THE ASTROPHYSICAL JOURNAL, 267:795–821, 1983 April 15

MULTISCALE MODEL EQUATIONS FOR TURBULENT CONVECTION AND CONVECTIVE OVERSHOOT[1]

PHILIP S. MARCUS
Department of Mathematics, Massachusetts Institute of Technology
WILLIAM H. PRESS
Harvard-Smithsonian Center for Astrophysics
AND
SAUL A. TEUKOLSKY
Physics Department, Cornell University; and Harvard-Smithsonian Center for Astrophysics
Received 1982 June 1; accepted 1982 September 3

ABSTRACT

The Navier-Stokes equations for a compressible fluid with radiative diffusivity and viscosity are analyzed in the anelastic limit (Mach number $\ll 1$), using window-function localized Fourier transforms. This technique allows the spectral dynamics of a turbulent eddy cascade to be studied in wavenumber space, while also modeling the spatial variation of the velocities and other mean quantities such as temperature, density, and opacity. Physical arguments which give the expected relative phase of various quantities are used instead of (e.g.) closure assumptions about higher moments. A set of time-dependent equations is obtained. These, in principle, model the multiscale and nonlocal effects of convective overshoot, location of the convective interface, time-dependent convection, and turbulent mixing.

Subject headings: convection — stars: interiors — Sun: interior — turbulence

I. INTRODUCTION

a) *Orientation*

Most calculations of stellar or solar evolution take the boundaries of a convection zone to be where the radiative temperature gradient changes from super to subadiabatic (or vice versa) and take the interior thermal structure of a convection zone to be that given by standard mixing-length theory (Vitense 1953; Bohm-Vitense 1958; Henyey, Vardya, and Bodenheimer 1965; Cox and Giuli 1968). These two prescriptions are matters of convenience only, since it is generally appreciated that, once initiated, the fluid motions of convection are capable of modifying their own boundary ("nonlocality" or "convective overshoot") and that the mixing-length temperature gradient is probably accurate only when the convection is highly *efficient* (so that the gradient is close to adiabatic), or highly *inefficient* (so that the gradient is close to radiative). Even for these simplifying regimes, no sensible dynamicist takes the mixing-length predictions of convective velocity as accurate to better than a factor of 2, say.

The shortcomings of mixing-length theory arise because it is not a full physical theory following from the Navier-Stokes equations, but, in essence, little more than a dimensional argument based on the assumptions that the fluid flow at a point depends only on its local thermal state ("locality") and that there is a single relevant length scale in the problem, a scale which can be identified indifferently as the pressure scale height, mixing distance, size of the largest eddy, or turbulent transport length. This "single-scale" assumption creates two distinct sorts of deficiencies: first, it washes out all *geometric details* of the largest scale structure. Planforms and radial eigenfunctions are obviated; spatial gradient operators are replaced by mere scalar numbers (inverse powers of the mixing length); boundary conditions and boundary layer flow anisotropies are disregarded.

Second, the single-scale assumption throws away any physical phenomena which arise from the *multi-length scale* nature of the underlying system: the Kolmogorov (1941*a*, *b*) turbulent cascade; dissipation of the cascade at the smallest scales; transport of mechanical (not thermal) energy by convective eddies. We will see below (§ IV) that these effects are of the same order as effects modeled in mixing length theory, a fact not always appreciated in previous

[1]Supported in part by the National Science Foundation, grants AST-8210933, PHY-80-08351, and PHY-80-07957, and in part by subcontract 9744809 with the University of California Lawrence Livermore National Laboratory. Supporting agencies not responsible for content or conclusions.

work. Further, we will argue that multiscale effects, as well as the nonlocality of eddies, can be the *determining* physical processes for the position of the convective boundary or extent of convective overshoot.

The purpose of this paper, then, is to construct a set of equations, deducible from the Navier-Stokes equations by formal manipulation, with no terms thrown out, which describe turbulent convection without the single-scale approximation, and one thus capable, in principle, of describing systems in which the effects of convective overshoot are important. In other words, we propose to go as far as we can in eliminating the physical uncertainties (the second set of deficiencies above) associated with mixing-length theory. We will not, however, do anything about the first (geometric) set of deficiencies; we consider it more important at this stage to have all relevant physical processes included, than to have all geometrical factors precisely correct. While this paper limits itself to the derivation and explication of the multiscale equations, subsequent papers in this series will apply the equations to both ideal and more realistic systems and explore the phenomenology of convective overshoot in some detail.

b) *Previous Work*

Various attempts to estimate the magnitude of the convective overshoot, and to express it in a form suitable for stellar evolutionary calculations, are found in the literature. Two rather distinct methodologies have emerged, which we can call the "modal" and "model" points of view.

In the modal case, one thinks about modes. One attempts to take the actual fluid equations (or at least an idealized set of equations which closely resembles the fluid equations), and then to solve for actual eigenfunctions or nonlinear solutions of the fluid flow. Pioneered by Veronis (1963) and Spiegel (1965), this approach also includes the more recent work by Gough *et al.* (1976), Latour *et al.* (1976), Toomre *et al.* (1976), Marcus (1979, 1980*a*, *b*), and others. Except in the most recent work it has usually been necessary to make the simplifying "single mode approximation" to find solutions in the nonlinear (and therefore realistic) regime. In this approximation the planform (or horizontal variation of the convection) is fixed *a priori*, and nonlinear ordinary differential equations are integrated to find the "lowest eigenvalue" radial dependence. One can summarize the results of this work by saying that the single-mode approach finds large convective overshoots, with mixing fluid flows extending of order a pressure scale height Λ_p or more into the radiatively stable region.

Unfortunately, there is reason to doubt the reliability of the single-mode methodology. A fluid constrained to single-mode motion can be viewed as having a high effective viscosity in some crucial respects: the Reynolds number $\mathcal{R}e$ of an unconstrained fluid flow measures the ratio of Reynolds stresses (dynamic stresses induced by the velocity) to viscous stresses. Single-mode theory explicitly removes those Reynolds stresses which would couple to other modes, and ultimately to a turbulent flow, from the modal equations. Therefore, in this sense, the effective $\mathcal{R}e$ of single-mode theory is small, corresponding to a large effective viscosity. Even though this effective viscosity does not appear as a dissipation term in the equations, it does constrain the solutions to be "stiff" on large length scales. The normal tendency of the fluid to shear off into a sharp turbulent boundary layer at the boundary of the radiatively stable region is not allowed by the single-mode equations. It is therefore not surprising that the eigensolutions should penetrate deeply into the stable region, but this result must be viewed cautiously, and may well be artifactual.

The "model" approach stays closer to the mixing-length theory in its spirit and constructs a physical picture based on the idea of rising (or falling) "bubbles" or "eddies" within the fluid. Model equations are not necessarily derived from the Navier-Stokes equations ab initio, but from energy and momentum arguments as applied to the bubbles. Although this approach is mathematically less precise than the modal viewpoint, it has the advantage that it can try to include the physical effects which are lost to the single-mode theory. In this category, one might include the work of Weyman and Sears (1965), Saslaw and Schwarzschild (1965), Shaviv and Salpeter (1973), Straus, Blake, and Schramm (1976), Ulrich (1976), and others. Generally, these treatments have attacked the "locality" assumption of standard mixing-length theory by recognizing the finite size of the largest scale eddies. However, none have faced up to the limitations of the single-scale assumption head on: the nonlocality length is scaled to approximately a pressure scale height along with every other length.

The results of Shaviv and Salpeter (1973) are now taken to exemplify the sort of results that come out of a model analysis. Previous to their work, Roxburgh (1965) and Saslaw and Schwarzschild (1965) had argued that the extent of overshoot was extremely small for realistic stellar conditions (as small as 10^{-2} pressure scale heights) due to the very high efficiency of the convection. High efficiency implies a very small superadiabaticity and a small convective velocity. These small velocities seemed to imply only small penetrations, going to zero, in fact, as the superadiabaticity goes to zero.

Shaviv and Salpeter showed that the penetration distance remains *finite*, of order one-tenth of a pressure scale height in their model, even as the superadiabaticity goes to zero. This result seems paradoxial at first acquaintance, but can be understood (as Shaviv and Salpeter point out) by considering the convective interface as a heat engine which transfers mechanical energy from the unstable region into the stable region, where it does work to keep the stable fluid mixed. In the limit of more efficient convection, a smaller mechanical flux is available, but the work required to keep the overshoot region mixed also becomes less in the same proportion, since the thermal diffusion time to restore a stable gradient also scales with convective efficiency.

A completely different approach to computing overshoots has been taken recently by Cloutman and Whitaker (1980). Their method is based on an engineering model of incompressible turbulent convection which includes, with some empirical parametrization, effects of turbulently diffusing both kinetic energy and entropy. We consider this to be a conceptual step forward. Unfortunately, the choice of free parameters becomes somewhat problematical in the compressible case of astrophysical interest, and the engineering model used cannot be directly related back to the Navier-Stokes equations for guidance on this choice. We consider the present paper to be, in some senses, a more formal attempt at treating convective overshoot from the Cloutman-Whitaker point of view, where mechanical (momentum transport) effects are included along with thermal ones (heat transport).

c) *Turbulent Erosion*

Most of this paper is formal manipulation of the fluid equations. The goal of that manipulation will be clearer, however, if we here go through a rough phenomenological example. Suppose that we have a convective interface predicted by mixing-length theory. Figure 1 shows, schematically, the variation of $\log T$ with $\log P$ in the fluid. The curve AOA′ (shown as a straight line) is an adiabat. The curve ROR′ is the radiative gradient needed to carry the total flux F. The radiative curve is subadiabatic to the left of O, superadiabatic to the right. In local mixing-length theory, with efficient convection, the actual gradient would be ROA′, with stability along RO, convection along OA′.

The mechanical necessity of convective overshoot or penetration to the left of O arises because the curves ROR′ and AOA′ are tangent at O. This implies that it requires arbitrarily small mechanical energy to mix a fluid element immediately to the left of O. Equivalently, the Brunt-Väisällä (or buoyancy) frequency goes to zero as O is approached from the left, so any finite fluid shear present (as O is approached from the right) generates at the interface a Richardson number which is much less than unity, hence unstable to fluid overturn (Miles and Howard 1964; for a general review, see Turner 1973, ch. 4).

Convective overshoot, then, will mix the fluid to the left of O, restoring it very nearly to the adiabatic curve. In the figure, this overshoot is shown as extending from O to O′. To the left of O′, we must return to a radiative gradient sufficient to carry the full flux. The variation in T must also be continuous. We therefore match to a curve R″O′ which is parallel to the curve RO (so that the two curves carry identical radiative fluxes). The actual gradient with overshoot is then R″O′OA′. The key point now is that there is inevitably a discontinuity of slope in R″O′OA at the point O′. This means that the Brunt-Väisällä frequency N jumps from (essentially) zero to the right of O′ to some finite value to the left of O′. This discontinuity in N gives the interface a finite stability against continued erosion by the turbulent overshoot. The criterion for stability is that the convective, turbulent shear to the right of the interface O′ (dimensions $[t^{-1}]$) be no larger than the value of N to the left of O′.

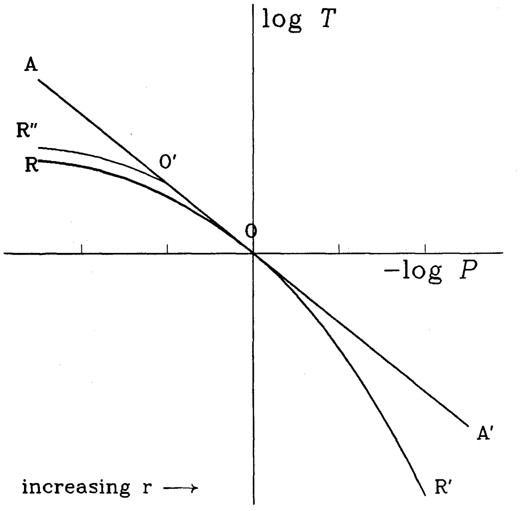

FIG. 1.—Variation of temperature with pressure near the base of a convection zone (schematic). The curve AOA′ represents the adiabatic variation, while ROR′ integrates the radiative gradient. According to a local convection theory, the actual variation (for efficient convection) would be ROA′; convective overshoot to some point O′ should instead yield R″O′OA′. See text for details.

Let us now put in numbers roughly appropriate for the base of the solar convection zone. The large-scale overturn frequency, which is about the same as the shear, is $\sim 10^{-6}\ \mathrm{s}^{-1}$, while the Brunt-Väisällä frequency in the solar radiative core is typically $\sim 10^{-3}\ \mathrm{s}^{-1}$ (see, e.g., Press 1981, Fig. 2). These numbers imply stability against erosion. There is, however, a turbulent Kolmogorov cascade coupled to the large-scale convective motion. In such a cascade, shear varies inversely with eddy size l as $l^{-2/3}$. Hence, there is a small scale $l \approx 10^{10}\ \mathrm{cm} \times (10^{-6}\ \mathrm{s}^{-1}/10^{-3}\ \mathrm{s}^{-1})^{3/2} \approx 3$ km whose eddies have sufficient shear to *destabilize* any part of the whole solar interior. Does the Kolmogorov cascade extend to this small a scale? The answer is found to be yes by computing the larger of the thermal and viscous microscales, $L/\mathscr{P}_e^{3/4} \approx 10^{10}\ \mathrm{cm}/(2\times 10^6)^{3/4} \approx 2$ km, and $L/\mathscr{R}_e^{3/4} \approx 10^{10}\ \mathrm{cm}/(5\times 10^{13})^{3/4} \approx 0.5$ cm, where $\mathscr{P}_e$ and $\mathscr{R}_e$ are, respectively, the Péclét and Reynolds numbers of the large-scale convective flow, and L is the pressure scale height or size of largest eddy (see, e.g., Landau and Lifshitz 1959, § 32, for a discussion of the viscous microscale; the thermal microscale is defined analogously with viscosity replaced by thermal diffusivity).

The conclusion is that small mechanical scales cannot be disregarded: they have sufficient shear to extend the solar convection zone to essentially arbitrary depth. The question is to what depth below the unstable region the fluid equations transfer sufficient mechanical energy into these small scales. To answer this and related questions, we need a convective model which allows us to compute, as a function of depth, not only mean thermal quantities of interest such as $\langle P \rangle$, $\langle \rho \rangle$, $\langle T \rangle$, and $\langle v^2 \rangle$, but also the detailed distribution of kinetic energy among different length scales in the cascade and the effect of such a distribution on a neighboring stable region. That is the task to which we now turn.

