

The rate of mixing in semiconvective zones

H.C. Spruit

Max Planck Institut für Astrophysik, Karl-Schwarzschildstrasse 1, D-W-8046 Garching, Federal Republic of Germany

Received September 5, 1990; accepted June 1, 1991

Abstract. Mixing in semiconvective zones is treated as a double-diffusive process. The mixing takes place by overturning cells in horizontal layers, separated by stable interfaces across which transport takes place by diffusion. Effective diffusion coefficients are derived and tested against laboratory results. Easily implementable expressions for stellar evolution calculations are given. They imply a much smaller mixing rate than assumed in other formulations.

Key words: semiconvection – stellar evolution – SN 1987A

1. Introduction

The problems introduced by semiconvective zones were first noted by R.J. Tayler around 1953. Since then, many different practical recipes have been used to deal with the mixing in such zones in stellar evolution calculations. In one of the more widely used recipes, the stabilizing μ -gradient is smoothed such that the resulting gradient is marginally stable to adiabatic instability (Ledoux criterion). This implies a fair amount of mixing. One of the more recent prescriptions, which yields less mixing, is that of Langer et al. (1983).

The fluid mechanical problem of a stably stratified composition gradient heated from below has been studied extensively, both experimentally and theoretically, in connection with geophysical observations (for reviews see Turner 1985; Huppert & Turner 1981). The studies have yielded detailed insight into the physics of this problem. In addition to presenting these insights where relevant for the astrophysical parameter range, I use them in the present paper to develop a simple theory of semiconvective mixing. Aspects of such a theory have previously been discussed by Spiegel (1972), and Gough & Toomre (1982). The problem is closely related to that of enhanced diffusion in steady flows (Batchelor 1956; Parker 1963; Weiss 1966). In the semiconvective case, heat is the destabilizing agent and the μ -gradient the stabilizing one; the opposite can also occur and has been studied in the astrophysical context by Ulrich (1972) and Kippenhahn et al. (1982).

The starting point of the considerations is that the overstable oscillations predicted by linear theory turn out to play only a very limited role in practice because a minor (nonlinear) rearrangement in the fluid can turn it into a sequence of overturning layers. This form of motion is stable and preferred above overstable

oscillations because it can occur at values of the temperature gradient lower than that required for linear instability¹.

2. Double diffusive mixing

2.1. The formation of layers

Though the overstable oscillations predicted by linear theory have been reproduced in the laboratory, they require very well controlled and stable conditions. The linear growth is usually swamped by the nonlinear development of overturning layers. The reason is that it requires only a very small amount of energy to overturn the initially smooth stratification into a set of layers if these layers are *thin* enough. Once the layers are formed, they are a stable form of convection. Let N be the (adiabatic) buoyancy frequency,

$$N^2 = g \frac{\partial \ln \mu}{\partial r} + \frac{g}{H} (\nabla_a - \nabla), \quad (1)$$

where g is the acceleration due to gravity, μ the mean weight per particle, H the pressure scale height, ∇ the logarithmic temperature gradient, and ∇_a the adiabatic gradient. Since the stratification is stable according to the Ledoux criterion, N^2 is positive. The energy required for overturning over a vertical distance d is of the order

$$e_{\text{overturn}} \sim d^2 N^2 \quad (2)$$

(per gram). A minimum value to d is set by viscous dissipation, $d_{\text{min}} = (\nu/N)^{1/2}$. Since the viscosity in a stellar interior is very small, we can effectively make e arbitrarily small by making the layers thin enough. This small amount of energy can derive, for example, from internal gravity waves generated by a nearby convective zone, or from the initial overstable oscillations. The reason why convection in layers is so clearly preferred in the experiments is

¹ The process of layer formation can be demonstrated easily at the coffee table. Fill a teaglass with (sufficiently hot) coffee, add two teaspoons of sugar and some condensed milk or coffee cream at room temperature (ordinary milk does not work as well). With teaspoon or straw carefully introduce some mixing in the lower part of the glass without mixing the entire fluid, to produce a smooth initial sugar and milk gradient. Wait until the internal waves have died out and observe how thin layers form and grow into thicker ones.

the fact that the oscillatory instability [Kato's (1966) oscillations] is *subcritical*. This means that (in the presence of some small amount of viscosity), convection in layers is possible at values of the temperature gradient which are still *stable* to the linear oscillatory instability. In fact, Proctor (1981) has proven that in the limit where the diffusivity of the stabilizing component is small compared to the thermal diffusivity, the minimum gradient for layers to be maintained is arbitrarily close to the minimum gradient in the absence of the stabilizing component. This result requires only that the diffusivity ratio is small, and therefore applies also to the astrophysical case.

