upper estimate of the optimal upper bound. Such a characterization makes it unclear how the methods might intersect.

Their correspondence can be established by showing that the optimal equation for each method derives from a single Lagrangian functional. As noted earlier, Kerswell (2001) proved this duality in the case of MHB and CDH methods applied to arbitrary-Prandtl-number convection. The following proof is for the specific case of $\sigma = \infty$.

LEMMA 1. *The CDH and MHB methods are dual variational problems estimating the highest stationary point, that with highest associated heat flux, of the following functional:*

$$N := \|\nabla T\|^2 - b\langle\theta\mathscr{H}\rangle - \left\langle q(\boldsymbol{x})\left(\nabla^4 w + Ra\nabla_H^2 T\right)\right\rangle \tag{2.16}$$

where $T = \tau(z) + \theta(\boldsymbol{x},t)$, and $\mathscr{H}$ is the heat equation:

$$\mathscr{H} := \frac{\partial\theta}{\partial t} + \boldsymbol{u}\cdot\nabla\theta - \nabla^2\theta + w\tau' - \tau'' = 0.$$

Proof. We begin the proof by deriving all of the variational derivatives of the functional N, which can be written in terms of τ and θ:

$$N(\tau,w,\theta,b,q) = \|\tau'\|^2 - \langle(b-1)|\nabla\theta|^2 - (b-2)\theta\tau'' + b\theta w\tau'\rangle - \left\langle q(\boldsymbol{x})\left(\nabla^4 w + Ra\nabla_H^2\theta\right)\right\rangle. \tag{2.17}$$

Variational equations for $\tau(z)$, $\theta(\boldsymbol{x})$, $w(\boldsymbol{x})$, $q(\boldsymbol{x})$, as well as the mean and zero-mean parts of θ, are given by

$$\frac{\delta N}{\delta\tau} = -2\tau'' + (b-2)\bar{\theta}'' + b(w\theta)' = 0, \tag{2.18}$$

$$\frac{\delta N}{\delta\theta} = 2(b-1)\nabla^2\theta + (b-2)\tau'' - bw\tau' - Ra\nabla_H^2 q = 0,$$

$$\frac{\delta N}{\delta w} = -b\theta\tau' - \nabla^4 q = 0, \tag{2.19}$$

$$\frac{\delta N}{\delta q} = \nabla^4 w + Ra\nabla_H^2\theta = 0, \tag{2.20}$$

$$\frac{\delta N}{\delta\theta} = 0 \qquad \left\{ \begin{aligned} \frac{\delta N}{\delta\bar{\theta}} &= 2(b-1)\bar{\theta}'' + (b-2)\tau'' = 0, \\ \frac{\delta N}{\delta\widehat{\theta}} &= 2(b-1)\nabla^2\widehat{\theta} - bw\tau' - Ra\nabla_H^2 q = 0. \end{aligned} \right. \tag{2.21}$$

In the following discussion the optimal fields (the solutions of the Euler–Lagrange equations) are denoted by a subscript *opt*. It will be demonstrated that both of the techniques can be derived by solving different subsets of the variational equations and rewriting the remaining equations appropriately.

Case 1: CDH method

Solving $\delta N/\delta\bar{\theta} = 0$ yields an expression for the mean of the optimal θ:

$$\bar{\theta}_{opt} = -\frac{b-2}{2(b-1)}[\tau + z - 1].$$

Substituting this into N we have

$$\begin{aligned} \check{N}(\tau, w, \widehat{\theta}, b, q) - 1 &= N(\tau, w, \bar{\theta}_{opt} + \widehat{\theta}, b, q) - 1 \\ &= \frac{b^2}{4(b-1)}(\|\tau'\|^2 - 1) - \langle (b-1)|\nabla\widehat{\theta}|^2 + bw\widehat{\theta}\tau'\rangle - \langle q(\boldsymbol{x})(\nabla^4 w + Ra\nabla_H^2\widehat{\theta})\rangle. \end{aligned} \tag{2.22}$$

This is the Lagrangian functional shown in (2.15).

Case 2: MHB method

Solving $\delta N/\delta\bar{\theta} = 0$ and $\delta N/\delta\tau = 0$ simultaneously allows us to deduce equations for the optimal background field and the mean of the fluctuation field in terms of the zero-mean fluctuation field $\widehat{\theta}$ and w, respectively:

$$\tau'_{opt} = \frac{2(b-1)}{b}[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] - 1, \qquad \bar{\theta}'_{opt} = -\frac{b-2}{b}[w\widehat{\theta} - \langle w\widehat{\theta}\rangle]. \tag{2.23}$$

Substituting these expressions into N followed by some algebra yields

$$\begin{aligned} \mathring{N}(w, \widehat{\theta}, b, q) - 1 &= N(\tau_{opt}, w, \bar{\theta}_{opt} + \widehat{\theta}, b, q) - 1 \\ &= \langle w\widehat{\theta}\rangle + (b-1)\{\langle w\widehat{\theta}\rangle - \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 - \|\nabla\widehat{\theta}\|^2\} - \langle q(\boldsymbol{x})(\nabla^4 w + Ra\nabla_H^2\widehat{\theta})\rangle. \end{aligned} \tag{2.24}$$

In this context $(b-1)$ is a Lagrange multiplier imposing the second-power integral balance (2.3) while $q(\boldsymbol{x})$ imposes the momentum constraint. The remaining variational equations for w and $\widehat{\theta}$ are

$$\frac{\delta N}{\delta\widehat{\theta}} = w + (b-1)\{w - 2w[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] + 2\nabla^2\widehat{\theta}\} - Ra\nabla_H^2 q = 0, \tag{2.25}$$

$$\frac{\delta N}{\delta w} = \widehat{\theta} + (b-1)\{\widehat{\theta} - 2\widehat{\theta}[w\widehat{\theta} - \langle w\widehat{\theta}\rangle]\} - \nabla^4 q = 0. \tag{2.26}$$

