J. Fluid Mech. (2001), *vol.* 441, *pp.* 223–253. Printed in the United Kingdom

Bound to improve: a variational approach to convective heat transport

By GLENN R. IERLEY[1] AND R. A. WORTHING[2]

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

[2]Breasco LLC, Ann Arbor, MI, USA

(Received 10 April 1999 and in revised form 19 February 2001)

To the long established idea of bounding turbulent convective heat transport by a variational method based on energetic constraints, we now add a richer class of 'z-constraints' with the hope of tightening bounds considerably. We establish that only certain moments of the governing equations are effective for this purpose. We explore the initial consequences of groups of such constraints by use of perturbation theory, which clarifies the need that a given set of elements be mutually congruent.

1. Introduction

The idea of characterizing statistically steady turbulence by looking for extrema of functionals of the flow subject to integral consequences of the Navier–Stokes equations was initiated in Malkus (1960). Following on this work, Howard (1963) and Busse (1969) deduced bounds on the heat transported in statistically steady planar Boussinesq convection. Doering & Constantin (1996), and more recently Kerswell (1997), reproduced, *ceteris paribus*, the same bound using, not the statistical stationarity hypothesis, but a long-time-average version of the equations. It is not clear which theory better models experimental observables measured over finite space and time. It is clear, though, that the bound obtained bears little resemblance to the data – at least over the experimentally accessible range of R, displayed in figure 1, where the solid curve labelled N_B indicates the best current upper bound on the Nusselt number Nu for finite but arbitrary Prandtl number, σ. It lies disappointingly far above the variety of experimental measurements (indicated by dotted lines).

Our main concern is the large gap between the bound and the data at intermediate Rayleigh numbers. At larger R, it has long been suspected (Kraichnan 1962) that the data veer upward more in accord with the upper bound (though possibly with logarithmic corrections). Interestingly, however, the latest experiment (Niemela *et al.* 2000), which ranges over $10^6 \leqslant R \leqslant 10^{17}$ in cryogenic helium gas, shows no such tendency. Rather, the data agree rather well with a single power-law with exponent approximately 0.31.

The bound plotted was obtained using only two integral consequences of the Boussinesq equations: the 'power integrals'. Speculation that these bounds could actually approach measured experimental values over a wide range of R as more and more statistically steady moments are included as constraints into the variational theory is tempting. Prescient adumbration both of this goal and of the impediments to its achievement is seen in the following quotation, drawn from Busse (1978).

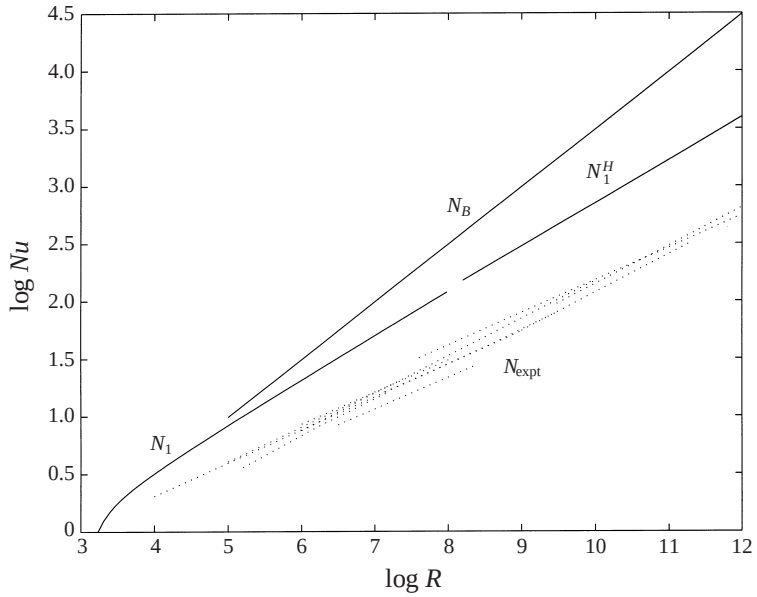

FIGURE 1. The Howard–Busse–Doering–Constantin upper bound (based on the constraints (2.16)–(2.20)) and experimental data. N_1, Howard's single-α bound, computed here; N_1^H, asymptotic estimate from Howard (1963); N_B, Busse's (1969) $R \to \infty$ infinite-α bound; N_{expt}, experimental data for various σ and domain sizes.

The major task in the further development of the optimum theory of turbulence is the introduction of additional constraints in order to improve the bounds. Although it is in principle possible to approach the actual solutions of the Navier–Stokes equations with the optimum transport by adding the hierarchy of moments of the equations as constraints, in practical applications it has proven to be rather difficult to proceed beyond the energetic constraints considered in this article.

However, except for singular situations in which the momentum equation can be reduced to a linear, time-independent, relation,† no improvements on these bounds have been obtained to date even though other statistically steady constraints have been suggested in the literature. Ayyaswamy and Busse (reported in Busse 1978) have looked into breaking down $\langle |\boldsymbol{u}|^2 \rangle_t = 0$ into components (poloidal and toroidal) for the case of Couette flow, but only describe the results as 'not very encouraging'. Using the idea of an adjacent trial field (also exploited in Worthing 1995) Kerswell & Soward (1996) have confirmed that the bound remains unchanged. Malkus & Smith (1989), also with the shear flow problem in mind, suggested the constraint $\langle \omega^2 \rangle_t = 0$ (where ω is a component of vorticity) due to its relationship with the Orr–Sommerfeld equation, but did not pursue its consequences.

It is naïve to think that each proposed constraint might be isolated and its ultimate contribution to improving a bound determined without regard to possible interaction with other constraints. Indeed, any single constraint, alone, may not limit Nu. Only collections of constraints can meaningfully be discussed as a variational problem. Thus, one can calculate the heat flux with and without a particular test constraint but the net reduction, or lack thereof, can only be spoken of with reference to the

† For example, Chan (1971) took the large Prandtl number limit of the convection equations and Gupta & Joseph (1973) took the large Prandtl–Darcy number limit of the porous layer convection equations from the outset.

particular nucleus of other constraints used. For want of a more powerful method, this inevitably somewhat heuristic approach is the route we shall follow at the start of § 3, although a limited basis for classification will emerge.

These ambiguities notwithstanding, there are some constraints about which we can make a definitive assessment. Remarkably enough, these 'inert constraints' abound and, by a process of their elimination, we are directed to a subclass of the horizontally averaged moments of the governing equations, so-called z-constraints. Using finite-amplitude techniques, the effectiveness of a few of the simplest of these is examined near the point of bifurcation in § 3. Based on this analysis, promising problems for future research are suggested and the simplest of these is solved numerically in § 4, where it is shown that even one added constraint, on enstrophy, makes a significant qualitative change in the asymptotic bound on heat flux.

1.1. *Governing equations and statistical steadiness*

The velocity and temperature of convecting fluids are often described well by the Boussinesq equations of motion (cf. Chandrasekhar 1961), which (after scaling) can be written

$$\sigma^{-1}[\boldsymbol{v}_t + \boldsymbol{v}\cdot\nabla\boldsymbol{v}] + \nabla\mathscr{P} = \nabla^2\boldsymbol{v} + R\mathscr{T}\hat{\boldsymbol{k}}, \tag{1.1a}$$

$$\nabla\cdot\boldsymbol{v} = 0, \quad \mathscr{T}_t + \boldsymbol{v}\cdot\nabla\mathscr{T} = \nabla^2\mathscr{T}. \tag{1.1b,c}$$

The Rayleigh number $R = \alpha g\Delta T d^3/\kappa\nu$ and the Prandtl number $\sigma = \nu/\kappa$ are the two dimensionless parameters describing convection between horizontal infinite plates separated by a distance d in the z-direction (denoted by the unit vector $\hat{\boldsymbol{k}}$).† For this geometry,

$$\overline{(\cdot)} \equiv \lim_{L\to\infty}\int_{-L}^{L}\int_{-L}^{L}(\cdot)\,\mathrm{d}x\,\mathrm{d}y\,/(4L^2) \quad \text{(a horizontal average)} \qquad \langle(\cdot)\rangle \equiv \int_0^1\overline{(\cdot)}\,\mathrm{d}z \quad \text{(a volume average)}$$

are useful averages and it is presumed that, even in large finite domains, similar averages of flow quantities are independent of time (the condition of statistical stationarity). With that, it is convenient to decompose the Boussinesq equations in the Reynolds manner,

$$\mathscr{T} = T(z) + \theta, \quad \boldsymbol{v} = \boldsymbol{U}(z) + \boldsymbol{u}, \quad \nabla\mathscr{P} = \nabla P + \nabla p,$$
$$\bar{\boldsymbol{u}} = \boldsymbol{0}, \quad \bar{\theta} = 0, \quad \overline{\nabla p} = 0,$$

and so obtain

$$\nabla\cdot\boldsymbol{U} = \nabla\cdot\boldsymbol{u} = 0, \tag{1.2a,b}$$

$$\sigma^{-1}[\boldsymbol{u}_t + \boldsymbol{u}\cdot\nabla\boldsymbol{u} + \boldsymbol{U}\cdot\nabla\boldsymbol{u} + w\boldsymbol{U}_z] + \nabla P + \nabla p = \boldsymbol{U}_{zz} + \nabla^2\boldsymbol{u} + R(T+\theta)\hat{\boldsymbol{k}}, \tag{1.2c}$$

$$\theta_t + \boldsymbol{u}\cdot\nabla\theta + \boldsymbol{U}\cdot\nabla\theta + wT' = T'' + \nabla^2\theta. \tag{1.2d}$$

Averaging (1.2) gives

$$\hat{\boldsymbol{k}}\cdot\boldsymbol{U} = 0, \quad (\overline{\theta w})_z = T'', \quad (\overline{w\boldsymbol{u}})_z + \nabla P = \boldsymbol{U}'' + RT\hat{\boldsymbol{k}},$$

† Of course, in a horizontally finite domain, the presence of lateral boundaries must be accounted for. It has often been supposed that this source of scaling dependence is lost as the domain is extended but only recently, in Worthing (2001), has the case actually been argued in rigorous detail.

which can be regarded as expressions for mean quantities in terms of fluctuating quantities. Though pressure-driven shear flows are important, here treatment is confined to the case $\hat{\boldsymbol{k}} \times \nabla P = \boldsymbol{0}$. Before the experimental work of Krishnamurti & Howard (1981) it had been assumed that under such conditions no mechanism existed to support large-scale mean velocities. Consequently, early bounds on Nu were derived under the added assumption $\boldsymbol{U} = \boldsymbol{0}$. Within the confines of the variational theory, if the constraint $\boldsymbol{U} = \boldsymbol{0}$ is now relaxed, the corresponding bound can only increase, should it change at all. In fact, at least for the simplest problem based solely on the power integrals, Howard (1990) later showed that his original bound (Howard 1963) remains unaltered when the constraint $\boldsymbol{U} = \boldsymbol{0}$ is dropped. Based on Kerswell's demonstrated coincidence of the Howard–Busse bound with that obtained via the Doering–Constantin formulation (Kerswell 1997), and the absence of any restrictions on $\boldsymbol{U}$ in the latter, evidently the results obtained, while including continuity (Howard 1963; Busse 1969) and the bounds at infinite σ (Chan 1971), also hold equally well (when maximizing the heat flux) without the assumption $\boldsymbol{U} = \boldsymbol{0}$. We verified this numerically for the single-α case subject to boundary conditions, continuity and the power integrals. As one might expect, maximum heat flux occurs for full θ–w correlation and zero u–w correlation.

2. Constraints – a general approach

The arguments in the succeeding three subsections are cumulative in their effect on estimates of the heat flux. Because the reader may find it hard to keep that larger context in mind, we have provided a roadmap in table 1 near the close of this section. That table outlines the cumulative hierarchy of constraints and the successive potential reductions in bounds on the Nusselt number, keying each pair to the subsection(s) in which it is discussed.

2.1. *Bounds without the continuity condition*

With the added assumption that $\boldsymbol{U} = \boldsymbol{0}$, Howard (1963) derived the bound

$$Nu - 1 \leqslant (3R/64)^{1/2} \tag{2.1}$$

based on the boundary conditions $\theta = w = 0$ at $z = 0, 1$ and the following form of the power integrals (Malkus, 1954*a*, *b*):

$$R\,\langle \theta w \rangle = \left\langle |\nabla \boldsymbol{u}|^2 \right\rangle, \tag{2.2a}$$

$$-\left\langle \theta w \{ \overline{\theta w} - \langle \theta w \rangle - 1 \} \right\rangle = \left\langle |\nabla \theta|^2 \right\rangle. \tag{2.2b}$$

No other information about the fluid dynamics was used. At any fixed R, his optimal fields (those that transport the most heat from this class) have the rather simple form

$$\theta = Af(z), \quad w = Bf(z), \quad u = v = 0, \tag{2.3a, b, c}$$

where A and B are scalar coefficients. Notice that these fields possess no vorticity and the full (incompressible) vorticity equation

$$\boldsymbol{\omega}_t + \boldsymbol{u} \cdot \nabla \boldsymbol{\omega} - \boldsymbol{\omega} \cdot \nabla \boldsymbol{u} = \nabla^2 \boldsymbol{\omega} + R \nabla \theta \times \hat{\boldsymbol{k}}, \quad \boldsymbol{\omega} = \nabla \times \boldsymbol{u}, \tag{2.4a, b}$$

is then satisfied outright! It is correct to say that the full time-dependent vorticity equation, alone, places no further restriction upon Howard's original variational problem and his upper bound (2.1) cannot be reduced by including it or any of

its simply related statistically steady moments into the theory. This is a wonderful example of how constraints may/must rely on each other for potency. Evidently, without the continuity constraint, the entire vorticity equation is limp.

