ICARUS 69, 387–422 (1987)

The Role of Turbulent Convection in the Primitive Solar Nebula

I. Theory

W. CABOT, V. M. CANUTO, AND O. HUBICKYJ

NASA Goddard Space Flight Center, Institute for Space Studies, 2880 Broadway, New York, New York 10025

AND

J. B. POLLACK

Space Science Division NASA–Ames Research Center, Moffett Field, California 94035

Received May 14, 1986; revised October 15, 1986

The theoretical framework for modeling the primordial solar nebula is presented in which convection is assumed to be the sole source of turbulence that causes the nebula to evolve. We use a new model of convective turbulence that takes into account the important effects of radiative dissipation, rotation, and anisotropy of convective motions. This model is based on a closure for the nonlinear interactions that employs the growth rates of hydrodynamic instabilities, a procedure that allows one to compute turbulence coefficients for instabilities other than convection. The vertical structure equations in the thin-disk approximation are developed for this new model, and a detailed comparison and critique of previous convective models of the solar nebula are presented. Numerical results are presented in a subsequent paper.

I. INTRODUCTION

In recent years, it has become increasingly clear that the origin, structure, and evolution of planetary systems are but the last in a sequence of events that are thought to begin with the collapse of a molecular cloud which fragments to form protostars surrounded by nebulae. The latter, due to rotation, assumes a disklike structure. Since the complete chain of events is clearly too complex to handle, one separates it into different parts in the hope that the real time sequence of events does not make this separation too unrealistic.

The "solar nebula" represents an intermediate phase in which the major period of infall from the collapsing cloud is over, but during which the gas and the grains forming the bulk of the nebula are still well mixed. The grains will ultimately drift toward midplane, where, upon becoming Jeans unstable, will give rise to the formation of protoplanets. While some insight has been recently gained for the latter stage, the evolution of the solar nebula is far from being understood. The basic problem is that of finding a process to "clear" the nebula, i.e., one that removes the gas by causing it to drift outward as well as inward toward the Sun. Since it is assumed that the solar nebula is not acted upon by external forces, one must search for an internally generated mechanism capable of initiating the drifting process. The problem is not an easy one, since the gas is locked in Keplerian orbits, which, if undisturbed, would remain so indefinitely.

To break the Keplerian motion, one may call upon the presence of large viscosities. Since, however, molecular forces are far too weak to produce the required effects, one usually resorts to "dynamical" processes to obtain an enhanced viscosity. This is where turbulence comes into play. It is known that if a fluctuating velocity is su-

0019-1035/87 $3.00

perimposed on a mean flow, the equations of the latter are affected by the former in a way that *may* be represented by an "enhanced viscosity." If $\mathbf{U}$ represents the mean flow velocity and $\mathbf{u}$ the turbulent or fluctuating velocity, the equations for U_i are ($\partial_j \equiv \partial/\partial x_j$) (Hinze, 1959; summation over repeated indices is implied)

$$\frac{\partial U_i}{\partial t} + U_j \partial_j U_i = -\frac{1}{\rho}\partial_i p + \partial_j(\nu S_{ij} + \tau_{ij}). \tag{1}$$

Here, S_{ij} is the ordinary stress tensor, ν is the molecular viscosity, and τ_{ij} (the Reynolds stress) is an additional stress caused by the fluctuating velocities, i.e.,

$$S_{ij} = \partial_i U_j + \partial_j U_i, \quad \tau_{ij} = -\overline{u_i u_j}. \tag{2}$$

Only if τ_{ij} and S_{ij} have the same sign will momentum be transmitted from the faster moving parts of the fluids to the slower ones. This can also be seen in a different way by considering the equations for the energy of the mean flow $U^2/2$, for the energy of the turbulent flow $u^2/2$, and for the total energy flux F. Calling $\langle \ldots \rangle$ the terms that have no bearing on the present argument, we derive from Monin and Yaglom (1971),

$$\frac{\partial U^2}{\partial t} = -\tau_{ij}S_{ij} + \langle \ldots \rangle, \tag{3}$$

$$\frac{\partial u^2}{\partial t} = +\tau_{ij}S_{ij} + \langle \ldots \rangle, \tag{4}$$

$$\frac{\partial F}{\partial z} = \frac{1}{2}\rho\tau_{ij}S_{ij}. \tag{5}$$

For turbulence to act as a source of energy (fed by the mean flow), one sees from Eq. (5) that one must have

$$\tau_{ij}S_{ij} > 0, \tag{6}$$

i.e., τ_{ij} and S_{ij} must have the same sign. Since for a thin disk $S_{r\phi}$ is the largest component of the tensor S_{ij} in cylindrical coordinates (r, ϕ, z),

$$S_{r\phi} = r\frac{\partial \Omega}{\partial r} < 0 \tag{7}$$

for any $\Omega \sim r^{-n}$ ($n = \frac{3}{2}$ for Keplerian motion), it follows that condition (6) obtains

$$\tau_{r\phi} < 0. \tag{8}$$

To make further progress, one must supplement Eq. (5) with independent information about $\tau_{r\phi}$; Eqs. (3) and (4) cannot be solved since the number of unknowns exceeds the number of equations, the well known closure problem.

Historically, Boussinesq (1877, 1897) was the first to propose that if one assumes that turbulence is generated by shear in the mean flow, then one may write

$$\tau_{ij} = \nu_t S_{ij}, \tag{9}$$

where ν_t plays the role of a turbulent viscosity. If one further assumes that

$$\nu_t > 0, \tag{10}$$

turbulence is fed by the mean flow. The original Boussinesq model is, however, unable to say anything about ν_t; all it does in practice is substitute one unknown τ_{ij}, with another, ν_t.

The next step was taken by Prandtl (1925; see, however, earlier ideas by Taylor, 1915) who introduced "the momentum transfer theory," whereby an explicit form for the viscosity ν_t was suggested, namely (a, b are two chosen components)

$$\nu_t = l^2|S_{ab}|, \tag{11}$$

where l represents the length traveled by an eddy before it releases its momentum to the surrounding medium. The relevant new aspect of this model is that fluctuating quantities have been substituted by a length that characterizes the scale of turbulence. Since the Prandtl model does not fix the value of this length, most of the discussions usually concentrate on ways to determine l. From a fundamental point of view, the value of l is not the most interesting part. The key point in Eq. (11) is that one has *assumed* the absolute value of the stress tensor S_{ij} so as to satisfy Eq. (10), *instead of proving it*. In fact, contrary to the kinematic viscosity ν, ν_t does not represent a physical property of the fluid, but rather characterizes statistical properties of fluctuations; it need not be positive (energy may be drained from turbulence to the mean flow). To postulate a positive ν_t is equivalent to postulating that one (or more) process exists that triggers turbulence. This basic fact should be proven rather than postulated, especially in the case of the solar nebula for which we do not have direct (or indirect) evidence that turbulence existed. Equation (11) is a useful "engineering" method to describe a turbulent flow when one knows that it exists, as is the case of laboratory turbulence, where one can measure ν_t directly. In the case of the solar nebula, one simply postulates its existence. Since turbulence is the putative mechanism for clearing the solar nebula, the lack of proof of its existence is highly unsatisfactory. The situation is, perhaps, even worse than that.

Let us go back and inquire whether Eq. (9) is justified in the case of rotating disks. The doubt arises because one notices that the situations for which Eq. (9) was originally intended correspond to plane parallel flows where the shear $S_{ij} = U_{i,j} + U_{j,i}$ is given in terms of the linear momenta in the i and j directions. How does one translate this to the case of circular motion? One may argue that the role played by the linear momentum ought to be replaced by the angular momentum in the circular case, which would lead one to propose (Von Karman, as cited in Safronov, 1969)

$$\tau_{r\phi} = \nu_t \frac{1}{r} \frac{\partial}{\partial r} (r^2\Omega) \tag{12}$$

rather than

$$\tau_{r\phi} = \nu_t r \frac{\partial \Omega}{\partial r} \tag{13}$$

as dictated by Eq. (9). The implications of Eqs. (12) and (13) may be drastically different. In fact, while Eq. (13) always satisfies Eq. (8) for $\Omega \sim r^{-n}$, (12) does so only if

$$\frac{\partial}{\partial r} (r^2\Omega) < 0, \tag{14}$$

which is the well-known Rayleigh condition for shear instability (Hunter, 1972). Keplerian disks do not satisfy condition (14) (Mestel, 1972; Pringle, 1981) and the use of Eq. (12) would lead to a violation of condition (6); more precisely, one would feed the mean motion at the expense of turbulence.

Ter Haar (1972) suggested that one should not actually apply Eq. (14) to disks since they do not satisfy the conditions underlying it, namely that the medium be infinite and homogeneous. Rather, he suggested that Rayleigh's criterion for shear instability in a plane parallel, inviscid fluid would be more appropriate (the velocity profile must have a point of inflection). For cylindrical geometry, the result is (Spiegel and Zahn, 1970)

$$\frac{d}{dr}\left[\frac{\rho(r)}{r} \frac{d}{dr} (r^2\Omega)\right] = 0 \tag{15}$$

or, for Keplerian rotation, $(\rho r^{-3/2})' = 0$. None of the published $\rho(r)$ vs r relations satisfies this condition.

This situation has however not prevented astrophysicists from carrying out extended calculations on accretion disks (Pringle, 1981). Turbulence is assumed to have arisen through some unspecified source(s) (Eqs. (9) and (6) are assumed to hold), and ν_t is written on purely dimensional grounds in terms of the local sound speed c_s and a vertical scale height h as

$$\nu_t = \alpha c_s h, \tag{16}$$

where α, *assumed* to be positive to satisfy Eq. (10), is fitted to the data in analogy with laboratory turbulence, in spite of the profound conceptual differences between the two situations.

In 1969, Cameron first suggested that convective instabilities caused by the nebula's opacity may have been the source of large turbulent viscosities (see also Cameron, 1978). Lin and Papaloizou (1980) carried out disk calculations employing an approximate version of the mixing length theory (MLT) to treat convective turbulence (for a critical analysis of this work, see Sect. VII).

Since the MLT may in principle provide a physical model to express properties of the turbulent flow in terms of temperature gradients, radiative losses, buoyancy, etc., one may hope to simultaneously be able to quantify both the candidacy of convection as a generator of turbulence as well as the ensuing disk properties (height, mass, etc.). *If* this process could be implemented, it would constitute a considerable improvement over the α model, which postulates turbulence that it then cannot describe. We must stress the "if," for there is no such thing as a "ready-made" MLT set of formulae that can dependably be applied to disks. Actually, the solution of the disk equations is the easiest part. The really difficult part consists of setting up (using the general prescription of the MLT philosophy) a set of formulae that account for all the physical effects that characterize the disk.

We shall show in Section VII that the straightforward application to disks of the MLT formulae used in stellar problems leads to conclusions about the role of convection *and* the values of disk properties significantly different from the ones that we have obtained. One of the main reasons will be shown to be our proper inclusion of anisotropy effects in the description of the large eddies, something missing in the standard MLT used in stellar structure calculations.

The main problem in using the MLT stems from the fact that, in spite of being called a "theory," it has been for many years no more than a phenomenological method without a derivation from first principles. That made the inclusion of new physical effects a difficult process at best, and one whose uniqueness is never certain. For example, as early as 1950, Öpik pointed out that radiative losses by traveling eddies would strongly reduce the efficiency of heat transport, and yet the phenomenological nature of the MLT formalism did not allow this important effect to be incorporated until very recently (e.g., Gough, 1976, 1978). Effects due to rotation are still largely unaccounted for, a problem that while not too serious for stellar structure computations becomes very important when dealing with a disk rotating at the Keplerian rate. Finally, anisotropy effects in the eddies' sizes, that can significantly reduce the efficiency of convection (which may not be very relevant in spherical geometries, but is so in disks), have been particularly difficult to incorporate into the MLT. It was only in 1978 that Gough provided the correct expressions.

While the above list touches upon the main effects that must be included (radiative losses, rotation, and anisotropy), it serves the purpose of illustrating a larger point: In the absence of an a priori derivation of the MLT, one cannot feel confident that effects germane to a particular problem have all been taken into account.

In conclusion, if one wants to reliably quantify the importance of convection as a source of turbulent viscosity, one cannot rely on an MLT formalism that, while correct for stellar structure, leaves out effects that are important in the specific case of disks. Moreover, the formalism employed should not contain free parameters (the model analyzed in Sect. VII contains four) because, by fixing them to yield the desired results, one defeats the very purpose of checking the role of convection as a source of viscosity. One actually reduces the problem to the level of the α model where one forces the strength of turbulence to fit the data. Finally, the formalism should contain all the effects that may be relevant to disks (rotation, radiative losses, eddy anisotropies, etc.) whose relative importance will be determined by the problem itself and not by a priori judgement of what is relevant and what is not; e.g., we shall show that a major role is played by the eddy anisotropy. It is only by following this protocol that one can hope to decide whether or not convection is an appropriate mechanism for generating the needed ν_t.

Recently, a model for large-scale turbulence that satisfies the above requirements has been proposed (Canuto and Goldman, 1985, referred to hereafter as CG). The model is based on a new treatment of the nonlinear interactions (the closure problem) that depend, like the energy source, on the specific instability assumed to generate the fluctuations. When applied to the case of convective instability, the results of the CG model were tested against laboratory and astrophysical data. In the first instance, the model reproduces measured convective fluxes (in water) up to a Rayleigh number of 10^{12}. In the astrophysical case, the CG model reproduces, as a particular case, the results of the MLT and can therefore be considered the first derivation of MLT from a model for nonlinear interactions among eddies, i.e., a model for large-scale turbulence that follows and extends the spirit of the Kolmogoroff–Heisenberg model (valid only for small-scale turbulence). Effects due to rotation, magnetic fields, conductive and radiative losses, eddy-size anisotropies etc., can be included in a natural way. While bulk properties (like fluxes) confirm the MLT results, the CG model is much richer in that it provides the energy spectral function that describes how turbulent energy is distributed among eddies of different sizes (MLT is a one mode model).

The availability of a new model of turbulence (Sect. III) and the availability of an improved model for grain opacity (Sect. III in Cabot *et al.*, 1987, Paper II, hereafter) were the basic motivation for the work presented in this paper. The construction of the disk model and the ensuing results are presented in Paper II. The differences between our model and results with those of Lin, Papaloizou, and Bodenheimer (LPB, hereafter) are discussed in Section VII.

II. THE ENERGY EQUATION FOR THE DISK

One of the basic equations characterizing the structure of a thin disk is the energy equation (Pringle, 1981),

$$\frac{dF}{dz} = \tau_{r\phi} S_{r\phi}, \tag{17}$$

where

$$\tau_{r\phi} = \rho\nu_t S_{r\phi}, \quad S_{r\phi} = r\frac{d\Omega}{dr}. \tag{18}$$

Here, ν_t is the turbulent viscosity and F the total flux. While Eq. (17) is used in all disk computations, few state that it is not the exact energy equation but rather a particular form of it. The approximation under which Eq. (17) holds consists of having changed the "ensemble average" (over the turbulent quantities) with a "volume average." It follows that once a model for turbulence has been chosen providing an expression for the turbulent viscosity, one must still carry out the "volume average" so as to obtain an expression for ν_t that can be meaningfully substituted in Eq. (17). Only then would one have an expression for ν_t compatible with the constraints under which Eq. (17) is valid. For example, the well-known α model for ν_t,

$$\nu_t = \alpha c_s h, \tag{19}$$

does comply with the "volume average" requirement although it is incapable of fixing the value of α. On the other hand, the mixing length theory does provide an expression for ν_t that is in principle superior to that of the α model, but which is not "volume averaged" and therefore cannot be directly employed in Eq. (17).