In the next section we introduce the equations and boundary conditions that determine the convective velocity, temperature, pressure, and density in a stellar convective zone. We review the anelastic (small Mach number) approximation which simplifies the dynamics and allows each term in the equations to be ordered by its power of the Mach number. We introduce the horizontally averaged pressure as the vertical coordinate. In § III we confront the problem that although the equations for stellar convection are well posed, their intrinsic nonlinearity prevents us from obtaining exact solutions. We abandon hope of finding the actual velocity and instead opt for finding the amplitudes of the convective eddies as a function of each eddy's vertical position and size. The term "eddy" is defined mathematically by introducing a local window Fourier transform. In § IV we examine the standard equations of mixing-length theory in the light of our definition of an eddy and of the actual equations of convective motion. We discuss the pitfalls and successes of the mixing-length theory and begin the transition to the eddy formalism. In the fifth and sixth sections we transform the actual equations of motion into a set of amplitude equations that govern the time dependence of the eddies and of the temperature, pressure, and density. We show the physical significance of each term in the resulting amplitude equations. The final section is a summary of the eddy-amplitude equations.

II. EQUATIONS AND BOUNDARY CONDITIONS

a) *Governing Equations*

The equations that govern convection are the equations for conservation of mass, momentum, and energy, and the equation of state. The continuity equation (conservation of mass) is

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \boldsymbol{v}). \tag{2.1}$$

The Navier-Stokes equation (conservation of momentum) for a flow in a constant gravitational field, g, (plane-parallel approximation, $g > 0$ for a downward force) is

$$\frac{\partial v_i}{\partial t} = -(\boldsymbol{v} \cdot \nabla) v_i - \frac{1}{\rho}\frac{\partial P}{\partial x_i} - g\delta_{iz} + \frac{1}{\rho}\frac{\partial}{\partial x_j}\Pi_{ij}, \tag{2.2}$$

where Π_{ij} is the viscous stress tensor, given by

$$\Pi_{ij} = \eta\sigma_{ij} + \xi(\nabla \cdot \boldsymbol{v})\delta_{ij}. \tag{2.3}$$

Here

$$\sigma_{ij} = \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} - \frac{2}{3}(\nabla \cdot \boldsymbol{v})\delta_{ij} \tag{2.4}$$

is the shear tensor and η and ξ are the first and second coefficients of viscosity. The kinematic viscosity, ν, is related to the dynamic viscosity by $\nu = \eta/\rho$. Instead of the energy conservation equation, one can use the equation for entropy generation,

$$\frac{\partial}{\partial t}(\rho s) = -\nabla \cdot \boldsymbol{F}_s + S_\nu + S_\sigma, \tag{2.5}$$

where the entropy flux is

$$\boldsymbol{F}_s = \rho s \boldsymbol{v} - c_p \frac{\rho}{T} \sigma \nabla T, \tag{2.6}$$

and the viscous and radiative entropy source terms are

$$S_\nu = \frac{\eta}{2T} \sigma_{ij} \sigma_{ij} + \frac{\xi}{T} (\nabla \cdot \boldsymbol{v})^2, \tag{2.7}$$

$$S_\sigma = \frac{c_p \rho \sigma}{T^2} (\nabla T \cdot \nabla T). \tag{2.8}$$

The entropy s and the specific heat at constant pressure c_p are per unit mass of fluid. The thermal diffusivity σ for an optically thick medium in which the thermal transport is due to radiative diffusion is

$$\sigma = \frac{4acT^3}{3\rho^2 \kappa c_p}, \tag{2.9}$$

where a is the radiation constant, κ is the opacity, and c is the speed of light.

We take the ideal gas equation as our equation of state,

$$P = R\rho T/\mu, \tag{2.10}$$

where μ is the mean molecular weight and R is the gas constant. We take c_v, the specific heat per unit mass at constant volume, to be constant. Then c_p is constant, and all adiabatic indices are constant and equal to $\gamma = c_p/c_v$. The internal energy per unit mass is $c_v T$, while the enthalpy per unit mass is $c_p T$. The first law of thermodynamics takes the form

$$c_v\, dT = T\, ds + (P/\rho^2)\, d\rho. \tag{2.11}$$

Equations (2.1), (2.2), (2.5), and (2.11) can be combined to give the energy conservation equation,

$$\frac{\partial}{\partial t}\left(\frac{1}{2}\rho v^2 + c_v \rho T + \rho \phi\right) = -\frac{\partial}{\partial x_i}\left(\frac{1}{2}\rho v^2 v_i - \Pi_{ij} v_j + c_p \rho T v_i - c_p \rho \sigma \frac{\partial T}{\partial x_i} + \rho \phi v_i\right), \tag{2.12}$$

where $\phi = gz$ is the gravitational potential.

A more useful quantity than entropy for our purposes is the potential temperature, θ, defined as

$$\theta \equiv T^{1/\gamma} \rho^{(1-\gamma)/\gamma} (\rho_0 T_0)^{(\gamma-1)/\gamma} \propto \exp(s/c_p). \tag{2.13}$$

The potential temperature of a piece of fluid in a star is just the temperature it would have if it were adiabatically transported from its true position in the star to some arbitrary reference position at pressure $P_0 = R\rho_0 T_0/\mu$. Thus, if the fluid were isentropic, the potential temperature would be constant and equal to T_0. Analogously, we define the potential density to be

$$\Gamma = T^{-1/\gamma} \rho^{(\gamma-1)/\gamma} (\rho_0 T_0)^{1/\gamma} \propto \exp(-s/c_p). \tag{2.14}$$

b) *Anelastic Approximation*

The equations of motion (2.1), (2.2), (2.5) can be simplified by assuming that the velocity is much less than the speed of sound. The equations of motion can be expanded as a series in even powers of the Mach number M. Retaining only the leading order terms, one obtains the anelastic approximation (see Ogura and Charney 1962; Gough 1969; Marcus 1978). The anelastic approximation is most easily implemented by writing each of the variables, Q, as a sum of its horizontally averaged mean part $\langle Q\rangle$ and its fluctuating part Q^F:

$$Q=\langle Q\rangle+Q^F, \tag{2.15}$$

$$\langle Q\rangle=\int Q\,dx\,dy \Big/ \int dx\,dy. \tag{2.16}$$

The following identities will prove useful:

$$\langle AB\rangle=\langle A\rangle\langle B\rangle+\langle A^FB^F\rangle, \tag{2.17}$$

$$(AB)^F=A^FB^F-\langle A^FB^F\rangle+A^F\langle B\rangle+\langle A\rangle B^F, \tag{2.18}$$

$$\langle 1/A\rangle=1/\langle A\rangle+O\left[(A^F)^2\right], \tag{2.19}$$

$$(1/A)^F=-A^F/\langle A\rangle^2+O\left[(A^F)^2\right]. \tag{2.20}$$

In the anelastic approximation the fluctuating part of all of the thermodynamic variables is first order in the expansion parameter Mach-squared:

$$\frac{P^F}{\langle P\rangle}\sim\frac{T^F}{\langle T\rangle}\sim\frac{\rho^F}{\langle\rho\rangle}\sim\frac{\theta^F}{\langle\theta\rangle}\sim\frac{\Gamma^F}{\langle\Gamma\rangle}\sim\frac{s^F}{c_p}=O(M^2). \tag{2.21}$$

Mean horizontal drifts generated by the fluctuating velocity are smaller by an additional Mach-squared:

$$\langle v_x\rangle\sim\langle v_y\rangle\lesssim M^2|\boldsymbol{v}^F|. \tag{2.22}$$

Since vertical motion, however, can be generated by readjustment to the hydrostatic equilibrium of a perturbed thermal state, such motion is first order in the perturbation:

$$\langle v_z\rangle\lesssim|\boldsymbol{v}^F|. \tag{2.23}$$

Only when the flow is steady-state does $\langle v_z\rangle$ also satisfy a relation of the form (2.22), as we will see below.

The mean and fluctuating parts of the equation of state are

$$\langle P\rangle=\frac{R}{\mu}\langle\rho\rangle\langle T\rangle, \tag{2.24}$$

and

$$\frac{P^F}{\langle P\rangle}=\frac{\rho^F}{\langle\rho\rangle}+\frac{T^F}{\langle T\rangle}. \tag{2.25}$$

The fluctuating part of the continuity equation is replaced by the condition that the fluctuating mass flux be divergence-free:

$$(\nabla\cdot\rho\boldsymbol{v})^F\approx\nabla\cdot(\langle\rho\rangle\boldsymbol{v}^F)=0. \tag{2.26}$$

The right-hand side of equation (2.26) is zero, rather than a time derivative, because (in anelastic approximation) $\partial\rho^F/\partial t$ can be shown to be one higher order in M^2.

The anelastic form of the horizontally averaged Navier-Stokes equation is just the hydrostatic pressure equation

$$\frac{\partial\langle P\rangle}{\partial z}=-g\langle\rho\rangle. \tag{2.27}$$

The horizontally averaged continuity equation determines the growth of the mean vertical velocity, $\langle v_z \rangle$. It is convenient to split $\langle v_z \rangle$ into two pieces, representing steady (s) and time-dependent (t) contributions,

$$\langle v_z \rangle = \langle v_z \rangle_s + \langle v_z \rangle_t , \tag{2.28}$$

where $\langle v_z \rangle_s$ is defined by

$$\langle v_z \rangle_s \equiv -\langle \rho^F v_z^F \rangle / \langle \rho \rangle = O(M^2 v^F) . \tag{2.29}$$

A first integral of the mean continuity equation then determines $\langle v_z \rangle_t$,

$$\langle v_z \rangle_t = \frac{\partial \langle P \rangle}{\partial t} \frac{1}{g \langle \rho \rangle} . \tag{2.30}$$

If the mean pressure is independent of time, $\langle v_z \rangle_t = 0$ and $\langle v_z \rangle = \langle v_z \rangle_s$. This shows that in a steady state $\langle v_z \rangle \approx M^2 v^F$, as we claimed after equation (2.23). In a non–steady-state fluid where the mean pressure is changing, $\langle v_z \rangle_t$ is not zero, and in the next subsection we show that $\langle v_z \rangle_t$ can be used to convert from an Eulerian coordinate system to a Lagrangian system.

c) Pressure Coordinates

It is convenient to work in a pseudo-Lagrangian coordinate system where the vertical coordinate is replaced by the mean pressure, $\langle P \rangle$. The Lagrangian time derivative $[\partial/\partial t]_{\mathrm{L}}$ is related to the Eulerian time derivative by

$$\left[\frac{\partial}{\partial t}\right]_{\mathrm{L}} = \frac{\partial}{\partial t} - \frac{\partial \langle P \rangle}{\partial t} \frac{\partial}{\partial \langle P \rangle} \tag{2.31}$$

$$= \frac{\partial}{\partial t} + \langle v_z \rangle_t \frac{\partial}{\partial z} . \tag{2.32}$$

Equation (2.32) shows that $\langle v_z \rangle_t$ is the speed at which the Lagrangian coordinate system moves past the Eulerian coordinates. The velocity that the pseudo-Lagrangian observer sees is $\boldsymbol{v}^{\mathrm{L}}$, where

$$v_i^{\mathrm{L}} \equiv v_i - \delta_{iz} \langle v_z \rangle_t . \tag{2.33}$$

Note that $(\boldsymbol{v}^{\mathrm{L}})^F = \boldsymbol{v}^F$. The mean continuity equation in Lagrangian coordinates is

$$\left[\frac{\partial \ln \langle \rho \rangle}{\partial t}\right]_{\mathrm{L}} = -\frac{\partial \langle v_z \rangle_t}{\partial Z} . \tag{2.34}$$

The fluctuating part of the anelastic Navier-Stokes equation in pressure coordinates is

$$\left[\frac{\partial v_i^F}{\partial t}\right]_{\mathrm{L}} = -\left[(\boldsymbol{v}^{\mathrm{L}} \cdot \nabla) v_i^{\mathrm{L}}\right]^F + v_z^F \left[\frac{\partial \ln \langle \rho \rangle}{\partial t}\right]_{\mathrm{L}} \delta_{iz} - \frac{1}{\langle \rho \rangle} \frac{\partial P^F}{\partial x_i} - g \frac{\rho^F}{\langle \rho \rangle} \delta_{iz} + \frac{1}{\langle \rho \rangle} \frac{\partial}{\partial x_j} \Pi_{ij}^F , \tag{2.35}$$

where we have used equation (2.34). We must make this replacement for $\langle v_z \rangle_t$ (involving, as it does, the unwelcome time derivative on the right-hand side of the equation) because we cannot directly compute $\langle v_z \rangle_t$ from the *mean* Navier-Stokes equation; the mean Navier-Stokes equation is just the hydrostatic pressure equation.

d) Boundary Conditions

We take our convecting fluid to extend over an infinite horizontal domain and to lie within a vertical region bounded above by a mean pressure $\langle P \rangle_1$ and below by $\langle P \rangle_2$. This Lagrangian box of fluid can rise or sink in our star, and the distance between the upper and lower boundaries can likewise change in time. The mass of fluid in the box is constant, since

$$\int_{z_2}^{z_1} \rho \, dx \, dy \, dz = \frac{1}{g} \left(\langle P \rangle_2 - \langle P \rangle_1 \right) \int dx \, dy . \tag{2.36}$$

Our box is big enough so that the convective region is wholly contained in the box and the velocity vanishes at the boundaries:

$$\boldsymbol{v}(\langle P\rangle_1) = \boldsymbol{v}(\langle P\rangle_2) = 0. \tag{2.37}$$

Vanishing velocity, together with equation (2.21), requires that the fluctuating parts of all of the thermodynamic variable vanish at the boundaries:

$$P^F = \rho^F = T^F = \theta^F = \Gamma^F = s^F = 0. \tag{2.38}$$

In anticipation of applying our calculation to the solar convective zone, where we will want to match the upper boundary of our fluid to the base of a solar atmosphere, we specify the mean temperature at $\langle P\rangle_1$ to be $\langle T\rangle_1$. At the base of our fluid we will want a solar luminosity of flux to enter our fluid. The vertical Eulerian energy flux on the right-hand side of equation (2.12) when put into Lagrangian form with $v = 0$ becomes

$$F_{\text{total}} = g\frac{\gamma}{\gamma-1}\langle\rho\rangle\sigma\frac{\partial\ln\langle T\rangle}{\partial\ln\langle P\rangle}. \tag{2.39}$$