In order to obtain a value for the optimal b we calculate $\langle\widehat{\theta}(\delta N/\delta\widehat{\theta})\rangle = 0$ and $\langle w(\delta N/\delta w)\rangle = 0$, which are respectively

$$(2-b)\langle w\widehat{\theta}\rangle - Ra\langle(\nabla_H^2 q)\widehat{\theta}\rangle = 0,$$

and

$$(2-b)\langle w\widehat{\theta}\rangle + 2(b-1)\|\nabla\widehat{\theta}\|^2 - \langle(\nabla^4 q)w\rangle = 0.$$

Given that $\langle(\nabla^4 q)w\rangle = \langle q(\nabla^4 w)\rangle = \langle q(-Ra\nabla_H^2\widehat{\theta}\rangle$ these equations can be added to give

$$b = \frac{\|\nabla\widehat{\theta}\|^2 - 2\langle w\widehat{\theta}\rangle}{\|\nabla\widehat{\theta}\|^2 - \langle w\widehat{\theta}\rangle}, \tag{2.27}$$

and the second-power integral can be employed to deduce a simple expression for $(b-1)$, which may be easily inserted back into equations (2.25) and (2.26):

$$b - 1 = \frac{\langle w\widehat{\theta}\rangle}{\|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2}. \tag{2.28}$$

After minor simplifications the w and $\widehat{\theta}$ variations turn out to be

$$\widehat{\theta}[\langle w\widehat{\theta}\rangle + \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2] - 2\widehat{\theta}\langle w\widehat{\theta}\rangle[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] - (\nabla^4 q)\ \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 = 0, \tag{2.29}$$

and

$$2(\nabla^2\widehat{\theta})\langle w\widehat{\theta}\rangle + w[\langle w\widehat{\theta}\rangle + \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2] - 2w\langle w\widehat{\theta}\rangle[w\widehat{\theta} - \langle w\widehat{\theta}\rangle] - Ra\left(\nabla_H^2 q\right)\ \|w\widehat{\theta} - \langle w\widehat{\theta}\rangle\|^2 = 0. \tag{2.30}$$

If we replace each $\langle w\widehat{\theta}\rangle$ multiplying $(w\widehat{\theta} - \langle w\widehat{\theta}\rangle)$ by the functional F in (2.4), then these equations are exactly the Euler–Lagrange equations derived in Chan (1971).

As a final comment note that, for the two problems to intersect, the MHB problem must additionally satisfy the spectral constraint to ensure that the field of highest heat flux is selected. Indeed, in general equations (2.29) and (2.30) will have multiple solutions, only one of which will be a global maximum for F. Note also that equation (2.28) implies that $(b-1) > 0$, which is consistent with the spectral constraint. This concludes the proof of duality. □

Next we seek to close the gap between the putative optimal scaling exponent of 1/3 and the original conservative-bound exponent of 2/5.

3. Conservative upper bounds

In previous studies the scaling exponents of CDH upper bounds have been saturated by the simplest possible test function, namely a single-parameter family of functions with constant-slope boundary layer and zero-slope interior. Those cases utilizing more complicated test functions yielded only an improvement in numerical prefactors. The rationale behind using simple boundary-layered test functions is that in order to prove that the positive definite term in the spectral constraint – $\langle|\nabla\theta|^2\rangle$ – is large enough to balance the sign-indefinite term – $2\langle w\theta\tau'\rangle$ – then τ' must be non-zero only in a small region close to the wall, where w and θ must vanish, thus limiting the sign-indefinite term.

In infinite-Prandtl-number convection these simple test functions have been used to prove conservative estimates of the form $Nu < c\,Ra^{2/5}$ both analytically, by Doering & Constantin (2001), and numerically, by Otero (2002). Thus, given Chan's result that the dual MHB theory leads to an upper bound of the form $Nu < c\,Ra^{1/3}$, it is clear that the single-parameter functions do not saturate the optimal scaling.

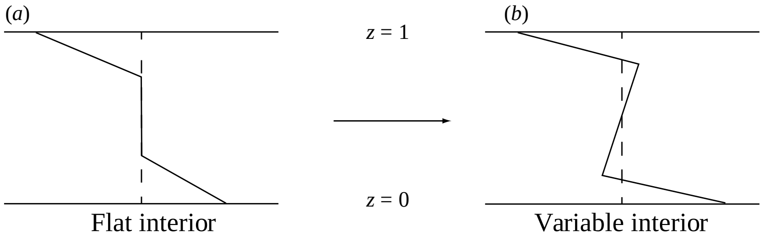

FIGURE 2. Enlarged set of test functions for the Lagrange multiplier $\tau(z)$: (*a*) from Otero (2002) (*b*) two-parameter form used here.

We therefore extend this class of τ-test functions to a two-parameter family having a constant-gradient boundary layer and constant, but arbitrary, interior slope. We are guided in our choice of test functions by observing that the optimal solutions to this problem (appearing in Part 2 of this paper) have overshoot, that is, regions where the optimal profile is stably stratified, just on the edge of the boundary layer.

The numerical details of a conservative estimate have also been described in detail in §4 of Otero *et al.* (2004) for the problem of convection in a porous medium. The only numerical subtlety is in the treatment of the spectral constraint. Optimization of the upper bound over the balance parameter b can be shown not to affect the scaling exponent (see Plasting 2004), hence we set $b=2$.

The two-parameter family of test functions, which we write in terms of τ' only, is

$$\tau' = \begin{cases} p - \dfrac{1+p}{2\delta} & \text{for } z \in [0, \delta], \\ p & \text{for } z \in [\delta, 1/2]. \end{cases} \tag{3.1}$$

This form reflects the boundary conditions on τ, namely $\tau(0)=1$ and $\tau(1)=0$, which imply that $\int_0^1 \tau' \,\mathrm{d}z = -1$. The interior slope is p; the test functions are odd about the mid-channel at $z=1/2$; and the boundary layer has thickness δ. Figure 2 depicts the enlargement of the class of test functions from those used in Otero (2002) (figure 2*a*) to the two-parameter form used here (figure 2*b*).