In the laboratory, the initially smooth gradient is indeed observed to break up into very thin layers with a thickness of the order d_{\min} . These then slowly grow by merging with one another (cf. Turner 1985). The process by which this merging takes place is more complicated, so that the mean layer thickness d is not a very well determined quantity. Since the overturning time scale within the layers is very short compared to the nuclear burning time scale on which the semiconvective zone as a whole evolves, I will assume in the following that the layers are, in fact, rather thick, may be even of the order of the depth of the semiconvective zone itself. As it turns out, the effective diffusion coefficient is *independent* of the layer thickness (at least as long as the layer thickness is not extremely thin), so that in practice this does not introduce a large uncertainty.

2.2. Transport at the interface between layers

The profile of entropy S and mean weight per particle μ is sketched in Fig. 1. Overturning motions that span the depth of the layer keep its interior well mixed. The initially smooth gradients are replaced by steep boundary layers at the interface between two overturning layers. The process by which these boundary layers form is a classical one referred to as *convective expulsion* of a composition gradient. The expulsion of vorticity from convective cells was studied by Batchelor (1956). A closely related

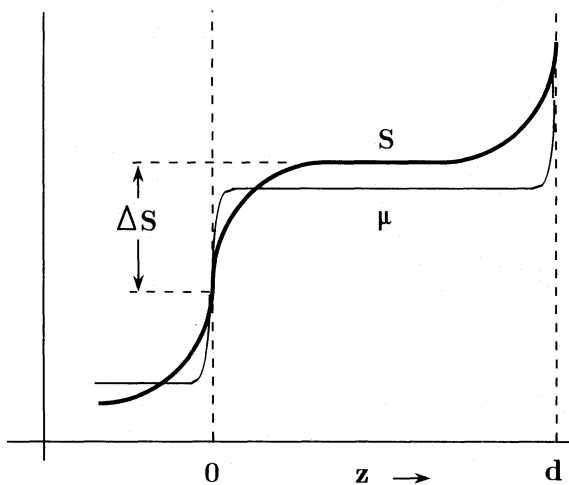


Fig. 1. Profiles of entropy S and mean weight per particle μ vs. depth across a double-diffusive layer in a semiconvective zone (schematic). The gradients of S and μ are concentrated into boundary layers at the interfaces ($z=0, d$) between successive layers. The transport inside the layers is by convection, at the interfaces by diffusion

problem is the convective expulsion of magnetic fields (Parker 1963; Weiss 1966).

The composition gradient at the interface is stable against overturning motion (Richardson criterion). This follows from the fact that the average stratification is stable according to the Ledoux criterion so that the kinetic energy of the flow (from the unstable thermal gradient) is less than that required for mixing across the interface.

The boundary layer thickness δ is determined by a balance between advection by the flow in the interior and diffusion across the interface, and is therefore different for the entropy and for each of the diffusing particle species. If v is the typical horizontal flow speed along the interface, and κ a diffusivity, the boundary layer thickness δ is

$$\delta = (\kappa\tau)^{1/2}, \quad (3)$$

where $\tau = d/v$ is the time scale of the convective cell. The diffusivity of the particles is often much smaller than that of heat (up to 6 orders of magnitude), so that the composition gradients have much smaller length scales.

I will use the name 'solute' for any of the constituents of the semiconvective zone, e.g. helium, hydrogen or carbon, and whose transport properties are the subject of this investigation.

Viscous coupling across the interface keeps the horizontal flow speeds the same on both sides, so that the flow pattern looks like Fig. 2. Laboratory experiments show this kind of pattern (in fact, the entire flow in such experiments is almost mirror-symmetric across the interface; see Fig. 1 in Turner 1985). An upward flow towards the interface turns horizontal, cools off by thermal conduction against the layer above it and drops away from the interface again. During the period of contact at the interface, the solute diffuses across as well. A part of the down-flowing material therefore carries a lower concentration than the mean. By advecting this fluctuation down to the base of the layer where it can diffuse across the next interface, the flow greatly enhances the effective rate of diffusion.