We specify the flux at $\langle P\rangle_2$ to be F_2:

$$F_2 = g\frac{\gamma}{\gamma-1}\left[\langle\rho\rangle\sigma\frac{\partial\ln\langle T\rangle}{\partial\ln\langle P\rangle}\right]_{\langle P\rangle=\langle P\rangle_2}. \tag{2.40}$$

III. MATHEMATICAL DEFINITION OF AN EDDY

a) Eddies in a Homogeneous Fluid

We cannot easily solve the exact equations of motion for stellar convection because there are too many degrees of freedom in a turbulent fluid. It is therefore natural to ask if we can throw away any information but still leave a sufficient description of the flow to calculate the gross features of stellar structure. One way to proceed is to think of the velocity field at each point in the star as being made up of a series of eddies of different sizes. Formally, we can decompose the velocity into eddies by examining the Fourier transform of the Navier-Stokes and entropy equations. The transformed equations still contain all of the original information, but we can reduce the information by looking only at the amplitude of the Fourier modes of the velocity and entropy and disregarding all phase information. We do this by multiplying the Navier-Stokes equation (2.35) by

$$\frac{1}{(2\pi)^{3/2}}v_i^F(\boldsymbol{k})e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}, \tag{3.1}$$

integrating over d^3x, and taking the real part. The Fourier transform of $\boldsymbol{v}(\boldsymbol{x})$ is denoted by $\boldsymbol{v}(\boldsymbol{k})$. The left-hand side of equation (2.35) becomes

$$\frac{1}{2}\left[\frac{\partial}{\partial t}\right]_{\mathrm{L}}|\boldsymbol{v}^F(\boldsymbol{k})|^2. \tag{3.2}$$

The resulting right-hand side governs the evolution of the amplitude of the Fourier modes (kinetic energy). We lose all information about the phase of $\boldsymbol{v}^F(\boldsymbol{k})$. Unfortunately, without phase information we are unable to compute many of the terms on the right-hand side. The most untractable such term is the nonlinear advective term

$$\mathcal{R}e\frac{-i}{(2\pi)^{3/2}}\int d^3k'\left[\boldsymbol{v}^F(\boldsymbol{k}-\boldsymbol{k}')\cdot\boldsymbol{k}'\right]\left[\boldsymbol{v}^F(\boldsymbol{k}')\cdot\boldsymbol{v}^F(\boldsymbol{k})\right]. \tag{3.3}$$

The convolution in equation (3.3) describes the cascade of energy among triads of wavevectors. One way to calculate the cascade term is to take higher moments of the Navier-Stokes equation and determine the time evolution of velocity-cubed terms like equation (3.3) as functions of convolutions of four powers of the velocity. Instead of developing a hierarchy of moment equations, our approach is to estimate the phases needed for calculating terms like equations (3.3) by appealing to the physics of turbulent eddies.

An eddy of turbulent flow of size L is made up of a wave packet of Fourier modes centered about wavenumber $2\pi/L$; it is not a single Fourier mode (see Tennekes and Lumley 1972). We consider packets of modes that extend over logarithmic intervals of wavenumber. The ith eddy extends over the bandwidth $B_i = 2^{i+1/2}k_0 - 2^{i-1/2}k_0$, where $i = 0$ to N. Here k_0 is the smallest physically important wavenumber, and $2^N k_0$ is the largest. We define $v^F(i)$ by

$$v^F(i)^2 \equiv \frac{1}{V}\int_{B_i} \left| \boldsymbol{v}^F(\boldsymbol{k}) \right|^2 k^2\, dk\, d\Omega_k. \tag{3.4}$$

Here the normalizing factor $V = \frac{1}{2}(2\pi/k_0)^3$ ensures that $v^F(i)$ has dimensions of velocity and that the sum over i of $v^F(i)^2$ is the mean square velocity. Henceforth we understand all integrals over B_i to be normalized with this factor.

We obtain the evolution equation for $v^F(i)^2$ by multiplying equation (2.35) by the expression (3.1), integrating over d^3x, taking the real part, integrating over $d\Omega_k$, and integrating with respect to $k^2\, dk$ over B_i. With this procedure we reduce the velocity to just $N+1$ degrees of freedom. However, we have lost all information on the spatial variation of these degrees of freedom, necessary if we want to study a star where the convection is bounded in space. We will now show how spatial information can be preserved.

b) *Eddies in an Inhomogeneous Fluid*

If we sample two patches of fluid at different positions but at the same radius in a stellar convective zone, we expect the time-averaged kinetic spectra of the two patches to be the same. If the two patches are from different radii, we expect them to differ. The spectrum in the center of a convective zone should look very different from the spectrum in an overshoot region. We want to know the kinetic energy spectrum as a function of height. The uncertainty principle prohibits us from simultaneously having too much spectral and spatial information. Therefore, we sample a finite section of fluid, convolved with a smooth window function of vertical size Λ centered at height z and compute the kinetic energy spectrum for wavenumbers greater than $2\pi/\Lambda$. We define

$$k_0 = 2\pi/\Lambda. \tag{3.5}$$

It will prove convenient to choose Λ to be approximately the pressure scale height. As we slide our window function along the vertical direction, we obtain a spectrum that is both a function of z and $\boldsymbol{k}$. In essence, this is a "two-timing" scheme which separates slowly varying spectral dependence on spatial scales $> \Lambda$ from rapidly varying eddy dynamics on scales $< \Lambda$. The window function $W(z-z')$ is normalized so that $W(0) = 1$. Formally we define the *window-Fourier transform* of the quantity $Q(\boldsymbol{x})$ to be

$$Q(z,\boldsymbol{k}) = \frac{1}{(2\pi)^{3/2}}\int W(z-z')\,Q(\boldsymbol{x}')\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}}\,d^3x'. \tag{3.6}$$

The window-Fourier transform is nonzero at nonzero $\boldsymbol{k}$ only for fluctuating quantities, so we will omit the superscript F henceforth.

With our choice of normalization of W we can recover $Q(\boldsymbol{x})$ from the local transform as follows:

$$Q(\boldsymbol{x}) = \frac{1}{(2\pi)^{3/2}}\int Q(z,\boldsymbol{k})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}\,d^3k. \tag{3.7}$$

Substituting equation (3.7) into equation (3.6), we obtain the useful relationship

$$Q(z,\boldsymbol{k}) = \frac{1}{2\pi}\int W(z-z')\,Q(z',k_x,k_y,k_z')\,e^{iz'(k_z-k_z')}\,dz'\,dk_z'. \tag{3.8}$$

We can see how the windowed transform separates the large-scale vertical variation from the small-scale variation by considering the local transform of the derivative $\partial Q(\boldsymbol{x}')/\partial z'$:

$$\frac{1}{(2\pi)3/2}\int d^3x'\,W(z-z')\,e^{i\boldsymbol{k}\cdot\boldsymbol{x}'}\left[\frac{\partial}{\partial z'}Q(\boldsymbol{x}')\right] = \frac{\partial}{\partial z}Q(z,\boldsymbol{k}) - ik_z Q(z,\boldsymbol{k}). \tag{3.9}$$

The first term in equation (3.9) represents the large-scale vertical structure of $Q(\boldsymbol{x})$, and the second term contains all of the three-dimensional small-scale information. Of course, the second term is the only one that would be present if we had taken the Fourier transform ($W=1$) and not the window-Fourier transform.

We obtain *local eddies* by integrating the window-Fourier modes over logarithmic intervals in k. The kinetic energy for the ith local eddy at height z is defined to be

$$\frac{1}{2}v(z,i)^2 = \frac{1}{2}\int_{B_i} |\boldsymbol{v}(z,\boldsymbol{k})|^2 k^2\, dk\, d\Omega_k . \tag{3.10}$$

The time evolution equation for $\frac{1}{2}v(z,i)^2$ is obtained by multiplying the Navier-Stokes equation (2.35) (which is now to be thought of as a function of $\boldsymbol{x}'$) by

$$\frac{1}{(2\pi)^{3/2}} W(z-z')\, v_i(z,\boldsymbol{k})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'} k^2 , \tag{3.11}$$

integrating over d^3x', taking the real part, integrating over $d\Omega_k$, and integrating k over the interval B_i.

c) *Eddies versus Internal Wave Packets*

In a stratified fluid, it would be wrong to identify every window-Fourier wave packet as an eddy. Our physical picture of an eddy is an element of fluid that turns over and loses its identity in some characteristic time. In one turnover time it breaks apart and forms smaller or larger eddies. In a stably stratified fluid (mean entropy increasing upward), buoyancy acts as a restoring force so a perturbed fluid element oscillates at the Brunt-Väisällä frequency around its equilibrium position.

If the fluid is stably stratified, and if the kinetic energy of the window-Fourier wave packets is not sufficient to overcome the potential energy of the restoring force, then the packets will act like internal waves instead of like eddies. They do not break apart and lose their identity, and they carry a relatively small energy flux when compared to convective eddies (Press and Rybicki 1981).

Recalling the analysis of the Kelvin-Helmholz instability, we propose identifying a window-Fourier packet as a local eddy if and only if the kinetic energy of the packet is greater than (within a factor of order unity) the potential energy needed to overturn the eddy. For example, in Boussinesq salt water where the mean density $\langle\rho(z)\rangle$ is a function of height, the energy needed for a local packet of vertical extent L to overturn and make its density distribution homogeneous is

$$\int_{z_0-L/2}^{z_0+L/2} d^3x \langle\rho(z)\rangle gz - \frac{1}{L}\int_{z_0-L/2}^{z_0+L/2} d^3x\, gz \int_{z_0-L/2}^{z_0+L/2} \langle\rho(z')\rangle\, dz' . \tag{3.12}$$

Analogously, an overturning eddy in an anelastic fluid tries to make its potential density homogeneous. We can therefore estimate the characteristic energy density needed for a packet of size $L=2\pi/2^i k_0$ and height z to overturn as (see eq. [2.14])

$$E_{ot}(\langle P\rangle,i) = \int_{z-L/2}^{z+L/2}\left[\langle\Gamma(z')\rangle - \frac{1}{L}\int_{z-L/2}^{z+L/2}\langle\Gamma(z'')\rangle\, dz''\right] z'g\, d^3x' \Big/ \int_{z-L/2}^{z+L/2} d^3x' . \tag{3.13}$$

In equation (3.13) z is considered a function of the independent variable $\langle P\rangle$. This measure of the potential energy senses the stabilizing effect of stratification due to both an adverse potential density gradient and due to discrete jumps in the potential density (where the gradient becomes infinite). So, if the velocity of a window-Fourier packet is greater than $v_{\text{crit}}(\langle P\rangle,i)$,

$$v_{\text{crit}}(\langle P\rangle,i) = \left[2E_{ot}(\langle P\rangle,i)/\langle\rho(\langle P\rangle)\rangle\right]^{1/2}\alpha_{\text{Ri}} , \tag{3.14}$$

we consider the packet to be an eddy; if it is less than $v_{\text{crit}}(\langle P\rangle,i)$, then it is an internal wave. The dimensionless constant α_{Ri} ("Richardson number"), which appears in equation (3.14), is of order unity and can, in principle, be measured by numerical simulations of turbulence.

IV. MIXING-LENGTH THEORY AND BELL-NELKIN CASCADE THEORY

We are now in a position to make contact with two established pieces of fluid theory: mixing-length theory, which purports to describe the behavior of the largest scale convective eddies; and Bell-Nelkin cascade theory, which (in the context of homogeneous turbulence) models the dynamics of a turbulent cascade. This contact is usefully made now, since the remainder of this paper will develop what amounts to a unification of these previously unrelated formalisms.

a) Mixing-Length Theory in the Present Notation

Stellar convection in a mixing-length description is characterized by the mean quantities $\langle \rho \rangle$, $\langle P \rangle$, $\langle T \rangle$, and $\langle s \rangle$, and by the characteristic size, velocity, and temperature fluctuation of some largest eddy (called the mixing length). All of the dynamics and transport properties of any smaller eddies are ignored. For simplicity, we will write the equations of the mixing-length model only in the simplest case, where the temperature gradient is nearly adiabatic and the convective flux F_c is much larger than the radiative flux F_r. The first equation of mixing-length theory is the hydrostatic equilibrium equation,

$$\frac{\partial \langle P \rangle}{\partial z} = -g\langle \rho \rangle, \tag{4.1}$$

which, as we have seen, is identical to the mean part of the Navier-Stokes equation, equation (2.27).

The second equation of mixing-length theory relates the convective flux F_c to the (rms) fluctuating velocity and temperature on the mixing-length scale:

$$F_c = c_p \langle \rho \rangle v(\langle P \rangle, 0) T(\langle P \rangle, 0). \tag{4.2}$$

Here the notation $v(\langle P \rangle, 0)$ is analogous to that of equation (3.10), denoting the fluctuating velocity at a position $\langle P \rangle$ (in pressure coordinates) on a scale defined by wavenumber band zero (taken to be the mixing length). The justification of equation (4.2) is that it is supposed to approximate the exact energy-flux equation (2.12) when the following assumptions hold: (i) Most of the flux is carried by the biggest eddy, either because its velocity dominates the transport, *or* because the velocity and temperature fluctuations of smaller eddies are not well correlated spatially. (ii) The kinetic energy flux (first term on the right-hand side of eq. [2.12]) is neglected in favor of the third term, the velocity-transported thermal flux. Below, we will see that this approximation is good only up to a factor of order unity, since these two fluxes are, in fact, generally comparable and tied together by a dynamical equation. (iii) The viscous flux, second term on the right-hand side of equation (2.12), is negligible. Generally this term is of order $\mathcal{R}e^{-3/4}$ times the size of the other terms, where $\mathcal{R}e$ is the Reynolds number. In the solar convection zone $\mathcal{R}e \sim 10^{14}$, so the approximation is very good. (iv) The convection is in steady state, so that there is no mean flux of gravitational potential energy ($\langle \rho v_z \rangle \phi$ vanishes). (v) The velocity and temperature fluctuations of the mixing-length scale have unity coefficient of correlation (colder fluid falling, hotter fluid rising). This is the crucial assumption which allows the term $c_p \rho T v_z$ to be replaced by the product of the rms fluctuations $c_p \rho T^F v^F$. The physical basis of this assumption is that both driving *of* the eddy and transport *by* the eddy are supposed to be dominated by the same mixing-length scale. There is no reason to think that this assumption holds any better than to a factor of order unity.