We shall now derive Eq. (17). Following Stewart (1976), we first write the exact equations representing conservation of mass, momentum, and energy:

$$\dot{\rho} + (\rho v_i)_{,i} = 0, \tag{20}$$

$$(\rho v_i) + (\rho v_i v_j)_{,j} = \rho\Psi_{,i} - p_{,i} + t_{ij,j}, \tag{21}$$

$$(\rho\varepsilon) + (\rho\varepsilon v_i)_{,i} = -p v_{i,i} + \tfrac{1}{2} t_{ij}\sigma_{ij} - F^{\text{cond}}_{i,i}, \tag{22}$$

where we employ the notation $\dot{A} = \partial A/\partial t$, $A_{,i} = \partial A/\partial x_i$. In Eqs. (20)–(22), Ψ is the gravitational potential, t_{ij} is the molecular viscous stress tensor, σ_{ij} is the strain tensor, ε is the internal energy per unit mass, v_i is the total velocity, and $F_i^{\rm cond}$ is the conductive flux,

$$F_i^{\rm cond} = -KT_{,i}, \tag{23}$$

where K is an effective thermal (including radiative) conductivity. We follow the standard procedure and separate the total velocity v_i into two components,

$$v_i = U_i + u_i, \tag{24}$$

corresponding to the mean flow, U_i, and the fluctuating or turbulent flow, u_i, defined by (an overbar denotes ensemble average)

$$\overline{\rho u_i} = 0, \quad U_i = \frac{\overline{\rho v_i}}{\bar{\rho}}, \tag{25}$$

appropriate for a compressible fluid. Taking the ensemble average of Eqs. (20) and (21), we obtain

$$\dot{\bar{\rho}} + (\bar{\rho} U_i)_{,i} = 0 \tag{26}$$

and

$$\bar{\rho} D_t U_i = \bar{\rho}\Psi_{,i} - \bar{p}_{,i} + \bar{\tau}_{ij,j}, \tag{27}$$

where $D_t \equiv D/Dt = \partial/\partial t + U_i\partial/\partial x_i$, and τ_{ij} is the Reynolds stress tensor,

$$\tau_{ij} = -\overline{\rho u_i u_j}. \tag{28}$$

Next, consider Eq. (22). Separating the variables as

$$\varepsilon = e + E, \quad E = \frac{\overline{\rho\varepsilon}}{\bar{\rho}}, \quad \sigma_{ij} = s_{ij} + S_{ij}, \quad t_{ij} = t_{ij}^* + T_{ij} \tag{29}$$

and taking the ensemble average, we obtain (we shall take $T_{ij} = 0$ since it is proportional to molecular viscosity)

$$\bar{\rho} D_t E = -\bar{p}\, U_{i,i} - \overline{p u_{i,i}} - \overline{(\rho e u_i)_{,i}} + \tfrac{1}{2} t_{ij}^* s_{ij} - \overline{F_{i,i}^{\rm cond}}. \tag{30}$$

Using (26) and $E = c_{\rm v}\bar{T}$, we obtain

$$\bar{\rho} c_{\rm v}\left(D_t\bar{T} - \frac{\bar{p}}{\bar{\rho}^2 c_{\rm v}} D_t\bar{\rho}\right) = -\overline{p u_{i,i}} - \overline{(\rho e u_i)_{,i}} + \frac{1}{2} t_{ij}^* s_{ij} - \overline{F_{i,i}^{\rm cond}}. \tag{31}$$

Next, let us add and subtract the quantity $\overline{u_i p_{,i}}$. Grouping terms, we obtain for the right-hand side of Eq. (31)

$$-\overline{[(p + \rho e)u_i + F_i^{\rm cond}]_{,i}} + \overline{u_i p_{,i}} + \tfrac{1}{2} t_{ij}^* s_{ij}. \tag{32}$$

Since

$$\varepsilon = c_{\rm v} T, \quad e = \varepsilon - E = c_{\rm v}(T - \bar{T}) = c_{\rm v}\theta, \tag{33}$$

we note that the first term in the square bracket is just the convective flux,

$$F_i^{\rm conv} = c_{\rm p}\overline{\rho u_i\theta}, \tag{34}$$

so that calling F the total flux, $F^{\rm cond} + F^{\rm conv}$, Eq. (32) becomes

$$\bar{\rho} c_{\rm v}\left(D_t\bar{T} - \frac{\bar{p}}{\bar{\rho}^2 c_{\rm v}} D_t\bar{\rho}\right) = -\overline{F_{i,i}} + \overline{u_i p_{,i}} + \frac{1}{2} t_{ij}^* s_{ij}. \tag{35}$$

This equation cannot be further simplified unless we use some additional information. For that, we go back to the equation of motion and, using the same general procedure, derive the equation satisfied by the turbulent kinetic energy. Calling

$$u'^2 \equiv \frac{\overline{\rho u^2}}{\bar{\rho}}, \tag{36}$$

the result is

$$\bar{\rho} D_t(\tfrac{1}{2}u'^2) = -\overline{u_i p_{,i}} - \tfrac{1}{2} t_{ij}^* s_{ij} + \tfrac{1}{2}\tau_{ij} S_{ij} - \overline{[(-t_{ij} + \tfrac{1}{2}\rho u_i u_j)u_i]_{,j}}. \tag{37}$$

At this point, one introduces the basic assumption alluded to earlier, namely one replaces the "ensemble average" with a "volume (i.e., a $z - \phi$) average." Once the divergence theorem is used to eliminate the last term in (37), one obtains, for stationary homogeneous turbulence,

$$\overline{u_i p_{,i}} = -\tfrac{1}{2} t_{ij}^* s_{ij} + \tfrac{1}{2}\tau_{ij} S_{ij}. \tag{38}$$

Once this relation is substituted in (35), one finally obtains

$$\bar{\rho} c_v\left(D_t T - \frac{\bar{p}}{c_v \bar{\rho}^2} D_t \bar{\rho}\right) = -F_{i,i} + \frac{1}{2} \tau_{ij} S_{ij}. \tag{39}$$

For stationary, incompressible, axisymmetric turbulence, the left-hand side of Eq. (39) vanishes, reducing to Eq. (17) in the case of a thin disk.

In conclusion, the correct way of carrying out a disk structure computation would require the use of the exact energy flux equation (35). For that we would need to know the turbulence terms appearing on the right-hand side of Eq. (35). This in turn calls for the solution of the dynamical equation for the turbulent kinetic energy (Eq. (37); or even more generally the equation for the Reynolds tensor, τ_{ij}), which is equivalent to solving the full problem of turbulence. In fact, Eq. (37) for $\overline{u^2}$ contains the nonlinear $\overline{u^3}$ term which depends on $\overline{u^4}$, and so forth. This infinite set of equations is the essence of the problem of turbulence whose solution calls for a truncation, the so-called closure problem. Such a program has never been undertaken in connection with the physics of disks. One avoids the need for a theory of turbulence by artificially introducing an average process not contained in the original equations. One makes use of the fact that the nonlinear terms, which determine the spectrum of turbulence by their transfer of energy among eddies of different sizes, do not contribute to the total energy budget. Their action is strictly that of redistributing energy.

In order to use this property, one forcibly alters the exact equation by replacing the local terms in Eq. (37) with their volume average. This is equivalent to going from a local dynamical equation, relating quantities at different points, to a relation essentially between numbers, Eq. (38). Alternatively, one substitutes a differential equation with its first integral, a procedure that impoverishes the problem considerably. Turbulent properties, like the eddy energy spectrum, are therefore lost. In practical terms, this means that a local ν_t provided by a possible theory of turbulence (which would also depend on the size of the eddy to which it refers, i.e., $\nu_t = \nu_t(k)$) is not to be identified with the ν_t appearing in Eq. (17). The volume average must first be carried out on any such local expression.

While the exact prescription for carrying out the averaging process is lacking, we shall see (Sect. VI) that a physically reasonable procedure can be devised. We shall also show (Sect. VII) that the LPB model uses an expression for ν_t that is not volume averaged and which is therefore not consistent with Eq. (17) in which it is used.

III. THE MODEL FOR TURBULENCE

The mathematical structure of the model of turbulence that will be used in this paper has been presented in detail elsewhere (Canuto and Goldman, 1985; referred to as CG) and only the main results will be presented here.

The fundamental quantity is the turbulent energy spectral function $E(k)$, defined in terms of the turbulent velocity v_t as

$$v_t^2(k) = \int_k^\infty E(q)dq. \tag{40}$$

$E(k)/2$ describes how turbulent energy (per unit mass) is distributed among eddies of different sizes. The size l of an eddy is defined as π/k; k_0 is the smallest wavenumber allowed by the system and corresponds to the largest eddy. For medium- to small-size eddies, the Kolmogoroff spectrum obtains $E(k) \sim k^{-5/3}$, i.e., $v_t \sim l^{1/3}$. The function $E(k)$ satisfies a nonlinear integral equation which expresses how the energy from the source is distributed, via nonlinear interactions, to eddies of different sizes, as well as dissipated into heat by molecular processes. To derive the form of $E(k)$, one must first quantify the nonlinear transfer process (the so-called "closure problem"). The closure proposed in the CG model, which applies to the large-scale eddies, allows the nonlinear integral equation for $E(k)$ to be solved analytically, with the result ($' = d/dk$)

$$-2\gamma_* k^2 E(k) = \left[kn^{1/2} \int_{k_0}^{k} kn^{1/2}(nk^{-2})' \, dk \right]', \tag{41}$$

where $n(k)$ is the growth rate of the instability that generates turbulence. The parameter γ_* is given from within the theory as

$$-2n(k_0)L_p^2 = \gamma_* k_0 (n(k)k^{-2})'_{k_0} \tag{42a}$$

where L_p is the longitudinal integral scale,

$$4L_p = 3\pi \left(\int_{k_0}^{\infty} k^{-1}E(k)dk \right) \left(\int_{k_0}^{\infty} E(k)dk \right)^{-1}. \tag{42b}$$

The model predicts an expression for ν_t in terms of the growth rate $n(k)$ calculated at the largest eddy, i.e. (see also Canuto *et al.*, 1984, cited hereafter as CGH),

$$\nu_t = n(k_0)k_0^{-2}. \tag{43}$$

Since large-scale turbulence is not expected to be isotropic, an anisotropy factor x is introduced (Yamaguchi, 1963),

$$x \equiv \frac{k_x^2 + k_y^2}{k_z^2}, \tag{44}$$

which is related to k_0 and d (the depth of the convective layer) by the expression

$$k_0 d = \pi(1 + x)^{1/2}. \tag{45}$$

As discussed in CG, the anisotropy x is determined by maximizing the growth rate computed at $k = k_0$,

$$\frac{dn(k_0)}{dx} = 0. \tag{46}$$

(See Appendix A for details.) Finally, the convective flux, $F_c = c_p\rho\langle\theta w\rangle$, where θ is the fluctuating temperature and w the z component of the turbulent velocity, is given by the expression (see Appendix B)

$$F_c = c_p\rho(g\alpha)^{-1} \int_{k_0}^{\infty} [n(k, \Omega) + \nu k^2]E(k)dk, \tag{47}$$

where α is the coefficient of thermal expansion, $-(\partial \ln \rho/\partial T)_p$, and where ν is the molecular viscosity. As shown in Appendix B, Eq. (47) is valid for *any degree* of rotation whose effects enter through the growth rate $n(k, \Omega)$.

In summary, once the form of the growth rate is known, both ν_t and F_c are fully determined.

(*a*) *The Growth Rate: No Self-Gravity*

A general treatment of the stability conditions for a differentially rotating disk has been carried out by Goldreich and Schubert (1967). In the case of Keplerian rotation, and in the absence of self-gravity, $n(k)$ is given by (see CGH), with $k^2 = k_r^2 + k_z^2$,

$$n^3 + \chi k^2 n^2 + \left(\frac{k_z^2}{k^2}\Omega^2 - \frac{k_r^2}{k^2} g\alpha\beta \right) n + \chi k_z^2 \Omega^2 = 0, \tag{48}$$

where β is the superadiabatic temperature gradient

$$\beta \equiv \left(\frac{dT}{dz}\right)_{\text{ad}} - \left(\frac{dT}{dz}\right), \tag{49}$$

and χ is the thermometric conductivity. For our purposes, only the radiative conductivity is used, as defined in Eq. (8) in Paper II. In the case of the solar nebula, one can neglect the kinematic viscosity ν, which is typically $\sim 10^{-8}\,\chi$ (Lang, 1980).

The cubic equation for $n(k)$ in Eq. (48) has two roots with positive real parts, each corresponding to unstable modes. To represent the total contribution of the convective instability, we follow previous authors in using the sum of the roots with positive real parts in computing F_c and ν_t. The unstable modes (n_1 and n_2) are either real or complex conjugates; thus their sum ($n = n_1 + n_2$) is real and given by the positive root of the cubic (see Appendix A)

$$n^3 + 2\chi k^2 n^2 - \left(g\alpha\beta \frac{x}{1 + x} - \frac{\Omega^2}{1 + x} - \chi^2 k^4 \right) n - g\alpha\beta \frac{x}{1 + x} \chi k^2 = 0. \tag{50}$$

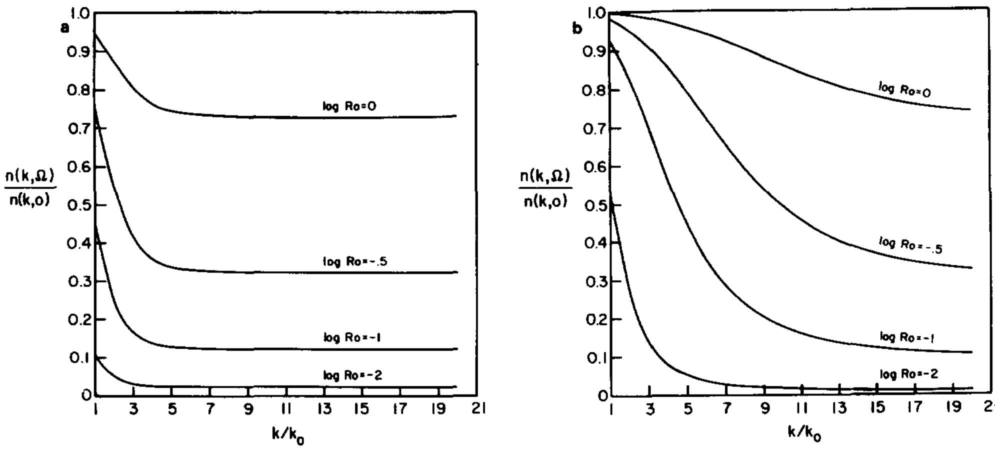

FIG. 1. The ratio of the convective growth rate $n(k)$ with rotation to the growth rate without rotation as a function of wavenumber for Rossby numbers ranging from 0 to 0.01 for (a) $S = 2.4 \times 10^5$ and (b) 2.4×10^{10}; S is defined in Eq. (52); Ro is defined in Eq. (55).

The ratio $n(k, \Omega)/n(k, 0)$ and $E(k)$ are shown in Figs. 1 and 2, respectively. In both cases, we note the stabilizing effect of Keplerian rotation.

(*b*) *Convective Flux for the* $\Omega = 0$ *Case*

In this case, Eq. (48) reduces to the well-known expression first derived by Rayleigh, i.e., with $k = k_0 q$,

$$n(k) = (1 + \sigma) \sqrt{g\alpha\beta \frac{x}{1 + x} \frac{1}{(bS)^{1/2}}} \times \left[\sqrt{1 - \mu + \frac{bS}{(1 + \sigma)^2} \frac{1}{q^4}} - 1\right] q^2, \quad (51)$$

where σ $(= \nu/\chi)$ is the Prandtl number, $\mu = 4\sigma\,(1 + \sigma)^{-2}$, and

$$S = g\alpha\beta d^4\chi^{-2}, \quad b = 4x(1 + x)^{-3}\pi^{-4}. \quad (52)$$

Here, $S = \sigma\mathrm{Ra}$, where Ra is the Rayleigh number, which represents the ratio between the thermal dissipation time, $t_\chi = d^2/\chi$, times the viscous dissipation time, $t_\nu = d^2/\nu$, and the free-fall time, $t_g \equiv (g\alpha\beta)^{-1/2}$, i.e., $t_\chi t_\nu / t_g^2$. When Eq. (51) is inserted in (41), the integration can be performed analytically (see CG). The resulting spectral function $E(k) \equiv E(k, 0)$ can then be inserted in Eqs. (40) and (47). For $\sigma \ll 1$, the integrations can again be performed analytically, with the results (CG)

$$F_c = c_p \rho \frac{a}{S} (\sqrt{1 + bS} - 1)^3 \beta\chi,$$
$$v_t = c\chi(\sqrt{1 + bS} - 1)d^{-1}, \quad (53)$$

where the coefficients a and c depend on x as

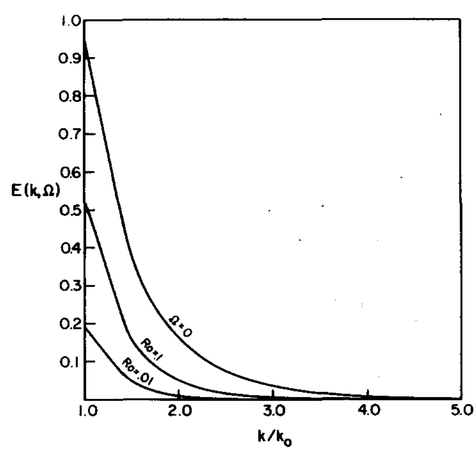

FIG. 2. The spectral function $E(k, \Omega)$ in units of $k_0\chi^2(bS)/4\gamma_*$ for turbulent convection as a function of wavenumber for the case of no rotation and different Rossby numbers and $S = 2.4 \times 10^5$.

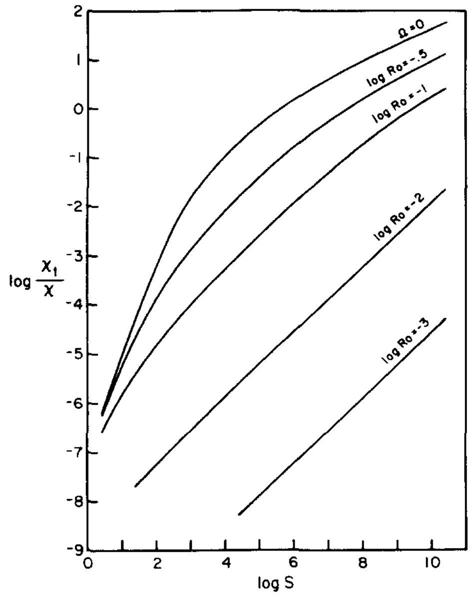

FIG. 3. The turbulent conductivity χ_t in units of the radiative conductivity χ (see Eq. (57)) as a function of S for the no rotation case and increasing rotation characterized by Rossby numbers $10^{-0.5}$, 10^{-1}, 10^{-2}, and 10^{-3}.

$$a = a_0(1 + x)^2, \quad c = c_0(1 + x)^{1/2}. \quad (54)$$

As discussed previously, *Eqs. (53) have the same structure as the expressions derived from the mixing length theory* (see (h) below).

(c) *The $\Omega \neq 0$ Case: Convective Fluxes and Velocities*

In this case, $n(k, \Omega)$ is given by Eq. (50). Since one can no longer solve Eqs. (41) and (47) analytically, numerical integrations are necessary. In Figs. 3 and 4, we present F_c and v_t vs S for different values of the Rossby number Ro, defined as

$$\mathrm{Ro} = \left(\frac{\mathrm{Ra}}{\sigma \mathrm{Ta}}\right)^{1/2} = \frac{\sqrt{g\alpha\beta}}{\Omega}, \quad (55)$$

where Ta is the Taylor number,

$$\mathrm{Ta} = d^4\Omega^2/\nu^2; \quad (56)$$

$(\mathrm{Ta})^{1/2}$ is the ratio of the dissipational time scale d^2/ν to the rotational time scale Ω^{-1}.

The effect of rotation in reducing F_c and v_t is clearly seen in Figs. 3 and 4. These general results, valid for arbitrary values of S and Ro, will be used in the disk calculations in Paper II where the values of S and Ro will be determined self-consistently.

(d) *Turbulent Viscosity for the $\Omega \neq 0$ Case*

The turbulent viscosity ν_t vs S is presented in Fig. 5 and Table I in units of the thermometric conductivity χ.

(e) *Turbulent Prandtl Number for the $\Omega \neq 0$ Case*

Equation (47) can be written in analogy with $F_{\mathrm{rad}} \sim \chi dT/dz$ as

$$F_c = c_p\rho\beta\chi_t \equiv c_p\rho\beta\chi\Phi, \quad (57)$$

where χ_t represents a "turbulent conductivity." The "turbulent Prandtl number,"

$$\sigma_t = \nu_t/\chi_t, \quad (58)$$

is presented in Fig. 6 vs S for different values of Ro. The quantity χ_t/χ is presented in Fig. 3 and Table II.