Many other constraints are however not satisfied by Howard's optimal fields (2.3). The continuity equation $\nabla \cdot \boldsymbol{u} = 0$, for example, is clearly violated. Other statistical moments, like $(\overline{\theta^2})_t = 0$, are neither included into the variational theory nor satisfied by its solution. Higher-order boundary conditions, i.e. $w_z = 0$ (no-slip) or $w_{zz} = 0$ (slip), are not maintained. The very definition, $\bar{\theta} = 0$, is violated. Offhand, it is not clear what effects these or other additional constraints might have on the bound (2.1). Certainly they provide further restrictions on the class of fields admissible when maximizing Nu. It seems encouraging that the 'old' optimal fields do not satisfy any of the new constraints. One wonders which of these many 'new' constraints might most profitably be appended to Howard's problem with the intent of improving upon his bound. Howard (1963) and Busse (1969) evidently felt the continuity condition was the logical next constraint and probably rightly so: it turns out that 'nearby' fields, transporting nearly the same amount of heat, can be constructed that do satisfy many of these other additional requirements. Besides the continuity condition, the work presented below has led to the emergence of $(\overline{\theta^3})_t = 0$ as a promising alternative.

Prior to this work, Howard (private communication) has kindly relayed to us that his original bound is not reduced when the additional constraint $\overline{\nabla \cdot \boldsymbol{u}} = 0$ (which amounts to $\bar{w} = 0$) is included into the variational theory. This is demonstrated below. In all, it is shown that the bound (2.1) cannot be reduced by the inclusion of the following constraints:

$$[\bar{\theta} = 0, \quad \bar{w} = 0, \quad \langle \theta^m \rangle_t = 0, \quad (\overline{\theta^{2m}})_t = 0, \quad m = 1, 2, \ldots.] \tag{2.5}$$

This is an interesting conclusion since, as shown shortly, only the volume-averaged moments $\langle \theta^m \rangle_t = 0$ are already satisfied by Howard's optimal fields (2.3).

2.1.1. *Preliminaries on Howard's work*

Using the power integrals (2.2), the convective heat transport can be written as a functional that is homogeneous in both θ and $\boldsymbol{u}$. Though formulated slightly differently below, Howard essentially solved for $N_H - 1$ where

$$N_H - 1 \equiv \max_{\theta,\boldsymbol{u}} \langle \theta w \rangle = \max_{\theta,\boldsymbol{u}} \frac{\langle \theta w \rangle^2 - R^{-1} \langle |\nabla \theta|^2 \rangle \langle |\nabla \boldsymbol{u}|^2 \rangle}{\left\langle \left(\overline{\theta w} - \langle \theta w \rangle \right)^2 \right\rangle} \tag{2.6}$$

and $\theta = w = 0$ at $z = 0, 1$. Because $N_H - 1$ is cast as a homogeneous functional of θ and $\boldsymbol{u}$ that satisfy the boundary conditions, one can satisfy (2.2) after the fact provided only that $N_H - 1 > 0$. The requisite rescaling of trial fields $\hat{\theta}$ and $\hat{\boldsymbol{u}}$ is simply

$$\left[\frac{(N-1) \langle |\nabla \hat{\boldsymbol{u}}|^2 \rangle}{R \langle \hat{\theta} \hat{w} \rangle} \right]^{1/2} \hat{\theta} \Rightarrow \theta, \quad \left[\frac{(N-1) R \langle \hat{\theta} \hat{w} \rangle}{\langle |\nabla \hat{\boldsymbol{u}}|^2 \rangle} \right]^{1/2} \hat{\boldsymbol{u}} \Rightarrow \boldsymbol{u}. \tag{2.7a,b}$$

With this observation, the power integrals need not be considered further. Inspection of (2.6) clearly reveals that the optimal fields will have $u = v = 0$ and, as stated earlier, Howard (1963) showed that they have the even simpler form (2.3). Incorporating this information into (2.6) gives

$$N_H - 1 = \max_{f(z)} \frac{\langle f^2 \rangle^2 - R^{-1} \langle f_z^2 \rangle^2}{\left\langle \left(f^2 - \langle f^2 \rangle \right)^2 \right\rangle} \tag{2.8}$$

with $f(0) = f(1) = 0$. Using boundary layer techniques, Howard solved (in the large-R limit) the Euler–Lagrange equations associated with (2.8) and obtained the asymptotic result (2.1).

2.1.2. *Including constraints* $\bar{\theta} = \bar{w} = 0$

Now consider the same problem but with the added requirements $\bar{\theta} = \bar{w} = 0$. The solution to the maximization problem with these constraints cannot be larger than the solution to the maximization problem without these constraints. Yet, a candidate field with heat flux approaching N_H is produced below that satisfies all of the requirements, including $\bar{\theta} = \bar{w} = 0$. Consequently, the upper bound on Nu remains unchanged.

Consider the single-α family

$$\theta = Af(z)\phi(x, y), \quad w = Bf(z)\phi(x, y), \quad u = v = 0$$

where

$$\nabla_H^2 \phi = -\alpha^2 \phi, \quad \bar{\phi} = 0, \quad \overline{\phi^2} = 1,$$

α is a free parameter and the amplitudes A and B are chosen so that the power integrals are satisfied.† Provided α is not identically zero, both $\bar{w} = 0$ and $\bar{\theta} = 0$ as desired. Direct substitution into the expression (2.6) for $Nu - 1$ yields

$$Nu(f;\alpha) - 1 \equiv \frac{\langle f^2 \rangle^2 - R^{-1} \langle f_z^2 + \alpha^2 f^2 \rangle^2}{\left\langle \left(f^2 - \langle f^2 \rangle \right)^2 \right\rangle}. \tag{2.9}$$

Clearly by setting this $f(z)$ to be Howard's 'old' $f(z)$ as derived from (2.8) and then letting α become smaller and smaller, the Nusselt number associated with this trial field approaches N_H from below. It must be concluded that, in the infinite domain, Howard's original bound cannot be improved upon solely by introducing the additional constraints $\bar{\theta} = \bar{w} = 0$ into the theory.

2.1.3. *Including moments of* θ

Results of the sort just presented are extended to encompass the entire class of moments,

$$\bar{\theta} = 0, \quad \bar{w} = 0, \quad \langle \theta^m \rangle_t = 0, \quad (\overline{\theta^{2m}})_t = 0, \quad m = 1, 2, \ldots. \tag{2.10}$$

Indeed, it is demonstrated that Howard's bound (2.1) based on the boundary conditions and power integrals alone, cannot be improved by the imposition of the constraints (2.10). The method here is similar to that of the last section. A family of trial fields is constructed that satisfies all of the constraints (2.10). It is observed that a member within this class possesses a heat flux arbitrarily close to the bound (2.1), and this proves the claim.

We begin with the Euler–Lagrange equations associated with maximizing $\langle \theta w \rangle$ subject to the power integrals (2.2), alone, which, after elimination of multipliers, reduce to

$$wT' - \nabla^2 \theta = 0, \quad \theta T' \hat{\boldsymbol{k}} - \frac{\langle |\nabla \theta|^2 \rangle}{R \langle \theta w \rangle} \nabla^2 \boldsymbol{u} = 0, \tag{2.11a,b}$$

where $T' = \overline{\theta w} - \langle \theta w \rangle - 1$. Any and all solutions of the above system will satisfy the

† As before, such values of A and B are always possible, for any $(f(z), \alpha)$ pair, provided $Nu(f;\alpha) - 1$ (as calculated by (2.9)) is positive, as will generally be true since it is the quantity being maximized.

power integrals. Equations (2.11) admit solutions of the form

$$w = \hat{w}(z;\alpha)\,\sqrt{2}\cos(\alpha x), \quad \theta = \hat{\theta}(z,\alpha)\,\sqrt{2}\cos(\alpha x), \quad u = v = 0. \tag{2.12}$$

From Howard's work, it is known that the solution of (2.11) with the maximum heat flux is symmetric about $z = 1/2$ and is the member of (2.12) with $\alpha = 0$. Based on the regularity of the perturbation in α as well as the variational genesis of the equations, it is assumed that even solutions of (2.11) exist at small α. This is all that is required to prove the results.

The class of moment constraints

$$\underbrace{\frac{1}{m}\langle\theta^m\rangle_t}_{\text{I}} + \underbrace{\frac{1}{m}\langle\nabla\cdot(\theta^m\boldsymbol{u})\rangle}_{\text{II}} + \underbrace{\langle\theta^{m-1}\left[wT' - \nabla^2\theta\right]\rangle}_{\text{III}} - \underbrace{\left\langle\overline{\theta^{m-1}}T''\right\rangle}_{\text{IV}} = 0 \tag{2.13}$$

and the somewhat stronger horizontally averaged versions

$$\underbrace{\frac{1}{m}(\overline{\theta^m})_t}_{\text{I}} + \underbrace{\frac{1}{m}(\overline{\theta^m w})_z}_{\text{II}} + \underbrace{\overline{\theta^{m-1}\left[wT' - \nabla^2\theta\right]}}_{\text{III}} - \underbrace{\overline{\theta^{m-1}}T''}_{\text{IV}} = 0 \tag{2.14}$$

follow directly from the Boussinesq 'θ' equation,

$$\theta_t + \boldsymbol{u}\cdot\nabla\theta + wT' = \nabla^2\theta + T''.$$

For the fields outlined in this section (namely, fields of the form (2.12), satisfying (2.11) and boundary conditions), each of the above terms I–IV in equation (2.13) vanishes independently for all m:

term I	vanishes by statistically steadiness;
term II	vanishes by the divergence theorem and boundary conditions;
term III	vanishes by (2.11);
term IV	vanishes by symmetry in the z-direction.

Terms I and III vanish, similarly, in equation (2.14). Yet, without the subsequent z-integration, terms II and IV do not vanish as their counterparts do in equation (2.13). However, one other mechanism by which these terms may be zero is x-orthogonality. It follows directly that if $\alpha \neq 0$ then odd multiples of single-α fields, e.g. (2.12), vanish under horizontal averaging. So, at least when m is even, equation (2.14) is indeed satisfied.

To sum up, a one-parameter (α) family of fields defined by (2.12) and (2.11) has been devised. As such, each member also satisfies the constraints (2.10) provided α is not identically zero. However, in the infinite domain α may be taken smaller and smaller. By considering the variational form of the heat flux (2.6) and the fact that Howard's solution is the exact limit of our class as $\alpha \to 0$, it is established that the heat transported by these fields comes arbitrarily close to Howard's bound (2.1). This proves that Howard's bound $Nu - 1 \leqslant (3R/64)^{1/2}$, which is based on maximizing the heat flux among all fields satisfying the power integrals and boundary conditions alone, cannot be reduced by including the additional constraints (2.10) in the variational theory. Based on these findings and the earlier observation that the full vorticity equation is satisfied by Howard's original fields,

$$\nabla\cdot\boldsymbol{u} = 0 \quad \text{and} \quad (\overline{\theta^3})_t = 0$$

emerge as perhaps the simplest, potentially immediately influential, constraints to consider including next into the theory.

2.2. *Bounds with the continuity condition*

This section mirrors the previous one in attempting to analyse the inclusion of 'new' constraints into the upper bound theory. The difference, here, is that the continuity condition, $\nabla \cdot \boldsymbol{u} = 0$, is always included in the nucleus of constraints used as a starting point for the assessment of others.

In the last section the maximization problems were solved exactly, even with the additional constraints. If $R \leqslant 32\,000$ and only some of the new constraints are used, then this remains true. However, for larger R and especially when more constraints are added, only a lower estimate of the solution to the maximization problem remains. Should this estimate lie below experimental data, it would be of little interest. However, as it resides well above the data, it provides an immediately useful partial assessment of the effectiveness of these constraints were they to be included into the full variational problem. Our underlying philosophy is pragmatic: if the best possible outcome is poor, try something else.

2.2.1. *Power integrals, continuity and boundary conditions*

Before continuing with the proposed theme, it is useful to review the remaining general results on determining bounds on the heat flux in Boussinesq convection.