It remains to write down a dynamical equation which determines $v(\langle P \rangle, 0)$. Although not usually written in this form, the dynamical equation assumed by the mixing-length theory is equivalent to positing energy equipartition between kinetic and thermal fluctuation energies on the mixing-length scale,

$$\tfrac{1}{2}\langle \rho \rangle v(\langle P \rangle, 0)^2 = \beta \langle \rho \rangle c_p T(\langle P \rangle, 0), \tag{4.3}$$

where β, a constant of order unity, is a parameter of the theory.

Equations (4.1), (4.2), (4.3), and a choice of β (e.g., $\beta = 1$), are a complete set of equations for the mixing-length theory (in the regime considered here). Notice that we never have to specify the actual value of the mixing length in this formulation! The reason that the conventional formulation of the theory does require that the mixing length be specified is that it purports to "derive" equation (4.3) from a dynamical buoyancy argument, as follows.

When a parcel of fluid of overdensity ρ^F falls through a mixing-length distance Λ, it acquires a velocity,

$$\tfrac{1}{2}\langle \rho \rangle v(\langle P \rangle, 0)^2 = g\rho(\langle P \rangle, 0)\, \Lambda. \tag{4.4}$$

Now, we need to relate ρ^F and T^F. The equation of state, equation (2.25), does not provide sufficient information alone. Mixing-length theory augments it by the assumption of pressure equilibrium,

$$P(\langle P\rangle,0) = 0, \tag{4.5}$$

from which follows for the rms quantities

$$\rho(\langle P\rangle,0)/\langle\rho\rangle = T(\langle P\rangle,0)/\langle T\rangle. \tag{4.6}$$

It then follows that equation (4.4) is equivalent to equation (4.3) if the mixing length is taken to be

$$\Lambda = \beta\frac{\gamma}{(\gamma-1)}\Lambda_P, \tag{4.7}$$

where $\Lambda_P \equiv -\langle P\rangle/(d\langle P\rangle/dz)$ is the pressure scale height.

There are various problems with equations (4.4) and (4.5). For example, equation (4.5) is demonstrably incorrect (see eq. [2.21]). The reason that pressure fluctuations are innocuous is *not* that they are small in magnitude, but rather that they can be shown to be spatially uncorrelated with the other quantities of interest (see Appendix D for further discussion of this point). Other deficiencies in equation (4.4) are discussed below. Our point of view in this paper is that equipartition (4.3) has a greater generality than the specific buoyancy model of the mixing-length theory. The eddy dynamics model which we develop in § V will yield (in the limit of a local theory) an equipartition relation like (4.3), with the value of β determined by the coherency length of the eddy velocity, rather than by any mixing argument. The mechanism for equipartition on this scale is then the pressure work done by the pressure fluctuation of the anelastic fluid in moving through this coherency distance.

b) *Need for Small Scales*

The mixing-length theory does not include the effects of smaller eddies or of viscosity. It is known in other contexts that the omission of small dissipative scales can do grievous damage to the *fluctuating* part of the Navier-Stokes equations. It is interesting to note that their omission in mixing-length theory also ruins the *mean* part of the entropy equation (2.5) and miscalculates the mean entropy flux $\boldsymbol{F}_s$ (eq. [2.6]). This miscalculation is somewhat alarming because mixing-length theory is supposed to treat the mean quantities correctly.

To see how the omission of the small dissipative scales leads to an erroneous mean entropy flux, consider the difference in entropy fluxes between the upper and lower boundaries of the convection zone (where velocities, and therefore the first term on the right of eq. [2.6], go to zero):

$$F_{s1} - F_{s2} = (\text{Area})^{-1}\int_2^1 \nabla\cdot\boldsymbol{F}_s\, d^3x = F(1/\langle T\rangle_1 - 1/\langle T\rangle_2). \tag{4.8}$$

This is positive and about as big as it dimensionally can be. The change in the entropy flux in equation (4.8) (with $\langle T\rangle_2 > \langle T\rangle_1$) occurs because, while photons enter the base of the convective zone hot, many more cold photons leave the top of the zone.

If, as in mixing-length theory, we neglect the dissipative terms S_σ and S_ν, then equation (2.5) shows that in steady state the entropy flux is divergence-free. This contradicts equation (4.8).

In Appendix A we show that, contrary to the mixing-length assumption, S_ν is not negligible, that it is, in fact, greater than the radiative source S_σ by a factor of F_c/F_r (which by our initial assumptions is a large number), despite the fact that the thermal diffusivity in a star is much greater than the kinematic viscosity. We show that the divergence of the entropy flux is

$$S_\nu \approx F/\Lambda_P T, \tag{4.9}$$

about as large as it could be dimensionally over one pressure scale height, and precisely what is required to resolve the apparent entropy paradox. In summary, the macroscopic equations of the mixing-length model—in being dissipationless—are inconsistent at zeroth order with the microscopic requirements of thermodynamics.

c) *Dynamics of Small Scales*

The key feature of homogeneous, isotropic, incompressible turbulence is its cascade of kinetic energy from the largest scale to all smaller scales. A quantitative, although simple and mechanistic, model of the cascade is that of Bell and Nelkin (1977). In the limit of constant temperature and density the Navier-Stokes equation (2.35) reduces to

$$\frac{\partial v_i^F}{\partial t} = -(\boldsymbol{v}\cdot\nabla v_i)^F - \frac{1}{\langle\rho\rangle}\frac{\partial P^F}{\partial x_i} + \nu\nabla^2 v_i^F. \tag{4.10}$$

The pressure in equation (4.10) is purely kinematic; it forces the velocity to be divergence-free. In our notation, Bell and Nelkin replace equation (4.10) with an amplitude equation for the eddies,

$$\begin{aligned}\frac{\partial}{\partial t}\left[\frac{1}{2}\langle\rho\rangle v(\langle P\rangle,i)^2\right] = -\nu k_i^2 v(\langle P\rangle,i)^2\langle\rho\rangle + \langle\rho\rangle v(\langle P\rangle,i)\\ \times\alpha_{\text{cas}}\Big(k_i\big[v(\langle P\rangle,i-1)^2 - 2v(\langle P\rangle,i)v(\langle P\rangle,i+1)\big]\\ -2^{1/3}Ck_i\big[v(\langle P\rangle,i-1)v(\langle P\rangle,i) - 2v(\langle P\rangle,i+1)^2\big]\Big),\end{aligned} \tag{4.11}$$

where

$$k_i \equiv 2^i k_0. \tag{4.12}$$

In this homogeneous model there is no dependence on $\langle P\rangle$. The term proportional to the viscosity is just the usual dissipation term that is proportional to the average wavenumber squared. The term cubic in velocity in equation (4.11) represents the transfer of energy among neighboring eddies in Fourier space. It models both the $\boldsymbol{v}\cdot\nabla\boldsymbol{v}$ and ∇P terms. The constant α_{cas} determines the ratio of the eddy turnover time to the time it takes an eddy to transfer its kinetic energy by the nonlinear cascade to its neighbors. We expect that $\alpha_{\text{cas}} \approx 1$.

The cubic velocity term in equation (4.11) has the following properties: (i) It is conservative, as is the original advective term in the Navier-Stokes equation that it models; that is, the sum of the nonlinear term over all eddies i is zero. (ii) It transfers energy only to nearest neighbors in Fourier space. There is no direct transfer of energy from large scales to small scales. A physical reason for this approximation is that eddies with vastly different size have very different turnover times and are therefore poorly temporally correlated with each other. (iii) The cascade model does not introduce any new length or time scales. (iv) The model equation (4.11) is "realizable," meaning that if the kinetic energies $v(\langle P\rangle,i)^2$ are initially nonnegative for all i, then they remain nonnegative. (v) There are two steady-state solutions to equation (4.11). In the limit of small viscosity, one of the solutions is

$$v(\langle P\rangle,i) \approx k_i^{-1/3}, \tag{4.13}$$

which is the Kolmogorov solution. The constant C determines the ratio of the rate at which an eddy transfers energy to smaller eddies to the rate at which it transfers energy to bigger eddies. Setting $C = 2^{-1/3}$ makes the second steady-state solution to equation (4.11) (in the limit of small viscosity) an equipartitioned state in which all of the $v(\langle P\rangle,i)$'s are equal. If turbulence is driven on some intermediate mesoscale, then the equipartitioned state will be the equilibrium on all larger scales, while the Kolmogorov cascade will hold on all smaller scales. Bell and Nelkin find that with $C < 1$, the cascade model behaves like three-dimensional turbulence. In a subsequent paper (Bell and Nelkin 1978) they show that the model is in best quantitative agreement with laboratory turbulence measurements for $C \approx 0.5$.

In the next two sections we extend the Bell-Nelkin model to a stratified anelastic fluid with buoyancy by window Fourier analyzing the equations of motion presented in § II. The reader who wishes to omit the derivation of the model can skip directly to § VII.

V. EDDY MODEL OF THE MOMENTUM EQUATION

We begin the eddy analysis of the Navier-Stokes equation (2.35) (which is considered a function of $\boldsymbol{x}'$) by multiplying both sides of the equation by the expression (3.11), integrating over all spatial volume d^3x', all Fourier angles $d\Omega_k$, and over the wavenumber interval B_i. Taking the real part of the resulting equation gives

$$\left[\frac{\partial}{\partial t}\right]_{\text{L}}\left[\frac{1}{2}v(\langle P\rangle,i)^2\right] = H_{\text{cas}} + H_{\text{pb}} + H_{\text{adv}} + H_{\text{vis}} + \frac{1}{3}v(\langle P\rangle,i)^2\left[\frac{\partial\ln\langle\rho\rangle}{\partial t}\right]_{\text{L}}, \tag{5.1}$$

where

$$H_{\rm vis}(\langle P\rangle, i) \equiv \frac{1}{(2\pi)^{3/2}} \operatorname{Re} \int W(z-z') v_i(z,\boldsymbol{k}) \frac{1}{\langle \rho(z')\rangle} \frac{\partial}{\partial x_j'} \Pi_{ij}^F(x') k^2 e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'} d\Omega_k\, dk\, d^3x', \tag{5.2}$$

$$H_{\rm pb}(\langle P\rangle, i) \equiv \frac{-1}{(2\pi)^{3/2}} \operatorname{Re} \int W(z-z') v_i(z,\boldsymbol{k}) \frac{k^2}{\langle \rho(z')\rangle} \left[\frac{\partial P^F}{\partial x_i'} + g\rho^F(x')\delta_{iz}\right] e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'} d\Omega_k\, dk\, d^3x'$$

$$- g\frac{\partial}{\partial\langle P\rangle} \operatorname{Re} \int P(z,-\boldsymbol{k}) v_z(z,\boldsymbol{k}) k^2\, d\Omega_k\, dk, \tag{5.3}$$

and

$$H_{\rm adv}(\langle P\rangle, i) + H_{\rm cas}(\langle P\rangle, i) \equiv \frac{-1}{(2\pi)^{3/2}} \operatorname{Re} \int W(z-z') v_i(z,\boldsymbol{k}) k^2 \left[\left(v^{\rm L}(\boldsymbol{x}')\cdot\nabla v_i^{\rm L}(\boldsymbol{x}')\right)\right] e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'} d\Omega_k\, dk\, d^3x'$$

$$+ g\frac{\partial}{\partial\langle P\rangle} \operatorname{Re} \int P(z,-\boldsymbol{k}) v_z(z,\boldsymbol{k}) k^2\, d\Omega_k\, dk. \tag{5.4}$$

$H_{\rm vis}$ is the rate at which energy leaves the ith eddy at height $\langle P\rangle$ due to *viscous* dissipation; $H_{\rm pb}$ is the rate of energy transfer due to the *pressure* and *buoyancy* forces; $H_{\rm cas}$ is the rate at which energy is transferred through the *cascade* of eddies of different sizes at the same height; and $H_{\rm adv}$ is the rate at which energy is *advected* among eddies of the same size but located at different positions. The k integrals in equations (5.2)–(5.4) extend over the band B_i. The term we have added and subtracted in equations (5.3) and (5.4) corresponds to the pressure term in equation (4.10). We do this so that $H_{\rm cas}$ will reduce to the Bell-Nelkin form in the homogeneous case.

In deriving equation (5.1) we have used an isotropic approximation for the kinetic energy in the last term:

$$v_z(\langle P\rangle, i)^2 = \tfrac{1}{3} v(\langle P\rangle, i)^2. \tag{5.5}$$

If we assume that the largest source of kinetic energy for each eddy is the nonlinear cascade of energy from other eddies (with the exception of the largest eddy which is directly driven by buoyancy in the vertical direction), and if we further assume that the cascade causes the fluid to lose its memory of the vertical direction, then the isotropic approximation (5.5) is reasonable. As one approaches steady-state, the last term in equation (5.1) becomes negligible anyway.