(f) *The Growth Rate for the Solar Nebula*

In most computations of the structure of the solar nebula, it is assumed that self-

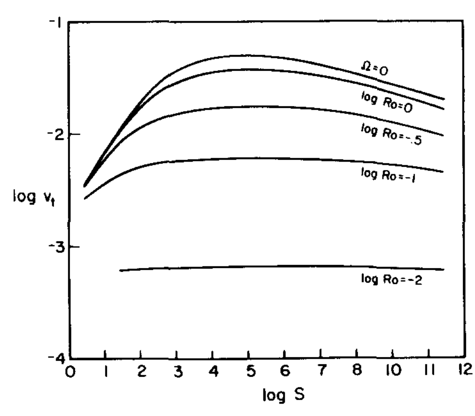

FIG. 4. The turbulent velocity v_t in units of $\sqrt{g\alpha\beta}d$, where d is the size of the convective layer as a function of S for the same rotation parameters as in Fig. 3.

TABLE I

LOG (ν_t/χ)

log Ro \ log λ_0	−4.5	−4.0	−3.5	−3.0	−2.5	−2.0	−1.5	−1.0	−0.5	0.0	0.5
0.5	2.0733	1.8191	1.5614	1.2976	1.0225	0.7259	0.3866	−0.0426	−0.6600	−1.5275	−2.5075
0.25	2.0339	1.7792	1.5209	1.2559	0.9788	0.6791	0.3365	−0.0911	−0.6892	−1.5338	−2.5082
0.0	1.9416	1.6859	1.4257	1.1576	0.8753	0.5681	0.2180	−0.2044	−0.7623	−1.5525	−2.5105
−0.25	1.7793	1.5211	1.2567	0.9814	0.6879	0.3662	0.0065	−0.4033	−0.9046	−1.6019	−2.5175
−0.5	1.5617	1.2987	1.0259	0.7373	0.4254	0.0860	−0.2778	−0.6684	−1.1163	−1.7071	−2.5379
−0.75	1.3136	1.0417	0.7547	0.4460	0.1132	−0.2365	−0.5952	−0.9674	−1.3760	−1.8797	−2.5909
−1.0	1.0469	0.7604	0.4527	0.1220	−0.2233	−0.5723	−0.9228	−1.2814	−1.6640	−2.1080	−2.7010
−1.25	0.7622	0.4548	0.1248	−0.2191	−0.5651	−0.9090	−1.2530	−1.6027	−1.9685	−2.3740	−2.8771
−1.5	0.4555	0.1257	−0.2177	−0.5628	−0.9047	−1.2442	−1.5841	−1.9280	−2.2830	−2.6639	−3.1072
−1.75	0.1260	−0.2173	−0.5621	−0.9033	−1.2415	−1.5783	−1.9156	−2.2559	−2.6037	−2.9687	
−2.0	−0.2172	−0.5619	−0.9029	−1.2406	−1.5764	−1.9117	−2.2476	−2.5856	−2.9286	−3.2832	
−2.25	−0.5618	−0.9028	−1.2403	−1.5759	−1.9105	−2.2449	−2.5799	−2.9163	−3.2563	−3.6039	
−2.5	−0.9027	−1.2402	−1.5757	−1.9101	−2.2441	−2.5781	−2.9125	−3.2479	−3.5857		
−2.75	−1.2402	−1.5756	−1.9100	−2.2438	−2.5775	−2.9113	−3.2453	−3.5801			
−3.0	−1.5756	−1.9099	−2.2438	−2.5774	−2.9109	−3.2445	−3.5783				

Note. For ease of numerical calculations, we present here the numerical values for the turbulent viscosity ν_t and turbulent conductivity χ_t (both in units of the thermometric conductivity χ). The defining Eqs. are (43) and (57). The independent variables are chosen to be the degree of rotation, via the Rossby number, Ro (Eq. (55)) and the degree of turbulence, via the Rayleigh number Ra: the quantity λ_0 is defined as $\lambda_0^2 = \pi^4/4S$, where $S = \sigma\mathrm{Ra}$ (Eq. (52)).

gravity effects are negligible (however, see Cameron, 1978). In the first part we adopt the same assumption and make use of Eqs. (48) and (50) for the growth rate $n(k)$. We shall then check a posteriori that the assumption is indeed satisfied. On the other hand, we shall also show that in the outer Solar System, self-gravity becomes important. A full discussion of the results will be presented in Paper II, Section VI.

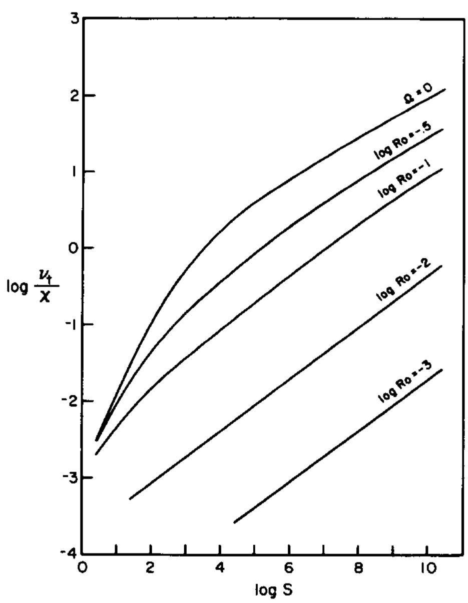

FIG. 5. The turbulent viscosity ν_t normalized to the radiative conductivity, χ, as a function of S for the same rotation parameters as in Fig. 3.

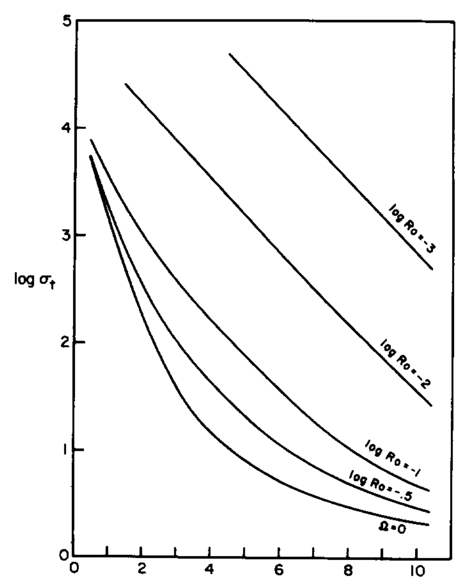

FIG. 6. The turbulent Prandtl number, $\sigma_t = \nu_t/\chi_t$, as a function of S for the same rotation parameters as in Fig. 3.

TABLE II

LOG (χ_t/χ)

log Ro \ log λ_0	−4.5	−4.0	−3.5	−3.0	−2.5	−2.0	−1.5	−1.0	−0.5	0.0	0.5
0.5	1.7521	1.4439	1.1165	0.7620	0.3672	−0.0910	−0.6572	−1.4243	−2.5827	−4.2764	−6.2287
0.25	1.7053	1.3946	1.0639	0.7050	0.3044	−0.1612	−0.7351	−1.5022	−2.6332	−4.2881	−6.2301
0.0	1.5946	1.2778	0.9389	0.5688	0.1532	−0.3315	−0.9239	−1.6900	−2.7615	−4.3230	−6.2344
−0.25	1.3956	1.0662	0.7103	0.3170	−0.1300	−0.6521	−1.2764	−2.0369	−3.0194	−4.4152	−6.2476
−0.5	1.1196	0.7692	0.3845	−0.0480	−0.5450	−1.1202	−1.7769	−2.5234	−3.4156	−4.6134	−6.2864
−0.75	0.7900	0.4082	−0.0194	−0.5079	−1.0669	−1.6907	−2.3652	−3.0929	−3.9140	−4.9421	−6.3869
−1.0	0.4160	−0.0101	−0.4958	−1.0496	−1.6630	−2.3152	−2.9941	−3.7047	−4.4758	−5.3826	−6.5964
−1.25	−0.0071	−0.4919	−1.0441	−1.6541	−2.2995	−2.9633	−3.6409	−4.3383	−5.0754	−5.9015	−6.9341
−1.5	−0.4907	−1.0424	−1.6513	−2.2945	−2.9535	−3.6209	−4.2960	−4.9840	−5.6983	−6.4713	−7.3797
−1.75	−1.0418	−1.6505	−2.2929	−2.9505	−3.6147	−4.2828	−4.9557	−5.6370	−6.3360	−7.0741	
−2.0	−1.6502	−2.2924	−2.9495	−3.6127	−4.2787	−4.9468	−5.6178	−6.2946	−6.9833	−7.6983	
−2.25	−2.2922	−2.9492	−3.6120	−4.2773	−4.9440	−5.6118	−6.2816	−6.9553	−7.6371	−8.3368	
−2.5	−2.9491	−3.6119	−4.2769	−4.9431	−5.6099	−6.2775	−6.9465	−7.6181	−8.2954		
−2.75	−3.6118	−4.2768	−4.9428	−5.6093	−6.2763	−6.9437	−7.6121	−8.2824			
−3.0	−4.2767	−4.9427	−5.6091	−6.2758	−6.9428	−7.6102	−8.2784				

Note. See Table I Note.

(*g*) *The Eddington Factor*

Perturbations with dimensions much less than the mean free path of the photons (i.e., which are optically thin) diffuse radiatively at rates independent of size rather than at the "optically thick" rate χk^2. Spiegel (1964) showed that the diffusion rate for both optically thick and thin perturbations can be represented by a renormalized "optically thick" case as

$$\chi k^2 f_E(k), \tag{59}$$

where the "Eddington factor" $f_E(k)$ is given by

$$f_E(k) = 3\left(\frac{\kappa\rho}{k}\right)^2\left[1 - \frac{\kappa\rho}{k}\tan^{-1}\frac{k}{\kappa\rho}\right]. \tag{60}$$

This expression is derived from the gray radiative transfer equations in the Eddington approximation. We have incorporated the Eddington factor by renormalizing χ with the function $f_E(k)$ computed at the largest scale k_0. While this procedure is valid for the computation of ν_t (see Eq. (43)), it is not entirely valid for the convective flux F_c which, as shown in Eq. (47), depends on the integration over a spectrum of wavelengths. However, the spectral function $E(k)$ is typically strongly peaked toward midplane of the largest scale (see Fig. 2) such that F_c depends mostly on the values of k near k_0. Thus, we believe that the renormalization of χ via $f_E(k_0)$ is an adequate approximation for computing F_c.

The effect of the Eddington factor is to limit the rate of radiative losses by convective fluid elements and increase the growth rates over the ones computed without it. It is found to be important in models in which a predominant part of the midplane convection zone has optically thin perturbations.

(*h*) *Assessment of the Turbulence Model*

The above model of turbulence was tested against laboratory and astrophysical data (CG) and a brief summary of the results is presented here.

(1) In the case of *laboratory* turbulent convection, the quantity of direct experimental interest is the measured heat flux, i.e., the Nusselt number, $N = F_c/F_{\mathrm{rad}} + 1$. Recent experimental data for convection in water, at a Rayleigh number up to 10^{11}, and for $\Omega = 0$, have confirmed the long-suspected relation

$$N = N_0\,\mathrm{Ra}^{1/3}. \tag{61}$$

The CG model reproduces the $\mathrm{Ra}^{1/3}$ dependence and predicts the coefficient N_0 to an accuracy of 20%; i.e., the model reproduces 80% of the measured energy flux.

(2) In the context of *astrophysical* turbulent convection, i.e., when one can take zero Prandtl number, the expressions for F_c and v_t derived from the model (Eq. (53)) *coincide with the ones of the mixing length theory* (see Cox and Giuli, 1968, Eqs. (14.108) and (14.110)). Moreover, the parameters a and b in Eq. (53) (which MLT does not determine and which had to be calibrated using stellar models; Gough and Weiss, 1976) are reproduced satisfactorily by the model. Moreover, the dependence of the convective flux F_c on the anisotropy parameter x, a problem that within the MLT context has only recently been solved by Gough (1978), is correctly reproduced by the CG model (see Sect. VII).

IV. THE PARAMETER α AND TURBULENT VELOCITIES

One of the advantages of the CG model of turbulence is that it no longer relies on empirical relations which introduce free parameters. In particular, since the CG model provides a well-defined expression for the turbulent viscosity, empirical relations like

$$\nu_t = \alpha_1 c_s h = \alpha_2 v_c h, \tag{62}$$

where v_c is the turbulent convective velocity, which have served a useful purpose but which have exhausted their fruitfulness, will hopefully be abandoned from here on.[1]

Due to the different notation used by different authors, we have found it useful to recast the "α-model" expression for ν_t in two forms. For example, Cameron (1977) has

$$\alpha_1 = 2/9 = (1/3)\alpha_2, \tag{63}$$

while LPB introduced a parameter α related to α_1 and α_2 by

$$\alpha_1 \equiv \alpha(v_c/c_s)^2, \quad \alpha_2 \equiv \alpha(v_c/c_s); \tag{64}$$

α is then taken to be of order unity (see, however, the discussion in Sect. VII).

Let us now discuss the ratio v_t/c_s. For that, let us recall that the z component of the Navier–Stokes equation for the mean flow is

$$U_{z,t} + U_r U_{z,r} + U_z U_{z,z} + \frac{1}{\rho}(p_{,z} - \tau_{zz,z}) = -\psi_{,z} + \frac{1}{\rho}\left(\tau_{zr,r} + \frac{1}{r}\tau_{zr}\right), \tag{65}$$

where

$$\tau_{ij} = -\rho\overline{u_i u_j} \tag{66}$$

is the generalized Reynolds stress generated by the turbulent velocities, $\mathbf{U}$ is the mean flow velocity, and ψ is the gravitational potential. In the thin-disk approximation, commonly used in disk calculations, Eq. (65) reduces to

$$\frac{d}{dz}(p + p_t) = -g\rho, \tag{67}$$

where p_t is the pressure contributed by the turbulent motion. In all disk calculations, it is further assumed that (67) simplifies to

$$\frac{dp}{dz} = -g\rho, \tag{68}$$

which obtains only if $p_t \ll p$ or equivalently if

$$v_t^2 < c_s^2 \equiv p/\rho. \tag{69}$$

The adoption of (68) is therefore inconsistent with a disk calculation where v_t approaches c_s (see Eq. (141)).

In spite of this consistency argument, the relation

$$v_t = \tfrac{1}{3}c_s \tag{70}$$

has been widely used in the literature. Is expression (70) correct for disks?

The general procedure for calculating turbulent velocities has been presented earlier: One must solve (50) to obtain $n(k)$; the result is then substituted in (41) and then in

[1] In Tables I and II, we present the values of the convective flux and of ν_t/χ in a way that is suitable for numerical calculations.

(40) to obtain v_t. Since in general Eq. (50) is a cubic equation, numerical integrations have to be carried out. The resulting velocities are presented in Fig. 4.

To gain physical insight, we shall carry out a *single-mode* calculation which, while preserving the main functional dependence on all quantities of interest, lends itself to an analytical treatment. (We recall that a single-mode computation is the basic assumption of the Mixing Length Theory. However, one *cannot* use the MLT expressions employed in stellar physics since they do not contain the anisotropy factor x, which plays a basic role in determining the value of the turbulent velocities.)

First, take the $k \to k_0$ limit of Eq. (41). The result is

$$-2\gamma_* E(k_0) = n(k_0)(n(k)/k^2)'_{k_0}. \quad (71)$$

Taking the same limit in Eq. (42), we obtain $L_p = 3\pi/4k_0$ which, together with the definition $4\gamma_*/3\pi \approx 1$, gives

$$E(k_0) = \frac{n^2(k_0)}{\gamma_*^2 k_0^3}. \quad (72)$$

From (40) we can now define the one-mode velocity as $v_t^2 = k_0 E(k_0)$, so that, finally,

$$v_t = \frac{1}{\gamma_*}\frac{n(k_0, \Omega)}{k_0}, \quad (73)$$

where we have explicitly written the possible dependence of n on Ω to stress that (73) is valid for a general $n(k)$. Solving (50), one obtains $n(k_0, \Omega)$ and, finally, v_t. For the purpose of establishing an upper limit to v_t, we recall that since rotation has a stabilizing effect, i.e., for any k (see Fig. 1),

$$n(k, \Omega) < n(k, 0),$$

thus the use of $n(k, 0)$ in Eq. (73) will yield an upper limit to v_t. Using (51), we have, for $\sigma \to 0$,

$$n(k_0, 0) \equiv n(k_0) = \left(g\alpha\beta \frac{x}{1+x}\right)^{1/2} f(S), \quad (74)$$

where

$$f(S) \equiv (bS)^{-1/2}(\sqrt{1+bS} - 1).$$

Since $f(S) \le 1$, taking $f(S) = 1$ maximizes v_t. Using

$$k_0 d = \pi(1+x)^{1/2}, \quad \mathrm{Ro} = \frac{\sqrt{g\alpha\beta}}{\Omega}, \quad (75)$$

where Ro is the Rossby number, we have from (73) and (75)

$$v_t < \frac{d\Omega}{\gamma_*\pi}\frac{x^{1/2}}{(1+x)}\mathrm{Ro}.$$

Dividing by c_s, we finally obtain, with $\Omega d \approx c_s$,

$$\frac{v_t}{c_s} \ll \frac{\mathrm{Ro}}{\gamma_*\pi}\frac{x^{1/2}}{(1+x)}. \quad (76)$$

Even if

$$x \approx 1, \quad (77)$$

Eq. (70) would still overestimate v_t. (The Rossby number is less than unity; see Fig. 3g in Paper II.) However, since a consistent treatment of the anisotropy factor x indicates that in disks

$$x \gg 1, \quad (78)$$

it follows from (76) that

$$\frac{v_t}{c_s} < \frac{\mathrm{Ro}}{\gamma_*\pi}\frac{1}{x^{1/2}} \ll 1. \quad (79)$$

In conclusion, while the first consistency argument shows that assuming Eq. (68) is equivalent to assuming $v_t < c_s$, this does not imply that turbulent velocities ought to be in general smaller than c_s: in principle v_t can have any relation with c_s. In particular, there may be physical situations when Eq. (70) holds true. For the specific case of thin disks we have, however, shown that the *large value* of the eddy anisotropy causes v_t to be indeed much smaller than c_s. (The same conclusion holds true for the z component of v_t, called v_c in LPB, where $v_t^2 = v_c^2(1+x)x^{-1} \approx v_c^2$, since $x \gg 1$.)