For the test functions described in equation (3.1) the value of the upper bound, $Nu \leqslant \int_0^1 \tau'^2$, is

$$Nu \leqslant N(\delta, p) := \frac{(1+p)^2}{2\delta} - p(2+p). \tag{3.2}$$

The spectral constraint is

$$Q = \langle |\nabla\theta|^2 \rangle + 2\langle w\theta\tau' \rangle \geqslant 0 \text{ for any } w, \theta \text{ satisfying (1.4)},$$

where w and θ must satisfy either the no-slip or free-slip boundary conditions and are solutions of the momentum constraint $\nabla^4 w + Ra\nabla_H^2\theta = 0$ (where ∇^2 is the Laplacian operator and the horizontal Laplacian is denoted $\nabla_H^2 = \partial_x^2 + \partial_y^2$). This spectral constraint is equivalent to insisting that the highest eigenvalue, henceforth denoted μ^0, of the following eigenvalue problem is negative semi-definite:

$$\left.\begin{aligned} 2\nabla^2\theta - 2w\tau' - Ra\nabla_H^2 q &= \mu\theta, \\ \nabla^4 q + 2\theta\tau' &= 0, \\ \nabla^4 w + Ra\nabla_H^2\theta &= 0, \end{aligned}\right\} \tag{3.3}$$

where the $q(\boldsymbol{x})$ is a Lagrange multiplier field for the momentum constraint satisfying the same boundary conditions as w. We call the optimal solution within the restricted set of test functions a 'semi-optimal' solution to this variation problem, since the form of the background profile is restricted to a two-parameter family of functions. The intent of the restriction is to facilitate the determination of a bound either analytically or, failing that, with an elementary numerical treatment.

We expand the fields w, θ and q in a Fourier series as

$$w(x, y, z) = \sum_{\boldsymbol{\alpha}} \mathrm{e}^{\mathrm{i}\boldsymbol{\alpha}\cdot\boldsymbol{x}} w_{\boldsymbol{\alpha}}(z)$$

where $\alpha^2 = |\boldsymbol{\alpha}^2|$. Since (3.3) is homogeneous in x and y the spectral constraint depends only on the magnitude of the wave-vector, namely α, and not its direction. The spectral constraint is imposed on a mode-by-mode basis because

$$Q \geqslant 0 \quad \forall w, \theta \iff Q_\alpha \geqslant 0 \quad \forall \alpha \in \mathbb{R} \text{ and } w_\alpha, \theta_\alpha \text{ satisfying (1.4)}, \tag{3.4}$$

where Q_α is just Q restricted to the horizontal wavenumber α. For each wavenumber α there is a convex set of background profiles characterized by $\mu_\alpha^0(\tau) \leqslant 0$ and a corresponding isospectral surface $\mu_\alpha^0(\tau) = 0$. Owing to this convexity property we need only consider the surface $\mu^0 = 0$, and seek the minimum Nusselt number $N(\delta, p)$ on it. Assuming this isospectral surface is smooth, then the minimum over (δ, p) can be computed with a continuation method. The eigenvalue problem is now

$$\left.\begin{aligned} 2(\mathrm{D}^2 - \alpha^2)\theta - 2w\tau' + Ra\alpha^2 q = \mu_\alpha \theta, \\ (\mathrm{D}^2 - \alpha^2)^2 q + 2\theta\tau' = 0, \\ (\mathrm{D}^2 - \alpha^2)^2 w - Ra\alpha^2\theta = 0\,, \end{aligned}\right\} \tag{3.5}$$

where $\mathrm{D} = \mathrm{d}/\mathrm{d}z$. For each α there exists a unique largest δ, say δ_α, for which $\mu_\alpha = 0$ and $N(\delta, p)$ is a minimum. The optimal δ, denoted δ^*, is then the minimum over all these δ_α.

Our intention is to fix both Ra and p in order to find δ^*. Since $N(\delta, p)$ is monotonic in δ it suffices to find the maximum δ consistent with spectral constraint $\mu_\alpha^0(\tau) \leqslant 0$, and for which there exists an α with $\mu_\alpha^0(\tau) = 0$. A further minimization of N over the slope parameter p then yields the semi-optimal estimate for this value of Ra.

Finding δ_α involves solving system (3.5) with $\mu_\alpha = 0$. The optimal boundary layer thickness is

$$\delta^* = \min_{\alpha < \alpha^+} \delta_\alpha, \tag{3.6}$$

where α^+ is the cut-off wavenumber above which the conduction solution $\tau' = -1$ is stable. Owing to the monotonicity of N with δ an equivalent statement is

$$N_1(p) = \max_{\alpha < \alpha^+} N(\delta_\alpha, p). \tag{3.7}$$

Noting that δ_α is an implicit function of p the next step in the optimization is a minimization over the interior gradient. The end result is a semi-optimal upper bound on Nu

$$Nu \leqslant \min_{p \in \mathbb{R}} \max_{\alpha < \alpha^+} N(\delta_\alpha, p). \tag{3.8}$$

Since τ' is piecewise constant, system (3.5) can be implemented as a condition that the determinant of a 10×10 matrix vanishes. The elements in this matrix are derived from the expansion of the eigenfunctions in complex exponentials. The following subsection describes the analytical solution and the matching and symmetry conditions which, once imposed, lead to the relevant determinant.

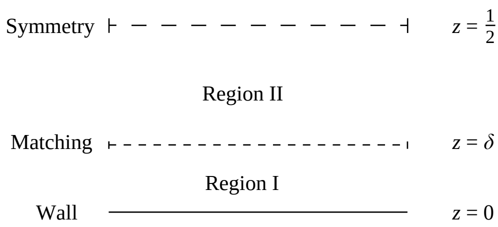

FIGURE 3. Regions of constant τ' for the two-parameter test functions.