In his first paper on the subject, Howard (1963) also considered the problem

$$\max_{\mathbf{u},\theta \in \mathscr{H}} \langle \theta w \rangle + 1 \tag{2.15}$$

with the set of constraints

$$\mathscr{H} \equiv \left\{ \begin{array}{c} \nabla \cdot \boldsymbol{u} = 0, \quad (\boldsymbol{u}, \theta)|_{z=0,1} = (\mathbf{0}, 0), \quad \bar{\boldsymbol{u}} = \bar{\theta} = \boldsymbol{U} = 0 \\ \left\langle |\boldsymbol{u}|^2 \right\rangle_t = 0, \quad \left\langle \theta^2 \right\rangle_t = 0, \end{array} \right\}.$$

He solved this problem for large R under the additional stipulation that it consisted of a single horizontal wavenumber, α. His single-α result, which behaves like

$$N_1 \sim (R/248)^{3/8}, \quad R \to \infty,$$

was discovered by Busse (1969) to be a correct representation of the solution to (2.15) only when $R \leqslant 32\,000$. Busse's work suggests that many bifurcations occur rather quickly after $R = 32\,000$, each associated with an optimal field having an additional horizontal wavenumber. Busse deduced that

$$N_\infty \sim (R/1035)^{1/2}, \quad R \to \infty,$$

which is an improvement over the bound (2.1) though still considerably conservative at experimental values of R.

Busse's successive boundary layers do not separate indefinitely but instead approach a fixed ratio (1/4), a result not obviously self-consistent with the assumed approximations. However, the numerical calculations of Straus (1976*a*, *b*), and the theoretical work of Chan (1971) both lend support to Busse's suggestion of a bifurcation structure in wavenumber. (And Chan's modified problem, in which he takes $\sigma = \infty$ and incorporates the resulting Stokes constraint, yields asymptotic multi-α type solutions directly.)

The most compelling study with reference to the problem of stress-free boundaries is the recent one by Vitanov & Busse (1997), which uses a direct Galerkin optimization with one-, two-, and three-wavenumber solutions. Interestingly, the authors' empirical results unambiguously support asymptotic power law scaling of the wavenumber. A similar conclusion on the optimized variational parameters is reported in Nicodemus,

Grossmann & Holthaus (1998) for the essentially similar problem of turbulent shear flow, but approached by means of the more recent background flow method of Doering & Constantin (1996).

At any rate, while the following 'negative' results might be seen to be yet more negative by utilizing Busse's findings, we have elected to present conclusions that do not hinge on the strict validity of Busse's (or any other's) multiple boundary-layer approximation, nor on any presumed existence (or non-existence) of a mean flow $\boldsymbol{U}$.

2.2.2. *Including other moments*

Here it is demonstrated that the solution of maximum heat flux subject to the constraints

$$\left.\begin{gathered} \langle\theta^2\rangle_t = 0, \quad \langle|\boldsymbol{u}^2|\rangle_t = 0, \\ (\theta, \boldsymbol{u}) = (0, \boldsymbol{0}) \quad \text{at} \quad z = 0, 1, \\ \bar{\theta} = 0, \quad \bar{\boldsymbol{u}} = \boldsymbol{0}, \\ \nabla \cdot \boldsymbol{u} = 0, \end{gathered}\right\} \tag{2.16}$$

and

$$\langle\theta^n\rangle_t = 0 \quad (\forall\, n), \tag{2.17}$$

$$\langle\boldsymbol{\omega}^m\rangle_t = 0 \quad (m \text{ odd}), \tag{2.18}$$

$$\langle\boldsymbol{\omega}^m\theta^n\rangle_t = 0 \quad (m + n \text{ odd}), \tag{2.19}$$

$$\left(\overline{\prod_{j=1}^{n} L_j(\theta)}\right)_t = 0 \quad (n \text{ even and } L_j \text{ arbitrary}), \tag{2.20}$$

cannot, at any R, lie below Howard's single-α, $Nu = N_1$ curve, provided for review in figure 1. (In fact, in the light of Busse's work, N_1 continues to be the exact bound for $R \leqslant 32\,000$.) As usual $\boldsymbol{\omega} = \nabla \times \boldsymbol{u}$ is the vorticity and the L_j are arbitrary linear spatial differential operators. For example, specific members of the class (2.20) whose time derivatives are set to zero are

$$\overline{\theta^2}, \qquad \overline{\theta^4}, \qquad \overline{\theta^3\theta_{zz}},$$

$$\overline{|\nabla\theta|^2}, \quad \overline{(\theta^4\nabla^2\theta_{xz})\nabla \times \hat{\boldsymbol{k}}\theta}, \quad \overline{|\nabla^3\theta|^{100}}.$$

The proof of this claim is quite simple – a specific Howard single-α field having an associated heat transport N_1 is shown to satisfy all of the constraints (2.16)–(2.20). The remainder of this section is devoted to demonstrating this fact.

Particular trial field (*TF*)

The Euler–Lagrange equations associated with just the constraints (2.16) and the restriction $\boldsymbol{U} = \boldsymbol{0}$ are the simple modification of (2.11),

$$\nabla \cdot \boldsymbol{u} = 0, \tag{2.21a}$$

$$wT' - \nabla^2\theta = 0, \tag{2.21b}$$

$$\theta T'\hat{\boldsymbol{k}} - \langle|\nabla\theta|^2\rangle\nabla^2\boldsymbol{u}/(R\,\langle\theta w\rangle) + \nabla p = 0, \tag{2.21c}$$

where $p(\boldsymbol{x})$ is the Lagrange multiplier associated with the divergence-free condition and, as always, $T' = \overline{\theta w} - \langle\theta w\rangle - 1$. Operating on the third of these equations with $\hat{\boldsymbol{k}} \cdot \nabla\times$ yields Laplace's equation with zero boundary conditions for the normal

vorticity $\eta = \hat{\boldsymbol{k}} \cdot \boldsymbol{\omega}$. It follows that solutions to (2.21) are purely poloidal. Still, by depending only on the combined wavenumber α^2, Howard's solution to (2.21) is really an entire class of solutions. For present purposes, the two-dimensional (flat), even (about $z = 1/2$) member of the class

$$w = \hat{w}(z;\alpha)\sqrt{2}\cos(\alpha x), \quad \theta = \hat{\theta}(z,\alpha)\sqrt{2}\cos(\alpha x), \quad \partial_y = v = \boldsymbol{U} = 0, \tag{2.22}$$

is singled out. Insertion of (2.22) into (2.21), along with the optimal choice of $\alpha(R)$, produces curve N_1 of figure 1. With R as a parameter, this constitutes our *trial field*. As shown below, this particular trial field (call it TF) satisfies all of the constraints (2.16)–(2.20).

TF satisfies the constraints

Naturally TF satisfies the original constraints (2.16). The θ-moments (2.17),

$$\begin{aligned} &\text{set} \\ 0 &= \langle \theta^n \rangle_t \\ &= -\underbrace{\langle \boldsymbol{\nabla} \cdot (\theta^n \boldsymbol{u}) \rangle}_{\text{I}} + \underbrace{\left\langle n\theta^{n-1}\left[\nabla^2\theta - wT'\right]\right\rangle}_{\text{II}} + \underbrace{\left\langle n\overline{\theta^{n-1}}T''\right\rangle}_{\text{III}} \end{aligned}$$

are also satisfied by TF:

term I vanishes by the divergence theorem and boundary conditions;
term II vanishes by (2.21);
term III vanishes by symmetry in the z-direction.

Having only a single component of vorticity, $\boldsymbol{\omega} = \gamma\hat{\boldsymbol{j}}$, considerable simplification occurs when the TF is substituted into the $\boldsymbol{\omega}$-moments (2.18),

$$\begin{aligned} &\text{set} \\ 0 &= \langle \boldsymbol{\omega}^m \rangle_t = \langle (\gamma\hat{\boldsymbol{j}})^m \rangle_t \\ &= \left\{ -\underbrace{\langle \boldsymbol{\nabla} \cdot (\gamma^m \boldsymbol{u}) \rangle}_{\text{I}} + \underbrace{\left\langle m\gamma^{m-1}\left[\nabla^2\gamma - R\theta_x\right]\right\rangle}_{\text{II}} \right\} (\hat{\boldsymbol{j}})^m. \end{aligned}$$

This equation is satisfied provided m is odd:

term I vanishes by divergence the theorem and boundary conditions;
term II vanishes when m is odd by orthogonality in x.

Similarly, manipulations on the mixed moments (2.19) yield (within a factor of $\hat{\boldsymbol{j}}$)

$$\begin{aligned} &\text{set} \\ 0 &= \langle \theta^n \boldsymbol{\omega}^m \rangle_t = \langle \theta^n \gamma^m \rangle_t \\ &= m\left\langle \theta^n \gamma^{m-1}\gamma_t \right\rangle + n\left\langle \theta^{n-1}\gamma^m\theta_t \right\rangle \\ &= -\underbrace{\langle \boldsymbol{\nabla} \cdot (\theta^n \gamma^m \boldsymbol{u}) \rangle}_{\text{I}} + \underbrace{\left\langle n\theta^{n-1}\gamma^m(\nabla^2\theta - wT')\right\rangle}_{\text{II}} \\ &\quad + \underbrace{\left\langle m\theta^n\gamma^{m-1}(\nabla^2\gamma - R\theta_x)\right\rangle}_{\text{III}} + \underbrace{\left\langle m\theta^{n-1}\gamma^m T''\right\rangle}_{\text{IV}}. \end{aligned}$$

TF also satisfies these constraints, provided $m+n$ is odd:

term I	vanishes by the divergence theorem and boundary conditions;
term II	vanishes by (2.21);
term III	vanishes when $n+m$ is odd by orthogonality in x;
term IV	vanishes by symmetry in the z-direction.

Lastly, the trial fields TF do indeed satisfy the remaining z-constraints (2.20),

$$\begin{aligned} &\text{set} \\ 0 \;&=\; \frac{\partial}{\partial t}\overline{\prod_{j=1}^{n} L_j(\theta)} \\ &=\; \sum_j \overline{L_j(\theta_t)\prod_{k\neq j} L_k(\theta)} \\ &=\; \underbrace{\sum_j \overline{L_j(\nabla^2\theta - wT')\prod_{k\neq j} L_k(\theta)}}_{\text{I}} + \underbrace{\sum_j \overline{L_j(T'' - \boldsymbol{u}\cdot\nabla\theta)\prod_{k\neq j} L_k(\theta)}}_{\text{II}} \end{aligned}$$

and

term I	vanishes by (2.21);
term II	vanishes when n is even by orthogonality in x.

To recapitulate, it has been demonstrated that any bound on the heat transport, derived subject only to the constraints (2.16)–(2.20), must, at all R, reside above the curve N_1 produced in figure 1 for review. This curve, lying well above the experimental measurements, proves that different and/or more constraints must be incorporated into the theory if bounds approaching the data are to result.

2.2.3. *Including more moments*

Continuing within the theme of the previous section, even more moments of the strongest variety (z-constraints) are considered here as additional constraints within the variational theory. Specifically, extensions of the θ-moments (2.20) are introduced which involve the vorticity. With the principal immediate goal being an assessment of potential, another 'negative' result is deduced.