How much solute is picked up at the interface by the flow is determined by two factors. First, it is proportional to the thickness of the solute boundary layer and therefore depends on the microscopic diffusion coefficient. Not all of this boundary layer is

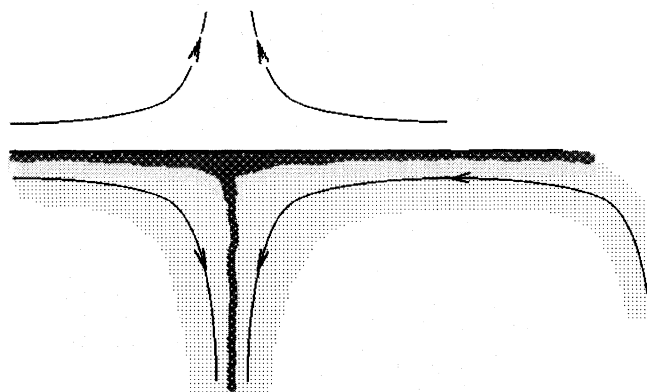


Fig. 2. Thermal (light shading) and solute (heavy shading) boundary layers at a diffusive interface. The solute boundary layer is much thinner than the thermal boundary layer due to the lower diffusivity. Descending and ascending plumes carry heat and solute away from the interface

carried down, because the downflowing fluid must have a negative buoyancy. The density contrast due to the solute can therefore not be larger than that due to cooling at the interface. This condition determines the fraction of the solute boundary layer that is 'peeled off' by the convecting cell.

In the following section these ingredients are developed into a quantitative theory for mixing in semiconvective zones. The theory is closely related to earlier work on convective expulsion (e.g. Shraiman 1987; Rhines & Young 1983; Merryfield & Knobloch 1990). In these theories, the diffusion of a passive scalar contaminant ('dye') in a prescribed, stationary or oscillatory flow is computed. The present theory differs, on the one hand, because the composition gradient cannot be considered as a passive ingredient (it is strong enough to keep the convective flows confined to layers between stable μ -jumps). This turns out to have important consequences for the derived effective mixing rate. On the other hand, the astrophysical case differs from laboratory situations by several physical effects (like radiation pressure) that cannot be omitted in a practically applicable theory. Related theoretical work (but addressing a typical laboratory case) is that of Gough & Toomre (1982).

3. Heat transport

Since the time of contact at the interface and the rate at which solute is advected depend on the convective flow speed, we first need a quantitative description of convection in thin layers. The main effect of the solute on the flow is to confine the convective cells to layers; this should be taken into account, but I will ignore the effect of solute inhomogeneities inside the cells themselves. This is particularly justifiable in the astrophysical (as opposed to the laboratory) case, because the diffusion coefficient of the solute is so small that only little solute is transported by the cells.

3.1. The convective flow

Assume that the convective flow is sufficiently vigorous. The entropy profile then has a boundary layer structure as sketched in Fig. 1. The thickness of these boundary layers is taken to be small compared to the layer thickness (the opposite case is considered in Sect. 5). The heat flux is due to the boundary layers, the horizontal flow in these turns into vertical plumes of cool downflowing fluid from the top of the layer and hot upflows from its base. The interior of the cell is relatively passive in the process.

Assume the gas to be fully ionized, nonrelativistic, nondegenerate and include the radiation pressure so that the equation of state is

$$P = P_g + P_r = \frac{R}{\mu} \rho T + \frac{1}{3} a T^4, \quad (4)$$

where R is the gas constant. Since the μ -gradient occurs in a thin boundary layer well inside the thermal boundary layer, μ can be taken constant for the heat flow problem. The entropy per unit mass is then (cf. Cox & Giuli 1968)

$$S = \frac{R}{\mu} \left(\ln \frac{T^{5/2}}{P_g} + 4 \frac{P_r}{P_g} \right) + \text{const.} \quad (5)$$

The entropy gradient can be written as

$$\nabla S = \frac{c_p}{H} (\nabla - \nabla_a), \quad (6)$$

where the specific heat at constant pressure c_p and the adiabatic gradient ∇_a are

$$c_p = \frac{R}{\mu} \left(\frac{16}{\beta^2} - \frac{12}{\beta} - \frac{3}{2} \right), \quad \nabla_a = \frac{4-3\beta}{16-12\beta-(3/2)\beta^2}, \quad (7)$$

and

$$\beta = P_g / (P_r + P_g). \quad (8)$$

Let ΔS be the entropy difference between the center of the layer and the top (Fig. 1). The flow brings fluid with the entropy of the cell center to the boundary layer and cools off against the overlying layer. It thereby develops an entropy difference