3.1. *Solution technique*

Within the regions of constant τ' – pictured in figure 3 – system (3.5) is a set of linear, constant-coefficient, differential equations soluble by complex exponentials. The system is equivalent to a tenth-order equation in w and therefore the solution will be a linear combination of ten exponentials. Substituting

$$[\theta,\ w,\ q] = [a,\ b,\ c]\ \mathrm{e}^{\beta\alpha z}$$

we find that the auxiliary equation has two real roots each of multiplicity two ($\beta = \pm 1$), and six complex roots satisfying $(\beta^2 - 1)^3 = 2Ra\tau'/\alpha^4$.

Note that: (i) for the repeated roots the particular solutions have

$$\theta(z) = 0 \quad \text{and} \quad q(z) = \frac{2\tau'}{Ra\alpha^2} w(z), \tag{3.9}$$

whereas (ii) for the complex roots the particular solutions have

$$\theta(z) = \frac{\alpha^2(\beta^2 - 1)^2}{Ra} w(z) \quad \text{and} \quad q(z) = -\frac{2\tau'}{Ra\alpha^2} w(z). \tag{3.10}$$

Building in boundary conditions and mid-channel symmetry† reduces the number of unknown coefficients in the problem by half. As an illustration the Region II solution, identical for no-slip and free-slip boundaries, is

$$w_2 = c_6 \cosh(\beta_6 \alpha \tilde{z}) + c_7 \cosh(\beta_7 \alpha \tilde{z}) + c_8 \cosh(\beta_8 \alpha \tilde{z}) \\ + c_9 \cosh(\alpha \tilde{z}) + c_{10} \tilde{z} \sinh(\alpha \tilde{z})$$

(where $\tilde{z} = z - 1/2$) and the Region I solution for free-slip boundaries is

$$w_1 = c_1 \sinh(\beta_1 \alpha z) + c_2 \sinh(\beta_2 \alpha z) + c_3 \sinh(\beta_3 \alpha z) \\ + c_4 \sinh(\alpha z) + c_5 z \cosh(\alpha z)$$

(a similar form is obtained for no-slip). The complex roots $\beta_1, \beta_2, \beta_3$ for Region I and $\beta_6, \beta_7, \beta_8$ for Region II are computed using the following formulae for $j = 1, 2, 3$:

$$\beta^2 - 1 = \begin{cases} \left(\dfrac{2|\tau'|Ra}{\alpha^4}\right)^{1/3} \mathrm{e}^{\mathrm{i}(2j+1)\pi/3} & \text{for } \tau' < 0, \\[2ex] \left(\dfrac{2|\tau'|Ra}{\alpha^4}\right)^{1/3} \mathrm{e}^{\mathrm{i}(2j)\pi/3} & \text{for } \tau' > 0. \end{cases} \tag{3.11}$$

† For a discussion of symmetry of the ground-state eigenfunction and why even symmetry is more potent than odd see Appendix C of Otero (2002).

The $c_1, \ldots, c_{10}$ are the unknown coefficients.

3.2. *Matching conditions*

It remains to specify ten jump conditions to match the solution of Region I to that of Region II at $z=\delta$:

$$\left.\begin{aligned} &[\theta]_\delta = [\mathrm{D}\theta]_\delta = 0, \\ &[w]_\delta = [\mathrm{D}w]_\delta = [\mathrm{D}^2 w]_\delta = [\mathrm{D}^3 w]_\delta = 0, \\ &[q]_\delta = [\mathrm{D}q]_\delta = [\mathrm{D}^2 q]_\delta = [\mathrm{D}^3 q]_\delta = 0, \end{aligned}\right\} \tag{3.12}$$

where the jump at $z=\delta$ has been denoted by $[f]_\delta = f_2(\delta) - f_1(\delta)$. The imposition of these jump conditions is then written in matrix form:

$$\boldsymbol{M}\,\boldsymbol{c} = \boldsymbol{0}, \qquad \text{for } \boldsymbol{c} = [c_1, \ldots, c_{10}]^T.$$

3.3. *Vanishing determinant method*

For fixed Ra, p and α the equivalent to seeking $\mu_\alpha = 0$ is finding a δ for which $\det \boldsymbol{M} = 0$ thus implying the existence of a non-trivial solution to equations (3.5) with $\mu = 0$. The smallest such δ for which $\det \boldsymbol{M} = 0$ is δ_α and thereby the least damped eigenvalue is zero ($\mu_\alpha^0 = 0$). Thus equations (3.6) to (3.8) can be rewritten with $\mu = 0$ replaced by $\det \boldsymbol{M} = 0$.

The method of solution is best represented by the mini-max problem for the upper bound:

$$Nu \leqslant \min_{p\in\mathbb{R}} \max_{\alpha<\alpha^+} \{N(\delta_\alpha) : \delta_\alpha \text{ smallest } \delta \text{ such that } \det \boldsymbol{M} = 0\} \tag{3.13}$$

For fixed Ra the primary task is to locate δ^*, the global minimum of δ_α over α (for which $\det \boldsymbol{M}$ and its first derivative with respect to α must both be zero), and then to follow this extremum by continuation. A secondary minimization over the interior slope p is performed in a separate loop. Starting at low Ra the extremum is traced to values sufficiently large that the asymptotic behaviour is clear. One subtlety is that the explicit Ra-dependence of the matrix entries needs carefully to be scaled out by suitable row and column operations in order to avoid a serious loss of accuracy. The more intransigent problem is one of precision: as Ra increases, there is a crucial cancellation of increasing stringency; this cancellation is the determining factor in computing δ^*. For $Ra \approx 10^{20}$ double-precision arithmetic fails to resolve this cancellation and multiple-precision numerical techniques are then employed to move further in Ra.