It is demonstrated that the solution of maximum heat flux subject to the constraints

$$\left.\begin{array}{c} \langle\theta^2\rangle_t = 0, \quad \langle|\boldsymbol{u}^2|\rangle_t = 0, \\ (\theta,\boldsymbol{u}) = (0,\boldsymbol{0}) \quad \text{at} \quad z = 0,1, \\ \bar{\theta} = 0, \quad \bar{\boldsymbol{u}} = \boldsymbol{0}, \\ \nabla\cdot\boldsymbol{u} = 0, \end{array}\right\} \tag{2.23}$$

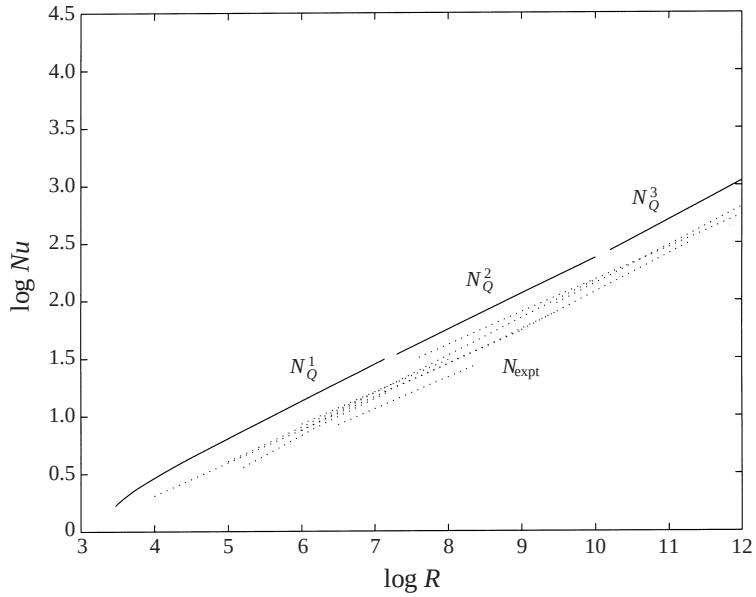

FIGURE 2. Least possible upper bound, N_Q, based on the constraints (2.23)–(2.26) and the measured experimental data. N_Q^1, (low R) present single-α calculations; N_Q^2, (med R) Chan's single-α result; N_Q^3, (high R) Chan's two-α result; N_{expt}, various experimental data.

and

$$\left.\begin{aligned} \langle\theta^n\rangle_t &= 0 \quad (\forall\, n), \\ \langle\boldsymbol{\omega}^m\rangle_t &= 0 \quad (\forall\, m), \\ \langle\boldsymbol{\omega}^m\theta^n\rangle_t &= 0 \quad (m+n \text{ odd}), \\ \left(\overline{\prod_{j=1}^{n} L_j(\theta)}\right)_t &= 0 \quad (n \text{ even and } L_j \text{ arbitrary}), \end{aligned}\right\} \tag{2.24}$$

and

$$\frac{\partial}{\partial_t}\left(\overline{\prod_{j=1}^{m} L_j(\boldsymbol{\omega})}\right) = 0(m \text{ even}), \tag{2.25}$$

$$\frac{\partial}{\partial_t}\left(\overline{\prod_{j=1}^{m}\prod_{i=1}^{n} L_{1,j}(\boldsymbol{\omega})L_{2,i}(\theta)}\right) = 0(m+n \text{ even}), \tag{2.26}$$

must lie above the curve N_Q given in figure 2. As before, the L represent arbitrary spatial linear differential operators.

The technique used to prove this is analogous to that employed earlier – a trial field (TF) is produced that satisfies all of the constraints and has heat flux N_Q. Motivation for the choice of trial function comes from inspection of the two-dimensional version of the simplest vorticity z-moment,

$$\tfrac{1}{2}(\overline{\gamma^2})_t + \tfrac{1}{2}(\overline{\gamma^2 w})_z = \overline{\gamma\left[\nabla^2\gamma - R\theta_x\right]}.$$

With single-α fields in mind, the vanishing of the cubic term suggests setting

$$\nabla^2\gamma - R\theta_x = 0. \tag{2.27}$$

Trial field via the quasi-linear approximation
The above suggests considering fields defined by

$$\left.\begin{aligned} \nabla\cdot\boldsymbol{u} &= 0,\\ wT' - \nabla^2\theta &= 0,\\ \nabla p = \nabla^2\boldsymbol{u} + R\theta\hat{\boldsymbol{k}},&\\ T' = \overline{\theta w} - \langle\theta w\rangle - 1, \quad \overline{uw} &= 0,\\ \bar{\boldsymbol{u}} = \bar{\theta} = 0, \quad (\boldsymbol{u},\theta)|_{z=0,1} &= (\boldsymbol{0},0). \end{aligned}\right\} \tag{2.28}$$

It should be clear that any single-α solution of the above system, with the necessary z-symmetry, will (by design) satisfy many of the integral consequences of the full Boussinesq equations. It is interesting, and relevant to what follows, that other types of solutions (multi-α) can also satisfy such constraints. Our trial field TF is defined as the particular solution of (2.28) having largest heat flux.

Fortunately, equations (2.28) are not new to the literature. They are the steady versions of the Herring (1963, 1964) quasi-linear approximation of the Boussinesq equations. To obtain the quasi-linear approximation, one simply drops fluctuating self-interaction terms like

$$\left\{\boldsymbol{u}\cdot\nabla\boldsymbol{u} - \overline{\boldsymbol{u}\cdot\nabla\boldsymbol{u}}\right\} \quad \text{and} \quad \left\{\boldsymbol{u}\cdot\nabla\theta - \overline{\boldsymbol{u}\cdot\nabla\theta}\right\}. \tag{2.29}$$

In fact, as a demonstration of the effectiveness of the multiple boundary layer techniques devised by Busse (1969) and extended by Chan (1971), Chan also considered exactly system (2.28) and solved asymptotically for those fields having maximum heat flux! Chan's solution has the form

$$w = \sum w_n(z)\phi_n(x,y), \quad \nabla_H^2\phi_n = -\alpha_n^2\phi_n,$$
$$\theta = \sum \theta_n(z)\phi_n(x,y), \quad \overline{\phi_n\phi_m} = \delta_{nm},$$

where u and v follow from continuity and the vanishing of normal vorticity, i.e.

$$\nabla_H^2 u = -w_{xz}, \quad \nabla_H^2 v = -w_{yz}. \tag{2.30}$$

A most remarkable property of his optimal fields is that successive wavenumbers separate asymptotically, i.e.

$$\frac{\alpha_n}{\alpha_{n+1}} \to 0 \quad \text{as} \quad R\to\infty. \tag{2.31}$$

The number of terms comprising the optimal solution is an increasing function of R. The optimal solution is single-α when $R \leqslant 10^{10}$ and carries a heat flux

$$N_Q \sim 0.148(\log R)^{1/5}\, R^{3/10}.$$

Near $R = 10^{10}$, Chan finds that a two-α field becomes optimal, which yields an increased flux of

$$N_Q \sim 0.070(\log R)^{11/50}\, R^{33/100},$$

and this remains optimal until $R \approx 10^{77}$. Similarly, other bifurcations occur and the limiting form is

$$N_Q \sim 0.152R^{1/3} \quad R\to\infty.$$

TF satisfies the constraints

Because the successive wavenumbers which characterize TF in Chan's expansion are well separated, odd products involving the multi-α forms of TF vanish under the horizontal average. This point is instrumental in extending the 'negative' results into the multi-α regime.

Going through all of the constraints, one by one, and showing that they are satisfied by TF seems unnecessarily repetitious. As a representative example, consider only (2.25):

set

$$
\begin{aligned}
0 &= \frac{\partial}{\partial t}\overline{\prod_{j=1}^{m} L_j(\boldsymbol{\omega})} \\
&= \sum_j \overline{L_j(\boldsymbol{\omega}_t)\prod_{k\neq j} L_k(\boldsymbol{\omega})} \\
&= \sigma \underbrace{\sum_j \overline{L_j(\nabla^2\boldsymbol{\omega} - R\hat{\boldsymbol{k}}\times\nabla\theta)\prod_{k\neq j} L_k(\boldsymbol{\omega})}}_{\text{I}} - \underbrace{\sum_j \overline{L_j(\nabla\times[\boldsymbol{u}\cdot\nabla\boldsymbol{u}])\prod_{k\neq j} L_k(\boldsymbol{\omega})}}_{\text{II}} \\
&\quad + \underbrace{\sum_j \overline{L_j(\nabla\times[\sigma\boldsymbol{U}_{zz} - \boldsymbol{U}\cdot\nabla\boldsymbol{u} - w\boldsymbol{U}_z])\prod_{k\neq j} L_k(\boldsymbol{\omega})}}_{\text{III}}
\end{aligned}
$$

which is satisfied by the outlined trial field TF since

term I	vanishes by (2.28);
term II	vanishes when m is even by orthogonality in the horizontal;
term III	vanishes since $\boldsymbol{U} = 0$ by (2.28).

Similar considerations show that (2.23)–(2.26) are also satisfied by TF, and this verifies the claims of this section.

Numerical treatment for moderate R

In the single-α regime, numerical calculations of TF are presented and provide accurate values of N_Q in the low-to-moderate R regime. These results are presented in figure 2 along with Chan's asymptotic estimates. Again, the problem of maximum heat transport subject to the set of constraints (2.23)–(2.26) must lie above N_Q.

2.3. *Bounds at infinite Prandtl number*

Chan solved an upper-bound problem constrained by the $\sigma = \infty$ limit of the momentum equation, the thermal power integral, the continuity equation and the boundary conditions (Chan 1971). In this section, other classes of moments are shown to be completely ineffective in reducing Chan's bound when included as additional constraints into the variational theory.

At $\sigma = \infty$, the momentum equation reduces to steady, linear, Stokes flow, which form can be directly incorporated into the variational theory for statistically steady

By sequentially including the following constraints:	The solution to the upper bound problem must lie above:
$\langle\theta^2\rangle_t = 0$, $\langle\lvert\boldsymbol{u}^2\rvert\rangle_t = 0$ $\theta = w = 0$ at $z = 0, 1$	$N_H \sim (3R/64)^{1/2} \quad R \to \infty$
$\bar{\theta} = 0$, $\bar{\boldsymbol{u}} = \boldsymbol{0}$ $\langle\theta^m\rangle_t = 0; \quad m = 1, 2, \ldots$ $(\overline{\theta^m})_t = 0; \quad m$ even	$N_H \sim (3R/64)^{1/2} \quad R \to \infty$
$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0$ $\langle\boldsymbol{\omega}^m\rangle_t = 0; \quad m$ odd $\langle\boldsymbol{\omega}^m\theta^n\rangle_t = 0; \quad m+n$ odd $\left(\overline{\prod_{j=1}^{m} L_j(\theta)}\right)_t = 0; \quad m$ even	$N_1 \sim (R/248)^{3/8} \quad R \to \infty$
$\left(\overline{\prod_{j=1}^{m} L_j(\boldsymbol{\omega})}\right)_t = 0; \quad m$ even $\left(\overline{\prod_{j=1}^{m}\prod_{i=1}^{n} L_{1,j}(\boldsymbol{\omega})L_{2,i}(\theta)}\right)_t = 0; \quad m+n$ even	$N_Q \sim 0.15R^{1/3} \quad R \to \infty$
Stokes ($\sigma = \infty$)	$N_Q \sim 0.15R^{1/3} \quad R \to \infty$
Change from rigid to free BCs	$N_C \sim 0.32R^{1/3} \quad R \to \infty$

TABLE 1. Summary of major results concerning adding constraints on the upper bound theory for heat transport are presented in this table. The L represent arbitrary linear spatial differential operators and $\boldsymbol{\omega} = \boldsymbol{\nabla}\times\boldsymbol{u}$.

fluid motions. Chan (1971) discovered that maximizing Nu subject to the set

$$\left.\begin{aligned} \boldsymbol{\nabla}\cdot\boldsymbol{u} &= 0, \\ \langle\theta(wT' - \nabla^2\theta)\rangle &= 0, \\ \boldsymbol{\nabla}p &= \nabla^2\boldsymbol{u} + R\theta\hat{\boldsymbol{k}}, \\ T' = \overline{\theta w} - \langle\theta w\rangle - 1, &\quad \overline{uw} = 0, \\ \bar{\boldsymbol{u}} = \bar{\theta} = 0, &\quad (\boldsymbol{u},\theta)|_{z=0,1} = (\boldsymbol{0}, 0), \end{aligned}\right\} \tag{2.32}$$

gives the bound N_Q, i.e. the same bound, to leading order, as obtained by maximizing from the quasi-linear set (2.28). This interesting point, as well as the observation that the quasi-linear set is a subset of Chan's set, are the essential ingredients needed to prove the claims of this section.

Consider an enriched Chan problem in which the set (2.32) is supplemented with other statistically steady constraints based on the remaining thermal evolution equation

$$\theta_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta + \boldsymbol{U}\cdot\boldsymbol{\nabla}\theta + wT' = T'' + \nabla^2\theta.$$

A principal class of such moments is

$$\left.\begin{aligned} \langle\theta^n\rangle_t &= 0 \quad (\forall\, n), \\ \left(\overline{\prod_{j=1}^{n} L_j(\theta)}\right)_t &= 0 \quad (n \text{ even and } L_j \text{ arbitrary}), \end{aligned}\right\} \tag{2.33}$$

and, while not obviously satisfied by Chan's original optimal fields, they are in fact satisfied by the optimal fields of the quasi-linear problem based on the constraints (2.28). Therefore, the optimal quasi-linear solutions remain admissible as trial fields for the appended Chan problem. It follows immediately that the bound for the appended Chan problem remains unchanged at N_Q.

3. Constraints – perturbation theory for slip

In this section some of the simplest z-constraints broached in the preceding section are explored in the regime accessible to finite-amplitude theory (Malkus & Veronis 1958). In addition to this implicit restriction on $R - R_0$, for analytic simplicity we adopt stress-free (slip) boundaries for their natural implementation in trigonometric functions in preference to the no-slip case which, even for the linear problem, has a solution expressed in terms of non-elementary functions. Our hope is that constraints found useful under these simplifying assumptions have greater general application. Also, by considering the Euler–Lagrange equations, we find the explicit rôle of a given set of constraints in determining the symmetry of the solution. We argue that symmetries are one fairly general means to determine in advance a minimal set of constraints that lead to a potent upper bound formulation.