Having presented the mathematical origin of the anisotropy factor x in the ratio $v_t/$

c_s, let us discuss the validity of Eq. (70). As explicitly stated by many authors (Cameron, 1977; Lin, 1981), the assumption underlying the use of Eq. (70) is the kinetic theory of gases, in analogy with molecular viscosity. Unfortunately, the molecular analogy fails when applied to large eddies since molecules are by definition isotropic ($x \approx 1$), while large eddies are by definition highly anisotropic ($x \gg 1$). To illustrate the limitations of the molecular analogy, consider a turbulent medium composed of eddies of widely different sizes.

The *smallest eddies* can be thought of as large molecules and therefore the above analogy is a useful guide. Next, consider the medium to small-size eddies, the Heisenberg–Kolmogoroff (HK) eddies, that populate the inertial part of the energy spectrum. By definition, the HK eddies have undergone sufficient interactions to have lost all of the stirring mechanism. The same interactions have homogenized and isotropized the system. Anisotropy effects are still unimportant. (This is borne out by the comparison of the isotropic HK model with the data.) Last, consider the *large*, energy-containing eddies, whose dynamics are largely dictated by the stirring mechanism. It is a general result of the theory of turbulence that unless large-scale turbulence is initially isotropic, it will remain anisotropic. For the large eddies, anisotropy is important and molecular analogies fail.[2]

Almost all theories that try to model turbulence from first principles have used some kind of molecular viscosity analogy (first suggested by Heisenberg) to describe the nonlinear interactions. These are thought to act as an enhanced (turbulent) viscosity generated by the eddies in the interval $k_2 - \infty$, and acting on the eddies in the interval $k_1 - k_2$. Whether this enhanced viscosity is affected by x depends on which part of the k spectrum one is considering. For example, if one is interested in medium- to small-size eddies, the interval $k_1 - k_2$ (the HK inertial regime) is sufficiently removed from the region of the largest eddies to be immune from x. Indeed, Heisenberg and Kolmogoroff constructed an expression for the turbulent viscosity that depends only on local properties, i.e., which is isotropic. The resulting HK spectrum has been amply confirmed by experiments.

However, the HK expression for ν_t cannot be extrapolated to low k's where the largest eddies reside (Canuto *et al.*, 1985). A new model is needed. For example, the CG model for large-scale eddies, while retaining Heisenberg's physical picture about the nonlinear interactions acting as an enhanced viscosity, abandons the HK *local* description and explicitly introduces anisotropy effects, $v_t = v_t(x)$ and $\nu_t = \nu_t(x)$.

In conclusion, models for ν_t that bodily extrapolate the molecular viscosity analogy all the way to the largest eddies completely miss what is perhaps the main feature of the large eddies, their anisotropy. These models overestimate ν_t and v_t by large factors.

V. MIDPLANE ($z = 0$) BEHAVIOR OF CONVECTIVE FLUXES AND TURBULENT VISCOSITIES: PHYSICAL, UNPHYSICAL, AND DEGENERATE SOLUTIONS

In this section, we discuss three models that provide expressions for F_c and ν_t. They are

$$\begin{aligned} \text{LPB: } \Omega = 0, \quad \chi = 0, \quad \nu_t \sim v_t^2, \\ \text{MLT: } \Omega = 0, \quad \chi \neq 0, \quad \nu_t \sim v_t^2, \\ \text{CG: } \Omega \neq 0, \quad \chi \neq 0, \quad \nu_t \sim n(k_0) \end{aligned} \tag{80}$$

(A rederivation of the LPB and MLT models is presented in Sect. VII). Introducing the standard notation

[2] The difference between small and large eddies can also be seen in the following way. Small eddies have characteristic lifetimes much shorter than the decay time of turbulence: consequently, they have ample time to reach some form of statistical equilibrium. On the other hand, large eddies have lifetimes of the same order as the decay time of turbulence itself, and therefore cannot reach equilibrium.

$$\nabla \equiv H_p T^{-1} \frac{dT}{dz}, \quad H_p = \frac{p}{\rho g}, \qquad (81)$$

where H_p is the pressure scale height, the z dependence of F_c and ν_t in the three cases above as $z \to 0$ is

$$\text{LPB: } F_c \sim z^2(\nabla - \nabla_{ad})^{3/2}, \quad \nu_t \sim (\nabla - \nabla_{ad}), \qquad (82)$$

$$\text{MLT: } F_c \sim z^{5+8\delta}(\nabla - \nabla_{ad})^3, \quad \nu_t \sim z^{4+6\delta}(\nabla - \nabla_{ad})^2, \qquad (83)$$

$$\text{CG: } F_c \sim z^{5+8\delta/3}(\nabla - \nabla_{ad})^3, \quad \nu_t \sim z^{2+4\delta/3}(\nabla - \nabla_{ad}), \qquad (84)$$

where in the MLT and in the CG models the "scale length" has been taken to have a possible z dependence of the form z^δ.

We can therefore conclude that in general Eq. (17) can be rewritten near the midplane as

$$\frac{dF}{dz} \sim \nu_t \sim z^b(\nabla - \nabla_{ad})^a. \qquad (85)$$

Moreover, since from Eq. (8) (in Paper II), $F_r \sim z\nabla$, we see that F_c goes to zero faster than F_r provided ∇ does not blow up as $z \to 0$; thus $F = F_r + F_c \simeq F_r$, and we can rewrite Eq. (85) as

$$\frac{dF}{dz} = \varepsilon_0 z^b \left(\frac{F}{f_0 z} - \nabla_{ad}\right)^a, \qquad (86)$$

where ε_0, f_0, and ∇_{ad} are constant near midplane (they are in fact thermodynamic quantities). The integration of the full-disk equations (see Sect. II in Paper II) is carried out from $z = H$ (H = photospheric height of the disk) downward toward midplane $z = 0$; *a physically consistent solution is defined* as an integration resulting in

$$F(0) = 0 \text{ (by symmetry), and } \nu_t(0) \text{ is well behaved.} \qquad (87)$$

Clearly, whether these requirements are satisfied depends on the specific values of the parameters a and b entering Eq. (86) and in some cases on the thermodynamic constants as well. We have made a study of Eq. (86) and the analysis is presented in Appendix C and Table III. It is convenient to classify the solutions into three categories: unphysical, U; physical but degenerate, D; and physical and nondegenerate, ND. They are defined as follows:

Unphysical solutions, U: either of conditions (87) is not satisfied.

Physical but degenerate, D: conditions (87) are satisfied for *arbitrary values* of the disk height, H. A unique solution of the disk is therefore impossible since one cannot define surface density, total mass of the nebula, etc.

Physical and nondegenerate, ND: conditions (87) are satisfied for a *unique value* of the disk height, H. A complete structure of the disk can be obtained.

TABLE III

SUMMARY OF THE RESULTS OF APPENDIX C

a/b	<0	0	>0
0	U	U *or* ND	U
$0 < a < 1$	U *or* D	U, D, *or* ND	U
1	D	D	U
>1	D	D	U

Note. The LPB model, using Eqs. (1)–(4) of Lin (1981) (see Eq. (90) above), is characterized by $a = 1$ and $b = 0$, corresponding to a degenerate situation in which $F(0) = 0$ for any value of height H. We have verified numerically that this is indeed the case. The method by which LPB ultimately chose H is discussed in Section VII.

VI. GLOBAL VS LOCAL VALUES OF TURBULENT VISCOSITY

The MLT and CG models of convective turbulence are both formulated for global convective properties averaged over a "mixing length." The use of local distributions of convective properties is not strictly valid within these models. Nonlinear transport terms have been globally averaged away in the turbulent energy equation, while, in fact, vertical redistribution of energy and momentum probably occurs to some degree within the vertical structure due to large-scale turbulent motions. We unfortunately cannot determine the extent to which the distribution of, say, turbulent heat dissipation is smoothed out in a realistic situation without detailed hydrodynamical studies. Though the structure equations are not valid for determining local vertical structure, some knowledge of the mean thermal structure is required to specify self-consistently the strength of the convective motions and amount of turbulent heat deposition in the convective region. We therefore make two important assumptions:

(1) The form of the globally averaged energy equation (17) specifies (in some mean sense) the local thermal structure due to turbulent heat dissipation, with the provision that

(2) the turbulent heat dissipation (measured by the turbulent viscosity ν_t) is redistributed evenly throughout the convective region by large-scale motions (i.e., ν_t is constant). The convective flux, however, is computed and incorporated locally in the vertical structure.

The thermodynamic structure that results from this is used to compute local values of ν_t from local convective growth rates (Eq. (50)). A numerical solution is deemed consistent when constant ν_t equals the density average of locally computed values of ν_t over the convective region. Note that our method seeks in a sense to preserve the net amount of heat deposition expected from the entire convective region while merely redistributing it in a somewhat ad hoc manner.

The method that we employ uses locally computed convective growth rates (in the Boussinesq approximation) to determine the mean turbulent viscosity by means of vertical averages. A computationally more time-consuming method, but a perhaps more consistent one, would be to use the local structure from the computation with a trial constant ν_t to estimate the growth rate of the lowest order (longest wavelength) mode in a boundary value problem of the linearized perturbation equations. This growth rate could then be used to calculate a global value of ν_t, which would then be compared to the trial ν_t in a consistency check. Such a method would automatically take into account the effects of compressibility, being an integration of eigenvalues over a few pressure scale heights. The more expedient method actually employed in this paper hopefully approximates the global growth rate by use of the vertically averaged local growth rate.

VII. ANALYSIS OF THE LIN–PAPALOIZOU–BODENHEIMER (LPB) MODEL

In this section, we discuss three topics: (1) the validity of the LPB expressions for turbulent fluxes and velocities, (2) the LPB treatment of the turbulent viscosity, and (3) the nature of the LPB disk solutions.

(*a*) *The MLT Expressions for Turbulent Fluxes and Velocities*

The two basic ingredients needed to quantify turbulent convection are the convective flux F_c and the turbulent viscosity ν_t,

$$F_c \sim \langle w\theta \rangle, \quad \nu_t \sim \langle w^2 \rangle, \tag{88}$$

where w is the z component of the turbulent velocity (called v_c in LPB) and θ the fluctuating temperature. The MLT formalism does provide expressions for F_c and w (see Cox and Giuli, 1968, Eqs. (14.108) and (14.110)), which, however, *do not* contain the anisotropy factor x, which, as we shall see, plays a fundamental role in our analysis. Furthermore, since MLT expressions for arbitrary x and S ($= g\alpha\beta d^4\chi^{-2}$ the "convective efficiency" Eq. (52)) are not available in the literature, we shall derive them first and then compare them with the expression used in LPB. We must note that Gough's (1976, 1977) expressions have the correct S dependence but an incorrect x dependence; Gough (1978) has the correct x dependence but the expressions are only valid in the $S \gg 1$ regime, which is manifestly invalid at the disk's midplane where $g \to 0$.

We shall employ the MLT formalism (Spiegel and Veronis, 1960; Spiegel, 1963, 1966) and in particular the kinetic theory of accelerating fluid elements. For a detailed review, see Gough (1978).

Consider the Navier–Stokes equations for the velocity field. Eliminating the pressure terms, making use of the compressibility relation, and, finally, taking the z component of the resulting equation, we obtain for w the equation

$$\frac{\partial w}{\partial t} = \frac{x}{1 + x} g\alpha\theta. \tag{89}$$

The fluctuation temperature θ is defined in terms of the "total temperature" $T'(xyz)$ and the "average temperature" $T(z)$ as (Spiegel, 1963)

$$T'(xyz) = T(z) + \theta(xyz). \tag{90}$$

(We have changed Spiegel's notation to conform to the LPB definition of T as the average temperature.) Taking $\partial/\partial t = w\partial/\partial\Lambda_\xi$, where Λ_ξ is the eddy displacement, integrating (89), and assuming that the fluctuations at the point (xyz) are caused by the arrival of a convective element from the point (xyz_0), one can show that (Spiegel, 1963)

$$\theta = \Lambda_z \left(\frac{dT'}{dz} - \frac{dT}{dz}\right), \quad \Lambda_z \equiv z - z_0. \tag{91}$$

We then obtain from (89),

$$w = \left(\frac{2x}{1 + x} \alpha g \Lambda_\xi \Lambda_z\right)^{1/2} \left(\frac{dT'}{dz} - \frac{dT}{dz}\right)^{1/2} \tag{92}$$

and from (88), in units of $c_p\rho$,

$$F_c = w\Lambda_z \left(\frac{dT'}{dz} - \frac{dT}{dz}\right) \tag{93}$$

(cf. Cox and Giuli, 1968, Eqs. (14.16) and (14.2)).

Next, we must eliminate dT'/dz. This is done using the MLT equation for θ (Spiegel, 1966),

$$\frac{\partial\theta}{\partial t} - \chi\nabla^2\theta - \beta w = 0, \tag{94}$$

where β is defined in Eq. (49). Proceeding as above and further taking $-\nabla^2 = k^2 \equiv \Lambda^{-2}$, where, using (44) and $k_z\Lambda_z = \pi$,

$$\pi\Lambda = \Lambda_z(1 + x)^{-1/2}, \tag{95}$$

we obtain from (94) and (91) the desired result,

$$\frac{dT'}{dz} - \frac{dT}{dz} = \left[\left(\frac{dT}{dz}\right)_{\rm ad} - \frac{dT}{dz}\right] \frac{w\Lambda^2\Lambda_z^{-1}}{\chi + w\Lambda^2\Lambda_\xi^{-1}}. \tag{96}$$

Since LPB introduced two length scales Λ_1 and Λ_2 which were taken to behave very differently, we have purposely left the three length scales Λ_z, Λ_ξ, and Λ different so as to check whether the LPB Λ_1 and Λ_2 are compatible with the basic formulation of the MLT. Substituting (92) into (96) and solving for dT'/dz yields

$$\frac{dT'}{dz} - \frac{dT}{dz} = \frac{\Lambda_\xi}{\Lambda_z}\left[\left(\frac{dT}{dz}\right)_{\rm ad} - \frac{dT}{dz}\right]\Sigma^{-1} \times (\sqrt{1 + \Sigma} - 1)^2 \tag{97}$$

with

$$\Sigma = 8x(1 + x)^{-1}\alpha\beta g\chi^{-2}\Lambda^4 \cong bS. \tag{98}$$

Substituting (97) into (92) and (93), we obtain

$$F_c = \frac{1}{2}\beta\chi\left(\frac{\Lambda_\xi}{\Lambda}\right)^2\Sigma^{-1}(\sqrt{1 + \Sigma} - 1)^3 c_p\rho,$$
$$w = \frac{1}{2}\chi\frac{\Lambda_\xi}{\Lambda^2}(\sqrt{1 + \Sigma} - 1). \tag{99}$$

The above procedure does not determine the correct x dependence, a problem that is by no means trivial within MLT. For example, Gough's (1976, 1977) expressions still contained an incorrect x dependence, that only later Gough (1978) was able to correct. The Canuto–Goldman (CG) model for large-scale turbulence (Canuto and Goldman, 1985) provides the correct x dependence, as one can see by comparing with

Gough's (1978) expressions. Since this is not the place to discuss this topic, it suffices to say that Eqs. (99) can be made to have the correct x dependence provided one takes

$$\Lambda_\xi = \Lambda_z x^{1/2}(1 + x)^{-1}. \tag{100}$$

Using (95) and (100), we finally obtain

$$F_c = \frac{\pi^2}{2}\frac{x}{1+x}\Sigma^{-1}(\sqrt{1+\Sigma} - 1)^3 c_p \rho \beta \chi, \tag{101}$$

$$w = \frac{\pi^2}{2}\chi^{1/2}(\sqrt{1+\Sigma} - 1)\chi\Lambda_z^{-1}, \tag{102}$$

together with

$$v_t = w\sqrt{(1+x)/x}. \tag{103}$$

Equations (101)–(103) yield the general MLT expressions for F_c and v_t for *arbitrary* Σ and x; to the best of our knowledge, this is the first such derivation of F_c and v_t for arbitrary x and Σ. Gough's (1976; 1977) work is valid for arbitrary Σ but has an incorrect x dependence. Gough's (1978) paper has the correct x dependence but is valid only for $\Sigma \gg 1$. Cox and Giuli's (1968) Eqs. (14.108) and (14.111) are a particular case of (101)–(103) for $x \approx 1$. The CG model reproduces exactly the above MLT expressions (Eq. (53)).

(b) The LPB Expressions for F_c and v_t

LPB expressions for F_c and w (called v_c) do not contain χ; i.e., it is assumed that throughout the entire nebula the eddies do not lose energy via radiative processes.[3] To obtain χ independent expressions, one must take the limit

$$\Sigma \gg 1, \tag{104}$$

in which case Eqs. (101) and (102) become (neglecting factors of order unity)

$$F_c = \frac{x^{3/2}}{(1+x)^{5/2}}(\alpha g)^{1/2}\beta^{3/2}\Lambda_z^2 c_p \rho, \tag{105}$$

$$v_c = \frac{x}{(1+x)^{3/2}}(\alpha g \beta)^{1/2}\Lambda_z. \tag{106}$$

Let us now compare (105) and (106) with the LPB expressions (Lin, 1981, Eqs. (2) and (3)),

$$F_c = \alpha_1(\alpha g)^{1/2}\beta^{3/2}\Lambda_1\Lambda_2 c_p \rho, \tag{107}$$

$$v_c = (\alpha g \beta)^{1/2}\Lambda_2. \tag{108}$$

This implies that

$$\Lambda_1 = \frac{x^{1/2}}{(1+x)}\Lambda_z, \quad \Lambda_2 = \frac{x}{(1+x)^{3/2}}\Lambda_z. \tag{109}$$

The adoption of (107) and (108) forced LPB to assume two independent forms for Λ_1 and Λ_2 (if $\Lambda_1 = \Lambda_2$, there is no solution for the disk equations), namely

$$z > H_p, \quad \Lambda_1 = \Lambda_2 = H_p \equiv p/\rho g, \tag{110a}$$

$$z < H_p, \quad \Lambda_1\Lambda_2 = \text{constant}, \quad \text{i.e., } \Lambda_1 \sim z,\ \Lambda_2 = H_p \sim z^{-1}. \tag{110b}$$

Inspection shows that LPB's choices (110a) and (110b) cannot be made compatible with the general MLT expressions (109).