An attempt was made to extract the limiting exponent from the analytic determinant, with various levels of truncation of the transcendental functions in $\boldsymbol{M}$. An expansion of the trigonometric terms out to $(\alpha\delta)^5$ is required to yield the correct limiting behaviour of the scaling exponent associated to the conservative bound (see the Appendix for details). Unfortunately, doing this analytically leads to so complex an algebraic set of manipulations we have not been able to distill from this the essence of the balance determination. This analytic difficulty is the precise complement of the numeric one above.

3.4. *Consistent numerics*

Two auxiliary tests were executed periodically: first, a test to check that the zero determinant corresponded to a zero crossing of the highest eigenvalue of Q; second, a test that multiple maxima in the envelope of the maximum eigenvalues of Q over α were not appearing. Fortunately, neither of these conditions obtained. For each

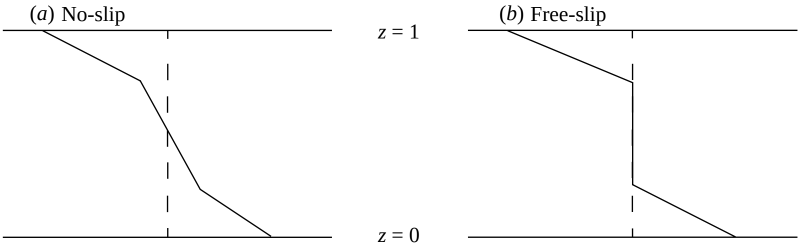

FIGURE 4. The form of the optimal test function at low Ra ($10^2 - 10^3$) for (*a*) no-slip and (*b*) free-slip boundary conditions.

boundary condition, we always found a single maximum between $\alpha = 0$ the cut-off point α^+.

4. Results

For notational brevity optimal parameters will be denoted by superscript *. We assume a polynomial dependence on Ra for N and α, so $N^* \sim Ra^{\gamma_1}$ and $\alpha^* \sim Ra^{\gamma_2}$. In the first instance, the exponents γ_1 and γ_2 are estimated by pointwise evaluation of $\mathrm{d}\log(\cdot)/\mathrm{d}\log(Ra)$. Owing to our choice of $b = 2$, the test-function bound bifurcates from the conduction profile at half the critical Rayleigh number $Ra_c/2$.

4.1. *Low-Ra behaviour*

A continuous transition from the conduction profile ($\tau' = -1$) to the (δ, p)-test function solution branch occurs in two possible ways. The two possible behaviours are pictured in figure 4. Our calculations show that for no-slip boundaries a negative interior slope leads to the smallest upper bound on Nu, while for free-slip boundaries the smaller upper bound is obtained with zero slope.

4.2. *High-Ra scalings*

In extending the test function approach of Otero (2002) to a two-parameter (δ, p)-family our goal was to bring into correspondence the conservative bound method, for which the previous best no-slip bound is $Ra^{2/5}$, with the putative optimal scaling of $Ra^{1/3}$ and, in so doing, reveal that particular feature of the background field critical to establishing the optimal behaviour of the upper-bound exponent.

We conclude that semi-optimal bounds are: for no-slip boundaries an improved scaling of $Ra^{7/20}$; and for free-slip boundaries a scaling of $Ra^{5/12}$. The exponents in decimal form are respectively $7/20 = 0.35$ and $5/12 = 0.41\bar{6}$. Without an analytic expansion of the determinant, some uncertainty must, of course, attach to these values. While the initial estimate of these from a logarithmic derivative suggestis that the values are converged to perhaps three decimal places, greater confidence can be had by appeal to extended Richardson extrapolation subject to a determination of the apparent form of higher correction terms. These appear as a series with algebraically decaying exponents, i.e. in the form $c_1 R^{\gamma_1} + c_2 R^{\gamma_1 - \zeta} + c_3 R^{\gamma_1 - 2\zeta} + \cdots$. On numerical evidence, $\zeta = 1/20$ for no-slip and $\zeta = 1/12$ for free-slip. The smallness of these is consistent with our having to reach $Ra = 10^{35}$ in the case of no-slip in order convincingly to show convergence for γ_1. Owing to the severity of cancellation as remarked above, this required up to 96 digits precision. Richardson extrapolation appears to yield ten significant digits in the determination of c_1, and fewer for the higher coefficients. The results of this fit are shown in figure 5.

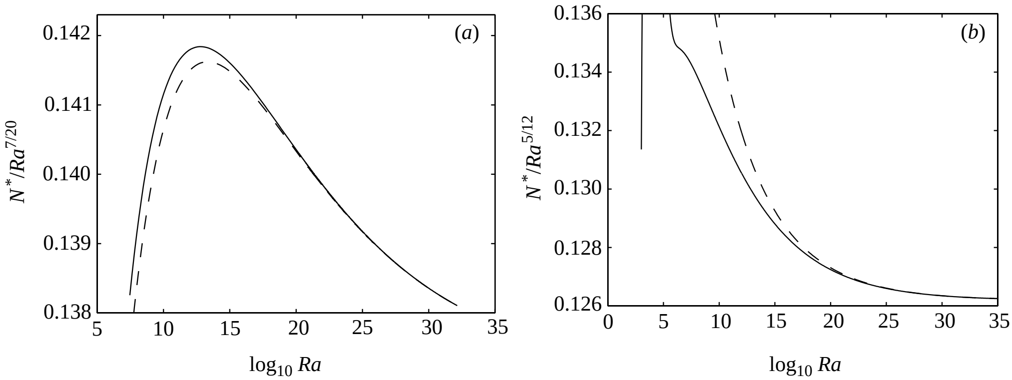

FIGURE 5. Comparison of the upper bound N^* for (*a*) no-slip and (*b*) free-slip with asymptotic series fitted to the data. To third order the no-slip fit to the semi-optimal bound is $N^* \sim 0.1371355872\, R^{7/20} + 0.0416286\, R^{3/10} - 0.096560\, R^{1/4}$ and the free-slip result is $N^* \sim 0.12618619557280\, R^{5/12} + 0.050506642\, R^{1/3} - 0.0717275\, R^{1/4}$. The number of decimal places stated is based on limits on accuracy indicated from Richardson extrapolation.