We now calculate the transfer rates $H_{\rm pb}$, $H_{\rm vis}$, $H_{\rm cas}$, and $H_{\rm adv}$.

a) Buoyancy and Pressure

Our strategy for evaluating $H_{\rm pb}$ is to reduce the terms on the right-hand side of equation (5.3) to terms that we can calculate. Since we do not know the phase of the eddies, we cannot calculate terms such as

$$J = \operatorname{Re} \int W(z-z') \frac{\rho^F(x')}{\langle\rho\rangle} k^2 e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'} v_z(z,\boldsymbol{k})\, d\Omega_k\, dk\, d^3x'. \tag{5.6}$$

For example, if $\rho^F/\langle\rho\rangle$ and v^F are in phase, then

$$J \approx \frac{1}{\langle\rho\rangle} \rho(\langle P\rangle, i) v(\langle P\rangle, i). \tag{5.7}$$

If $\rho^F/\langle\rho\rangle$ and v^F are completely out of phase, however, then

$$J \approx 0. \tag{5.8}$$

The relative phase of v^F and $\rho^F/\langle\rho\rangle$ changes for each eddy size, i, and also changes as a function of height. We write the right-hand side of equation (5.3) in terms that either do not require phase information or in which the phase information is obvious for physical reasons. In Appendix B we show that to the same order as the previous approximations

$$H_{\text{pb}}(\langle P\rangle,i)=\frac{g}{\langle\theta\rangle}\operatorname{Re}\int v_z(z,\boldsymbol{k})\theta(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk+\frac{1}{\langle\rho\rangle\langle\theta\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\operatorname{Re}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk. \tag{5.9}$$

The first term in equation (5.9) is proportional to the flux of potential temperature carried by the ith eddy at height $\langle P\rangle$:

$$F_\theta(\langle P\rangle,i)\equiv\operatorname{Re}\int v_z(z,\boldsymbol{k})\theta(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk. \tag{5.10}$$

The second term on the right-hand side of equation (5.9) vanishes in an adiabatic fluid since $\partial\langle\theta\rangle/\partial z=0$. Even if the fluid is not adiabatic, however, the second term on the right-hand side of equation (5.9) is smaller than the first term by order M^2. (This is proved in Appendix C.) The pressure-buoyancy source of energy therefore reduces to

$$H_{\text{pb}}(\langle P\rangle,i)=\frac{g}{\langle\theta\rangle}F_\theta(\langle P\rangle,i). \tag{5.11}$$

Equation (5.11) shows that if the fluid does not transport potential temperature (entropy), then there is no source of kinetic energy. Furthermore, if the entropy flux is transported upward, then H_{pb} is a source of kinetic energy, but if the flux is downward, then H_{pb} is a sink of kinetic energy. The fact that H_{pb} depends on the sign of the entropy transport is intimately related to the Schwarzschild stability criterion: the fluid is stable if $\partial s/\partial z>0$. This relationship is examined in § VI.

b) *Viscosity*

The viscous term, H_{vis}, has two contributions. One can be written as the divergence of a flux; the other corresponds to a negative definite dissipation. Following Bell and Nelkin we ignore the flux (which is of order $\mathcal{R}_e^{-3/4}$ times H_{cas}). The dissipative term is only large for the smallest eddies. Since the smallest eddies in the flow are much smaller than the local density scale height, they behave like the eddies of a constant density or incompressible fluid. Therefore, the dominant term in the dissipation is identical to the dissipative term of incompressible flow. A quantitative restatement of this fact is obtained by expanding the local Fourier transform of $1/\langle\rho\rangle\,\partial\Pi_{ij}/\partial x_j$:

$$\frac{1}{(2\pi)^{3/2}}\int\frac{1}{\langle\rho(z')\rangle}\frac{\partial}{\partial x_j'}\Pi_{ij}(\boldsymbol{x}')W(z-z')e^{i\boldsymbol{k}\cdot\boldsymbol{x}'}\,d^3x'=-\nu k^2v(z,\boldsymbol{k})+O\left[\nu kv(z,\boldsymbol{k})/\Lambda_\rho\right]+O\left[\nu kv(z,\boldsymbol{k})/\Lambda_v\right], \tag{5.12}$$

where Λ_ρ is the mean density scale height and Λ_v is the scale height of $v(z,\boldsymbol{k})$. To leading order, therefore, we can use the Bell-Nelkin approximation or dissipation in an incompressible flow,

$$H_{\text{vis}}(\langle P\rangle,i)=-\nu k_i^2v(\langle P\rangle,i)^2. \tag{5.13}$$

c) *Cascade and Advection*

The cascade and advection terms are due to the nonlinear term and a piece of the pressure gradient term of the Navier-Stokes equation. In Appendix E we show that

$$\sum_i\left[H_{\text{cas}}(\langle P\rangle,i)+H_{\text{adv}}(\langle P\rangle,i)\right] \tag{5.14}$$

can be written as the divergence of a flux. When this divergence is integrated over the entire convective zone, the surface terms vanish. Therefore, the term $(H_{\text{cas}}+H_{\text{adv}})$ is neither a source nor a sink of kinetic energy. It represents the rate of transfer of kinetic energy into the ith eddy at height $\langle P\rangle$ from eddies at different heights and from eddies of different wavenumber k_i.

We now separate the term for the exchange of energy from eddies at the same height but different wavenumber, H_{cas}, from the terms for the exchange of energy from eddies of the same wavenumber but at different height, H_{adv}. (Our local eddy analysis does not permit the transfer of kinetic energy from an eddy at a different height and a different wavenumber.) By manipulations analogous to those of Appendix E, we find

$$H_{\text{cas}}(\langle P\rangle,i)+H_{\text{adv}}(\langle P\rangle,i)=\frac{-1}{(2\pi)^{3/2}}\operatorname{Re} i\int v_j(z,\boldsymbol{k})W(z-z')v_j^F(\boldsymbol{x}')\boldsymbol{k}\cdot\boldsymbol{v}^F(\boldsymbol{x}')k^2e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}dk\,d\Omega_k\,d^3x'$$

$$-\frac{1}{(2\pi)^{3/2}}\operatorname{Re}\int v_j(z,\boldsymbol{k})\frac{1}{\langle\rho\rangle}\frac{\partial}{\partial z}[W(z-z')\langle\rho\rangle]v_jF(\boldsymbol{x}')v_z^F(\boldsymbol{x}')$$

$$\times k^2e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}dk\,d\Omega_k\,d^3x'+g\frac{\partial}{\partial\langle P\rangle}\operatorname{Re}\int P(z,-\boldsymbol{k})v_z(z,\boldsymbol{k})k^2\,dk\,d\Omega_k. \quad (5.15)$$

The first integral on the right-hand side of equation (5.15) is the only term that would be present if we were using the full Fourier transform [i.e., $W(z-z')=1$]. If the flow were homogeneous, this integral would be exactly the same as the nonlinear cascade term of the incompressible flow treated by Bell and Nelkin. We identify this first integral as H_{cas}:

$$H_{\text{cas}}(\langle P\rangle,i)\equiv\frac{-1}{(2\pi)^{3/2}}\operatorname{Re} i\int v_j(z,\boldsymbol{k})W(z-z')v_j^F(\boldsymbol{x}')\boldsymbol{k}\cdot\boldsymbol{v}^F(\boldsymbol{x}')k^2e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}dk\,d\Omega_k\,d^3x'. \quad (5.16)$$

We use the Bell-Nelkin model discussed in equation (4.11) to evaluate the right-hand side of equation (5.16):

$$H_{\text{cas}}(\langle P\rangle,i)=H_{\text{cas}}^{\text{up}}(\langle P\rangle,i)+H_{\text{cas}}^{\text{down}}(\langle P\rangle,i). \quad (5.17)$$

We see that H_{cas} is the sum of two contributions. One contribution,

$$H_{\text{cas}}^{\text{up}}(\langle P\rangle,i)\equiv 2v(\langle P\rangle,i)k_i\alpha_{\text{cas}}\left[-v(\langle P\rangle,i)v(\langle P\rangle,i+1)+2^{1/3}Cv(\langle P\rangle,i+1)^2\right], \quad (5.18)$$

links the kinetic energies of the i and $i+1$ eddies. The other contribution,

$$H_{\text{cas}}^{\text{down}}(\langle P\rangle,i)\equiv -H_{\text{cas}}^{\text{up}}(\langle P\rangle,i), \quad (5.19)$$

links the kinetic energies of the i and $i-1$ eddies. The relation (5.19) guarantees energy conservation.

The remaining integrals on the right-hand side of equation (5.15) are H_{adv}. These terms have the following properties: (i) They are zero if the fluid is spatially homgeneous. (ii) They represent the large-scale vertical derivative of the enthalpy and kinetic energy flux and are therefore conservative. (iii) An eddy's enthalpy and kinetic energy are most efficiently advected by larger eddies or by the eddy itself. (iv) The effective "diffusivity" for a large eddy to advect a smaller eddy is the velocity of the large eddy times the length of a large eddy.

Based on these properties, we model H_{adv} by

$$H_{\text{adv}}(\langle P\rangle,i)=-g^2\frac{\partial}{\partial\langle P\rangle}\langle\rho\rangle\sigma^*\frac{\partial}{\partial\langle P\rangle}\langle\rho\rangle v(\langle P\rangle,i)^2. \quad (5.20)$$

The effective diffusivity, σ^*, that appears in equation (5.20) is a linear sum of the effective diffusivities of all of the eddies that are larger than or equal to the ith eddy:

$$\sigma^*(\langle P\rangle,i)=\alpha_{\text{adv}}\sum_{j<i}2\pi v(\langle P\rangle,j)/k_j. \quad (5.21)$$

The constant α_{adv} that appears in equation (5.21) is of order unity.

The model equation (5.20) for H_{adv} has the four desired properties: (i) The terms are realizable—if the initial value of $v(\langle P\rangle,i)$ is nonnegative, then the H_{adv} term never produces a negative value for $v(\langle P\rangle,i)$ at a later time. (ii) Like the terms that it models, H_{adv} is cubic in the velocity. (iii) H_{adv} introduces no new length or times scales. (iv) H_{adv} does not permit energy transfer among eddies with different wavelengths; it only transfers energy among same size eddies at different heights.

d) *Effects of Stable Stratification*

In obtaining the terms H_{cas} and H_{adv} we have ignored the possibility that for some wavenumbers k_i and some heights $\langle P \rangle$, the velocity $v(\langle P \rangle, i)$ may be less than $v_{\mathrm{crit}}(\langle P \rangle, i)$, and the wave packets are not eddies but internal waves. If the ith wave packet at height $\langle P \rangle$ is an internal wave, then the potential temperature flux carried by the packet will oscillate in time, and the time-averaged flux will be nearly zero (Press 1981). The total kinetic energy pumped into or out of the fluid will also oscillate in time, and the time average of H_{pb} will also be zero. Equation (5.11), which shows that H_{pb} is proportional to the potential temperature flux, remains valid. However, we anticipate that when (in the next section) we evaluate F_θ we will find that internal waves carry no flux.

Since internal waves retain their identity and do not break apart like eddies, H_{cas} must be modified. We set the links which remove energy from the ith wave packet equal to zero if the packet is an internal wave. These are the terms with negative coefficients in equation (5.17).

Since internal waves oscillate, they do not contribute significantly to the time-averaged enthalpy flux of an eddy; they must not be included in the sum of effective diffusivities that make up σ^* (see eq. [5.21]).

VI. MOMENTS OF THE THERMAL EQUATION

The evolution equation (5.1) for $v(\langle P \rangle, i)$ and the equations for H_{cas}, H_{vis}, H_{pb}, and H_{adv} given in § V are not a complete set of equations. We still need equations for $\langle \rho \rangle$ and the potential temperature fluxes, $F_\theta(\langle P \rangle, i)$.

a) *Mean Temperature, Density, and Potential Temperature*

In Lagrangian coordinates where $\langle P \rangle$ is the independent variable, once we determine $\langle T \rangle$ we can immediately find $\langle \rho \rangle$ from the mean equation of state (2.24) and $\langle \theta \rangle$ from its definition (eq. [2.13]). We therefore concentrate on finding an evolution equation for $\langle T \rangle$. From equations (2.5) and (2.11), the mean Eulerian thermal energy equation is

$$\frac{\partial c_v \langle \rho \rangle \langle T \rangle}{\partial t} = -\frac{\partial}{\partial z}\langle c_p \rho T v_z \rangle + \langle v \cdot \nabla P \rangle + \frac{\partial}{\partial z}\left\langle c_p \rho \sigma \frac{\partial T}{\partial z}\right\rangle + \left\langle \frac{\eta}{2}\sigma_{ij}\sigma_{ij}\right\rangle + \xi\langle (\nabla \cdot v)^2 \rangle . \tag{6.1}$$

The Lagrangian evolution equation for $\langle T \rangle$ (or thermal energy per unit mass) is

$$\left[\frac{\partial c_p \langle T \rangle}{\partial t}\right]_{\mathrm{L}} = \frac{\partial}{\partial \langle P \rangle}\left(g c_p \langle \rho T v_z \rangle - g c_p \langle \rho \rangle \langle T \rangle \langle v_z \rangle_t\right) + \frac{1}{\langle \rho \rangle}\langle v^F \cdot \nabla P^F \rangle + \frac{g}{\langle \rho \rangle}\langle v_z^F \rho^F \rangle$$
$$+ \frac{\partial}{\partial \langle P \rangle}\left[c_p g^2 \sigma \langle \rho \rangle^2 \frac{\partial \langle T \rangle}{\partial \langle P \rangle}\right] + \frac{1}{\langle \rho \rangle}\left\langle \frac{\eta}{2}\sigma_{ij}\sigma_{ij}\right\rangle + \frac{1}{\langle \rho \rangle}\left\langle \xi (\nabla \cdot v)^2 \right\rangle . \tag{6.2}$$

In obtaining equation (6.2) we have used the relationship

$$\frac{\partial \langle P \rangle}{\partial t} = -\frac{\partial \langle P \rangle}{\partial z}\langle v_z \rangle_t = -\langle v \rangle \cdot \nabla \langle P \rangle + \langle v_z \rangle_s \frac{\partial \langle P \rangle}{\partial z} , \tag{6.3}$$

which follows from operating on $\langle P \rangle$ with equation (2.32). The fourth term on the right-hand side of equation (6.2) corresponds to dissipation via the radiative diffusion of thermal energy. The viscous terms in equation (6.2) give the positive definite increase in thermal energy due to the viscous losses of kinetic energy, $H_{\mathrm{vis}}(\langle P \rangle, i)$. To be consistent with our approximation for H_{vis} in the momentum equations (and to insure that total energy is exactly conserved) we replace the viscous dissipation terms (last two terms) in equation (6.2) with equation (5.13) summed over i. The biggest contribution to the sum will generally come from the smallest eddies.