The $\chi \to 0$ limit and the midplane behavior. The "high-efficiency limit" (104) is often used in the application of MLT to problems of stellar structure, where it is valid because the ingredients of Σ do not vanish within the convective regions of stars. This is, however, no longer true in the solar nebula where at midplane

$$g = z\Omega^2 \to 0, \quad \text{i.e., } \Sigma \to 0, \tag{111}$$

in contradiction to (104). The proper procedure consists of taking the limit (111) first in the general expressions (101)–(103). This results in

$$F_c \sim g^2\chi^{-3}, \quad v_t \sim g\chi^{-1}, \tag{112}$$

which shows that the $\chi \to 0$ limit can no

[3] It may be recalled that Öpik's (1950) suggestion that radiative losses may significantly reduce the convective efficiency was the major motivation for the work on MLT in the late sixties and seventies that finally led to the general expressions (101) and (102). Without χ, the MLT expressions used by LPB reduce to the ones by Prandtl (1925) that were criticized by Öpik as inadequate for astrophysical purposes.

longer be taken. We conclude that Eqs. (107) and (108) are not compatible with the correct MLT expessions at midplane where the major contribution to convection arises in LPB's model.

Rotational effects: The $x \to \infty$ limit. The LPB expressions for F_c and v_t do not contain rotation, whose effect is to lower both F_c and v_t. To see under which conditions this approximation holds, we use a general result of the MLT formalism (Spiegel, 1963; Gough, 1978), namely that[4]

$$F_c \sim n^3(k_0), \quad w \sim n(k_0), \tag{113}$$

where $n(k_0)$ is the growth rate computed for the largest eddy, $k = k_0$. Since $n(k)$ depends on the rotation Ω (Eq. (50)), the only way to have F_c and w independent of Ω is by taking

$$x \to \infty, \tag{114}$$

in which case Eq. (50) shows that rotational effects become unimportant. *Since the LPB expressions for F_c and w do not contain Ω, we must conclude that the limit (114) is implicitly built into their formalism.* As a consequence of (114), it follows from (109) that

$$\Lambda_1 = \Lambda_2. \tag{115}$$

This shows again that assumptions (110) are not internally consistent with (114). Furthermore, even if (115) was adopted, Eqs. (107) and (108) would not be compatible with MLT Eqs. (105) and (106) unless the coefficient α_1 is chosen to be

$$\alpha_1 \sim \frac{x^{3/2}}{(1+x)^{5/2}} \sim \frac{1}{x} \ll 1 \tag{116a}$$

instead of

$$\alpha_1 \sim 1 \tag{116b}$$

[4] Equations (113) can easily be derived using the CG model for turbulence. Since the MLT is by definition a one-mode theory, one can take the limit $k \to k_0$ in Eq. (41) and derive the function $E(k)$ for a general $n(k)$. When this is substituted in (40) and (47), one obtains

$$F_c \sim \frac{n^3(k_0)}{1+x}\frac{d^2}{g\alpha}, \quad w \sim \frac{x^{1/2}}{1+x} n(k_0) d,$$

i.e., Eqs. (113) above. We have used Eqs. (44) and (103) above.

as in LPB. This implies that the LPB strength of the convective flux F_c has been overestimated by a factor $x \gg 1$.

(c) *The LPB Treatment of ν_t: The Effect of Anisotropy*

LPB adopted the expression

$$\nu_t \sim \frac{v_c^2}{\Omega}. \tag{117a}$$

It may appear that the presence of Ω in Eq. (117a) represents, or at least partially accounts for, the effects of rotation on the turbulent eddy viscosity ν_t. This is, however, not the case as we shall now show.

If one adopts the line of reasoning outlined in the Introduction, the proper and only way to interpret (117a) is to recognize that it *defines*

$$\tau_{r\phi} \sim \nu_t S_{r\phi} \sim \nu_t \Omega; \tag{117b}$$

i.e., (117a) is not an equation for ν_t, which is an artificial auxiliary quantity. The stratagem of not solving the full equations for the Reynolds stress tensor τ_{ij} but of guessing the solution written in the form (117b) has meaning only if one can provide independent ways to compute both ν_t and $S_{r\phi}$. The Ω appearing in (117b) comes from $S_{r\phi}$, *but nothing yet has been said about the function ν_t, which remains an unknown.*

No further progress can be made unless one adopts a model for turbulence. One cannot hope to make real progress unless one says something concrete about how the non-linear interactions should be treated. As discussed in Section III, a model for large scale turbulence has recently been proposed. If we restrict ourselves to a one-mode analysis, which for the present purpose is sufficient, the expression for ν_t is given by (CGH, 1984)

$$\nu_t = \frac{n(k_0)}{k_0^2} = \frac{n v_t^2}{k_0^3 E(k_0)}, \tag{118}$$

where all quantities are evaluated at $k = k_0$. This same model also provides a way to compute the spectral function $E(k)$. It can be seen from Eq. (72) that

$$E(k_0) = n^2(k_0)k_0^{-3} \quad (119)$$

so that, finally,

$$\nu_t = \frac{v_t^2}{n}, \quad (120)$$

where

$$v_t \equiv v_t(k_0, \Omega), \quad n \equiv n(k_0, \Omega); \quad (121)$$

i.e., both v_t and n depend on rotation. Since n is obtained by solving Eq. (50), we shall write

$$n(k_0, \Omega) = \sqrt{g\alpha\beta}\, f(k_0, \Omega), \quad (122)$$

where *all* the rotational effects are included in the function f. The factor in front is the natural frequency of convection. When rotational effects are absent, the function $f(k_0)$ is given by Eq. (51).

Suppose now that we *purposely* neglect all rotational effects on the previous formulae. We obtain

$$\nu_t = \frac{v_t^2(k_0)}{\sqrt{g\alpha\beta}} \frac{1}{f(k_0)} \sim \frac{v_t^2(k_0)}{\Omega} \quad (123)$$

since

$$g = z\Omega^2. \quad (124)$$

Equation (123) is exactly Eq. (117a) above. We therefore see that the Ω in (117a) is *not* due to rotational effects on the turbulent eddies, i.e., the Ω in $f(k_0, \Omega)$, but rather to the Ω dependence of the local gravity g.

Having clarified the meaning of (117a), let us study the effects of anisotropy. To be consistent with MLT, one must use the full MLT expression for v_c, Eq. (102). Because we need to compare the final result with the LPB expression that does not contain χ, we use the limit (106). The result is

$$\nu_t^{\text{MLT}} \sim \frac{1}{\Omega} \frac{x^2}{(1+x)^3} \alpha\beta g \Lambda_z^2, \quad (125a)$$

to be compared with the LPB expression

$$\nu_t^{\text{LPB}} = \frac{\alpha_2}{\Omega} \alpha\beta g \Lambda_2^2. \quad (125b)$$

Expression (125b) is consistent with MLT only if α_2 is taken to be

$$\alpha_2 \sim \frac{x^2}{(1+x)^3} \sim \frac{1}{x} \ll 1, \quad (126a)$$

while LPB chose

$$\alpha_2 \sim 1. \quad (126b)$$

We conclude that the LPB treatment overestimates ν_t by a large factor ($x \gg 1$), and it overestimates the amount of heat generated at the nebula's midplane since

$$\frac{dF}{dz} \sim \tau_{r\phi} S_{r\phi} \sim \nu_t \Omega^2 \sim \frac{1}{x} \nu_t^{\text{LPB}} \Omega^2.$$

From a qualitative point of view, Eq. (126b) may lead one to conclude that "convection" alone is sufficient to generate all the viscosity that is required for the dynamical evolution of nebula, while in reality that may not be the case. From a quantitative point of view, it is known that the larger the turbulent viscosity, the thinner and less massive the resulting nebula for a given $\dot{M}$ or T_e. To get a feeling of how the LPB results ought to be correctd to account for *at least* the anisotropy factor x, one may use Eqs. (16) of Lin (1981)[5] with the results

$$(H, T_c, \rho_c, \text{Re}) \sim (x^{1/6}, x^{1/3}, x^{1/2}, x^{2/3}) \times (\text{LPB results}) \quad (127)$$

for the nebula's height, central temperature, central density, and Reynolds number, respectively. The renormalization shows that $x \geq 1$, giving a smaller viscosity, leads to a thicker and hotter nebula. For the mass of the nebula, M_N, we can use a result of Lin and Papaloizou (1980, Eq. (31)) whereby

$$M_N \sim v_c^{-4} \text{Re}^{-3/2}. \quad (128)$$

Since, from Eqs. (106) and (108),

$$v_c \sim x^{-1/2} v_c\ (\text{LPB}), \quad (129)$$

we conclude that

$$M_N \sim x M_N\ (\text{LPB}), \quad (130)$$

thus resulting in a more massive nebula.

[5] In that paper, the two strength parameters α_1 and α_2 introduced above are called $\alpha_1 = \alpha_2 \equiv \alpha$.

(d) The LPB Treatment of the Scale Heights

Next, we shall analyze the LPB choice of scale heights. For that, we first introduce the definitions (81) and

$$c_s^2 = \Gamma_1 \frac{p}{\rho}, \quad Q = -\left(\frac{\partial \ln \rho}{\partial \ln T}\right)_p \equiv \alpha T. \tag{131}$$

The convective flux F_c and the turbulent velocity in the LPB treatment, given by Eqs. (107) and (108), are then written as

$$F_c = \frac{1}{2} \alpha_1 c_p \rho T \left(\frac{\Lambda_1}{H_p}\right) v_c(\nabla - \nabla_{ad}), \tag{132}$$

$$v_c^2 = \frac{Q}{8\Gamma_1} c_s^2 \left(\frac{\Lambda_2}{H_p}\right)^2 (\nabla - \nabla_{ad}). \tag{133}$$

The turbulent viscosity and the energy equation are written as

$$\nu_t = \alpha_2 v_c^2 \Omega^{-1}, \quad \frac{dF}{dz} = \frac{9}{4} \rho \nu_t \Omega^2. \tag{134}$$

Let us now consider the behavior at small z of the three quantities dF/dz, F_c, and ν_t. We have

$$\frac{dF}{dz} \sim \left(\frac{\Lambda_2}{H_p}\right)^2 (\nabla - \nabla_{ad}),$$

$$F_c \sim \frac{\Lambda_1 \Lambda_2}{H_p^2} (\nabla - \nabla_{ad})^{3/2},$$

$$\nu_t \sim \left(\frac{\Lambda_2}{H_p}\right)^2 (\nabla - \nabla_{ad}). \tag{135}$$

Consider dF/dz first. If we write in general

$$\left(\frac{\Lambda_2}{H_p}\right)^2 \sim z^b, \tag{136}$$

then Eqs. (135) acquire the form (86) with $a = 1$. We then see from Table III that the best one can hope for is a physical but degenerate solution. Since a negative value of the parameter b is hard to visualize, the best available alternative is $b = 0$, which implies that one must choose

$$\Lambda_2 = H_p = \frac{p}{\rho g} \sim \frac{1}{z}, \tag{137}$$

which is precisely LPB's choice. With this, the third of (135) yields

$$\nu_t(0) = \text{constant}, \tag{138}$$

while the second of (135) gives

$$F_c \sim \frac{\Lambda_1}{H_p} \sim z\Lambda_1. \tag{139}$$

Since a physical solution is defined as one with $F(0) = 0$, LPB could not choose $\Lambda_1 = \Lambda_2$ near midplane. They chose instead

$$\Lambda_1 = z, \quad \text{for } z < H_p. \tag{140}$$

How acceptable are these choices of Λ_1 and Λ_2? We note that:

(1) The MLT, even in its general form does not allow $\Lambda_1 \neq \Lambda_2$, as we have shown above.

(2) The scale height Λ_2 diverges at midplane, which is unphysical.

(3) LPB's choice of a Λ_2 is physically equivalent to creating an artificial energy source. This is illustrated in Figs. 7a and 7b. Figure 7a shows $\nu_t(z)$ from Eqs. (133) and (134) with Λ_2 = constant, as it would follow from local MLT. One sees that $\nu_t(z)$ vanishes at $z = 0$, which is a reasonable behavior since the convective buoyancy vanishes there. A similar behavior of $\nu_t(z)$ near midplane occurs in the CG model, as discussed in Section V. The type of distribution of turbulent viscosity in the disk shown in Fig. 7a does not, however, provide physical solutions when employed in the energy generation equation. As we have discussed in section VI, one can remedy the situation by spreading the available $\nu_t(z)$ under the area in Fig. 7a uniformally throughout the nebula. This procedure conserves the total amount of turbulence.

On the other hand, rather than redistribute ν_t, LPB artificially enhanced ν_t by a substantial amount near midplane by their choice (137). The net result is shown in Fig. 7b, where it is seen that LPB's procedure

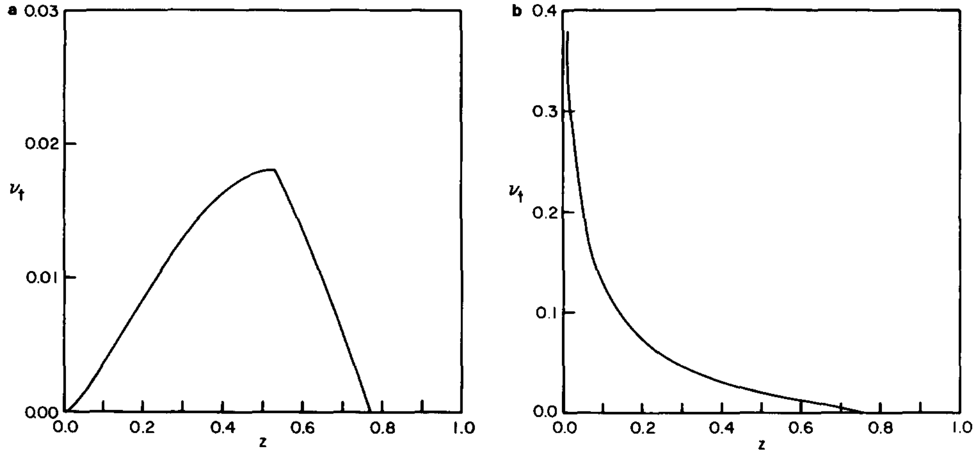

FIG. 7. (a) The local turbulent viscosity (normalized to 10^{16} cm^2 sec^{-1}) as from the application of MLT with Λ_2 = constant near midplane. To obtain solutions to the disk equations, one must average the area under the ν_t curve in the midplane convection zone. (b) The turbulent viscosity ν_t vs z as from the choice of Λ_1 and Λ_2 made by LPB. This corresponds to introducing a large artificial contribution near midplane. Same normalization as in (a).

"creates" an energy source near midplane over and above that prescribed by MLT.

(4) Even with the above choices of Λ_1 and Λ_2 to get $F(0) = 0$, the best LPB can achieve are degenerate solutions, since they do not allow the determination of the nebula's height. How H, the height of the nebula, was chosen by LPB is discussed below.

(5) The choice (137) causes the convective velocity v_c (Eq. (133)) to exceed the sound speed c_s at midplane. Since a theory of supersonic turbulence does not exist, LPB were forced to introduce an ad hoc prescription, which we shall discuss below.

(e) The Nature of the LPB Solutions for the Disk

When v_c becomes equal to c_s, LPB replace "v_c^2" with "c_s^2", so that the energy equation is now written as

$$\frac{dF}{dz} = \frac{9}{4}\alpha_2\rho\Omega \min\{v_c^2, c_s^2\}. \quad (141)$$

Although the study of the most general equation has been presented in Appendix C, here we shall solve Eq. (141) in the neighborhood of midplane[6] to show explicitly the appearance of degenerate solutions and the manner in which LPB chose a value of the nebula's height out of the degenerate set.

Explicitly, we shall show that Eq. (141) gives

$$F(0) = 0, \quad \text{for } any\ H, \text{ if } v_c < c_s,$$
$$F(0) \neq 0, \quad \text{for } any\ H, \text{ if } v_c > c_s. \quad (142)$$

In the first case, the desired zero flux condition at midplane is fulfilled but at the expense of obtaining degenerate solutions, since in fact the nebula's height H is undetermined. In the second case, one is left with a nonzero residual flux and the solutions are clearly unphysical. We shall deal with the subsonic and supersonic cases separately.

(*i*) *Subsonic convection*: $v_c < c_s$. In this case, Eq. (141) has the solutions

[6] By which we mean the region near $z = 0$ where thermodynamic quantities (ρ, T, p, c_v, etc.) are effectively constant and where the convective flux is negligible compared with the radiative flux; in practice, the solutions discussed are valid for $z \leq 10^{-2}h$, where $h^2 = zH_p$.

$$F = C_1 z^{\theta_0} + \left(\frac{\theta_0}{\theta_0 - 1}\right) f_0 \nabla_{ad} z, \quad \theta_0 \neq 1,$$
$$= -f_0 \nabla_{ad} z \ln(z/C_2), \qquad \theta_0 = 1, \quad (143)$$

where C_1 and C_2 are integration constants depending on matching conditions with solutions beyond the neighborhood of the midplane and where

$$\theta_0 \equiv \frac{9Q\alpha_2}{32} \frac{3p^2\kappa}{4acT^4\Omega}, \quad f_0 \equiv \frac{4acT^4\Omega^2}{3\kappa p}. \quad (144)$$

The convective velocity given by

$$v_c^2/c_s^2 = (Q/8\Gamma_1 f_0)[C_1 z^{\theta_0 - 1} + f_0 \nabla_{ad}/(\theta_0 - 1)], \quad \theta_0 \neq 1$$
$$= -(Q/8\Gamma_1)\nabla_{ad}[1 + \ln(z/C_2)], \quad \theta_0 = 1 \quad (145)$$

must be less than unity. Both the integration constants and the thermodynamic quantities depend on H. Since the thermodynamic constant θ_0 is positive definite, it follows that

$$F(0) = 0 \quad (146)$$

for all values of H that generate midplane conditions satisfying (145). The above condition is always satisfied at a point arbitrarily close to midplane, if

$$\theta_0 > \theta_0^{cr} \equiv 1 + \nabla_{ad} Q/8\Gamma_1. \quad (147)$$

The case $\theta_0 = \theta_0^{cr}$ belongs to the supersonic case if the convection entering the neighborhood of the midplane is supersonic, i.e., $C_1 > 0$, since $v_c^2/C_s^2 < 1$ everywhere near midplane; $\theta_0 = \theta_0^{cr}$ belongs to the subsonic case if the convection entering midplane is subsonic, i.e., $C_1 < 0$, since $v_c^2/c_s^2 < 1$ everywhere near midplane. In order that convection be supersonic entering midplane, condition (145) requires that $\nabla - \nabla_{ad} > 8\Gamma_1/Q = O(10)$; however, the requirement that θ_0 be $\theta_0^{cr} \approx 1$ in practice constrains $\nabla - \nabla_{ad}$ to be $O(1)$ entering the neighborhood of the midplane. We therefore conclude that

$$\theta_0 \geq \theta_0^{cr} \quad (148)$$

is the general criterion in practice for satisfying condition (145) at points arbitrarily close to midplane and guarantee solutions with $F(0) = 0$.