Our approach is mainly to deploy the familiar tools of perturbation theory. While this choice limits our investigation to a small neighbourhood of the initial point of bifurcation (out to perhaps $\epsilon = 1/3$ for quantitative accuracy), it has the virtue of a reasonable degree of rigour. The considerable complexity of the governing equations, to say nothing of the resulting solutions, is such that any immediate attempt to exploit an asymptotic line of reasoning for $R \to \infty$ seems premature in the light of our discovery that arbitrarily constructed sets of constraints may not be 'congruent', a topic on which we shall say more shortly. Similarly, before one embarks on an ambitious numerical program of discovery, such issues merit study because of the likelihood otherwise of mistaking each parametric occurrence of small least-square error as a new 'solution'. Linking, as we do here, the meticulous expansion of the Euler–Lagrange solutions to the classic roll solution that sets in at $R = 27\pi^4/4$, one knows exactly how to interpret the results. These are only a precursor, of course, to the far more interesting issue of behaviour at large R, but they establish the groundwork needed. We conclude with just a brief numerical foray for the simplest possible extension of the Howard–Busse study as a foretaste of more numerical results to come in a future communication.

Here we summarize in table 2 all the constraints that we shall consider in application to a perturbation expansion of convective solutions. In all cases we impose $\nabla \cdot \boldsymbol{u} = \bar{\theta} = \bar{\boldsymbol{u}} = 0$. The latter pair account for the overbars appearing below in (3.1). An issue that certainly requires more scrutiny than we are prepared to give here is proof that solutions of upper-bound problems discussed can, as we assume from here on, be taken in the form of two-dimensional rolls. Near the initial bifurcation this does not seem likely to be a problem but, for larger R, we have no way to exclude the possibility

	Constraint	Lagrange Multiplier
(1)	$\langle \|\boldsymbol{u}\|^2\rangle_t = 0$	$\mu_a[R\langle\theta w\rangle - \langle\|\nabla\boldsymbol{u}\|^2\rangle]$
(2)	$\langle\theta^2\rangle_t = 0$	$\mu_b[\langle\theta\nabla^2\theta\rangle - \langle\theta w T_z\rangle]$
(3)	$\overline{\theta^2_t} = 0$	$\langle\mu_c(z)\{\frac{1}{2}(\overline{\theta^2 w})_z + \overline{\theta w}T_z - \overline{\theta\nabla^2\theta}\}\rangle$
(4)	$\overline{\theta^3_t} = 0$	$\langle\mu_d(z)\{\frac{1}{3}(\overline{\theta^3 w})_z + \overline{\theta^2 w}T_z - \overline{\theta^2\nabla^2\theta} - \overline{\theta^2}T_{zz}\}\rangle$
(5)	$\overline{\gamma^2_t} = 0$	$\langle\mu_e(z)\{(\overline{\gamma^2 w})_z/(2\sigma) + R\overline{\gamma\theta_x} - \overline{\gamma\nabla^2\gamma}\}\rangle$
(6)	$\langle\gamma^2\rangle_t = 0$	$\hat{\mu}_e[\langle\gamma\theta_x\rangle - \langle\gamma\nabla^2\gamma\rangle]$
(7)	$\overline{\gamma^3_t} = 0$	$\langle\mu_f(z)\{(\overline{\gamma^3 w})_z/(3\sigma) + R\overline{\gamma^2\theta_x} - \overline{\gamma^2\nabla^2\gamma}\}\rangle$
(8)	Exact vorticity	$\langle\mu_g(x,z)(R\theta_x - \nabla^2\gamma)\rangle$

TABLE 2.

of bifurcation to a more richly structured upper-bounding flow. Whether and when this may happen is certainly a function of the set of constraints employed. For certain of the z-constraints the solutions may remain two-dimensional to arbitrarily high R. For others, and especially at small σ, fully three-dimensional solutions may emerge.

Application of one or more of these leads to Euler–Lagrange equations of the general form

$$-\Phi\theta_x + \mathscr{G}_1 - \overline{\mathscr{G}_1} = 0, \tag{3.1a}$$

$$\Phi w + \mathscr{G}_2 - \overline{\mathscr{G}_2} = 0, \tag{3.1b}$$

along with the corresponding constraints from table 2. Here

$$\Phi = 1 + \mu_a R - \mu_b(1 + 2T_z) + \mu_c(z)\overline{\theta w} - \langle\mu_c(z)\theta w\rangle + \mu_d(z)\overline{\theta^2 w} - \langle\mu_d(z)\theta^2 w\rangle + (\mu_d\overline{\theta^2})_z \tag{3.2}$$

(incorporating Lagrange multipliers as appropriate) and contributions to the other factors are given by

$$\mathscr{G}_1 = \begin{cases} 2\mu_a\nabla^2\gamma \\ -- \\ \mu_c'(\theta^2)_x/2 - \mu_c T_z\theta_x \\ \frac{1}{3}\mu_d'(\theta^3)_x - \mu_d T_z(\theta^2)_x \\ [\mu_e'\gamma\gamma_x + \nabla^2(\mu_e'\gamma w)]/\sigma - R\nabla^2(\mu_e\theta_x) + \nabla^2(\mu_e\nabla^2\gamma + \nabla^2(\mu_e\gamma)) \\ \hat{\mu}_e(2\nabla^4\gamma - R\nabla^2\theta_x) \\ \left[\mu_f'\gamma^2\gamma_x + \nabla^2(\mu_f'\gamma^2 w)\right]/\sigma - 2R\nabla^2(\mu_f\gamma\theta_x) + \nabla^2(2\mu_f\gamma\nabla^2\gamma + \nabla^2(\mu_f\gamma^2)) \\ \nabla^4\mu_g(x,z) \end{cases} \tag{3.3}$$

and

$$\mathscr{G}_2 = \begin{cases} -- \\ 2\mu_b \nabla^2\theta \\ -\mu_c'\theta w + \mu_c T_z w - \nabla^2(\mu_c\theta) - \mu_c\nabla^2\theta \\ -\mu_d'\theta^2 w + 2\mu_d T_z\theta w - \nabla^2(\mu_d\theta^2) - 2\mu_d\theta\nabla^2\theta - 2\mu_d\theta T_{zz} \\ -R\mu_e(z)\gamma_x \\ -R\hat{\mu}_e\gamma_x \\ -R\mu_f(z)(\gamma^2)_x \\ -R\partial\mu_g/\partial x. \end{cases} \tag{3.4}$$

(The explicit coordinate dependence of the Lagrange multipliers is shown only for those entries where it is not obvious from the context.)

The Euler–Lagrange equations do not have solutions for arbitrary $\mu_{a,b,c,\ldots}$. Rather these must satisfy certain consistency relations in terms of the physical variables. The complicated forms do not usually permit one to isolate simple expressions for each μ hence in all but a few cases these auxiliary variables must typically be carried through as a part of the solution procedure. Below we present suitable forms for perturbation expansion about the initial bifurcation:

$$\mu_{a,b} = \sum_{k=-1} \mu_{a,b}^{(k)} \epsilon^{2k}, \tag{3.5a}$$

$$\mu_{c,e}(z) = \sum_{k=-1} \sum_{j=0}^{k+1} \mu_{c,e}^{(k,j)} \epsilon^{2k} \cos(2j\pi z), \tag{3.5b}$$

$$\mu_d(z) = \sum_{k=0} \sum_{j=1}^{k+2} \mu_d^{(k,j)} \epsilon^{2k} \cos((2j-1)\pi z)/\sin(\pi z). \tag{3.5c}$$

The subordinate multiplier, $\hat{\mu}_e$, is just the restriction of the general sum above for $\mu_e(z)$ to $j = 0$. In lieu of expanding μ_g, it is easier to eliminate it from the Euler–Lagrange equations by differentiation and substitution, leaving a single higher-order equation in θ.

The constraint of incompressibility is enforced by introduction of a stream function, ψ, where $\boldsymbol{u} = -\nabla \times (\psi(x,z)\hat{\boldsymbol{j}})$. Finally, we give the most general form we shall require in the expansion of ψ and θ, for which it is useful to note that the steady Boussinesq problem admits the two-dimensional symmetry ($x \to -x, z \to -z, \psi \to \psi$, $\theta \to -\theta$):

$$\psi = \epsilon\psi_{1,1}^{(1)} \sin(kx)\sin(\pi z) + \sum_{q=1}^{\infty}\sum_{m=1}^{q}\sum_{n=1}^{q+1} \epsilon^{2q+2}\psi_{m,n}^{(2q+2)} \sin(2mkx)\sin(2n\pi z),$$

$$+ \sum_{q=1}^{\infty}\sum_{m=1}^{q}\sum_{n=1}^{q+1} \epsilon^{2q+1}\psi_{m,n}^{(2q+1)} \sin[(2m-1)kx]\sin[(2n-1)\pi z], \tag{3.6a}$$

$$\theta = \epsilon\theta_{1,1}^{(1)}\cos(kx)\sin(\pi z) + \sum_{q=1}^{\infty}\sum_{m=1}^{q}\sum_{n=1}^{q+1}\epsilon^{2q+2}\theta_{m,n}^{(2q+2)}\cos(2mkx)\sin(2n\pi z)$$

$$+\sum_{q=1}^{\infty}\sum_{m=1}^{q}\sum_{n=1}^{q+1}\epsilon^{2q+1}\theta_{m,n}^{(2q+1)}\cos[(2m-1)kx]\sin[(2n-1)\pi z]. \quad (3.6b)$$

(Results of interest will require that we carry out the expansion of ψ and θ to $O(\epsilon^7)$.) This is, of course, the form for the canonical roll solution, as first elucidated in the influential and still commonly cited Malkus & Veronis (1958). Without loss of generality, $\psi_{1,1}^{(q)}$ can be set to zero for $q > 1$ (this amounts to a definition of ϵ).

A minor qualification: the need to carry out expansions to $O(\epsilon^7)$ to understand the interaction among constraints has prompted us to fix our attention on a periodic box of length $L = 2\sqrt{2}$ (i.e. $\alpha_c = k_c/\pi = 1/\sqrt{2}$), the value at the first onset of convection. Over a limited range of $R - R_0$, it is reasonable that the extremum is realized on the same length scale. At fixed α (or L), each set of constraints yields a bound for Nu in the form of a power series, $\sum c_j r^j$, whose coefficients we present shortly. If we allowed for variation in α as well, the series would assume the more general form $\sum d_{j,k} r^j(\alpha-\alpha_c)^k$, representing a surface instead of simply a slice through that surface. Though all the $d_{j,k}$ are determinate, the algebra required for their determination when expansions are carried to $O(\epsilon^7)$ is forbidding even with the aid of *Maple*. To illustrate what we mean by 'forbidding', note that a typical result (here carried only to $O(\epsilon^5)$ in (ψ, θ) for the shortly to be discussed Case I) is

$$R/R_0 = 1 + \frac{\alpha^2\epsilon^2}{\alpha^2+1} - \frac{\epsilon^4(\alpha^4+8\alpha^2+23)\alpha^4}{2(3\alpha^4+30\alpha^2+91)(\alpha^2+1)}, \quad (3.7a)$$

$$Nu < 1 + 2\frac{\alpha^2\epsilon^2}{\alpha^2+1} - 2\frac{\epsilon^4\alpha^4}{(\alpha^2+1)^2}$$

$$+\frac{\alpha^6(\alpha^2+5)(\alpha^{10}+49\alpha^8+698\alpha^6+4882\alpha^4+18101\alpha^2+30253)\epsilon^6}{8(\alpha^2+1)^3(3\alpha^4+30\alpha^2+91)^2}, \quad (3.7b)$$

where $\alpha = k/\pi$.

Finally, we make a few remarks on the rôle of the Prandtl number. As the reader can see in the Euler–Lagrange equations for the Howard–Busse problem, σ does not enter in the solution and thus the results remain valid at all σ without restriction. The same is true for any combination of constraints (1–4) and (6) chosen from table 2. In both the fifth constraint, $\overline{\gamma_t^2} = 0$, and the seventh constraint, $\overline{\gamma_t^3} = 0$, σ appears explicitly and this will be reflected in the solutions to the Euler–Lagrange equations. The last constraint is applicable only to the case of infinite Prandtl number. In first comparing various results below we confine ourselves to infinite Prandtl number. As with wavenumber, our choice here too is partly a concern for economy and clarity, but in addition there are some lacunae for certain expansions at finite σ which are best reserved for § 3.1 for separate discussion.

Table 3 shows the solution for $Nu(r)$ for ten model variational problems (where $r = R/R_0 - 1$). The eleventh line is an exact solution of the Boussinesq problem. We have reduced all the exact rational expressions appearing as coefficients to decimal form to facilitate quantitative comparison.