We now manipulate the remaining nonlinear terms in equation (6.2) (the first three terms on the right-hand side) into a form in which we can evaluate them with our limited phase information. To conserve the total energy, $c_p \langle T \rangle + \frac{1}{2} v(\langle P \rangle, i)^2$, equations (5.1) and (6.2) require that these terms satisfy

$$\frac{\partial}{\partial \langle P \rangle}\left(g c_p \langle \rho T v_z \rangle - g c_p \langle \rho \rangle \langle T \rangle \langle v_z \rangle_t\right) + \frac{1}{\langle \rho \rangle}\langle v^F \cdot \nabla P^F \rangle + \frac{g}{\langle \rho \rangle}\langle v_z^F \rho^F \rangle$$
$$= -\sum_i H_{\mathrm{pb}}(\langle P \rangle, i) + \text{the divergence of some flux.} \tag{6.4}$$

To find the unknown flux in equation (6.4), we expand the temperature and density in terms of the pressure and potential temperature and obtain (see Appendix F)

$$\frac{\partial}{\partial\langle P\rangle}\left(gc_p\langle\rho T v_z\rangle - gc_p\langle\rho\rangle\langle T\rangle\langle v_z\rangle_t\right) + \frac{1}{\langle\rho\rangle}\langle v^F\cdot\nabla P^F\rangle + \frac{g}{\langle\rho\rangle}\langle v_z^F\rho^F\rangle$$
$$= -\frac{g}{\langle\theta\rangle}\langle v_z^F\theta^F\rangle + \frac{\partial}{\partial\langle P\rangle}\left[\frac{\gamma}{\gamma-1}g\frac{\langle P\rangle}{\langle\theta\rangle}\langle v_z^F\theta^F\rangle\right]. \tag{6.5}$$

We recognize the term $\langle v_z^F\theta^F\rangle$ as the sum of the potential temperature fluxes defined in equation (5.10):

$$\langle v_z^F\theta^F\rangle \approx \mathrm{Re}\int v_z(\langle P\rangle,\boldsymbol{k})\theta(\langle P\rangle,-\boldsymbol{k})\,d^3k = \sum_i F_\theta(\langle P\rangle,i). \tag{6.6}$$

Substituting equation (6.5) into equation (6.2), we obtain the Lagrangian evolution equation for $\langle T\rangle$ in terms of familiar quantities,

$$\left[\frac{\partial c_p\langle T\rangle}{\partial t}\right]_{\mathrm{L}} = -\sum_i\left[H_{\mathrm{vis}}(\langle P\rangle,i) + H_{\mathrm{pb}}(\langle P\rangle,i)\right] + \frac{\partial}{\partial\langle P\rangle}\left[c_p g^2\sigma\langle\rho\rangle^2\frac{\partial\langle T\rangle}{\partial\langle P\rangle}\right] + \frac{\partial}{\partial\langle P\rangle}\left[\frac{\gamma}{\gamma-1}g\frac{\langle P\rangle}{\langle\theta\rangle}\sum_i F_\theta(\langle P\rangle,i)\right]. \tag{6.7}$$

It is now apparent that the unspecified flux in equation (6.4) is the nonlinear flux of thermal energy, F_{TE}, and is related to the potential temperature flux, F_θ, by

$$F_{\mathrm{TE}}(\langle P\rangle,i) = \frac{\gamma}{\gamma-1}g\frac{\langle P\rangle}{\langle\theta\rangle}F_\theta(\langle P\rangle,i). \tag{6.8}$$

b) *Potential Temperature Flux*

We can compute the potential temperature flux from second-order moment equations. Just as we found the Lagrangian evolution equation for the square of the velocity $v(\langle P\rangle,i)^2$, we can use equations (2.2) and (2.5) to find the evolution equations for the potential temperature variance, $\theta(\langle P\rangle,i)^2$, and the potential temperature flux of equation (6.6). In Appendix G we show how the moment equations lead to an expression for the potential temperature flux. However, there is also a simpler physical argument based on a gradient-transport model that gives the same result.

The potential temperature is an adiabatic invariant so that in the absence of any dissipation (viscosity or thermal diffisivity) the advective derivative of θ is zero:

$$\frac{D\theta^F}{Dt} = 0. \tag{6.9}$$

Consider a flow that initially has no potential temperature fluctuations, $\theta^F = 0$, but has a mean potential temperature gradient $\partial\langle\theta\rangle/\partial z$. If we introduce a velocity field that is characterized by eddies of size $2\pi/k_i$, then the fluctuating potential temperature will no longer be zero. The fluid particles carried by the velocity retain their potential temperature, and potential temperature fluctuations of size $2\pi/k_i$ form. We can easily estimate the magnitude of these fluctuations. An eddy centered at height z will carry fluid from height $z-\pi/k_i$ where the potential temperature is $\langle\theta\rangle - (\pi/k_i)\,\partial\langle\theta\rangle/\partial z$ to the height $z+\pi/k_i$ where the surrounding potential temperature is $\langle\theta\rangle + (\pi/k_i)\,\partial\langle\theta\rangle/\partial z$. The difference between the potential temperature in the eddy and its surroundings is the fluctuation

$$\theta(\langle P\rangle,i) = \left|\frac{\partial\langle\theta\rangle}{\partial z}\right|\frac{2\pi}{k_i}. \tag{6.10}$$

Since the velocity field created the potential temperature fluctuation, $v(\langle P\rangle,i)$ and $\theta(\langle P\rangle,i)$ are well correlated. Therefore,

$$\mathrm{Re}\int_{B_i} v_z(\langle P\rangle,\boldsymbol{k})\theta(\langle P\rangle,-\boldsymbol{k})k^2\,d\Omega_k\,dk = -\alpha_\theta\frac{\partial\langle\theta\rangle}{\partial z}\frac{2\pi}{k_i}v(\langle P\rangle,i), \tag{6.11}$$

or

$$\frac{F_\theta(\langle P\rangle, i)}{\langle\theta\rangle} = \alpha_\theta g \frac{2\pi}{k_i} v(\langle P\rangle, i)\left[\frac{1}{\gamma}\frac{\langle\rho\rangle}{\langle P\rangle} - \frac{\partial\langle\rho\rangle}{\partial\langle P\rangle}\right] = -\alpha_\theta \frac{2\pi}{k_i} v(\langle P\rangle, i)\frac{1}{c_p}\frac{\partial\langle s\rangle}{\partial z}, \tag{6.12}$$

where α_θ is a positive constant of order unity. Equations (6.12) and (6.10) are valid for an adiabatic fluid in which the only source of potential temperature fluctuation, $F_\theta(\langle P\rangle, i)$, is eddies of the same size interacting with a mean potential temperature gradient. Any other source of potential temperature fluctuation such as the nonlinear cascade of potential temperature from eddies of one size to eddies of another size, invalidates equation (6.10). However, the flux only senses potential temperature that is correlated with the velocity. If the other sources of fluctuating potential temperature of wavenumber k_i are uncorrelated with the velocity at wavenumber k_i, then equation (6.12) is valid. This is true for adiabatic convection, since the only sources of potential temperature fluctuation are the velocity interaction with a mean potential temperature gradient and the nonlinear cascade, and a turbulent cascade loses all phase information between eddies of different sizes. Notice that we cannot use a gradient-transport equation like (6.12) to estimate the density or temperature flux since the density and temperature are not adiabatic invariants.

In equation (5.11) we showed that the energy source (or sink) due to buoyancy depended on the sign of $F_{\text{pb}}(\langle P\rangle, i)$. Equation (6.12) now shows us that buoyancy acts as an energy source (sink) if $\partial\langle s\rangle/\partial z$ is negative (positive). Therefore, equations (5.11) and (6.12) imply the Schwarzschild stability condition in an adiabatic fluid.

In the presence of dissipation the gradient-transport equation for the potential temperature needs one additional modification. Let us restrict ourselves to fluids where $\nu/\sigma \ll 1$ (for the Sun $\nu/\sigma \approx 10^{-7}$), so that we can ignore viscous dissipation compared to the thermal dissipation. The characteristic time for thermal diffusion to leak an eddy's potential temperature into the surrounding fluid is $(2\pi/k_i)^2/\sigma$. The characteristic time of the eddy to advect the potential temperature across itself is $2\pi/[v(\langle P\rangle, i)k_i]$. The ratio of these two times is the Péclét number for eddies of size k_i at height $\langle P\rangle$:

$$\mathcal{P}e(\langle P\rangle, i) \equiv \frac{2\pi v(\langle P\rangle, i)}{\sigma k_i}. \tag{6.13}$$

Thermal dissipation is important for little eddies where the Péclét number is small, i.e., where the thermal dissipation time is short compared to the advection time. We modify equation (6.12) for the potential temperature flux by (see Appendix G)

$$\frac{F_\theta(\langle P\rangle, i)}{\langle\theta\rangle} = \alpha_\theta g \frac{2\pi}{k_i} v(\langle P\rangle, i)\left[\frac{1}{\gamma}\frac{\langle\rho\rangle}{\langle P\rangle} - \frac{\partial\langle\rho\rangle}{\partial\langle P\rangle}\right][1 + 1/\mathcal{P}e(\langle P\rangle, i)]^{-1}. \tag{6.14}$$

If the Péclét number is large, then the flux is not modified; if the Péclét number is unity so that the thermal and advective times are equal, then eddies leak half of their flux; if the Péclét number is small, then the potential temperature flux goes to zero.

Finally, if $v(\langle P\rangle, i) < v_{\text{crit}}(\langle P\rangle, i)$, then the ith wave packet is stably stratified and is an internal wave, not an eddy. Since an internal wave oscillates in time, the zeroth order time average of the potential temperature flux is zero. Thus, the final form of equation (6.14) is

$$\frac{F_\theta(\langle P\rangle, i)}{\langle\theta\rangle} = \begin{cases} \alpha_\theta g \dfrac{2\pi}{k_i} v(\langle P\rangle, i)\left[\dfrac{1}{\gamma}\dfrac{\langle\rho\rangle}{\langle P\rangle} - \dfrac{\partial\langle\rho\rangle}{\partial\langle P\rangle}\right][1 + 1/\mathcal{P}e(\langle P\rangle, i)]^{-1}, & v(\langle P\rangle, i) > v_{\text{crit}}(\langle P\rangle, i), \\ 0, & v(\langle P\rangle, i) \le v_{\text{crit}}(\langle P\rangle, i). \end{cases} \tag{6.15}$$

We should note that although the nonlinear diffusion has a huge effect (of order the Péclét number, 10^7 for the Sun) on H_{adv} in equation (5.1), the nonlinear eddy diffusion has only a small effect (which we can absorb into α_θ) on the potential temperature flux. Nonlinear advection by large eddies is effective at transporting scalar quantities such as kinetic energy or heat; we need to include nonlinear effects in H_{adv}. Nonlinear advection by a big eddy of vector quantities such as an eddy's potential temperature flux tends to orient randomly the directions of the vector quantities (but preserves their magnitudes). The random orientation of all the advected fluxes of the small eddies averages out to zero.

VII. DISCUSSION

Finally, we emerge from the notational thicket with our multiscale model equations in hand. Since the final model is in many ways more transparent than its detailed derivation, it seems desirable to summarize it here.

The independent variables are time t and mean (i.e., horizontally averaged) pressure $\langle P \rangle$, the latter variable related to the depth coordinate z by the standard, time-independent equation of hydrostatic equilibrium (2.27), where g, the acceleration of gravity, is assumed known.

The dependent variables, functions of $\langle P \rangle$ and t, are the mean density $\langle \rho \rangle$, the mean temperature $\langle T \rangle$, and the rms fluctuating velocities $v(\langle P \rangle, i)$ corresponding to scales $2\pi/k_i$, $k_i = 2^i k_0$, $i = 1, 2 \ldots N$, where $2\pi/k_0$ is some largest eddy size on the order of a density scale height. An equation of state (2.24) relates $\langle \rho \rangle$, $\langle P \rangle$, and $\langle T \rangle$, and it is assumed that a microscopic energy diffusivity σ and viscosity ν can be calculated from $\langle P \rangle$, $\langle \rho \rangle$, and $\langle T \rangle$.

At a fixed time t, the dependent variables yield, in order, the following derived quantities: $\mathcal{P}e(\langle P \rangle, i)$, the Péclét number (thermal transparency) of the flow on scale i (eq. [6.13]); $F_\theta(\langle P \rangle, i)$, the potential temperature flux on scale i (eq. [6.15]); $\sigma^*(\langle P \rangle, i)$, the eddy diffusivity seen by scale i (eq. [5.21]).

The total energy flux is the sum of the standard radiative flux $(c_p\langle \rho \rangle \sigma \nabla T$ (see eqs. [2.12] and [6.7]), and two parts which make up the convective flux (and which can be of comparable size): a convective thermal flux, given by equation (6.8) summed over all scales i; and a convective flux of mechanical energy given by the sum over scales i of

$$\sigma^*(\langle P \rangle, i) \nabla \left[\langle \rho \rangle v(\langle P \rangle, i)^2 \right] \text{ (see eq. [5.20]).}$$

An eddy on scale i at depth $\langle P \rangle$ is taken to be overturning if it satisfies $v(\langle P \rangle, i) > v_{\text{crit}}(\langle P \rangle, i)$ given by equations (3.14) and (3.13) (using the definition of potential density [2.14]); otherwise it is taken to be an oscillating and nonmixing internal wave packet.

There are now two time evolution equations in the model, for $\langle T \rangle$ and for $v(\langle P \rangle, i)$. The time evolution equation for $\langle T \rangle$ is equation (6.7), which includes terms on the right-hand side which correspond to viscous dissipation H_{vis} (written in detail as eq. [5.13]), pressure-buoyancy work H_{pb} (given in detail as eq. [5.11]), microscopic heat diffusion, and thermal convective flux. The time evolution equation for $v(\langle P \rangle, i)$ is equation (5.1), which includes on its right-hand side the terms H_{pb} and H_{vis} already noted, and also a turbulent cascade term H_{cas} (eqs. [5.17]–[5.19]), a turbulent advection term H_{adv} (eq. [5.20]), and a "decompressive" term which vanishes in steady state.

It is interesting to see how this set of equations conserves energy. If we sum equation (5.1) over i and add equation (6.7), we get

$$\left[\frac{\partial}{\partial t}\right]_{\text{L}} \left[\sum_i \frac{1}{2} v(\langle P \rangle, i)^2 + c_p \langle T \rangle \right] = \sum_i \frac{1}{3} v(\langle P \rangle, i)^2 \left[\frac{\ln \langle \rho \rangle}{\partial t} \right]_{\text{L}} + \frac{\partial}{\partial \langle P \rangle} \left[c_p g^2 \sigma \langle \rho \rangle^2 \frac{\partial \langle T \rangle}{\partial \langle P \rangle} \right]$$
$$+ \sum_i H_{\text{adv}} + \frac{\partial}{\partial \langle P \rangle} \left[\frac{\gamma}{\gamma - 1} g \frac{\langle P \rangle}{\langle \theta \rangle} \sum_i F_\theta(\langle P \rangle, i) \right]. \tag{7.1}$$

This is in the form of a conservation equation, with all the terms on the right-hand side perfect divergences, except for the term involving $\partial(\ln \langle \rho \rangle)/\partial t$. This term occurs because we are not including in equation (7.1) the kinetic energy of the mean velocity, $\frac{1}{2}\langle |\boldsymbol{v}| \rangle^2$. One can show that the equation for the time evolution of this mean kinetic energy contains a term which exactly cancels the term proportional to $\partial(\ln \langle \rho \rangle)/\partial t$. While this term can be important in highly dynamical situations, it goes to zero as one approaches a steady state.

The remaining terms on the right-hand side of equation (7.1) represent the radiative flux and the mechanical and thermal contributions to the convective flux, respectively. In a steady state, the sum of these fluxes is zero.