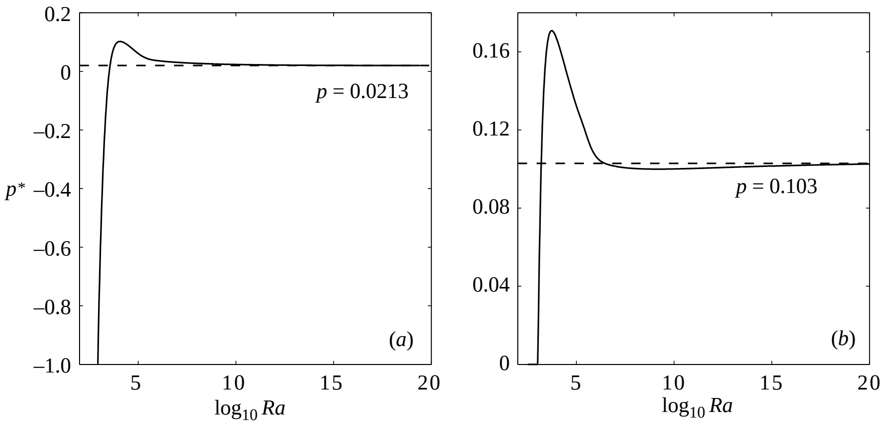

FIGURE 6. The optimal p for (*a*) no-slip and (*b*) free-slip boundaries. The limiting value for no-slip is $p^* = 0.0213$ and for free-slip $p^* = 0.103$. Owing to the slow rate of convergence we have used Shanks extrapolation to resolve the limiting value of p^*.

The optimal interior gradient is plotted in figure 6. For each boundary condition the optimal slope p settles, after some transient differences, to a positive value, which means that the optimal test function retains a small residual positive slope in the high-Ra limit. The behaviour of the optimal wavenumber is $\alpha^* \sim Ra^{1/4}$ for both no-slip and free-slip boundaries.

5. Conclusions

The original conservative Howard bound for finite Prandtl number Doering & Constantin (1996), which uses piecewise linear profiles with an interior slope of $+1$, and a standard constant-slope boundary layer, takes the form of $Nu \leqslant cRa^{1/2}$. This bound holds for either no-slip or free-slip boundary conditions, and applies uniformly in Prandtl number. Thus the limiting exponents calculated here are constrained to be less than 1/2.

The simplest class of test functions (pictured on figure 2a) was previously applied to the problem of infinite-Prandtl-number convection. The best analytic result for no-slip boundary conditions was found to be $Ra^{2/5}$ (Doering & Constantin 2001)† and was corroborated by Otero (2002) using numerical implementation of the spectral constraint, by means of the determinant method, as in § 3 of this article.

We have presented compelling evidence that a semi-optimal analysis of the CDH variational problem for infinite-Prandtl-number convection results in (i) $Nu \leqslant cRa^{7/20}$ for no-slip boundaries and (ii) $Nu \leqslant cRa^{5/12}$ for free-slip boundaries. It is our hypothesis that the exponents in these two problems take the simple rational values indicated, the only evidence for which to emerge thus far is numerical.

We note, again, that multiple-precision arithmetic is required before a limiting behaviour for the free-slip case is observed for $Ra \sim 10^{30}$. The no-slip calculation is yet more delicate, requiring 96 (versus 32) digits to reach the needed degree of convergence in the exponent, and achieving this only for $Ra \sim 10^{35}$. This numerical delicacy arises from exquisite cancellation of intermediate operands. There is an analogous difficulty one encounters in attempting to derive the scaling exponents from an analytic expansion of the determinant, shown by the need to carry the relevant series to high order in δ as discussed in the Appendix. Unfortunately, the ensuing determinant appears virtually intractable to deal with owing to the factorial profusion of terms. Though expressed in one instance numerically, and the other analytically, clearly these difficulties derive from a common cause.

These results cast light onto the non-standard nature of this variational problem in contrast to equivalent applications to Couette flow or finite-Prandtl-number convection (see for example Doering & Constantin 1994, 1996). In the latter problems the simple asymptotic scaling of the optimal solution is easily attained by semi-optimal estimates using piecewise linear background fields with a fixed slope in the channel interior, for example, the profile depicted in figure 2(a) – (see Busse 1970; Doering & Constantin 1994). Functional estimates of sign-indefinite terms in the spectral constraint enable easy calculation of semi-optimal upper bounds analytically. Here, however, we find an embarrassing gap between the estimates achieved using simple test functions, in conjunction with functional inequalities or numerical implementation of the spectral constraint, and the putative optimal scaling of $Ra^{1/3}$ reported in the Chan (1971) multi-α boundary layer solution following Busse's Couette flow solution.

To support the assumption that this variational problem is non-standard and that more complex test functions must be brought to bear, we present in figure 7 a comparison of the mean temperature field ($\overline{T}$) for free-slip boundaries, taken from three-dimensional direct numerical simulations (DNS) presented in Sotin & Labross (1999), with optimal fields presented in Part 2. The DNS and optimal profiles are in very good qualitative agreement; significantly both have an inversion in the sign of the gradient within the thermal boundary layer. However, the optimal mean has a non-zero interior gradient. The localized overshoot in the boundary layer region does not appear in the optimal solution of the arbitrary-Prandtl-number CDH problem and we therefore speculate that this is one key feature which prevents both the one-parameter and the two-parameter test functions described here from attaining the optimal scaling.