$$\delta S = \varepsilon \Delta S = \varepsilon \overline{\nabla S} = \varepsilon \frac{d c_p}{2 H} (\nabla - \nabla_a), \quad (9)$$

where $\overline{\nabla S}$ is the mean gradient across the layers and ε a coefficient of order unity, for which a value will be assigned later. The flow is strongly subsonic so that the pressure fluctuation is negligible and the temperature difference corresponding to δS is $\delta \ln T = \delta S / c_p$. With the equation of state the corresponding density fluctuation is

$$\delta \rho / \rho = - \left(\frac{4}{\beta} - 3 \right) \delta \ln T = - \varepsilon \left(\frac{4}{\beta} - 3 \right) \frac{d}{2H} (\nabla - \nabla_a). \quad (10)$$

The downward velocity of the plume is estimated, as in mixing length convection, by equating the kinetic energy to the potential energy associated with the density contrast $\delta \rho$:

$$\frac{1}{2} v^2 = \eta g d |\delta \rho / \rho|, \quad (11)$$

where η , of order unity, measures the efficiency of conversion of potential into kinetic energy. The heat flux is given by the entropy contrast and the mass flux. If the plumes are separated by a horizontal distance equal to the layer depth d and each carries a mass flux F_m , the average convective heat flux per unit horizontal area is

$$F_c = \frac{1}{d} F_m T \delta S. \quad (12)$$

The plume that carries this flux is fed from the boundary layer, and therefore has a width of the order of the boundary layer thickness δ_l . The mass flux is then

$$F_m = v \delta_l \rho. \quad (13)$$

For the length of time during which diffusion takes place I take the horizontal flow time across a length of the order d :

$$\tau = d / v_h = \zeta d / v, \quad (14)$$

where v_h is the horizontal flow speed in the boundary layer and ζ is another coefficient of order unity. The boundary layer width is then

$$\delta_l = (\zeta d \kappa_t / v)^{1/2}. \quad (15)$$

The thermal diffusivity κ_t is

$$\kappa_t = \frac{16 \sigma T^3}{3 \kappa_R \rho^2 c_p}, \quad (16)$$

where κ_R is the Rosseland opacity, and $\sigma = ac/4$ the

Stefan–Boltzmann constant. Substituting into (7) one gets

$$F_c = \frac{1}{2} \varepsilon (\eta \varepsilon)^{1/4} \zeta^{1/2} \left(\frac{4}{\beta} - 3 \right)^{1/4} g^{1/4} d H^{-5/4} (\nabla - \nabla_a)^{5/4} \kappa_t^{1/2} \rho T c_p, \quad (17)$$

whereas the radiative flux is

$$F_r = \rho c_p \kappa_t \frac{T}{H} \nabla. \quad (18)$$

Hence,

$$F_c/F_r = f u^{1/4} (\nabla - \nabla_a)^{5/4} \nabla^{-1}, \quad (19)$$

where

$$f = \frac{1}{2} \varepsilon (\eta \varepsilon)^{1/4} \zeta^{1/2} \quad (20)$$

is a number of order unity and

$$u = \frac{g}{H} \frac{d^4}{\kappa_t^2} \left(\frac{4}{\beta} - 3 \right). \quad (21)$$

With the definition of the radiative gradient $\nabla_r/\nabla = F_c/F_r + 1$ we have

$$f u^{1/4} (\nabla - \nabla_a)^{5/4} + \nabla - \nabla_a = \nabla_r - \nabla_a. \quad (22)$$

Since I have assumed that the radiative flux is small compared to the convective flux, the second term on the left-hand side is negligible compared to the right-hand side, so that

$$(\nabla - \nabla_a)^{5/4} = f^{-1} u^{-1/4} (\nabla_r - \nabla_a). \quad (23)$$

This replaces the usual mixing length expression for $\nabla - \nabla_a$. Due to the fact that the entire heat flux is carried by the plumes, which cover only a part of the cell, the $3/2$ power of $\nabla - \nabla_a$ in the mixing length formalism is replaced by the weaker $5/4$ power.