Note that the conserved quantity on the left-hand side of equation (7.1) does not have an explicit gravitational contribution and involves the enthalpy $c_p\langle T \rangle$ rather than the internal energy $c_v\langle T \rangle$. This is because in the anelastic approximation the rate of doing gravitational work on unit mass of fluid, $-g\langle v_z \rangle$, is equal to $\partial(\langle P \rangle / \langle \rho \rangle)/\partial t|_{\text{L}}$ plus a divergence. The enthalpy and internal energy differ by exactly $\langle P \rangle / \langle \rho \rangle$.

In a future paper we will describe how the multiscale model can be implemented computationally; this will involve translation of the model's analytic spatial derivatives, $\partial/\partial\langle P \rangle$, into a finite difference scheme. We will then use the model to investigate a variety of time-dependent and time-independent phenomena associated with convective overshoot and mixing.

One of us (S. T.) acknowledges support from the Harvard-Smithsonian Center for Astrophysics Visiting Scientist Program, and from a John Simon Guggenheim Memorial Foundation Fellowship. Another of us (P. M.) acknowledges the support of NSF postdoctoral fellowship SPI-8009181.

APPENDIX A

EVALUATION OF ENTROPY SOURCES AND SINKS

Comparing equations (2.7) and (5.2), we see that S_ν is proportional to the rate at which kinetic energy is viscously dissipated,

$$S_\nu \sim -\frac{\langle \rho \rangle}{\langle T \rangle} \sum_i H_{\text{vis}}(\langle P \rangle, i). \tag{A1.1}$$

Now equations (5.17)–(5.19) show that kinetic energy cascades out of the largest eddies at a rate proportional to $\langle \rho \rangle [v(\langle P \rangle, i=0)]^3 k_0$. The rate at which kinetic energy cascades into the smallest dissipative scales is also of order $\langle \rho \rangle [v(\langle P \rangle, i=0)]^3 k_0$ if we assume that the kinetic energy cascade is not "leaky"; i.e., the pressure and buoyancy sources and sinks of energy in the intermediate eddies change the kinetic energy cascade rate only by factors of order unity. We can now write equation (A1.1) as

$$S_\nu \sim \frac{\langle \rho \rangle}{\langle T \rangle} [v(\langle P \rangle, i=0)]^3 k_0. \tag{A1.2}$$

Finally, using equations (4.2) and (4.3) to estimate $v(\langle P \rangle, i=0)$ and setting $1/k_0$ equal to the pressure scale height, we obtain

$$S_\nu \sim \frac{F_c}{\langle T \rangle \Lambda_p}. \tag{A1.3}$$

To estimate the radiative diffusion entropy source we note that the radiative flux is

$$F_r = \langle c_p \sigma \rho \nabla T \rangle. \tag{A1.4}$$

Therefore, by equation (2.8)

$$S_\sigma \sim \frac{F_r}{T} \frac{\nabla T}{T}. \tag{A1.5}$$

If the temperature gradient is nearly adiabatic, we have

$$\frac{\nabla T}{T} \sim \frac{1-\gamma}{\gamma} \frac{1}{\Lambda_p}. \tag{A1.6}$$

Combining equations (A1.6) and (A1.5), we obtain

$$S_\sigma \sim \frac{F_r}{\langle T \rangle \Lambda_p}. \tag{A1.7}$$

Comparing equations (A1.3) and (A1.7), we find

$$S_\sigma / S_\nu \sim F_r / F_c. \tag{A1.8}$$

If the convective flux of energy dominates the radiative flux, then S_σ is negligible compared to S_ν.

APPENDIX B

PROOF OF EQUATION (5.9)

Throughout this paper, we shall make frequent use of the assumption that all of the horizontally averaged thermodynamic quantities, $\langle Q(z) \rangle$, have such slow variation in height compared to any fluctuating quantity, $A^F(x)$, that we can use the approximation

$$\int W(z-z')\langle Q(z')\rangle e^{-ik\cdot x'}A^F(x')k^2\,d\Omega_k\,dk\,d^3x' \approx \langle Q(z)\rangle \int W(z-z')e^{-ik\cdot x'}A^F(x')k^2\,d\Omega_k\,dk\,d^3x'. \quad \text{(A2.1)}$$

For equation (A2.1) to be valid, the eddies that make up the fluctuating quantities A^F must be smaller than the scale height of $\langle Q\rangle$. Therefore, the equations derived in this Appendix will lose validity for eddies larger than a pressure scale height.

Now consider the term $\rho^F(x')/\langle\rho(z')\rangle$ in the expression (5.3) for H_{pb}. Using the definition of θ and the hydrostatic pressure equation, we obtain

$$\frac{\rho^F}{\langle\rho\rangle} = -\frac{\theta^F}{\langle\theta\rangle} + \frac{1}{\gamma}\frac{P^F}{\langle P\rangle} \quad \text{(A2.2)}$$

$$= -\frac{\theta^F}{\langle\theta\rangle} - \frac{P^F}{g\langle\rho\rangle}\left(\frac{1}{\gamma}\frac{\partial\langle P\rangle}{\partial z}\frac{1}{\langle P\rangle}\right) \quad \text{(A2.3)}$$

$$= -\frac{\theta^F}{\langle\theta\rangle} - \frac{P^F}{g\langle\rho\rangle}\left(\frac{1}{\langle\rho\rangle}\frac{\partial\langle\rho\rangle}{\partial z} + \frac{1}{\langle\theta\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\right). \quad \text{(A2.4)}$$

With equations (A2.1) and (A2.4), equation (5.3) becomes

$$\begin{aligned} H_{\text{pb}} &= \text{Re}\,\frac{g}{\langle\theta(z)\rangle}\int v_z(z,\boldsymbol{k})\theta(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk \\ &- \text{Re}\,\frac{1}{\langle\rho(z)\rangle}\frac{1}{(2\pi)^{3/2}}\int W(z-z')v_i(z,\boldsymbol{k})\frac{\partial P^F}{\partial x_i'}(x')k^2 e^{-ik\cdot x'}\,d\Omega_k\,dk\,d^3x' \\ &+ \text{Re}\,\frac{1}{\langle\theta\rangle\langle\rho\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk + \text{Re}\,\frac{1}{\langle\rho\rangle}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})\frac{1}{\langle\rho(z)\rangle}\frac{\partial\langle\rho(z)\rangle}{\partial z} \\ &\times k^2\,d\Omega_k\,dk + \text{Re}\,\frac{1}{\langle\rho\rangle}\frac{\partial}{\partial z}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk. \end{aligned} \quad \text{(A2.5)}$$

Now we replace the expression $v_z(z,\boldsymbol{k})/\langle\rho\rangle\,\partial\langle\rho\rangle/\partial z$ which appears in the fourth integral on the right-hand side of equation (A2.5). Start with

$$\nabla\cdot v^F = \nabla\cdot[(\langle\rho\rangle v^F)/\langle\rho\rangle] = \frac{-1}{\langle\rho\rangle}\frac{\partial\langle\rho\rangle}{\partial z}v_z^F. \quad \text{(A2.6)}$$

Take the local transform of equation (A2.6) and use equations (3.9) and (A2.1) to obtain

$$\frac{v_z(z,\boldsymbol{k})}{\langle\rho\rangle}\frac{\partial\langle\rho\rangle}{\partial z} = i\boldsymbol{k}\cdot\boldsymbol{v}(z,\boldsymbol{k}) - \frac{\partial}{\partial z}v_z(z,\boldsymbol{k}). \quad \text{(A2.7)}$$

Substituting equation (A2.7) into equation (A2.5), we obtain

$$\begin{aligned} H_{\text{pb}} &= \text{Re}\,\frac{g}{\langle\theta\rangle}\int v_z(z,\boldsymbol{k})\theta(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk + \text{Re}\,\frac{1}{\langle\theta\rangle\langle\rho\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk \\ &- \text{Re}\,\frac{1}{\langle\rho\rangle}\frac{1}{(2\pi)^{3/2}}\int W(z-z')v_i(z,\boldsymbol{k})\frac{\partial P^F}{\partial x_i'}(x')k^2 e^{-ik\cdot x'}\,d\Omega_k\,dk\,d^3x' \\ &+ \text{Re}\,\frac{1}{\langle\rho\rangle}\int v_z(z,\boldsymbol{k})\frac{\partial P(z,-\boldsymbol{k})}{\partial z}k^2\,d\Omega_k\,dk + \text{Re}\,\frac{1}{\langle\rho\rangle}\int P(z,-\boldsymbol{k})i\boldsymbol{k}\cdot\boldsymbol{v}(z,\boldsymbol{k})k^2\,d\Omega_k\,dk. \end{aligned} \quad \text{(A2.8)}$$

The local transform of $\partial P^F/\partial x_i'$ appears in the third integral on the right-hand side of equation (A2.8). We expand the transform using equation (3.9) to obtain a new expression for this third integral,

$$-\mathrm{Re}\,\frac{1}{\langle\rho\rangle}\frac{1}{(2\pi)^{3/2}}\int W(z-z')v_i(z,\boldsymbol{k})\frac{\partial P^F}{\partial x_i'}(x')k^2 e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}\,d\Omega_k\,dk\,d^3x'$$

$$=-\mathrm{Re}\,\frac{1}{\langle\rho\rangle}\int i\boldsymbol{k}\cdot\boldsymbol{v}(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk\,d^3x'-\mathrm{Re}\,\frac{1}{\langle\rho\rangle}\int v_z(z,\boldsymbol{k})\frac{\partial P(z,-\boldsymbol{k})}{\partial z}k^2\,d\Omega_k\,d^3x'. \tag{A2.9}$$

Substituting equation (A2.9) into (A2.8) we immediately obtain equation (5.9).

APPENDIX C

We here prove the claim following equation (5.10) that

$$\frac{g}{\langle\theta\rangle}\left|\mathrm{Re}\int v_z(z,\boldsymbol{k})\theta(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk\right|\gg\left|\frac{1}{\langle\theta\rangle\langle\rho\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\right|\left|\mathrm{Re}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk\right|. \tag{A3.1}$$

Using the gradient transport approximation (6.14) for the potential temperature flux, we find that the ratio ζ of the right-hand side of equation (A3.1) to the left-hand side is

$$\zeta=\frac{\left|\mathrm{Re}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk\right|}{\alpha_\theta(2\pi/k_i)v(\langle P\rangle,i)g\langle\rho\rangle}[1+1/\mathcal{P}e(\langle P\rangle,i)]. \tag{A3.2}$$

Defining the correlation δ between P^F and v_z^F as

$$\delta(\langle P\rangle,i)\equiv\frac{\mathrm{Re}\int v_z(z,\boldsymbol{k})P(z,-\boldsymbol{k})k^2\,d\Omega_k\,dk}{v_z(\langle P\rangle,i)P(\langle P\rangle,i)}, \tag{A3.3}$$

where $|\delta(\langle P\rangle,i)|\le 1$, we find

$$\zeta=\frac{P(\langle P\rangle,i)}{\alpha_\theta(2\pi/k_i)g\langle\rho\rangle}\delta(\langle P\rangle,i)[1+1/\mathcal{P}e(\langle P\rangle,i)]. \tag{A3.4}$$

Now from the Navier-Stokes equation (2.2), the gradient of the fluctuating pressure should be of the same order as $g\rho^F$, or

$$\frac{k_i}{2\pi}P(\langle P\rangle,i)\sim g\rho(\langle P\rangle,i). \tag{A3.5}$$

Thus,

$$\zeta\sim\frac{\rho(\langle P\rangle,i)}{\langle\rho\rangle}\frac{1}{\alpha_\theta}\delta(\langle P\rangle,i)[1+1/\mathcal{P}e(\langle P\rangle,i)]\sim\frac{M^2\delta(\langle P\rangle,i)}{\alpha_\theta}[1+1/\mathcal{P}e(\langle P\rangle,i)]. \tag{A3.6}$$

Therefore, for eddies large enough so that their Péclét number is greater than unity, ζ will be at most of order M^2. For small scales the eddies act like an incompressible flow, and $\mathcal{P}e(\langle P\rangle,i)<1$. In Appendix D we show that the pressure and velocity are out of phase in this case and $\delta(\langle P\rangle,i)\approx 0$.

APPENDIX D

CORRELATION OF $v_z F$ AND P^F

An eddy acts like incompressible flow if the eddy velocity is divergence-free; i.e., if

$$\left|\frac{1}{(2\pi)^{3/2}}\int W(z-z')e^{i\boldsymbol{k}\cdot\boldsymbol{x}'}(\nabla\cdot\boldsymbol{v})^F d^3x'k^2\,d\Omega_k\,dk\right| \ll k_i v(\langle P\rangle,i). \tag{A4.1}$$

Using equation (A2.6) in equation (A4.1), we get

$$\left|\frac{1}{(2\pi)^{3/2}}\int W(z-z')e^{i\boldsymbol{k}\cdot\boldsymbol{x}'}\frac{1}{\Lambda_\rho}v_z^F(\boldsymbol{x}')\,d^3x'k^2\,d\Omega_k\,dk\right| \ll k_i v(\langle P\rangle,i), \tag{A4.2}$$

where Λ_ρ is the mean density scale height. Equation (A4.2) shows that if

$$\Lambda_\rho \gg 2\pi/k_i, \tag{A4.3}$$

then the ith eddy is incompressible.

The velocity at wavenumber k_i is created by density and pressure fluctuations at wavenumber k_i and also by cascade and advection from other eddies. A velocity fluctuation will be temporally correlated with the density or pressure fluctuation that created it. However, the piece of the velocity created by cascade or advection will have a random temporal phase with respect to $\rho(\langle P\rangle,i)$ and $P(\langle P\rangle,i)$. From the Navier-Stokes equation (2.2) we see that a density fluctuation proportional to $\cos kz$ produces a velocity fluctuation proportional to $-\cos kz$. A pressure fluctuation proportional to $\cos kz$ produces a velocity fluctuation proportional to $(1/\langle\rho\rangle)\sin kz$. If $\Lambda_\rho \gg 2\pi/k_i$, then the pressure and velocity will be 90° out of phase, and $\langle v_z^F P^F\rangle$ integrated over the whole volume averages to zero. However, $\langle v_z^F \rho^F\rangle$ integrated over the volume is a large negative quantity.

These correlations can also be understood in mechanical terms. Consider a rising bubble whose center is where v_z is a maximum. At the center of the bubble ρ^F is a minimum—light bubbles rise. The maximum of the pressure, P^F, is not at the center of the bubble. The maximum of P^F is at the base, and the minimum of P^F is at the top of the bubble. The pressure gradient across the bubble forces the bubble to rise. We conclude that the time average of the correlation function δ is zero for (incompressible) eddies where $2\pi/k_i \ll \Lambda_\rho$. Note that v_z^F and $\partial P^F/\partial z$ *are* correlated. This is why the pressure work term $\boldsymbol{v}\cdot\nabla P$ is nonzero and leads to energy equipartition.