(*ii*) *Supersonic convection:* $v_c > c_s$. For

$$\theta_0 < \theta_0^{cr}, \quad (149)$$

condition (145) is violated at some positive $z < z_*$ in the upper plane and convection is supersonic at midplane. As stated above, we find in practice that the convection is subsonic entering the neighborhood of the midplane. Thus the solution switches at $z = z_*$ from Eq. (135) for $z > z_*$, to the supersonic solution

$$F = F_0 + \theta_0 f_0 (8\Gamma_1/Q) z, \quad z < z_*, \quad (150)$$

where F_0 is an integration constant which is determined by matching the fluxes at $z = z_*$ and eliminating C_1 or C_2 via Eq. (145) for $v_c^2/c_s^2 = 1$. The result is

$$F_0 = (8\Gamma_1/Q) f_0 z_* (\theta_0^{cr} - \theta_0). \quad (151)$$

Because of condition (149), $F_0 > 0$. Since the residual flux at midplane given by Eq. (150) is $F(0) = F_0$, we find for values of H satisfying condition (149) that

$$F(0) > 0, \quad (152)$$

and *so no physical solution can be obtained.*

The solutions to LPB's structure equations thus fall into two categories (see Figs. 8a–8c):

(1) When $v_c \leq c_s$ at midplane, $F(0) = 0$ and the solution is physically consistent;

(2) When $v_c > c_s$, $F(0) > 0$ and the solution is physically inconsistent.

However, even in the first case, there is still in general a continuum of values of H able to satisfy $F(0) = 0$. The solutions depend on H only to the extent that the boundary(ies) where $\theta_0 = \theta_0^{cr}$ is (are) a function of H. Figures 8a–8c show θ_0 are $F(0)$ plotted against H for three values of T_e using the same opacities employed by LPB. Because θ_0 depends linearly on opacity, the θ_0 curves reflect peaks in midplane opacities

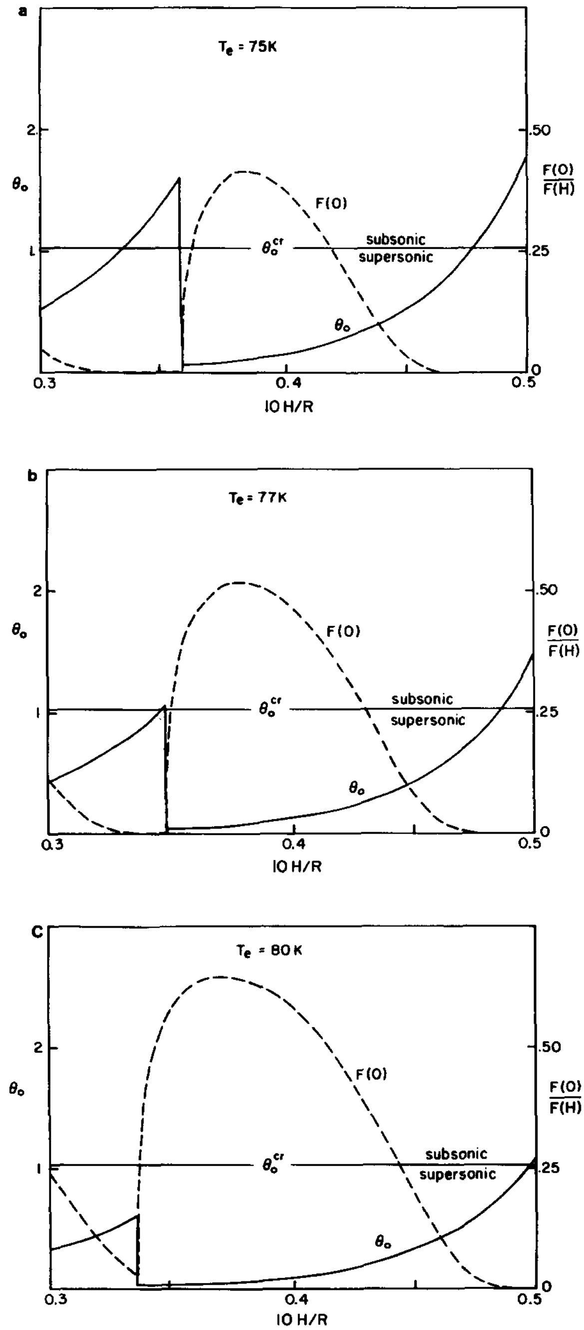

FIG. 8. The behavior of $F(0)$, the flux at midplane, in the LPB model. Physical solutions are possible only when $F(0) = 0$ is satisfied. (a) For $T_e = 75°K$, there are two degenerate regions where $F(0) = 0$), the first is bounded, the second is unbounded. (b) For $T_e = 77°K$, the first region of (a) is reduced to a single value, whereas the second degenerate region still exists. (c) For $T_e = 80°K$, the first region in (a) has disappeared and only the second degenerate region remains.

that determine, along with T_e, where and how many $\theta_0 = \theta_0^{cr}$ boundaries occur. For sufficiently large H, θ_0 always becomes greater than θ_0^{cr}. At smaller H, the peak in θ_0 (due to water-ice grain opacities) surpasses θ_0^{cr} provided T_e is less than some critical value (in the case of Figs. 8, the critical value is seen to be 77°K).

For $T_e < T_e^{cr}$, there is a continuous, *bounded* range of H with solutions satisfying $F(0) = 0$ and, at larger values of H, another such continuous, but *unbounded* range exists; for $T_e > T_e^{cr}$, only the latter range exists. Any value of H in these two ranges has solutions satisfying the midplane boundary condition of vanishing flux. These solutions are "degenerate" with respect to boundary conditions at the optical surface (see Appendix C).

One can force a "unique" solution by requiring that

$$T_e = T_e^{cr} \tag{153}$$

at *each* value of R, the distance from the Sun. We have found numerically that the values of T_e so obtained depend on R as

$$T_e \sim R^{-1/6}, \tag{154}$$

whereas for a quasistatic disk, one has (Pringle, 1981)

$$T_e \sim R^{-3/4}. \tag{155}$$

Therefore condition (153) cannot be imposed, or conversely (153) holds at *only one* value of R.

We surmise that LPB have chosen H as the point where θ_0 first becomes $\theta_0^{cr} \approx 1.03$ with increasing H for a fixed T_e. The point where $T_e = T_e^{cr}$ defines where LPB switch between solutions with small and large H. We must stress, however, that there is nothing inherent in the LPB formulation to constrain the value of θ_0 to θ_0^{cr} or any larger value.

The solution $\theta_0 = \theta_0^{cr}$ above is equivalent to having the turbulent viscosity parameter

$$\alpha = \alpha_2 \frac{v_c^2}{c_s^2}, \tag{156}$$

with $\alpha_2 \approx 1$, be locally equal to unity at midplane. Since $\alpha \sim \nu_t$ is strongly peaked toward midplane (Fig. 7a), the vertically averaged value of α is somewhat less than unity (typically about $10^{-1.5}$).

What size of α would LPB have obtained, however, had they used a consistently vertically averaged ν_t to determine the vertical structure? The largest value of $\nabla - \nabla_{ad}$ in Eq. (133) is $\nabla_r - \nabla_{ad}$ for a negligible convective flux[7]; here ∇_r is the temperature gradient if all the flux were radiative,

$$\nabla_r \equiv \frac{1}{\rho\chi}\frac{p}{c_p\rho T}\frac{F}{g}. \tag{157}$$

If ν_t is replaced in the second of Eqs. (134) by a constant averaged $\langle\nu_t\rangle$, and because the density does not vary drastically near midplane, we estimate

$$F \simeq \frac{9}{4}\rho\langle\nu_t\rangle\Omega^2 z = \frac{9}{4}\rho\langle\nu_t\rangle g. \tag{158}$$

Also using $p/c_p\rho T = R_g/\mu c_p \approx \nabla_{ad}$, we find

$$\nabla_r - \nabla_{ad} \approx \left(\frac{9}{4}\frac{\langle\nu_t\rangle}{\chi} - 1\right)\nabla_{ad}. \tag{159}$$

For a consistent solution, we require $\langle\nu_t\rangle \approx \chi$ in the midplane region; thus $\nabla_r - \nabla_{ad} \approx \nabla_{ad}$. Equation (156) becomes

$$\alpha \leq \alpha_2 \frac{Q\nabla_{ad}}{8\Gamma_1}\left(\frac{\Lambda_2}{H_p}\right)^2. \tag{160}$$

Since $Q \approx 1$ and $\Gamma_1^{-1} \approx 1 - \nabla_{ad}$, $Q\nabla_{ad} \leq \frac{1}{4}\Gamma_1$. Using LPB's choice of $\Lambda_2 = H_p$ and $\alpha_2 = 1$, we find a maximal local value of

$$\alpha \lesssim \frac{1}{32}. \tag{161}$$

We expect the vertically averaged values of α to be somewhat lower, suggesting a maximal global value of $\alpha_c \sim O(10^{-2.5})$, which is consistent with our numerical results.

[7] When radiative dissipation is correctly taken into account, as in Cox and Giuli (1968), the maximal superadiabatic temperature gradient experienced by convecting elements is $\frac{1}{4}(\nabla_r - \nabla_{ad})$.

(*f*) *Summary*

(1) *The LPB model contains four free parameters* (two lengths Λ_1 and Λ_2 and two strengths α_1 and α_2), which were assumed to be

$$\Lambda_1 \neq \Lambda_2, \tag{162}$$

$$\alpha_1 = \alpha_2 = 1. \tag{163}$$

(2) *Radiative losses by the eddies were assumed to be zero:* LPB's convective formulae do not contain χ. This overestimates the convective efficiency. At midplane, where gravity goes to zero, $F_c \sim \chi^{-3}$ and $v_c \sim \chi^{-1}$ and the $\chi \to 0$ limit leads to divergences.

(3) *Coriolis forces were neglected in LPB convective formulae* in spite of the Ω appearing in ν_t. This holds only if $x \gg 1$, which must therefore be considered implicitly assumed in the LPB treatment.

(4) *Eddy anisotropy effects were neglected.* LPB's convective formulae do not depend on x, or equivalently, they are valid for the $x \approx 1$ case. Furthermore:

(5) When (2) and (3) are assumed, the MLT formalism implies that

$$\Lambda_1 = \Lambda_2, \tag{164}$$

in disagreement with (162). On the other hand, if (164) is accepted, no physical solutions of the disk equations can be found.

(6) The parameters α_1 and α_2 can be shown to satisfy

$$\alpha_1 \sim \alpha_2 \sim \frac{1}{x} \ll 1, \tag{165}$$

in disagreement with (163).

(7) Points (3) and (4) above are not mutually compatible.

(8) The choice $\Lambda_2 \sim z^{-1}$ (which diverges at $z \to 0$) and $\alpha_2 \sim 1$ imply that artificial turbulent viscosity is being created at midplane, thus resulting in an unphysical source of heat (Figs. 7).

(9) With (1)–(4), the solutions for the disk height H are *degenerate*, i.e., the model cannot provide a unique solution for the nebula's height, allowing for an infinite set of values of H. LPB chose the smallest value of this degenerate set. The arbitrariness in the choice of H is reflected in the value of the nebula's surface density (Σ) and mass (M).

(10) The use of a local ν_t is not compatible with the general energy equation. To be consistent with Eq. (39), one must use an averaged ν_t. If so, we find that *even without* the effects of eddy anisotropy, radiative losses, and rotation, the resulting α_c would be about $10^{-2.5}$ rather than about $10^{-1.5}$ as in LPB.

VIII. CONCLUSIONS

In this paper we develop a new approach to constructing turbulent disk models of the inner solar nebula in which turbulence is driven by thermal convection. We assume a thin-disk geometry, as is done in the seminal work by LPB, with which we make extensive comparisons. We use the CG method to specify the relations needed to describe convective turbulence; one advantage of the model is its flexibility in allowing us to include rotational effects on convective motions. The way in which we propose to implement our turbulence model in numerical modeling of the vertical disk structure differs markedly from LPB in some respects. The main difference is that convective–turbulent properties in LPB's models are wholly local in application and feature a turbulent viscosity that is strongly peaked at midplane. In the method developed in this paper, the turbulent viscosity is created predominantly in regions above and below midplane, but it is assumed to be redistributed evenly in the convective region by nonlinear transport; it is this smoothed turbulent viscosity and heat dissipation that is used to determine the vertical thermal structure of the disk.

The numerical results of the model developed herein are presented in Paper II of this work (Cabot *et al.*, 1987). Similarities and differences in approach between ourselves and LPB lead to results that are both similar and different. Unexpectedly, the differences prove to be especially striking with respect to the stability of the models.

APPENDIX A

Maximization of the Sum of Growth Rates with Positive Real Parts

The cubic that governs the growth rates with zero kinematic viscosity (Eq. (70)) is

$$C(n) = n^3 + \chi k^2 n^2 - \left(g\alpha\beta \frac{x}{1+x} - \frac{\Omega^2}{1+x}\right) n + \frac{\Omega^2}{1+x}\chi k^2 = 0. \quad \text{(A1)}$$

For $g\alpha\beta > 0$, Eq. (A1) has two roots with positive real parts (n_1 and n_2, say) satisfying $C(n_1) = C(n_2) = 0$; n_1 and n_2 can both be real or a complex conjugate pair, such that their sum is always real and positive. From Eq. (A1), we can write

$$C(n_2) - C(n_1) = (n_2 - n_1)\Big[(n_1 + n_2)^2 + \chi k^2(n_1 + n_2) - \left(g\alpha\beta \frac{x}{1+x} - \frac{\Omega^2}{1+x}\right) - n_1 n_2\Big] = 0. \quad \text{(A2)}$$

Since $n_2 \neq n_1$ in general, Eq. (A2) gives

$$n_1 n_2 = (n_1 + n_2)^2 + \chi k^2(n_1 + n_2) - \left(g\alpha\beta \frac{x}{1+x} - \frac{\Omega^2}{1+x}\right). \quad \text{(A3)}$$

We also have

$$C(n_2) + C(n_1) = (n_1 + n_2)^3 + \chi k^2(n_1 + n_2)^2 - \left(g\alpha\beta \frac{x}{1+x} - \frac{\Omega^2}{1+x}\right)(n_1 + n_2) - 2\chi k^2 \frac{\Omega^2}{1+x} - n_1 n_2[3(n_1 + n_2) + 2\chi k^2] = 0. \quad \text{(A4)}$$

Eliminating $n_1 n_2$ from Eq. (A4) by (A3), we arrive at

$$C_\Sigma(n_1 + n_2) = (n_1 + n_2)^3 + 2\chi k^2(n_1 + n_2)^2 - \left(g\alpha\beta \frac{x}{1+x} - \frac{\Omega^2}{1+x} - \chi^2 k^4\right)(n_1 + n_2) - g\alpha\beta \frac{x}{1+x}\chi k^2 = 0, \quad \text{(A5)}$$

the cubic governing the *sum* of the growth rates with positive real parts (Eq. (72)). $C_\Sigma(n_1 + n_2)$ has either three real roots or one real and one pair of complex conjugate roots. Only one root has a positive real part. Ergo, the root with the positive real part is always a real root. This is the root of interest in Eq. (A5).

The maximization of the sum of the growth rates at the largest scale is determined implicitly from Eq. (A5) by

$$\frac{\partial C_\Sigma(n_1 + n_2)}{\partial x} = 0, \quad \text{(A6)}$$

with

$$k^2 = k_0^2(1 + x), \qquad k_0 = \text{constant}, \quad \text{(A7)}$$

and the condition

$$\frac{\partial(n_1 + n_2)}{\partial x} = 0. \quad \text{(A8)}$$

Equations (A6)–(A8) yield

$$2\chi k_0^2(n_1 + n_2)^2 - \left[\frac{g\alpha\beta + \Omega^2}{(1+x)^2} - 2\chi^2 k_0^4(1 + x)\right](n_1 + n_2) - g\alpha\beta\chi k_0^2 = 0. \quad \text{(A9)}$$

Through a series of algebraic manipulations, the $(n_1 + n_2)^3$ and $(n_1 + n_2)^2$ terms can be eliminated from Eqs. (A5) and (A9), leaving

$$\frac{n_1 + n_2}{\chi k_0^2} = \frac{B}{A}, \quad \text{(A10)}$$

where

$$A \equiv \frac{1}{2}\Gamma^2 \frac{(1 + r)^2}{(1 + x)^4} + 2\frac{1 + r}{1 + x} - 1,$$

$$B \equiv x - 1 - \frac{1}{2}\Gamma^2 \frac{1 + r}{(1 + x)^2},$$

and

$$r \equiv \frac{\Omega^2}{g\alpha\beta}, \quad \Gamma^2 \equiv \frac{g\alpha\beta}{\chi^2 k_0^4}.$$

Using Eq. (A10) to eliminate $(n_1 + n_2)$ from Eqs. (A5) and (A9) gives the equation governing the value of x that maximizes the sum of the growth rates:

$$f(x) = B^2(1 + A) + 2BA - \tfrac{1}{2}\Gamma^2 A^2 = 0. \tag{A11}$$

We solve Eq. (A11) numerically through a Newton's iteration with a reasonable guess for x. The correction to trial values of x is

$$\delta x = -f(x)/f'(x),$$

where

$$f'(x) = 2\left[B(1 + A) + A + \frac{1 + r}{(1 + x)^2}(\Gamma^2 A - 2B - B^2)\right] B',$$

$$B' \equiv 1 + \Gamma^2 \frac{1 + r}{(1 + x)^3}.$$

APPENDIX B

The spectral functions $E(k)$, $G(k)$, and $H(k)$ are defined in terms of ensemble averages of the square of the velocity, square of the temperature, and of the velocity–temperature correlation by

$$\langle v^2 \rangle = \int_{k_0}^{\infty} E(k)dk, \tag{B1}$$

$$\langle \theta^2 \rangle = \int_{k_0}^{\infty} G(k)dk, \tag{B2}$$

$$\langle v_3\theta \rangle = \int_{k_0}^{\infty} H(k)dk, \tag{B3}$$

where v_i is the ith component of the velocity v, θ is the fluctuating part of the temperature field, and k_0 is the wavenumber corresponding to the largest eddy available to the system. It is desired to express the spectral functions $G(k)$ and $H(k)$ in terms of the velocity spectral function $E(k)$.