F/F_r is called the Nusselt number Nu . If we define a modified Rayleigh number Ra_* by

$$Ra_* = \frac{g d^4}{H \kappa_t^2} \left(\frac{4}{\beta} - 3 \right) (\nabla - \nabla_a), \quad (24)$$

then (19) can also be written as

$$Nu = f Ra_*^{1/4} \frac{\nabla - \nabla_a}{\nabla}. \quad (25)$$

Closer inspection shows that it is hard to justify more complicated choices for the numerical constants than simply $\varepsilon = \eta = \zeta = 1$, which yield

$$f = 0.5. \quad (26)$$

3.2. Comparison with laboratory convection

Convection experiments at Rayleigh numbers up to 10^{13} have been done by Libchaber & Maurer (1982) and Castaing et al. (1989). These experiments use helium gas in a box at 5 K. The physical conditions in semiconvective layers are similar to this in that the density stratification is not important (unlike in other astrophysical cases), and that the convective cells are confined between well-defined horizontal boundaries. Different is the Prandtl number, which is close to unity in the experiment but very small in the astrophysical case. Different are also the boundaries which are non-slip in the experiment whereas they are essentially free in the semiconvection case (in part because of the

low viscosity, in part because of the symmetry of the flow across the interface, see Sect. 2.2). The experiments are fitted by (Castaing et al. 1989)

$$Nu = 0.23 Ra^{0.282} \quad (Ra > 10^7), \quad (27)$$

where Ra is the Rayleigh number. For a meaningful comparison with the results given above, I rederive them for the slightly different case of a Boussinesq fluid with thermal expansion coefficient α , which is closer to the experiment. The same steps as in Sect. 3.1 then yield

$$Nu = 0.5 Ra_*^{1/4}, \quad (28)$$

where the modified Rayleigh number Ra_* is now defined as

$$Ra_* = Pr Ra = g \alpha \Delta T d^3 / \kappa_t^2. \quad (29)$$

Since the Prandtl number of the experiment is close to unity, Ra in (27) is equal to Ra_* in (28). The two expressions are then equal within 30% up to $Ra = 10^{13}$. Other scalings are discussed by Castaing et al. (1989), which take into account the viscosity of the gas, and these yield somewhat different exponents than our value 0.25. The essential ingredient in these scalings, as in my estimate, is the experimentally observed fact that the transport of heat is due to the boundary layers, so that only a part of the fluid takes an active part transporting heat. This is the cause of the weaker dependence on Rayleigh number.

4. Mixing

The mixing can now be calculated analogously to the energy transport. The flow pattern is assumed to be given by the previous calculation of the energy transport. That is, the gradient in solute is assumed to be present only in the form of steps at the interfaces between layers, the backreaction of concentration inhomogeneities inside the layer on the flow is neglected. This is justified in the case of very small solute diffusivity, since the steps are then sharp and the amount of solute inhomogeneities inside the cell is small.

4.1. The solute flux

The amount of solute transported across the layer is limited by two factors. The width of the solute boundary layer at the interface determines how deep into the overturning cell the solute diffuses. In addition, the solute diffusing into the cell affects the buoyancy of the fluid; the thermal buoyancy driving the flow is able to overcome this stabilizing effect only up to a certain maximum value of the solute inhomogeneity. This means that only a fraction of the solute boundary layer will be carried with the flow across the layer. The larger the overall concentration gradient is compared to the driving thermal gradient, the more pronounced is this effect.

If more than one solute component is present with different diffusivities, the situation is more complicated. I assume for the remainder of this section that there is only one component. This is a fair approximation since the products of nuclear burning in a particular zone (at least those that dominate the mean weight per particle) are characterized by roughly equal mass and charge, so that they also have roughly equal microscopic diffusivities.

Let c be the concentration of the stabilizing solute, related to the mean weight per particle by

$$\partial \ln \mu / \partial \ln c = \chi. \quad (30)$$

Since the solute boundary layer is very thin compared to the thermal boundary layer, the temperature variation across it can be ignored. The temperature difference of the solute boundary layer with respect to the middle of the layer is therefore just (cf. Eq. (9) and Fig. 1)

$$\delta \ln T = \frac{d}{2H} (\nabla - \nabla_a). \quad (31)$$

From the equation of state, and still making use of the fact that the pressure fluctuation is negligible:

$$\delta \ln \rho = \delta \ln \mu - \left(\frac{4}{\beta} - 3 \right) \delta \ln T. \quad (32)$$