APPENDIX E

To show that the expression (5.14) is a divergence, we note that from equation (5.4) that

$$\text{equation (5.14)} = \frac{-1}{(2\pi)^{3/2}}\,\mathrm{Re}\int W(z-z')v_i(z,\boldsymbol{k})\left[\boldsymbol{v}^{\mathrm{L}}(\boldsymbol{x}')\cdot\nabla v_i^{\mathrm{L}}(\boldsymbol{x}')\right]e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}d^3x'\,d^3k$$

$$+\,g\frac{\partial}{\partial\langle P\rangle}\,\mathrm{Re}\int P(z,-\boldsymbol{k})v_z(z,\boldsymbol{k})\,d^3k, \tag{A5.1}$$

where the integration in equation (A5.1) is over all d^3k and we are allowed to drop the superscript F in the first term.

The second term on the right-hand side of equation (A5.1) is already a divergence. Performing the integration over $\boldsymbol{k}$ and replacing $\boldsymbol{v}^{\mathrm{L}}$ with $\boldsymbol{v}^F$ [since $\boldsymbol{v}^{\mathrm{L}} - \boldsymbol{v}^F = O(M^2)\boldsymbol{v}^F$], we obtain for the first term on the right-hand side

$$-\mathrm{Re}\int W(z-z')v_i^F(x')\left[\boldsymbol{v}^F(x')\cdot\nabla v_i^F(x')\right]d^3x'. \tag{A5.2}$$

Integrating by parts and using $\nabla\cdot\langle\rho\rangle\boldsymbol{v}^F \approx 0$, we obtain

$$\mathrm{Re}\int\frac{1}{2}\left(v_i^F\right)^2 v_z^F\langle\rho\rangle\frac{\partial}{\partial z'}\frac{W(z-z')}{\langle\rho(z')\rangle}d^3x'. \tag{A5.3}$$

Changing $\partial W/\partial z'$ to $-\partial W/\partial z$ and using the same approximation that we used in equation (A2.1) to remove $\partial\langle\rho\rangle/\partial z'$ from the integrand, we obtain

$$\frac{\partial}{\partial\langle P\rangle}\,\mathrm{Re}\,g\int\frac{1}{2}\langle\rho(z')\rangle v_i^F(x')^2 v_z^F(x')W(z-z')\,d^3x', \tag{A5.4}$$

which is a perfect divergence.

APPENDIX F

PROOF OF EQUATION (6.5)

From equation (A2.4) we have

$$\frac{g}{\langle\rho\rangle}\langle v_z^F\rho^F\rangle=-\frac{g}{\langle\theta\rangle}\langle\theta^F v_z^F\rangle-\frac{\langle P^F v_z^F\rangle}{\langle\rho\rangle}\left(\frac{1}{\langle\rho\rangle}\frac{\partial\langle\rho\rangle}{\partial z}+\frac{1}{\langle\theta\rangle}\frac{\partial\langle\theta\rangle}{\partial z}\right). \tag{A6.1}$$

Furthermore, from equation (A2.6)

$$\frac{1}{\langle\rho\rangle}\langle v^F\cdot\nabla P^F\rangle=-g\frac{\partial}{\partial\langle P\rangle}\langle v_z^F P^F\rangle+\frac{1}{\langle\rho\rangle^2}\frac{\partial\langle\rho\rangle}{\partial z}\langle P^F v_z^F\rangle. \tag{A6.2}$$

Using equations (A6.1), (A6.2), and (A3.1) we have, to leading order in M^2,

$$\frac{1}{\langle\rho\rangle}\langle v^F\cdot\nabla P^F\rangle+\frac{g}{\langle\rho\rangle}\langle v_z^F\rho^F\rangle=-\frac{g}{\langle\theta\rangle}\langle\theta^F v_z^F\rangle-g\frac{\partial}{\partial\langle P\rangle}\langle v_z^F P^F\rangle. \tag{A6.3}$$

Now to leading order in M^2,

$$\frac{\partial}{\partial\langle P\rangle}\left(gc_p\langle\rho T v_z\rangle-gc_p\langle\rho\rangle\langle T\rangle\langle v_z\rangle_t\right)=\frac{\partial}{\partial\langle P\rangle}\left(gc_p\langle\rho\rangle\langle T\rangle\langle v_z\rangle_s+g\frac{\gamma}{\gamma-1}\langle P^F v_z^F\rangle\right), \tag{A6.4}$$

where we have used equations (2.24) and (2.25). Using the definition of $\langle v_z\rangle_t$ in equation (A6.4), we obtain

$$\begin{aligned}\frac{\partial}{\partial\langle P\rangle}\left(gc_p\langle\rho T v_z\rangle-gc_p\langle\rho\rangle\langle T\rangle\langle v_z\rangle_t\right)&=\frac{\partial}{\partial\langle P\rangle}\left(-gc_p\langle T\rangle\langle v_z^F\rho^F\rangle+g\frac{\gamma}{\gamma-1}\langle P^F v_z^F\rangle\right)\\&=\frac{\partial}{\partial\langle P\rangle}\left(gc_p\frac{\langle T\rangle\langle\rho\rangle}{\langle\theta\rangle}\langle v_z^F\theta^F\rangle+g\langle P^F v_z^F\rangle\right),\end{aligned} \tag{A6.5}$$

where we have used equation (A2.2) to obtain the last line. Combining equations (A6.3) and (A6.5), we obtain equation (6.5).

APPENDIX G

ALTERNATIVE "DERIVATION" OF THE GRADIENT TRANSPORT APPROXIMATION FOR THE POTENTIAL TEMPERATURE FLUX

The potential temperature equation is

$$\left[\frac{\partial\theta^F}{\partial t}\right]_\mathrm{L}=-[v^F\cdot\nabla\theta^F]^F-v_z^F\frac{\partial\langle\theta\rangle}{\partial z}+\left[\frac{\theta}{\rho T}\nabla\cdot(\rho\sigma\nabla T)\right]^F+\nu\frac{\langle\theta\rangle}{2c_p\langle T\rangle}\left[\sigma_{ij}\sigma_{ij}+\left(\frac{2\xi}{\eta}\right)(\nabla\cdot v)^2\right]^F. \tag{A7.1}$$

The equation for the variance $\frac{1}{2}[\theta(\langle P\rangle,i)]^2$ is found by multiplying equation (A7.1) by

$$\frac{1}{(2\pi)^{3/2}}\theta(\langle P\rangle,k)W(z-z')e^{-ik\cdot x'}k^2$$

and integrating over all $d\Omega_k\, d^3x'$ and over dk in the range B_i:

$$\left[\frac{\partial}{\partial t}\right]_{\mathrm{L}}\left[\frac{1}{2}\theta(\langle P\rangle,i)^2\right]=-\frac{\partial\langle\theta\rangle}{\partial z}F_\theta(\langle P\rangle,i)+T_{\mathrm{cas}}(\langle P\rangle,i)+T_{\mathrm{adv}}(\langle P\rangle,i)+T_\sigma(\langle P\rangle,i)+T_\nu(\langle P\rangle,i), \quad \text{(A7.2)}$$

where

$$T_{\mathrm{cas}}(\langle P\rangle,i)+T_{\mathrm{adv}}(\langle P\rangle,i)=-\frac{1}{(2\pi)^{3/2}}\,\mathrm{Re}\int\theta(\langle P\rangle,\boldsymbol{k})W(z-z')e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}k^2\left[\boldsymbol{v}^F(\boldsymbol{x}')\cdot\nabla'\theta^F(\boldsymbol{x}')\right]d\Omega_k\,dk\,d^3x', \quad \text{(A7.3)}$$

$$T_\sigma(\langle P\rangle,i)=\frac{1}{(2\pi)^{3/2}}\,\mathrm{Re}\int\theta(\langle P\rangle,\boldsymbol{k})W(z-z')e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}k^2\left[\frac{\theta}{\rho T}\nabla'\cdot(\rho\sigma\nabla'T)\right]d\Omega_k\,dk\,d^3x', \quad \text{(A7.4)}$$

and

$$T_\nu(\langle P\rangle,i)=\frac{\nu}{(2\pi)^{3/2}}\,\mathrm{Re}\int\theta(\langle P\rangle,\boldsymbol{k})W(z-z')e^{-i\boldsymbol{k}\cdot\boldsymbol{x}'}k^2\frac{\langle\theta\rangle}{2c_p\langle T\rangle}\left[\sigma_{ij}\sigma_{ij}+\left(\frac{2\xi}{\eta}\right)(\nabla'\cdot\boldsymbol{v})^2\right]d\Omega_k\,dk\,d^3x'. \quad \text{(A7.5)}$$

We crudely estimate the advection and cascade terms as

$$T_{\mathrm{cas}}(\langle P\rangle,i)+T_{\mathrm{adv}}(\langle P\rangle,i)\approx v(\langle P\rangle,i)\theta(\langle P\rangle,i)^2k_i. \quad \text{(A7.6)}$$

If $\nu/\sigma \ll 1$, and if we consider the range of scales where $\mathcal{P}_e(\langle P\rangle,i)\sigma/\nu \gg 1$, then

$$|T_\nu(\langle P\rangle,i)| \ll |T_\sigma(\langle P\rangle,i)|. \quad \text{(A7.7)}$$

The thermal source term is estimated as

$$T_\sigma(\langle P\rangle,i)\approx k_i^2\sigma\theta(\langle P\rangle,i)^2. \quad \text{(A7.8)}$$

Now, the source of variance due to the potential temperature flux will produce potential temperature that is correlated with the velocity. The other sources of variance in equation (A7.2) will not produce potential temperature correlated with the velocity. Define a measure of the correlation by

$$\delta_{v\theta}=\frac{\int v(\langle P\rangle,\boldsymbol{k})\theta(\langle P\rangle,-\boldsymbol{k})k^2\,d\Omega_k\,dk}{v(\langle P\rangle,i)\theta(\langle P\rangle,i)}=\frac{F_\theta(\langle P\rangle,i)}{v(\langle P\rangle,i)\theta(\langle P\rangle,i)}. \quad \text{(A7.9)}$$

For small δ, we estimate δ^2 as the ratio of the rate at which correlated potential temperature variance is produced to the rate at which uncorrelated variance is produced:

$$\delta_{v\theta}^2=\frac{-(\partial\langle\theta\rangle/\partial z)F_\theta(\langle P\rangle,i)}{k_iv(\langle P\rangle,i)\theta(\langle P\rangle,i)^2+\sigma k_i^2\theta(\langle P\rangle,i)^2}. \quad \text{(A7.10)}$$

Combining equation (A7.9) with equation (A7.10) and solving for $F_\theta(\langle P\rangle,i)$, we obtain the gradient transport equation,

$$F_\theta(\langle P\rangle,i)=\frac{-(\partial\langle\theta\rangle/\partial z)v(\langle P\rangle,i)/k_i}{1+1/\mathcal{P}_e(\langle P\rangle,i)}. \quad \text{(A7.11)}$$

REFERENCES

Bell, T. L., and Nelkin, M. 1977, *Phys. Fluids*, **20**, 345.
______. 1978, *J. Fluid, Mech.*, **88**, 369.
Bohm-Vitense, E. 1958, *Zs. Ap.*, **46**, 108.
Cloutman, L. D., and Whitaker, R. W. 1980, *Ap. J.*, **237**, 900.
Cox, J. P., and Giuli, R. T. 1968, *Principles of Stellar Structure*, Vol. **1** (New York: Gordon & Breach).
Gough, D. O. 1969, *J. Atmos. Sci.*, **26**, 448.
Gough, D. O., Moore, D. R., Spiegel, E. A., and Weiss, N. O. 1976, *Ap. J.*, **154**, 1005.
Henyey, L. G., Vardya, M. S., and Bodenheimer, P. L. 1965, *Ap. J.*, **142**, 841.
Kolmogorov, A. N. 1941*a*, *C. R. Acad. Sci., U.R.S.S.*, **30**, 299.
______. 1941*b*, *C. R. Acad. Sci., U.R.S.S.*, **31**, 538.
Landau, L. D., and Lifshitz, E. M. 1959, *Fluid Mechanics* (London: Pergamon).
Latour, J., Spiegel, E. A., Toomre, J., and Zahn, J.-P. 1976, *Ap. J.*, **207**, 233.
Marcus, P. S. 1978, Ph.D. thesis, University of Michigan, Ann Arbor.
______. 1979, *Ap. J.*, **231**, 176.
______. 1980*a*, *Ap. J.*, **239**, 622.
______. 1980*b*, *Ap. J.*, **240**, 203.
Miles, J. W., and Howard, L. N. 1964, *J. Fluid Mech.*, **20**, 331.
Ogura, Y., and Charney, J. G. 1962, *Proc. Internat. Symposium Numerical Weather Prediction*, Tokyo, p. 431.
Press, W. H. 1981, *Ap. J.*, **245**, 286.
Press, W. H., and Rybicki, G. B. 1981, *Ap. J.*, **248**, 751.
Roxburgh, I. W. 1965, *M.N.R.A.S.*, **130**, 223.
Saslaw, W. C., and Schwarzschild, M. *Ap. J.*, **142**, 1468.
Shaviv, G., and Salpeter, E. E. 1973, *Ap. J.*, **184**, 191.
Spiegel, E. A. 1965, *Ap. J.*, **141**, 1068.
Straus, J. M., Blake, J. B., and Schramm, D. N. 1976, *Ap. J.*, **204**, 281.
Tennekes, H., and Lumley, J. L. 1972, *A First Course in Turbulence* (Cambridge: MIT Press).
Toomre, J., Zahn, J.-P., Latour, J., and Spiegel, E. A. 1976, *Ap. J.*, **207**, 545.
Turner, J. S. 1973, *Buoyancy Effects in Fluids* (Cambridge: Cambridge University Press).
Ulrich, R.K. 1976, *Ap. J.*, **207**, 564.
Veronis, G. 1963, *Ap. J.*, **137**, 641.
Vitense, E. 1953, *Zs. Ap.*, **32**, 135.
Weyman, R., and Sears, R. L. 1965, *Ap. J.*, **142**, 174.

P. S. MARCUS: Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139

W. H. PRESS: Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138

S. A. TEUKOLSKY: Newman Laboratory, Cornell University, Ithaca, NY 14853