† This result uses only functional analytic inequalities to bound the quadratic functional, however, Constantin & Doering (1999) deduce a logarithmic upper bound of the form $Ra^{1/3}(\log Ra)^{2/3}$ using extra PDE information not contained in the standard CDH problem.

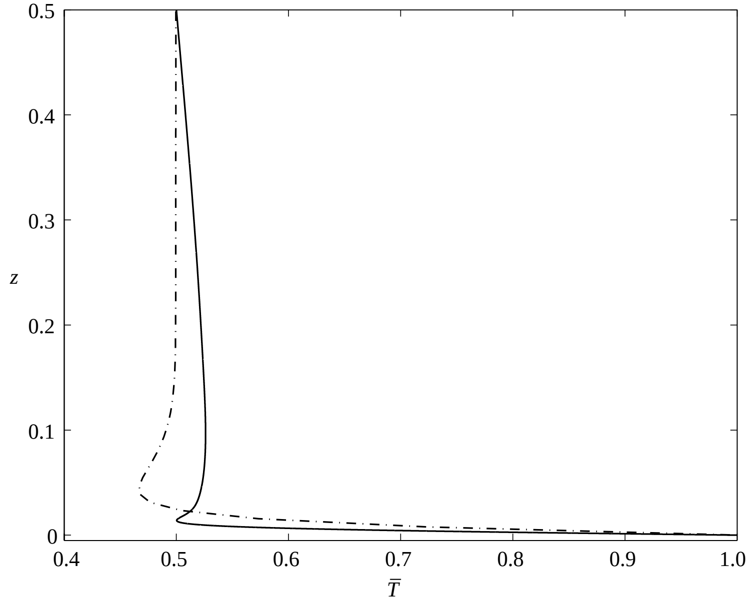

FIGURE 7. Comparison plot of the mean temperature profile for free-slip boundaries at $Ra = 10^7$. The dot-dashed curve depicts data taken from a three-dimensional DNS of the governing equations generously provided by S. Labrosse.

In this paper we have explored the effect of adding more structure to the test functions. We find that the (δ, p)-family also fails to saturate the optimal scaling presented by Chan, and therefore we can neither support, nor refute, the Chan analysis of the optimal solution. The subtlety of the variational solution is such that a simple test function may never saturate the optimal asymptotic scaling. For this reason, in Part 2, we explore the full optimal solution of equations (2.18)–(2.21) with a mixture of numerical and analytic means in order to assess both boundary conditions.

This work was undertaken while S. C. P. was a PhD student of Richard Kerswell. We thank him for his support and encouragement during this project, which included numerous constructive suggestions. We would also like to thank Charles Doering and Jesse Otero for many helpful discussions as well as the 2002 organisers of the GFD summer program at the Woods Hole Oceanography Institution, where this work was begun. G. R. I. especially wishes to acknowledge the generous support from EPSRC grant GR/S02204/01 received during the course of his sabbatical year. He also wishes to thank the Mathematics Department of the University of Bristol for its gracious hospitality during his stay there.

Appendix

Symbolic expansion of the exact 10×10 determinant is well beyond the range of computation. A natural response is to use an approximate form for the matrix, but even this has so far proved intractable on the basis of two plausible approaches. First, consider that we seek only to confirm a specific scaling hypothesis. Taking the case of no–slip, for example, we propose on the basis of numerical experiment that

$$\alpha \sim \alpha_0 \, Ra^{1/4}, \quad \delta \sim \delta_0 \, Ra^{-7/20}, \quad p^* \sim p_0.$$

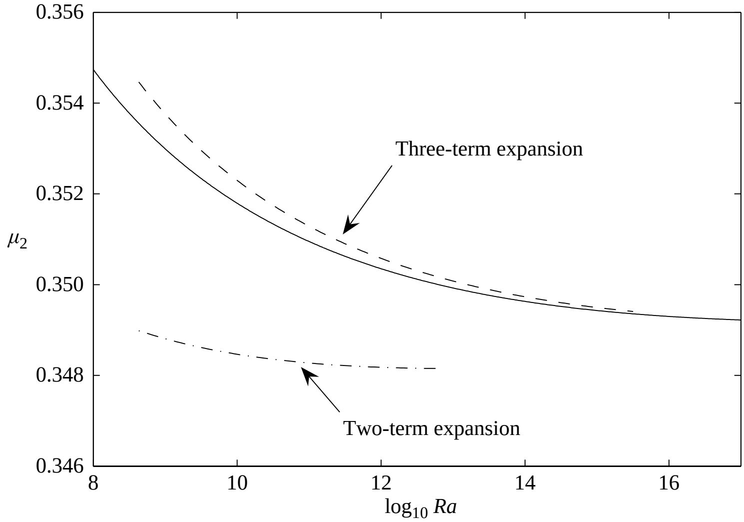

FIGURE 8. Comparison of the numerical exponent $\mu_2 = \mathrm{d}\log(Nu)/\mathrm{d}\log(Ra)$ of the conservative bound on Nu for the no-slip problem (solid line) with a three-term expansion of hyperbolic functions of small argument in $\boldsymbol{M}$, that is expansion out to $([1, \beta_i]\alpha\delta)^5$ order, and a two-term expansion out to $([1, \beta_i]\alpha\delta)^3$.

These already yield the desired answer, $Nu \sim c\,Ra^{7/20}$, and the problem reduces to one of showing that vanishing of the determinant suffices to fix the unknown constants $[\alpha_0, \delta_0, p_0]$, the needed relations for which should follow from application of the fundamental theorem of algebra to the first three terms in Ra appearing in the expansion of the determinant. While simple in principle, in practice expansion of just the intermediate factors which compose the matrix entries already introduces terms of order Ra^{μ} with μ taking values of

$$[1/4, 7/30, 7/60, 7/120, -1/24, -1/12, -1/5, -1/3].$$

If individual matrix entries are then each expanded to leading order only – which leads to a profusion of 25 distinct exponents – the resulting determinant vanishes identically, hence even the first non-trivial relation of the three required arises from products of cross-terms in higher order expansions (symptomatic of the extreme numerical cancellation we observed). Carrying all the needed intermediate expressions in terms of the three unknowns is certainly possible but the resulting calculation – amounting to a proof that 7/20 is a consistent scaling – almost certainly can have no useful insight to convey apart from the bare fact that it can be done.