The results above are in one respect misleading: the coincidence of coefficients in $Nu(r)$ masks that expansion coefficients of ψ and θ can differ qualitatively. Table 4

Case	Constraints	$Nu(r)$
I*	$\langle\|\boldsymbol{u}\|^2\rangle_t = \langle\theta^2\rangle_t = 0$	$1 + 2r - 1.617096019r^2 + 1.635930447r^3$
II*	$\langle\|\boldsymbol{u}\|^2\rangle_t = \overline{\theta_t^2} = 0$	$1 + 2r - 1.617096019r^2 + 1.635930447r^3$
III*	$\langle\|\boldsymbol{u}\|^2\rangle_t = \langle\theta^2\rangle_t = \overline{\theta_t^3} = 0$	$1 + 2r - 1.617096019r^2 + 1.584452076r^3$
IV	$\langle\|\boldsymbol{u}\|^2\rangle_t = \langle\theta^2\rangle_t = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.666957515r^2 + 1.669589876r^3$
IV(a)*	$\langle\|\boldsymbol{u}\|^2\rangle_t = \langle\theta^2\rangle_t = \langle\gamma^2\rangle_t = 0$	$1 + 2r - 1.666957515r^2 + 1.669589876r^3$
IV(b)**	$\langle\theta^2\rangle_t = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.666957515r^2 + 1.669589876r^3$
V	$\langle\theta^2\rangle_t = R\theta_x - \nabla^2\gamma = 0$	$1 + 2r - 1.666957515r^2 + 1.669589876r^3$
VI**	$\langle\theta^2\rangle_t = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.666957515r^2 + 1.599629855r^3$
VII	$\langle\|\boldsymbol{u}\|^2\rangle_t = \langle\theta^2\rangle_t = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.666957515r^2 + 1.597087539r^3$
VIII	$\langle\theta^2\rangle_t = \overline{\theta_t^3} = R\theta_x - \nabla^2\gamma = 0$	$1 + 2r - 1.666957515r^2 + 1.592612917r^3$
IX**	$\overline{\theta_t^2} = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.667154567r^2 + 1.669845200r^3$
X**	$\overline{\theta_t^2} = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0$	$1 + 2r - 1.667154567r^2 + 1.637365763r^3$
XI**	Malkus–Veronis	$1 + 2r - 1.667154567r^2 + 1.592322379r^3$

TABLE 3. Cases marked with a single asterisk apply for all σ as they stand. The double asterisk indicates results that have supplementary σ-dependent terms (listed in table 5). All other cases apply only at infinite Prandtl number.

m	I	II	III	IV	IV(a)	IV(b)	V	VI	VII	VIII	IX	X	XI
1	•	•	•	•	•	•	•	•	•	•	•	•	•
2			◇			○		•	•	•	•	•	•
3			•			○		•	•	•	•	•	•
4													•
5													•

TABLE 4. The harmonic content (mth multiple of $k_c x$) present for each perturbation expansion to $O(\epsilon^7)$. Open circles denote terms that are absent for $\sigma \to \infty$, the diamond indicates terms occurring in θ only.

partly clarifies this by indicating which x-harmonics are non-zero in each solution set. (The last column is the exact Malkus–Veronis reference solution.) Further detail remains hidden: that two solutions share the same pattern of harmonics does not indicate identity in the values of the coefficients. But it would consume many pages to exhibit the exact expansion coefficients even if limited to fixed α and infinite Prandtl number, and to little purpose.

One can see by comparison with the exact solution which degrees of freedom remain yet unexploited by any particular set of constraints. That the relevant coefficients in the upper bounding ψ or θ go unused points to the pertinence of other z-dependent constraints as an agent needed in producing any further qualitative, as opposed to merely quantitative, reductions.

Cases I and XI provide bounds for the remainder of results: Case I the result of applying only the power integrals – the Howard–Busse model – and Case XI the exact Malkus–Veronis solution.

Case II is included to verify the result in § 2 that, perhaps counter-intuitively, the $\overline{\theta_t^2}$ correction taken in conjunction with only the power integrals is nugatory. Note that one power integral is in fact redundant, so the proper variational formulation omits it.

Case III marks the first improvement to the upper bound. Drawing on the arguments of § 2, we use the the cubic z-constraint on θ to augment the power integrals. Notice that the r^2 coefficient is unaffected. That the first correction is $O(r^3)$ can be anticipated from the Euler–Lagrange equations but in the interest of brevity we omit the argument. Rather, we draw the reader's attention to a more critical point: the contribution to $\mathscr{G}_1$ manifestly does not permit solutions of a single wavenumber in x. This seems at first glance also to be true of Case II, but there is a subtle difference: notice that the cubic moment, $\frac{1}{2}(\overline{\theta^2 w})_z$, in the associated Lagrange multiplier vanishes identically for a Howard–Busse single-wavenumber solution. Thus there is no need for an associated constraint and, for Case II, μ_c is not a function of z. In contrast, the single-wavenumber solution gives a non-zero contribution to $\frac{1}{3}(\overline{\theta^3 w})_z$ arising from $\overline{\theta_t^3}$, hence is not admitted as a possible solution. The multiplier μ_d thus always has a non-trivial z-dependence and the Euler–Lagrange equations not only preclude a single-wavenumber solution, they also break the two-fold symmetry of Cases I and II. Once the $\overline{\theta_t^3}$ constraint is introduced, the associated variational field θ cannot have any higher symmetry than that of the original equations, that is, the Boussinesq symmetry. As one can readily imagine, it takes the application of a z-constraint derived from the momentum equation to comparably restrict the symmetry of the stream function. In purely quantitative terms, the addition of the $\overline{\theta_t^3}$ constraint alone could be considered slight since it enters only at the r^3 coefficient. But for values of R well above R_0 (but not necessarily in the asymptotic regime where shear flow instabilities begin to dominate the scaling) there is good reason to believe that the θ-restriction to the Boussinesq symmetry will exert an influence at leading order (cf. § 4) that qualitatively alters the simpler Howard–Busse bound.

Cases IV, IV(a), IV(b) and V have increasingly tight constraints on the velocity field, with V incorporating the Stokes equation which, for $\sigma \to \infty$, gives an exact, not merely integral, representation for the velocity field in the variational theory. Notice that entries in table 3, however, are identical to $O(r^3)$, that is to say, that $Nu(r)$ is not necessarily a very sensitive indicator near the point of bifurcation given that the detailed forms of ψ and θ are distinct in the four cases. Oddly enough, as we shall see later, this insensitivity in Nu persists to large R, with Cases IV(a) and V (the Chan problem) giving essentially similar results. Case IV, which is more tightly constrained than IV(a) and less than V, must therefore exhibit a similar convergence.

Cases VI and VII differ only in that the former lacks one power integral. A fuller discussion of the consequences is deferred to the next section. The immediate consequence is a slight change in the r^3 coefficient. The natural case with which to compare VI is Case X, which substitutes $\overline{\theta_t^2} = 0$ in place of $\langle \theta^2 \rangle_t = 0$. Notice that the quadratic coefficient then matches the Malkus–Veronis result exactly, but the cubic coefficient is thrown considerably off. Either Case VI or Case X is a natural pivot for any numerical investigation having as its object the impact of z-constraints. Case III is also possible for that purpose, but notice how improvements at $O(r^2)$ are most strongly affected by the simple volume constraint on the enstrophy. In the instance of heat transport in a porous medium satisfying Darcy's law, Doering & Constantin (1998) find that the enstrophy constraint plays a critical rôle for two-dimensional convection at arbitrary Darcy–Prandtl number; with its inclusion, the earlier single-

wavenumber result of Busse & Joseph (1972) can be shown to be a rigorous (if not optimal) upper bound for all R, not just $R < 113$, the limit of validity when only the energy integrals are used.

The purpose in presenting Case VIII is that all of the error is attributable to defects in θ-constraints since the velocity field is exactly represented by the Stokes solution. Compare Case V, the Chan problem: all other things being held fixed, the constraint that $\overline{\theta_t^3} = 0$ tightens the r^3 coefficient. But it does so in a more complicated way – the single-α solution is gone, as one can see from table 4.

Case IX and Case I are natural points of comparison: both enforce two quadratic constraints, but the first solely through volume integrals, the second only via z-constraints. The price, as with all the cases starting with VI, is structurally more complicated solutions.

3.1. *Congruent constraints*

One is free to invoke any mixture of constraints, but it is not always the case that added constraints actually produce a tighter bound. In some instances, it may be that the added constraint is already satisfied owing to a (perhaps overlooked) identity. But a more interesting case is one in which the reason that the added constraint proves ineffective is that the solution falls outside the class of smooth solutions, typically through the induction of singularities that affect the constraint but make an arbitrarily small change in the variational form. The failure of standard perturbation theory easily to accommodate certain combinations of elements is, we believe, because the variational solution is non-smooth. The cases here are sufficiently complex that we are not yet able to prove this assertion. (What we have established in the cases noted below is the appearance of an inconsistent overdetermined system of algebraic equations in the perturbation coefficients.) In order, however, that the reader can at least see how a non-smooth solution can arise, we turn later in this section to a simpler setting, where a more constructive approach is possible.

In at least one respect, one can anticipate why a mixed bag of constraints might be problematic. Exact solutions of the p.d.e. satisfy an infinite number of constraints and, in particular, one can find as many quadratic moments of either equation as desired, say $\langle \gamma^2 \rangle_t = 0$ but also $\langle |\nabla \psi|^2 \rangle_t = 0$. From the standpoint of perturbation theory, however, if one imposes two z-constraints of a given order, one in each field (as in Case IX), a unique expansion follows. Any such an expansion will, in general, fail to satisfy any other quadratic z-constraints derived from the same differential equation or indeed any other volume constraints that do not follow as an identity from the full z-constraints already used. However, the violations, while stemming from, say, a quadratic constraint, emerge at higher order in ϵ, for example, Case IX above with its $O(\epsilon^8)$ discrepancy in satisfying $\langle |\boldsymbol{u}|^2 \rangle_t = 0$. The remedy is not simply to add $\langle |\boldsymbol{u}|^2 \rangle_t = 0$ as yet another constraint. Rather, further z-constraints of higher order (cubic and quartic moments, for example) will yield an orderly perturbation expansion, modify the necessary higher-order coefficients, and thus cure the apparent various quadratic constraint misfits retroactively. A heuristic suggestion is to use just one z-constraint of a given order in each field and no volume constraints that are not otherwise automatically satisfied. Unfortunately, the answer is not quite so easy to implement, since, as the reader may have noticed, the seventh entry of the constraint table 2, $\overline{\gamma_t^3} = 0$, is nowhere employed in the various cases. Solutions with the Boussinesq symmetry satisfy this automatically, so the expansion coefficients of μ_f turn out to be zero order by order. (Similar remarks apply to $\overline{\gamma_t^{2j+1}} = 0$.) We have not

Case	Constraints	σ corrections to r^3 term in $Nu(r)$
IV(b)	$\langle\theta^2\rangle_t = \overline{\gamma_t^2} = 0$	$0.000369452\,\sigma^{-2}$
VI	$\langle\theta^2\rangle_t = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0$	$-0.001860525\,\sigma^{-1} + 0.000356787\,\sigma^{-2}$
IX	$\overline{\theta_t^2} = \overline{\gamma_t^2} = 0$	$0.000146879\,\sigma^{-1} + 0.000388586\,\sigma^{-2}$
X	$\overline{\theta_t^2} = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0$	$-0.001907692\,\sigma^{-1} + 0.000353656\,\sigma^{-2}$
XI	Malkus–Veronis	$-0.011034667\,\sigma^{-1} + 0.002790460\,\sigma^{-2}$

TABLE 5.

yet discovered an elementary cubic constraint derived from the general form of the momentum equation which, along with $\overline{\theta_t^3} = 0$, complements Case X and reproduces the Malkus–Veronis result exactly to $O(r^3)$. The various forms that might work all seem to induce surface terms, at least formally, a complication we have generally tried to avoid. Evidently it is appropriate to impose $\overline{\theta_t^k} = 0$ for $k = 2, 3, 4, \ldots$ in the search for successively tighter bounds but a comparable recipe deriving from the momentum equation is still not clear.

3.1.1. *Problematic expansions*

The first and most overwhelming impression of the results in table 5 is that the σ-dependent bounds are terrible. Note that in Case IX, even the sign is wrong on the first term. (Note that we are assuming that for R sufficiently close to the bifurcation, it should be possible to approach the Malkus–Veronis solution with a sufficiently constrained upper bound.) To be sure, the absolute corrections would, for many common fluids, be so small that one might set this issue aside as a minor curiosity. The exception, of course, is exploration of the limit of zero Prandtl number, an upper-bound problem which, in the instance of z-constraints, is of considerable interest.

For infinite σ, one can pose the problem

$$\langle|\boldsymbol{u}|^2\rangle_t = \langle\theta^2\rangle_t = \overline{\theta_t^3} = \overline{\gamma_t^2} = 0 \tag{3.8}$$

as indicated by the result for Case VII in table 3. For finite σ, however, the expansion fails. The only resolution is to discard the first power integral (Case VI). Doing so gives the following σ-dependent bound:

$$1 + 2r - \frac{513913}{308294}r^2 + \left(\frac{41386105598591}{25872301312368} - \frac{388854312}{209002434071}\sigma^{-1} + \frac{18269508312}{51205596347395}\sigma^{-2}\right) r^3. \tag{3.9}$$

Notice that the limit of this for $\sigma \to \infty$ shown in table 3 is a less tight bound on the r^3 coefficient than for Case VII, as one might expect given that a scalar constraint is lost.