Expanding v^2 in terms of its components, Eq. (B1) can be written

$$\langle v_3^2 \rangle (1 + \bar{x}) = \int_{k_0}^{\infty} E(k)dk, \tag{B4}$$

where $\bar{x} \equiv \langle v_1^2 + v_2^2 \rangle / \langle v_3^2 \rangle$. In the usual fashion, ratios of Eqs. (B3) and (B2), and of (B3) and (B4) are formed, giving

$$\left\langle \frac{v_3}{\theta} \right\rangle = \frac{H(k)}{G(k)} \equiv \phi(k), \tag{B5}$$

$$\left\langle \frac{\theta}{v_3} \right\rangle \frac{1}{1 + \bar{x}} = \frac{H(k)}{E(k)} = \frac{1}{\phi(k)} \frac{1}{1 + \bar{x}}, \tag{B6}$$

hence, solving for $G(k)$ and $H(k)$,

$$G(k) = \frac{1}{\phi^2(k)} \frac{E(k)}{1 + \bar{x}}, \tag{B7}$$

$$H(k) = \frac{1}{\phi(k)} \frac{E(k)}{1 + \bar{x}}. \tag{B8}$$

The approximation used to effect a solution is that the *ratios* $\bar{x}$ and ϕ are given by the linear analysis. This assumption allows a complete solution to be developed.

Consider an infinite horizontal layer of fluid. The components of the velocity in the horizontal plane can be written quite generally in terms of v_3 and the z component of vorticity ζ by

$$v_1 = \frac{1}{a_\perp^2}\left(\frac{\partial^2 v_3}{\partial x \partial z} + d\frac{\partial \zeta}{\partial y}\right), \tag{B9}$$

$$v_2 = \frac{1}{a_\perp^2}\left(\frac{\partial^2 v_3}{\partial y \partial z} - d\frac{\partial \zeta}{\partial x}\right), \tag{B10}$$

in which all distances are measured in units of the depth of the convective layer d and $a_\perp^2 = a_1^2 + a_2^2 = k^2 d^2 = (k_1^2 + k_2^2)d^2$. Since the geometry is plane parallel, the appropriate forms for v_3 and ζ are

$$v_3 = W(z) \cos a_1 x \cos a_2 y, \tag{B11}$$

$$\zeta = Z(z) \cos a_1 x \cos a_2 y. \tag{B12}$$

Using (B11) and (B12) in (B9) and (B10), squaring the velocity components, and averaging, one finds

$$\langle v_1^2 + v_2^2 \rangle = \frac{1}{4a_\perp^2} [\langle (DW)^2 \rangle + d^2 \langle Z^2 \rangle], \tag{B13}$$

where $D \equiv d/dz$. For this geometry, assume $W(z) = W_0 \cos a_3 z$, where W_0 is a constant amplitude and $a_3 = k_3 d$. Calculating $\langle (DW)^2 \rangle$, it is easily seen that $\langle (DW)^2 \rangle = 4a_3^2 \langle v_3^2 \rangle$. Substituting this expression in Eq. (B13), dividing by $\langle v_3^2 \rangle$, and adding 1, one finds from the definition of $\tilde{x}$ that ($a^2 = a_\perp^2 + a_3^2$)

$$1 + \tilde{x} = \frac{a^2}{a_\perp^2}\left[1 + \frac{a_3^2}{a^2} d^2 \frac{\langle Z^2\rangle}{\langle (DW)^2\rangle}\right]. \quad \text{(B14)}$$

Using the expressions for Z and W provided by the linear analysis, we have

$$\frac{\langle Z^2\rangle}{\langle (DW)^2\rangle} = \frac{T}{d^2 n_a^2}, \quad n_a \equiv a^2 + n_*, \quad \text{(B15)}$$

where the symbols T and n_* stand for

$$T = \left(\frac{\Omega d^2}{\nu}\right)^2, \quad n_* = nd^2/\nu. \quad \text{(B16)}$$

Here T is the Taylor number and $n(k)$ the growth rate. Finally,

$$1 + \tilde{x} = \frac{a^2}{a_\perp^2}\left(1 + \frac{a_3^2}{a^2}\frac{T}{n_a^2}\right). \quad \text{(B17)}$$

Next, let us estimate ϕ. From the equation relating $W(z)$ and θ, and the use of the dispersion relation

$$(D^2 - a_\perp^2 - \sigma n_*)[(D^2 - a_\perp^2 - n_*)^2(D^2 - a_\perp^2) + TD_2]W = -Ra_\perp^2(D^2 - a_\perp^2 - n_*)W, \quad \text{(B18)}$$

where $D \to a_3^2$, $\sigma = \nu/\chi$ is the Prandtl number, and R is the Rayleigh number $g\alpha\beta d^4/\nu\chi$, we obtain after some algebra

$$\frac{1}{1+\tilde{x}}\frac{1}{\phi} = \frac{1}{g\alpha}(n + \nu k^2), \quad \text{(B19)}$$

which together with (B8) and (B3) gives the desired result (Eq. (69) of the text), since $F_c = c_p\rho\langle v_3\theta\rangle$. Let us note that Eq. (B19) is valid for any degree of rotation.

APPENDIX C

Midplane Solutions of the Energy Equation: General Case

Introducing the dimensionless variables

$$x \equiv z/h,$$

$$f \equiv F(f_0 h \nabla_{ad})^{-1}, \quad \text{(C1)}$$

$$\theta_0 = \varepsilon_0 h^b f_0^{-1} \nabla_{ad}^{a-1},$$

Eq. (86) can be written as

$$\frac{df}{dx} = \theta_0 x^{b-a}(f - x)^a, \quad \text{(C2)}$$

where f is the radiative flux and where ε_0, F_0, and ∇_{ad} are thermodynamic quantities taken to be constant near midplane ($x = 0$).

Alternatively, one can write Eq. (C2) in terms of the superadiabatic temperature gradient y,

$$y \equiv f/x - 1, \quad \text{(C3a)}$$

$$x\frac{dy}{dx} + y + 1 = \theta_0 x^b y^a. \quad \text{(C3b)}$$

Solutions are sought for small positive x in the limit $x \to 0$. *Physically realistic solutions have $f \geq 0$ for $x > 0$ and $f = 0$ at $x = 0$;* in order to have convection to midplane, one requires $y > 0$ for $x > 0$; the energy generation rate should be zero or finite everywhere, so $x^b y^a$ should be zero or finite for $x \geq 0$. The superadiabatic temperature gradient in dimensional form, $-dT/dz + (dT/dz)_{ad}$, should vanish at midplane, thus xy should vanish at $x = 0$; this is equivalent to $f = 0$ at $x = 0$.

In the following an "inward integration" will refer to an integration from $x = x_i > 0$, with a corresponding initial value of $y = y_i > 0$, toward $x = 0$ (i.e., with decreasing values of x).

(*a*) *Special Case*: $b = 0$

Equation (C3) becomes

$$x\frac{dy}{dx} = \theta_0 y^a - (y + 1) \equiv g(y), \quad \text{(C4)}$$

which is separable. The points satisfying $g(y) = 0$ are special solutions of Eq. (C4) defining neutral equilibria. $g(y)$ has an extremum at

$$y_m = (a\theta_0)^{1/(1-a)}, \quad \text{(C5)}$$

with

$$g(y_m) = (a\theta_0)^{1/(1-a)}\left(\frac{1-a}{a}\right) - 1. \quad \text{(C6)}$$

(*i*) $a > 1$. For large y, $g(y) \approx y^a > 0$; for small y, $g(y) \approx -1$; $g(y_m) < 0$, for $a > 1$; therefore, $g(y)$ decreases from positive values at large y, through zero at $y = y_0$, decreases to a minimum at $y = y_m$, and increases to a value of -1 at $y = 0$. For an inward integration with $y_i > y_0$, $g(y) > 0$ and y decreases toward y_0.

For $y_i < y_0$, $g(y) < 0$ and y increases toward y_0. Thus y_0 is an attracting point of neutral equilibrium for inward integrations (i.e., asymptotically stable); for any initial $y_i > 0$, y tends asymptotically to y_0 at $x = 0$. As a consequence the flux, $f = (y + 1)x$, tends to $(y_0 + 1)x \rightarrow 0$. *One can therefore find a physical solution—in fact, a whole family of physical solutions—regardless of the initial integration point.*

(*ii*) $a = 1$. Equation (C4) has an attracting point of neutral equilibrium at $y_0 = (\theta_0 - 1)^{-1}$. If $\theta_0 > 1$, $y \rightarrow y_0$ as $x \rightarrow 0$. For $\theta_0 < 1$, $y_0 < 0$, such that y becomes negative for some $x > 0$, which is not a physical solution. The exact solution is

$$
\begin{aligned}
y &= Cx^{\theta_0 - 1} + (\theta_0 - 1)^{-1}, && \theta_0 \neq 1, \\
&= C - \ln x, && \theta_0 = 1, \\
f &= Cx^{\theta_0} + x\theta_0(\theta_0 - 1)^{-1}, && \theta_0 \neq 1, \\
&= (C + 1)x - x \ln x, && \theta_0 = 1,
\end{aligned}
$$

where C is the integration constant. For $\theta_0 = 1$, y blows up at $x = 0$, which causes the energy generation rate to blow up unphysically. Note that all solutions for $\theta_0 \geq 0$, physical or not, have $f = 0$ at $x = 0$. As for $a > 1$, all inward integrations lead to $f = 0$. *The physical solutions are degenerate in the sense that physical constraints at $x = 0$ ($f = 0$, etc.) cannot distinguish between initial integration points.*

(*iii*) $0 < a < 1$. For large values of y, $g(y) \approx -y < 0$. For small values of y, $g(y) \approx -1$. At an intermediate value, y_m, given by Eq. (C5), $g(y)$ has a maximum, $g(y_m)$, given by Eq. (C6). For

$$\theta_0 < \theta_{0c} \equiv \frac{1}{a}\left(\frac{a}{1 - a}\right)^{1-a}, \qquad \text{(C7)}$$

$g(y_m) < 0$, such that $g(y) < 0$ for all $y \geq 0$, and *all* inward integrations result in ever-increasing values of y. Since $g(y) \approx -y$ for large enough values of y, y will have the approximate solution of $y \approx C/x$ near $x = 0$, where C is the integration constant. The flux, $f = (y + 1)x$, goes to $x + C$; thus the flux goes to a finite positive constant at midplane.

For $\theta_0 > \theta_{0c}$, $g(y_m) > 0$, and so $g(y)$ has two zeros at, say, y_1 and y_2, with y_2 defined as the greater of the zeros. If $y_i > y_2$ in an inward integration, then $g(y) < 0$, and y increases without bound, going as C/x as $x \rightarrow 0$. If $y \leq y_i < y_2$, then $g(y) > 0$, and y decreases asymptotically toward y_1 as $x \rightarrow 0$. If $0 < y_i \leq y_1$, then $g(y) < 0$, and y increases asymptotically toward y_1 as $x \rightarrow 0$. Thus for $0 < y_i < y_2$, inward integrations are attracted to the neutral equilibrium y_1; for $y_i > y_2$, they are repelled by the neutral equilibrium y_2. For $\theta_0 = \theta_{0c}$, y_1 collapses onto $y_2 = y_m = a/(1 - a)$. Values of $y_i > y_m$ are repelled, giving $y \rightarrow C/x$ as $x \rightarrow 0$; values of $y_i \leq y_m$ are attracted, giving $y \rightarrow y_m$ as $x \rightarrow 0$.

In summary, *unphysical solutions* of the form $y \rightarrow C/x$ and $f \rightarrow C$ as $x \rightarrow 0$ occur for all values of $\theta_0 < \theta_{0c}$, for $\theta_0 = \theta_{0c}$ when $y_i > y_2 = y_1$, and for $\theta_0 > \theta_{0c}$ when $y_i > y_2$. *Physical, but indistinguishable* (*degenerate*) *solutions* of the form $y \rightarrow y_1 =$ constant and $f \rightarrow 0$ as $x \rightarrow 0$ occur for $\theta_0 \geq \theta_{0c}$ when $y_i < y_2$. A *physical solution* that is "unique" (*nondegenerate*) in the sense that the midplane conditions are peculiar to one particular initial condition occurs for $y_i = y_2$, which is one of the special solutions to Eq. (C4). In this case $y = y_2$ and $f = (1 + y_2)x = 0$ at $x = 0$.

(*iv*) $a = 0$. The one point of neutral equilibrium, $y_2 = \theta_0 - 1$, is repelling, and the special solution $y = y_2$ is *physically unique* (*nondegenerate*). The exact solution is

$$
\begin{aligned}
y &= \theta_0 - 1 + C/x, \\
f &= \theta_0 x + C,
\end{aligned}
$$

with $C = 0$ corresponding to the special solution. Values of $\theta_0 \leq 1$ give unphysical values of $y (\leq 0)$ for $x > 0$.

In summary for $b = 0$, there are a variety of types of solutions depending on the values of a and θ_0. *The only physically nondegenerate* solutions occur for $0 \le a < 1$.

It is now possible to study some more general cases with $b \neq 0$ using the foregoing techniques in this section.

(*b*) *General Cases*

For $b \neq 0$ and $a \ge 1$, consider Eq. (C3) in the form

$$x\frac{dy}{dx} = G(x, y) \equiv \theta(x)y^a - (y + 1),$$
$$\theta(x) \equiv \theta_0 x^b. \tag{C8}$$

There will be no neutral equilibria for Eq. (C8), but it will be instructive to consider lines in (x, y) corresponding to $G(x, y) = 0$.

(*i*) $b > 0$, $a > 1$. Consider the function $y_0(x)$ which is defined by $G(x, y_0) = 0$. When x is very large, $\theta(x)$ is very large, and $y_0 \approx \theta(x)^{-1/a} \sim x^{-b/a}$, which is very small. For x very small, $y_0 \approx \theta(x)^{-1/(a-1)} \sim x^{-b/(a-1)}$, which becomes very large as $x \to 0$, and blows up at $x = 0$. For $y(x) > y_0(x)$, $G(x, y) > 0$; for $y(x) < y_0(x)$, $G(x, y) < 0$. If $y_i > y_0(x_i)$, an inward integration causes y to decrease through some $y_0(x)$ into the region $y(x) < y_0(x)$, where y will increase toward $y_0(x)$, which is blowing up. If $y_i < y_0(x_i)$, y increases toward $y_0(x)$, as in the previous case. Since all integrations lead to large values of y at small x, consider the limit $y \gg 1$, in which case Eq. (C8) is approximately

$$x\frac{dy}{dx} \approx \theta_0 x^b y^a - y,$$

with solutions

$$(yx)^{1-a} \approx C + \theta_0\left(\frac{1-a}{b+1-a}\right)x^{b+1-a}, \qquad b + 1 \neq a,$$
$$\approx C - (a-1)\theta_0 \ln x, \quad b + 1 = a,$$

where C is the integration constant. If $b + 1 > a$, then $y \to C^{1/(1-a)}/x$ as $x \to 0$, and $f \to C^{1/(1-a)}$. If $b + 1 < a$, then

$$y \to [\theta_0(a-1)/(a-1-b)]^{-1/(a-1)}x^{-b/(a-1)},$$
$$f \to (y+1)x = x + [\theta_0(a-1)/(a-1-b)]^{-1/(a-1)}x^{1-b/(a-1)};$$

since $b/(a-1) < 1$, y blows up, but $f \to 0$ as $x \to 0$. If $b + 1 = a$, then

$$y \to \frac{1}{x[-(a-1)\theta_0 \ln x]^{1/(a-1)}},$$
$$f \to x + [-(a-1)\theta_0 \ln x]^{-1/(a-1)};$$

thus y blows up as $x \to 0$, but $f \to 0$. In summary, $f \to 0$ for all inward integrations when $b + 1 \le a$, and $f \to$ constant for $b + 1 > a$. In all cases, y blows up at $x = 0$; for $b + 1 \le a$, the energy generation rate goes as

$$x^b y^a \sim x^{-b/(1-a)}, \qquad b + 1 < a,$$
$$\sim x^{-1}(-\ln x)^{-a/(a-1)}, \quad b + 1 = a,$$

which blows up at $x = 0$. *Therefore no strictly physical solutions exist for* $b > 0$, $a > 1$.