While it will be apparent that a more basic starting point, where even the exponents governing $[\alpha, \delta, p^*]$ are to be regarded as unknowns, can serve no constructive purpose, we can nonetheless explore a comparably general expansion in purely a numerical setting and see complementary, and more complete, evidence of the cancellations involved. For this purpose, we turn to approximation of the hyperbolic and trigonometric terms, whose arguments are $\alpha\,\delta$ and $\beta_i\,\alpha\,\delta$, assuming only that these products tends to zero.

The result of such an exercise reveals that determination of the saturated limit of the no-slip exponent, namely 7/20, is contingent on the number of terms kept in these expansions, in support of which we refer the reader to figure 8. The vertical axis is the local exponent estimate, $\mu_2 = \mathrm{d}\log Nu/\mathrm{d}\log R$ (notation in keeping with that employed in Part 2). We contrast $\mu_2(Ra)$ for the no-slip conservative bound based on the exact hyperbolic functions with the values of μ_2 calculated using two or three non-zero terms in the transcendental approximations, i.e. (ignoring prefactors) to orders δ^3 and δ^5 respectively. It is manifestly clear that the three-term approximation and the exact form share a common limit of 7/20, while the two-term approximation singles out a different limit. Once again, the implied algebraic burden in approaching the determinant analytically appears disproportionate to the marginal gain in search of rigour.

REFERENCES

Busse, F. H. 1969 On Howard's upper bound for heat transport by turbulent convection. *J. Fluid Mech.* **37**, 457–477.

Busse, F. H. 1970 Bounds for turbulent shear flow. *J. Fluid Mech.* **41**, 219–240.

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

Chan, S. K. 1971 Infinite Prandtl number turbulent convection. *Stud. Appl. Maths* **50**, 13–49.

Chandrasekhar, S. 1961 *Hydrodynamic and Hydromagnetic Stability*. Clarendon.

Constantin, P. & Doering, C. R. 1999 Infinite Prandtl number convection. *J. Statist. Phys.* **94**, 159–172.

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. & Constantin, P. 2001 On upper bounds for infinite Prandtl number convection with or without rotation. *J. Math. Phys.* **42**, 784–795.

Doering, C. R. & Gibbon, J. D. 1995 *Applied Analysis of the Navier–Stokes Equations*. Cambridge University Press.

Foias, C., Manley, O., Rosa, R. & Temam, R. 2001 *Navier–Stokes Equations and Turbulence*. Cambridge University Press.

Hopf, E. 1941 Ein allgemeiner endlichkeitssatz der hydrodynamik. *Mathematische Annalen* **117**, 764–775.

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

Howard, L. N. 1972 Bounds on flow quantities. *Annu. Rev. Fluid Mech.* **4**, 473–494.

Ierley, G. R., Kerswell, R. R. & Plasting, S. C. 2005 Infinite-Prandtl-number convection. Part 2. A singular limit of upper bound theory. *J. Fluid Mech.* (Submitted).

Kerswell, R. R. 1997 Variational bounds on shear-driven turbulence and turbulent Boussinesq convection. *Physica* D **100**, 355–376.

Kerswell, R. R. 1998 Unification of variational principles for turbulent shear flows: the background method of Doering–Constantin and Howard–Busse's mean-fluctuation formulation. *Physica* D **121**, 175–192.

Kerswell, R. R. 2001 New results in the variational approach to turbulent Boussinesq convection. *Phys. Fluids* **13**, 192–209.

Malkus, W. V. R. 1954*a* Discrete transitions in turbulent convection. *Proc. R. Soc. Lond.* **225**, 185–195.

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

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

Nicodemus, R., Grossmann, S. & Holthaus, M. 1998 The background flow method. Part 1. Constructive approach to bounds on energy dissipation. *J. Fluid Mech.* **363**, 281–300.

Otero, J. 2002 Bounds for the heat transport in turbulent convection. PhD thesis, University of Michigan.

Otero, J., Dontcheva, L. A., Johnston, H., Worthing, R. A., Kurganov, A., Petrova, G. & Doering, C. R. 2004 High-Rayleigh-number convection in a fluid-saturated porous layer. *J. Fluid Mech.* **500**, 263–281.

Plasting, S. C. 2004 Turbulence has its limits: *a priori* estimates of transport properties in turbulent fluid flows. PhD thesis, University of Bristol.

Plasting, S. C. & Kerswell, R. R. 2003 Improved upper bound on the energy dissipation in plane Couette flow: The full solution to Busse's problem and the Constantin–Doering–Hopf problem with one-dimensional background field. *J. Fluid Mech.* **477**, 363–379.

Protter, M. H. & Weinberger, H. F. 1984 *Maximum Principles in Differential Equations*. Springer.

Sotin, C. & Labross, S. 1999 Three-dimensional thermal convection in an iso-viscous, infinite Prandtl number fluid heated from within and from below: applications to the transfer of heat through planetary mantles. *Phys. Earth Planet. Inter.* **112**, 171–190.

Vitanov, N. K. 1998 Upper bound on the heat transport in a horizontal fluid layer of infinite Prandtl number. *Phys. Lett.* A **248**, 338–346.

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

Wang, X. M. 2004 Infinite prandtl number limit of rayleigh-bénard convection. *Commun. Pure Appl. Maths* **57**, 1265–1282.

Yan, X. 2004 On limits to convective heat transport at infinite prandtl number with or without rotation. *J. Math. Phys.* **45**, 2718–2743.