Case IV is similar to VII – it also has a smooth expansion only for $\sigma \to \infty$. The restriction to the volume integral of γ^2, rather than the z-constraint, Case IV(a), reverts to a σ-independent constraint as shown in table 2. This, like the original Howard–

Busse result, is a bound that applies at all σ. To obtain a well-posed formulation with the z-constraint, again we are forced to drop one power integral. This is recorded as Case IV(b). Notice in table 4 that the symmetry-breaking terms that result all vanish in the limit of $\sigma \to \infty$ and, in contrast with the VI/VII pair above, the expansions for $Nu(r)$ for IV/IV(b) are then identical.

Interestingly, the appearance of breakdown is not solely an artifact of finite Prandtl number: already for $\sigma \to \infty$ Case IX in the table does not admit the addition of $\langle |\boldsymbol{u}|^2 \rangle_t = 0$. Thus the selection rule that says when a power integral (or other constraint) is permitted, when it is incongruent, and when it is simply redundant, is not elementary. Fortunately, the rule is not hopelessly opaque either: the success or failure of the perturbation expansion is a useful, if tedious, operational guide. Useful because the formalism tells one unambiguously that the solution does not lie in the class of smooth solutions and tedious because, in several instances, the expansion breaks down only at $O(\epsilon^7)$.

3.1.2. *Discontinuous vs. smooth solutions: a model problem*

The introduction of integral moments having boundary terms, like the vorticity integral suggested by Malkus & Smith (1989), as constraints into the variational theory may not lead to tighter bounds. One reason for this is that such constraints can often be satisfied by local adjustments near or at the boundary, to which effect the other relevant global quantities are insensitive. As an example, the form of the boundary term associated with the enstrophy integral $\langle |\nabla \times \boldsymbol{u}|^2 \rangle_t = 0$ is considered. In two dimensions ($v = \partial_y = 0$) this boundary component emerges from the term $\langle \gamma \nabla^2 \gamma \rangle$ where $\gamma = \boldsymbol{j} \cdot \boldsymbol{\omega} = u_z - w_x$ is the remaining component of vorticity. Integrating by parts while observing no-slip boundary conditions gives

$$\overline{u_z u_{zz}}|_{z=0}^{z=1} - \langle u_{zz}^2 \rangle + \text{other volume terms.}$$

3.1.3. *Model*

A simple one-dimensional model containing a similar constraint is perhaps instructive. Consider the minimisation problem

$$\lambda = \min_{u(z)} \frac{\langle u_z^2 \rangle}{\langle u^2 \rangle}$$

subject to the boundary conditions $u(0) = u(1) = 0$. With no other constraints, the global minimum is, of course, $\lambda = \pi^2 = 9.8696044\ldots$ with $u(z) = \sin(\pi z)$ being the optimal form.

The change in λ is now investigated when the problem is modified to include the additional constraint

$$\langle u_z u_{zzz} \rangle = u_z u_{zz}|_0^1 - \langle u_{zz}^2 \rangle = 0. \tag{3.10}$$

Clearly (3.10) is not trivially satisfied by the previous extremal. However two pieces of evidence are provided below suggesting that the minimum λ remains unaltered.

First, numerical solutions are obtained within the space of polynomials of degree $\leqslant N$. For different N the values of λ are contained in table 6 and the optimal polynomials are given in figure 3 along with informative higher derivatives.

From this sequence, it is apparent that the solution remains essentially $u(z) = \sin(\pi z)$ except near the boundaries where large changes in higher-order derivatives

N	λ
4	9.87996
8	9.87016
12	9.86968

TABLE 6.

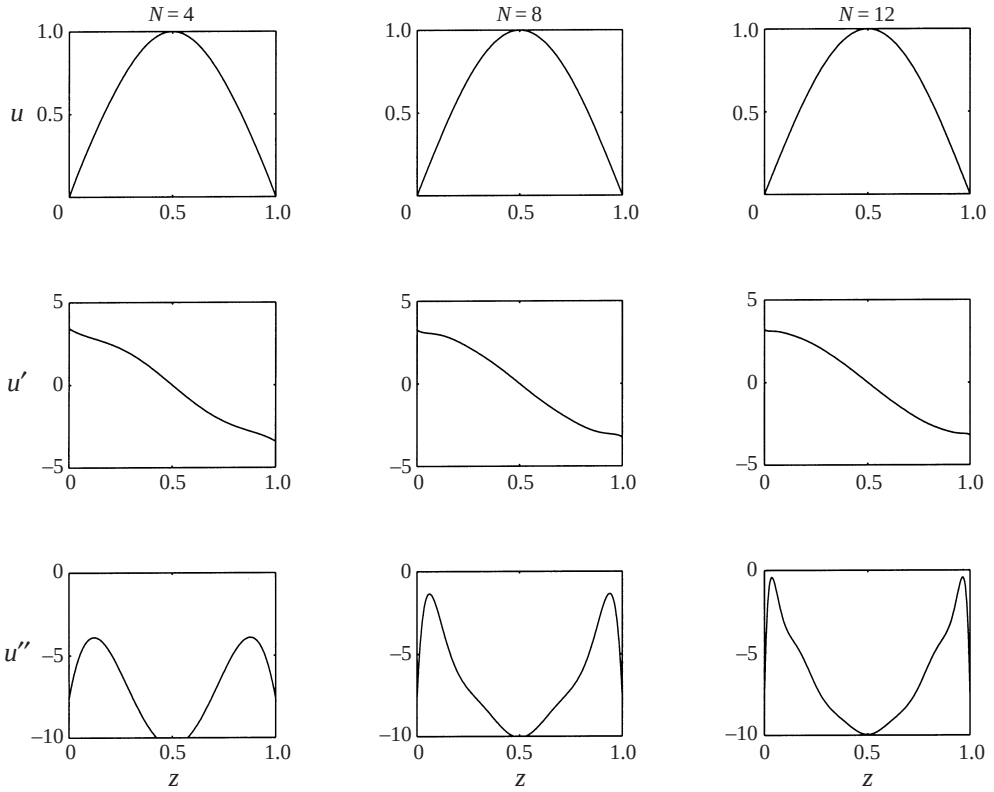

FIGURE 3. The solutions of minimum $\langle u_z^2 \rangle / \langle u^2 \rangle$ subject to $u(0) = u(1) = 0$ and $\langle u_z u_{zzz} \rangle = 0$ are computed using polynomials of degree $\leqslant N$, where $N = 4$, 8 and 12, respectively.

occur. These deviations seem designed to satisfy $u_z u_{zz}|_0^1 - \langle u_{zz}^2 \rangle = 0$ without incurring significant alterations of the global integrals contained in the calculation of λ.

This same idea is captured by the trial field

$$u_{zz}(z) = \begin{cases} -A, & 0 < z < \epsilon \\ -\pi^2 \sin(\pi z), & \epsilon < z < 1, \end{cases}$$

in which a jump is allowed in the curvature at $z = \epsilon$. Integrating twice gives

$$u(z) = \begin{cases} -Az^2/2 + Bz + C, & 0 < z < \epsilon \\ \sin(\pi z), & \epsilon < z < 1, \end{cases}$$

and the coefficients A, B and C are found to be

$$A = \frac{\pi^3\left(1-\epsilon+(2\pi)^{-1}\sin(2\pi\epsilon)\right)}{2\cos(\pi\epsilon)} \sim \pi^3/2 + \pi^5\epsilon^2/4 + O(\epsilon^3),$$
$$B = \epsilon A + \pi\cos(\pi\epsilon) \sim \pi + \pi^3\epsilon/2 - \pi^3\epsilon^2/2 + O(\epsilon^3),$$
$$C = A\epsilon^2/2 - B\epsilon + \sin(\pi\epsilon) \sim -\pi^3\epsilon^2/4 + O(\epsilon^3),$$

by continuity of u and u_z at $z = \epsilon$ and forcing directly the constraint $u_z u_{zz}|_0^1 - \langle u_{zz}^2\rangle = 0$. In this limited formulation, $u(0) = C = O(\epsilon^2)$ and so that particular boundary condition is formally only satisfied in the limit $\epsilon \to 0$. The value of the functional provided by this trial field is

$$\frac{\langle u_z^2\rangle}{\langle u^2\rangle} \sim \pi^2 + \pi^4\epsilon^2 + O(\epsilon^3)$$

and so the earlier value of $\lambda = \pi^2$, as well as all other conditions, are approached by this trial set as $\epsilon \to 0$.

3.2. *More on* $\langle\gamma^2\rangle_t = 0$

To give a formulation closer to the problems considered in § 3, we now turn to Case IV(a), the addition, that is, of $\langle\gamma^2\rangle_t = 0$, into Howard's original problem, except that here we impose no-slip conditions. In § 1 it was shown that heat flux of the no-slip solution must lie between the curves N_b of figure 1 and N_Q in figure 3. Interestingly, the variational problem now leads to the emergence of non-smooth solutions.

To recapitulate, the problem is to maximize $\langle\theta w\rangle + 1$ subject to

$$\langle\gamma^2\rangle - R\langle\theta w\rangle = 0, \tag{3.11}$$

$$\langle\theta\nabla^2\theta\rangle - \langle\overline{\theta w}T'\rangle = 0, \tag{3.12}$$

$$\langle\gamma\nabla^2\gamma\rangle - R\langle\gamma\theta_x\rangle = 0, \tag{3.13}$$

where $T' \equiv \overline{\theta w} - \langle\theta w\rangle - 1$, $\gamma \equiv u_z - w_x$, and $u_x + w_z = 0$. In this formulation the independent fields become the streamfunction $\psi(x,z)$, where $u = \psi_z$ and $w = -\psi_x$, and the temperature fluctuations $\theta(x,z)$.

Rather than deriving the differential Euler–Lagrange equations, the approach taken here is to consider ψ and θ of the form

$$\psi(x,y) = \phi(x)\left[z(1-z)\sum_{k=0}^{K} a_k T_k(2z-1)\right],$$
$$\theta(x,y) = \phi_x(x)\left[z(1-z)\sum_{k=0}^{K} b_k T_k(2z-1)\right]$$

where $\phi_{xx} = -\alpha^2\phi$, $\overline{\phi_x^2} = 1$, and the $\{T_k\}$ are Chebyshev polynomials. Note that the $w = 0$ and $\theta = 0$ boundary conditions are enforced by the pre-factor of $z(z-1)$. Other boundary conditions, for example $\psi_z(0,x) = 0$ (or $\psi_{zz}(0,x) = 0$ for slip), are included as additional constraints – on the same footing as the three integral constraints. The 'equivalent' algebraic constrained optimization problem for the set $\{a_k, b_k, \alpha\}$ is then easily addressed using Lagrange multipliers and Newton's method. While the optimal solution at larger R is likely to assume a multi-α form, for this exploration we consider only the single-α solution.

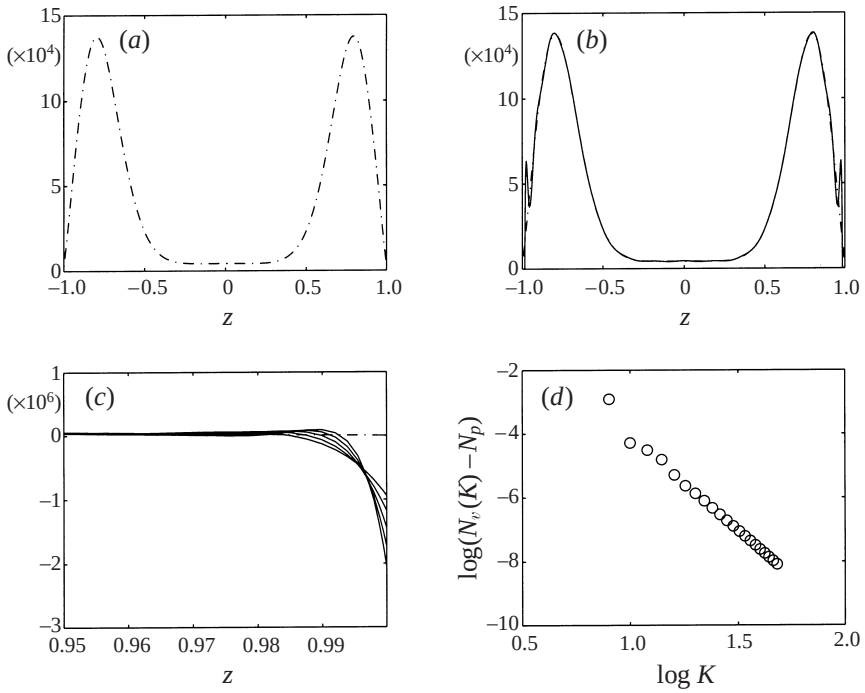

FIGURE 4. Defining $\gamma = \hat{\gamma}(z)\phi(x)$ and $L(\hat{\gamma}) = (\partial_z^2 - \alpha^2)\hat{\gamma}$, these panels illustrate: (*a*) $L(\hat{\gamma})$ without the enstrophy constraint (3.13) applied. (*b*) $L(\hat{\gamma})$ with the enstrophy constraint (3.13) applied at a resolution of $K = 34$ (solid) along with Howard's original (dashed). (*c*) A close-up near the boundary of $L(\hat{\gamma})$ for the problem including the enstrophy constraint. Different solid curves correspond to different values of K, the degree of the trial space. Note that these seem to approach a non-smooth function as K increases. (*d*) As K increases we see that the Nusselt number of this problem approaches that of the problem not constrained by the enstrophy integral. (Note, all computations have $R = 10^4$.)