(*ii*) $b < 0$, $a > 1$. For x very large, $y_0(x) \approx \theta^{-1/(a-1)} \sim x^{-b/(a-1)}$, and for x very small, $y_0(x) \approx \theta(x)^{-1/a} \sim x^{-b/a}$; thus $y_0(x)$ decreases from large values at large x toward zero at small x. For $y > y_0(x)$, $G(x, y) > 0$; for $y < y_0(x)$, $G(x, y) < 0$. If $y_i > y_0(x_i)$, then y decreases in an inward integration toward $y_0(x) \to 0$. If $y_i < y_0(x_i)$, then y increases through y_0 at some x, and then decreases toward $y_0(x) \to 0$. Since y always tends to small values as $x \to 0$, consider the limit where $y \ll 1$: $y(x)$ clearly has a special solution of $\theta(x)^{-1/a}$ as $x \to 0$ since y and xdy/dx will vanish for this solution at $x = 0$. Letting $y = \theta(x)^{-1/a} + v$, one has

$$x\frac{dv}{dx} = -v - (1 - b/a)\theta(x)^{-1/a} + av\theta(x)^{1/a},$$

which has the solution

$$v = Cv^*(x) - (1 - b/a)\int^x [\theta(x)^{-1/a}/v^*(x)]dx,$$
$$v^* \equiv x^{-1}\exp[(a^2/b)\theta(x)^{1/a}].$$

In a set of manipulations, that will not be shown here (but see case b (iii) following), it can be shown that

$$\lim_{x\to 0} v = \frac{1}{a}(1 - b/a)\theta(x)^{-2/a} \to 0,$$

independent of the integration constant, since $v^*(x)$ goes to zero more rapidly as x^{-1} exp[$-$ constant $x^{b/a}$]. In summary, *all inward integrations for $b < 0$, $a > 1$ give $y \to 0$ and $f \to 0$ as $x \to 0$, corresponding to physical, "degenerate" solutions.*

(*iii*) *$a = 1$, $b \neq 0$.* Equation (C8) is linear in y and has the exact solution

$$y = \frac{1}{x} e^{\theta_0 x^b/b} \left(C - \int^x e^{-\theta_0 x^b/b}\, dx\right).$$

Evaluating C at x_i where $y = y_i$, one has

$$y = y_i \left(\frac{x_i}{x}\right) e^{\theta_0/b(x^b - x_i^b)} + \frac{1}{x}\int_x^{x_i} e^{\theta_0/b(x^b - x'^b)}\, dx'.$$

For $b > 0$, the integral term is positive definite, therefore $y \to$ constant/x as $x \to 0$ unless the constant is zero; for that to happen, y_i must be negative, which is not physically allowed. *Thus, for $b > 0$, the flux $f \to$ constant as $x \to 0$.* For $b < 0$, the first term vanishes for $x = 0$; one must now evaluate the integral term for small x. Let $u(x') = \exp \theta_0/b(x^b - x'^b)$, then

$$\frac{1}{x}\int_x^{x_i} u(x')\, dx' = -\frac{1}{x}\int_{u(x)}^{u(x_i)} \theta_0^{-1} x'^{(1-b)}\, du(x'),$$

where $u(x) = 1$ and $u(x_i) \to 0$ as $x \to 0$. Thus

$$\frac{1}{x}\int_x^{x_i} u(x')\, dx' \approx \int_0^1 \frac{x^{-b}}{\theta_0} \left(1 - \frac{b \ln u}{\theta_0 x^b}\right)^{(1-b)/b} du.$$

Since $x^{-b} \to 0$ for $x \to 0$ faster than ln u blows up for $u \to 0$, $x^{-b} \ln u \to 0$ for all $u \in [0, 1]$ as $x \to 0$, and

$$y \approx \frac{1}{x}\int_x^{x_i} e^{\theta_0/b(x^b - x'^b)}\, dx' \to \frac{x^{-b}}{\theta_0},$$

$$f = (y + 1)x \to x^{1-b}/\theta_0.$$

Thus for all $b < 0$, y vanishes and *f* vanishes at $x = 0$ in a family of *physical, but degenerate solutions.*

(*iv*) *$0 < a < 1$, $b > 0$.* This case superficially resembles case a(iii) for $b = 0$. For sufficiently large values of x, $\theta(x) = \theta_0 x^b > \theta_{0c}$ and $G(x, y)$ has two zeros, $y_2(x)$ and $y_1(x)$, with $y_2(x) > y_1(x)$; however, $\theta(x)$ decreases with decreasing x through θ_{0c} at x_c, where $y_2(x) = y_1(x)$, and to values less than θ_{0c} for $x < x_c$, where $G(x, y)$ has no zeros. For $y > y_2(x)$ and $y < y_1(x)$ when $x > x_c$, and for all y when $x < x_c$, $G(x, y) < 0$; for $y_1(x) < y < y_2(x)$ when $x > x_c$, $G(x, y) > 0$. Consider integration paths with $x_i > x_c$: When $y_i > y_2(x_i)$, y increases without bound for an inward integration, such that $G(x, y) \to -y$ and $y \to C/x$. For $y_i < y_2(i)$, y tends toward $y_1(x)$ with decreasing x; but when x is less than x_c, y_1 (and y_2) cease to exist, and since $G(x, y) < 0$, y again increases without bound with $y \to C/x$. In this case all inward integrations lead to $y \to C/x$ and $f \to C$ as $x \to 0$, *leading to unphysical solutions.*

(*v*) *$0 < a < 1$, $b < 0$.* For sufficiently large values of x, $\theta(x) < \theta_{0c}$ and $G(x, y)$ has no zeros; in fact, $G(x, y)$ is negative for all $x > x_c$. At $x = x_c$, there is one value of y where $G(x, y) = 0$, i.e., $\theta(x) = \theta_{0c}$. For $x < x_c$, $G(x, y)$ has two zeros, $y_2(x)$ and $y_1(x)$, with $y_2(x) > y_1(x)$. For $y > y_2(x)$ or $y < y_1(x)$, $G(x, y) < 0$; for $y_1 < y < y_2$, $G(x, y) > 0$. As x approaches 0, $y_2(x) \approx \theta(x)^{1/(1-a)} \sim x^{b/(1-a)}$, which blows up at $x = 0$; $y_1(x) \approx \theta(x)^{-1/a} \sim x^{-b/a}$, which goes to zero at $x = 0$. The value of x_c is given by Eq. (C7):

$$x_c = \left[\frac{1}{a\theta_0}\left(\frac{a}{1-a}\right)^{1-a}\right]^{1/b};$$

$G(x_c, y_m) = 0$ for $y_m = a/(1 - a)$. The behavior of solutions on an inward integration is determined by the value of y at x_c; therefore, consider y_i at $x_i = x_c$. For $y_i \leq y_m$, y will increase through $y_1(x)$, and then decrease and be attracted to $y_1(x)$ as $x \to 0$. The behavior of y is more complicated for $y_i > y_m$ at $x_i = x_c$. $G(x_c, y)$ is initially less than zero, but it is possible for y to increase slower than $y_2(x)$ increases, move through

y_2 at some x, and decrease asymptotically to the $y_1(x)$ solution. In other cases y may increase, above the $y_2(x)$ line, and blow up as $x \to 0$ with some negative power of x less than or equal to $b/(1 - a)$. To explore this asymptotic behavior, consider $y \gg 1$, where Eq. (C8) has the approximate solution

$$(yx)^{1-a} = C + \left(\frac{1-a}{b+1-a}\right) \theta_0 x^{b+1-a}, \quad b + 1 \neq a,$$

$$= C + (1-a)\theta_0 \ln x, \quad b + 1 = a.$$

If y has the value $y_A \gg 1$ at some $x_A \ll x_{c}$, then

$$C = (yx)_A^{1-a} - \left(\frac{1-a}{b+1-a}\right) \theta_0 x_A^{b+1-a}, \quad b + 1 \neq a,$$

$$= (yx)_A^{1-a} + (1-a)\theta_0 \ln x_A, \quad b + 1 = a.$$

For $b < -(1-a)$, x^{b+1-a} blows up with $x \to 0$. The constant $(1-a)/(b+1-a)$ is negative, so $C > 0$ and

$$y^{1-a} = Cx^{-(1-a)} + (1-a)/(b+1-a)\theta_0 x^b,$$

which will asymptotically *decrease* as $x \to 0$. Let $\gamma \equiv -(1-a)/(b+1-a)$; then $y = y_2 \approx (\theta_0 x^b)^{1/(1-a)}$ at

$$(x/x_A)^{-b-(1-a)} = \frac{\gamma + 1}{\gamma + (y/y_2)_A^{1-a}}.$$

Once y has decreased to values less than y_2, it will continue decreasing toward $y_1(x)$. For $b + 1 = a$,

$$(yx)^{1-a} = (yx)_A^{1-a} + (1-a)\theta_0 \ln(x/x_A),$$

where clearly the logarithmic term will cause y to go through y_2 for some value of x; y then decreases toward $y_1(x)$.

For $b > -(1-a)$, x^{b+1-a} goes to zero as $x \to 0$ and the constant $\gamma \equiv (a-1)/(b+1-a) > 0$, so C can have positive, negative, or zero values. If $C < 0$, y will pass through y_2 at

$$(x/x_A)^{b+1-a} = \frac{\gamma - (y/y_2)_A^{1-a}}{\gamma - 1},$$

where $(y/y_2)_A^{1-a} \leq \gamma$. If $C > 0$, i.e., $(y/y_2)_A^{1-a} > \gamma$, then y increases away from $y_2(x)$, going asymptotically as $y \approx C^{1/(1-a)}/x$ as $x \to 0$. If $C = 0$, i.e., $(y/y_2)_A^{1-a} = \gamma$, then

$$y = \gamma^{1/(1-a)} y_2 = \left[\left(\frac{1-a}{b+1-a}\right) \theta_0 x^b\right]^{1/(1-a)}. \tag{C9}$$

It is therefore possible to have a physically unique solution ($C = 0$), for which y blows up as $x^{b/(1-a)}$ and f vanishes as $x^{1/\gamma}$. However, the energy generation rate, proportional to $y^a x^b$, goes as $x^{b/(1-a)}$, which blows up unphysically as $x \to 0$. The value of y_i at x_c for $y(C = 0)$ is less than that predicted by the asymptotic form in Eq. (C9) due to terms neglected in its approximation. Letting

$$y \equiv y_0(x)\omega(x),$$

where $y_0(x)$ is the solution y in Eq. (C9), Eq. (C8) becomes

$$x\frac{d\omega}{dx} = \gamma^{-1}(\omega^a - \omega) - y_0^{-1}(x).$$

Assume that $\omega = 1 + \varepsilon$ where $|\varepsilon| \ll 1$; then

$$x\frac{d\varepsilon}{dx} \approx -\gamma^{-1}(1-a)\varepsilon - y_0^{-1}(x),$$

with solution

$$\varepsilon = Cx^{-(1-a)/\gamma} - x^{-(1-a)/\gamma} \int^x y_0^{-1} x^{-1+(1-a)/\gamma}\, dx.$$

Since $\varepsilon = 0$ at $x = 0$, $C = 0$, hence

$$\varepsilon = -[(1-a)/\gamma - b/(1-a)]^{-1} y_0(x),$$

$$= -[1 - a/\gamma]^{-1}[\gamma\theta_0 x^b]^{-1/(1-a)},$$

and

$$y \approx y_0(x) - (1 - a/\gamma)^{-1}$$

were $|\varepsilon(x_c)| \ll 1$.

(*vi*) *a = 0, b ≠ 0.* Eq. (C3) becomes

$$x\frac{dy}{dx} = \theta_0 x^b - (y + 1)$$

with exact solution

$$y = \frac{C}{x} + \frac{1}{1+b}\,\theta_0 x^b - 1, \quad b \neq -1,$$

$$\frac{C}{x} + \frac{1}{x}\,\theta_0 \ln x - 1, \quad b = -1.$$

For *$b > 0$, $y \to C/x$* $- 1$ as $x \to 0$, so all solutions with positive (physical) values of y go to $y = C/x$ ($C > 0$) at $x \to 0$. This is an extension of case b(iv) above. For $b < 0$, both the x^{-1} and x^b or $x^{-1} \ln x$ terms blow up as $x \to 0$. If $b \leq -1$, the x^b or $x^{-1} \ln x$ term dominates, sending y to negative values for some $x > 0$. Therefore there are *no physical solutions for $b \leq -1$*. For $-1 < b < 0$, the $1/x$ term will dominate as $x \to 0$ unless $C = 0$. If $C < 0$, y will go to negative values for some $x > 0$, which is physically disallowed; if $C > 0$, y will go as C/x, blowing up with positive values as $x \to 0$, and the flux $f = (y + 1)x \to C$, if $C = 0$, y blows up as x^b, but $f \to 0$ as x^{b+1}. The case $a = 0$, $b < 0$ is an extension of case (v) above, except that $y_2(x)$ extends down to $y_m = 0$, and the $y_1(x)$ branch (and asymptotic solution) no longer exists. In summary, only $-1 < b < 0$ in this case allows solutions with $f = 0$ at $x = 0$; however, the energy generation rate, proportional to x^b, blows up at $x = 0$, so *there are no strictly physical solutions.*

(c) *Summary*

Physical solutions are those with midplane flux $f(0) = 0$ and midplane superadiabatic temperature gradient $y(0) = 0$ or a finite positive value.

Physical solutions do not exist for $b > 0$ and $a \geq 0$; for $b < 0$ and $a = 0$, and for some cases with $0 < a \leq 1$ and $-(1 - a) < b \leq 0$.

Physical solutions exist for $a > 0$ and $b \leq \min(0, a - 1)$ excluding $a = 1$; $b = 0$; and for some cases with $0 < a \leq 1$ and $-(1 - a) < b \leq 0$; however, only the physical solutions with $b = 0$ and $0 \leq a < 1$ are nondegenerate in the sense that one and only one initial integration point in an inward integration will give the value $y(0) = y_2$, a thermodynamic constant. *It is therefore the solutions with $b = 0$ and $0 \leq a < 1$ that provide a unique boundary value problem in the vicinity of the midplane.*

ACKNOWLEDGMENTS

We are grateful to Al Cameron, John Cannizzo, Pat Cassen, Doug Lin, Steve Ruden, Frank Shu, David Stevenson, Mike Torbett, and Stu Weidenschilling for useful and enlightening discussions and comments. Part of this work was done while one of the authors (W.C.) held a NRC-NASA Research Associateship. Doris Smith skillfully typed the manuscript.

REFERENCES

Boussinesq, J. (1877, 1897). Cited in Monin and Yaglom, 1971.

Cabot, W., V. M. Canuto, O. Hubickyj, and J. B. Pollack (1987). The role of turbulent convection in the primitive solar nebula: II. Results. *Icarus* **69,** 423–457.

Cameron, A. G. W. (1969). Physical conditions in the primitive solar nebula. In *Meteorite Research* (P. M. Millman, Ed.), pp. 14–15. Springer-Verlag, New York.

Cameron, A. G. W. (1977). The primitive solar accretion disk and the formations of the planets. In *The Origin of the Solar System* (S. F. Dermott, Ed.), pp. 49–74. Wiley, New York.

Cameron, A. G. W. (1978). Physics of the primitive solar accretion disk. *Moon Planets* **18,** 5–40.

Canuto, V. M., W. Cabot, G. J. Hartke, and A. Battaglia (1985). Is the Kolmogoroff model applicable to large scale turbulence? *Astrophys. Space Sci.* **116,** 367–375.

Canuto, V. M., and I. Goldman (1985). Analytical model for large scale turbulence. *Phys. Rev. Lett.* **54,** 430–433.

Canuto, V. M., I. Goldman, and O. Hubickyj (1984). A formula for the Shakura–Sunyaev turbulent viscosity parameter. *Astrophys. J.* **280,** L55–L58.

Cox, J. P., and R. T. Giuli (1968). *Principles of Stellar Structure.* Gordon & Breach, New York.

Goldreich, P., and G. Schubert (1967). Differential rotation in stars. *Astrophys. J.* **150,** 571–587.

Gough, D. O. (1976). The current state of stellar mixing length theory. In *Problems in Stellar Convection* (E. A. Spiegel and J. P. Zahn, Eds.), pp. 15–56. Springer-Verlag, New York.

Gough, D. O. (1977). Mixing length theory for pulsating stars. *Astrophys. J.* **214,** 196–213.

Gough, D. O. (1978). The estimates of fluxes due to small scale turbulence. In *Proceedings of the Workshop on Solar Rotation* (G. Belvedere and L. Paterno, Ed.), pp. 337–351. Oss. Astr. Catania, No. 162.

Gough, D. O., and N. O. Weiss (1976). The calibration of stellar convection theories. *Mon. Not. R. Astron. Soc.* **176,** 589–607.

HINZE, J. O. (1959). *Turbulence*. McGraw-Hill, New York.

HUNTER, C. (1972). Self-gravitating gaseous disks. *Annu. Rev. Fluid Mech.* **4,** 219–242.

LANG, K. R. (1980). *Astrophysical Formulae*, p. 226. Springer-Verlag, New York.

LIN, D. N. C. (1981). Convective accretion disk model for the primordial solar nebula. *Astrophys. J.* **246,** 972–984.

LIN, D. N. C., AND P. BODENHEIMER (1982). On the evolution of convective accretion disk models of the primordial solar nebula. *Astrophys. J.* **262,** 768–779.

LIN, D. N. C., AND J. PAPALOIZOU (1980). On the structure and evolution of the primordial solar nebula. *Mon. Not. R. Astron. Soc.* **191,** 37–48.

MESTEL, L. (1972). Magneto-hydrodynamics, hydrodynamics, dynamics of the solar system in different models. In *Symposium on the Origin of the Solar System* (H. Reeves, Ed.), pp. 21–27. CNRS, Paris.

MONIN, A. S., AND A. M. YAGLOM (1971). *Statistical Fluid Mechanics*. MIT Press, Cambridge.

ÖPIK, E. J. (1950). Transport of heat and matter by convection in stars. *Mon. Not. R. Astron. Soc.* **110,** 559–589.

PRANDTL, L. (1925). Bericht Über Untersuchungen zur Ausgebildeten Turbulenz. *Zs. Angew. Math. Mech.* **5**(2), 136–139.

PRINGLE, J. E. (1981). Accretion disks in astrophysics. *Annu. Rev. Astron. Astrophys.* **19,** 137–162.

SAFRONOV, V. S. (1969). *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*. NASA TT F-677, 1972.

SPIEGEL, E. A. (1963). A generalization of the mixing length theory of turbulent convection. *Astrophys. J.* **138,** 216–225.

SPIEGEL, E. A. (1964). The effect of radiative transfer on convective growth rates. *Astrophys. J.* **139,** 959–974.

SPIEGEL, E. A. (1966). Energy transport by turbulent convection. In *Stellar Evolution* (R. F. Stein and A. G. W. Cameron, Eds.), pp. 143–173. Pergamon, New York.

SPIEGEL, E. A., AND G. VERONIS (1960). On the Boussinesq approximation for a compressible fluid. *Astrophys. J.* **131,** 442–447.

SPIEGEL, E. A., AND J. P. ZAHN (1970). Instabilities of differential rotation. *Commun. Astrophys. Space Sci.* **2,** 178–183.

STEWART, J. M. (1976). The hydrodynamics of accretion disks II: Turbulent models. *Astron. Astrophys.* **49,** 39–51.

TAYLOR, G. I. (1915). Eddy motion in the atmosphere. *Philos. Trans. R. Soc. A* **215,** 1–26.

TER HAAR, D. (1972). Some remarks on solar nebula type theories of the origin of the Solar System. In *Symposium on the Origin of the Solar System* (H. Reeves, Ed.), pp. 71–79. CNRS, Paris.

YAMAGUCHI, S. (1963). On turbulent thermal convection. *Publ. Astron. Soc. Japan* **15,** 412–427.