The numerical solution is found to depend on the rank K of the trial space in such a way that solutions of limited differentiability are expected in the limit of large K. Numerically we find that relevant integral properties of these fields, like N, approach the values obtained by the (smooth) optimal solutions obtained in the absence of the enstrophy constraint (3.13), as illustrated in figure 4. The discontinuous higher derivatives near the boundaries allow the enstrophy integral to be satisfied while essentially preserving all moments composed of lower-order derivatives. This behaviour can be traced back to the boundary term

$$\overline{\gamma\gamma_z}|_{z=0}^{z=1}$$

hidden in (3.13). As this term is not known *a priori*, it provides the requisite freedom to make adjustments to higher derivatives near or at the boundaries while not significantly altering the balance of the power integrals. While perhaps a physically unsatisfying means of satisfying the added constraint, from a mathematical point of view the operative point is that the bound does remain unchanged.

Subject to a slip boundary condition, Cases IV and VII, which suffer a breakdown in the perturbation analysis at finite σ, do not have boundary terms. They do, however, have one or more z-constraints, which we suggest can act in a similar way. Numerical

investigation should furnish the needed evidence of this. We hope to report such results in the near future.

4. Conclusions

As Spiegel (1971) noted, steady solutions of the governing equations, regardless of stability, necessarily satisfy any and all statistically steady integral moments of the equations. Consequently, steady solutions are always admissible when seeking extrema and moreover provide a hard limit for any upper bound, no matter how refined. In the case of convection, the oft-cited pair of papers, Roberts (1977, 1979), would seem to limit the prospects for upper bounds as his asymptotic prediction for the $Nu(R)$ relation of the steady roll solution of maximum heat flux at infinite Prandtl number still lies disappointingly far above representative data taken from experiment and numerical simulation (it is clear that the gap cannot be explained away as a consequence of the effects of finite Prandtl number). But in a subsequent, unjustly neglected, paper Jimenez & Zufiria (1987) present a subtle and elegant chain of reasoning that corrects a serious flaw in Roberts' boundary-layer analysis. Where Roberts predicts a heat flux of $c(\alpha)R^{1/3}$ peaking at $c = 0.321$ for a cell aspect ratio α of 0.75, Jimenez & Zufiria (1987) derive a corrected $c(\alpha)$, which crests at 0.20. While the inference of a scaling law of $R^{0.315}$ is simply inappropriate, Hansen, Yuen & Malevsky (1992) is an otherwise sound vindication of Jimenez & Zufiria (1987). The Hansen *et al.* data for their numerically determined two-dimensional steady-state solutions, when plotted as a function of $Nu/R^{1/3}$ up to $R = 10^9$, are clearly consistent with Jimenez & Zufiria (1987) for $\alpha = 0.55$. Our own recent spectral computations, also carried to $R = 10^9$ but optimized over α, further confirm the Jimenez & Zufiria (1987) result, not those of Roberts. Moreover, the exceedingly tight coincidence of the heat flux computed in their time-dependent (also two-dimensional) solutions with the heat flux of the corresponding steady (unstable) solution leaves open that a sufficiently constrained upper bound result may come very close to observation.

To give the reader preliminary evidence that the potential for improvement in the upper bound actually can be realized, we re-examine Case IV(a) but now addressing the solution of the associated Euler–Lagrange equations (3.1)–(3.2) by direct numerical means, not perturbative. This is an especially elementary extension of the Busse–Howard upper-bound problem and lends itself to ready computation using the numerical algorithm previously discussed in § 3.2. For the purpose of comparison, we show the large-R solution of Case I (recomputed for this figure, and in agreement with both the previous computation by Straus 1976*a* and the more recent direct attack on the multi-α problem appearing in Vitanov & Busse 1997). Figure 5(*a*) shows that the perturbative reduction in Nu of table 3, if continued to large R, yields not only a quantitative, but a qualitative change in the bound. Panel (*b*) makes this explicit with a comparison of the numerically inferred exponents of the two curves. (The upper curve tends to a value somewhat greater that 0.4, the line at $2/5$ is simply an aid to the eye.) The reduction to $1/3$ is tantalizing, though it must be borne in mind that this single-α result may give way to a multi-α solution at intermediate values of R.

This surprising, indeed gratifyingly, large influence of a single new constraint on enstrophy is good evidence that the upper-bound problem is ripe for renewed attack. Indeed, we have proposed a broader line of attack, with a new class of so-called 'z-constraints'. Based on symmetries, the assumption of statistical steadiness, and kinematic and dynamic identities, not all such z-constraints are effective; many are vacuous. We have presented arguments in favour of specific forms. These affect the

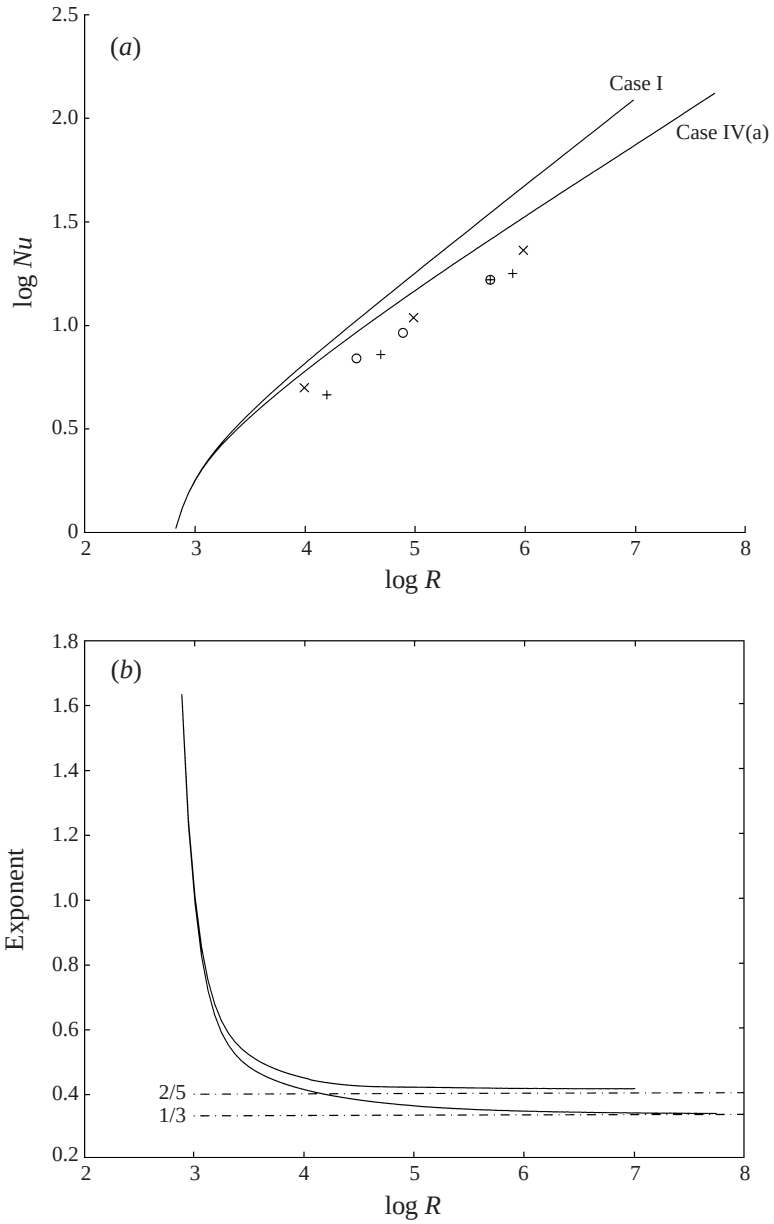

FIGURE 5. The influence of $\langle \gamma^2 \rangle_t = 0$ in reducing the bound on Nu. Symbols in (a) are experimental results.

symmetry of the upper-bound solution, in particular precluding the appearance of 'single-α' solutions. While perturbation theory has pointed out both potential peril and potential promise, it remains to explore the numerical consequences of this exciting new field.

Our inspiration for this exploration is the forty-five year ongoing odyssey of Willem Malkus in search of deductive and quantitatively useful results on turbulence. Over the course now of many years, he has patiently and generously shared his time and

his enthusiasm in conversations with each of us. Without this stimulus, the present work would never have come to fruition. Willem has spoken often of establishing a 'language of inquiry'. To such an end we hope that with this work we have perhaps added a few new words.

R. A. W. wishes to thank the NSF for support he received through grant ATM92-08373, while a graduate student at MIT. (Indeed, the first two sections of this paper are a minor revision of material completed by R. A. W. in his thesis.)

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. 1978 The optimum theory of turbulence. *Adv. Appl. Mech.* **18**, 77.

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

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

Chandrasekhar, S. 1961 *Hydrodynamic and Hydromagnetic Stability*. Oxford University Press.

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

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

Hansen, U., Yuen, D. A. & Malevsky, A. V. 1992 Comparison of steady-state and strongly chaotic thermal-convection at high Rayleigh number. *Phys. Rev.* A **46**, 4742–4754.

Herring, J. R. 1963 Investigation of problems in thermal convection. *J. Atmos. Sci.* **20**, 325–338.

Herring, J. R. 1964 Investigation of problems in thermal convection: rigid boundaries. *J. Atmos. Sci.* **21**, 277–290.

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

Howard, L. N. 1990 Limits on the transport of heat and momentum by turbulent convection with large-scale flow. *Stud. Appl. Maths* **83**, 273–285.

Jimenez, J. & Zufiria, J. A. 1987 A boundary-layer analysis of rayleigh-benard convection at large Rayleigh number. *J. Fluid Mech.* **178**, 53–71.

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

Kerswell, R. R. & Soward, A. M. 1996 Upper bounds for turbulent Couette flow incorporating the poloidal power constraint. *J. Fluid Mech.* **328**, 161–176.

Kraichnan, R. H. 1962 Turbulent thermal convection at arbitrary Prandtl number. *Phys. Fluids* **5**, 1374–1389.

Krishnamurti, R. & Howard, L. N. 1981 Large-scale flow generation in turbulent convection. *Proc. Natl Acad. Sci.* **4**, 1981–1985.

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

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

Malkus, W. V. R. 1960 Lectures on turbulence. *Woods Hole Oceanographic Institution Technical Report:* WHOI–60–46, vol. 2, pp. 1–68.

Malkus, W. V. R. & Smith, L. M. 1989 Upper bounds on functions of the dissipation rate in turbulent shear flow. *J. Fluid Mech.* **208**, 479–507.

Malkus, W. V. R. & Veronis, G. 1958 Finite amplitude cellular convection. *J. Fluid Mech.* **4**, 225–260.

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.

Niemela, J. J., Skrbek, L., Sreenivasan, K. S. & Donnelly, R. J. 2000 Turbulent convection at very high Rayleigh numbers. *Nature* **404**, 837–840.

Roberts, G. O. 1977 Fast viscous convection. *Geophys. Astrophys. Fluid Dyn.* **8**, 197–233.

Roberts, G. O. 1979 Fast viscous Bénard convection. *Geophys. Astrophys. Fluid Dyn.* **12**, 235–272.

SPIEGEL, E. A. 1971 Convection in stars I. basic Boussinesq convection. *Ann. Rev. Astron. Astrophys.* **9**, 323–352.

STRAUS, J. M. 1976*a* A note on the multi-α solutions of the upper bounding problem for thermal convection. *Dyn. Atmos. Oceans* **1**, 71–76.

STRAUS, J. M. 1976*b* On the upper bounding approach to thermal convection at moderate Rayleigh numbers II. rigid boundaries. *Dyn. Atmos. Oceans* **1**, 77–90.

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.

WORTHING, R. A. 1995 Contributions to the variational theory of convection. PhD thesis, Massachusetts Institute of Technology.

WORTHING, R. 2001 Bounds on convective heat transport in containers. *J. Fluid Mech.* **431**, 427–432.