The maximum change in μ for which the fluid still has a net buoyancy of the right sign is therefore

$$(\delta \mu / \mu) = \left(\frac{4}{\beta} - 3 \right) \delta \ln T, \quad (33)$$

and the corresponding concentration is

$$\delta c_{\max} = c / \chi (\delta \mu / \mu)_{\max}. \quad (34)$$

This part of the boundary layer follows the overturning cell and forms a solute plume, inside the thermal plume (Fig. 2). The average solute flux carried by the cell is estimated similar to the heat flux [cf. Eqs. (12) and (13)]:

$$F_s = \frac{1}{d} \gamma v \delta_c \delta c_{\max}, \quad (35)$$

where

$$\delta_c = (\kappa_s \tau)^{1/2} \quad (36)$$

is the solute boundary layer thickness, with τ the cell overturning time (14) and γ a factor of order unity analogous to ε . Substituting, we have

$$F_s = \frac{1}{2} \gamma \zeta^{1/2} (v \kappa_s d)^{1/2} \frac{c}{\chi H} \left(\frac{4}{\beta} - 3 \right) (\nabla - \nabla_a), \quad (37)$$

or, with (11):

$$F_s = \frac{1}{2} \gamma \zeta^{1/2} (\eta \varepsilon)^{1/4} g^{1/4} d H^{-5/4} \left(\frac{4}{\beta} - 3 \right)^{5/4} (\nabla - \nabla_a)^{5/4} \kappa_s^{1/2} \frac{c}{\chi}. \quad (38)$$

If ∇c is the mean concentration gradient across the layers, the diffusive solute flux in this gradient would be

$$F_{sd} = \kappa_s \nabla c. \quad (39)$$

The effective diffusivity

$$\kappa_{s \text{ eff}} = \kappa_s F_s / F_{sd} \quad (40)$$

is therefore given by

$$\kappa_{s \text{ eff}} = f_s (g/H)^{1/4} d \left(\frac{4}{\beta} - 3 \right)^{5/4} (\nabla - \nabla_a)^{5/4} \kappa_s^{1/2} / \nabla \mu, \quad (41)$$

where

$$f_s = \frac{1}{2} \gamma \zeta^{1/2} (\eta \varepsilon)^{1/4} \quad (42)$$

and

$$\nabla \mu = \frac{d \ln \mu}{d \ln P} = \frac{\chi}{c} H \nabla c. \quad (43)$$

This gives the effective diffusivity if the mean thermal gradient is given. Taking the mean thermal gradient from (23), this finally gives the simple expression

$$\kappa_{s \text{ eff}} = \frac{\gamma}{\varepsilon} (\kappa_s \kappa_t)^{1/2} \left(\frac{4}{\beta} - 3 \right) \frac{\nabla_r - \nabla_a}{\nabla \mu}. \quad (44)$$

The effective diffusivity is thus proportional to the square root of the microscopic diffusivity.

The most pleasing aspect of expression (44) is that it is independent of the layer thickness d . This can be traced to the fact that the maximum concentration inhomogeneity that can be transported is related to the thermal buoyancy (Eq. 33), resulting in a very direct connection between the two transport processes. In deriving (44) however I have assumed that the thermal boundary layers at the interfaces are thin compared with the layer thickness. In Sect. 5 a more general result is derived; unfortunately, it does not have the property of being independent of d any more.

4.2. Comparison with observations and other theories

An important difference between the astrophysical case on the one side, and laboratory experiments, geophysical observations, and theories designed to interpret these on the other (for a review see Turner 1985), is the very small viscosity. For a comparison of our result with observations, we must first examine to what extent the result would have to be modified to include a large viscosity, of the order of the thermal diffusivity. The viscosity will not have a large effect on the convective velocities in the interior of the layers (as long as the Reynolds number is large), but near the interfaces viscous boundary layers might develop in which the velocity is reduced. Since the solute flux is directly proportional to the horizontal velocity, this could significantly reduce the effective diffusivity. This is the case for example in the calculations of Gough & Toomre (1982), which addressed typical laboratory conditions. In these calculations, the interfaces between successive layers were treated as non-slip boundaries. Since the viscosity is high (Gough & Toomre use $Pr = \nu / \kappa_t \approx 7$, $\nu / \kappa_s \approx 700$), the solute boundary layer is a region of much reduced flow. From experiments one knows, however, that the interface behaves much more like a free surface, since the flow is nearly symmetric with respect to it (see e.g. Fig. 1 in Turner 1985). For a free boundary, the horizontal velocity would be much less affected by the viscosity. As long as a free boundary may be assumed, the results of this section hold for all Prandtl numbers of order unity or less, and therefore can be compared also with the experimental results.

An important experimental result concerns the ratio of the heat and solute flux as expressed by the 'ratio of buoyancy fluxes'. In our variables, this ratio is

$$r = \frac{c_p \rho T \chi}{c} \frac{F_s}{F_c}. \quad (45)$$

With (17) and (38) one finds (with the standard choice $\gamma = \varepsilon = 1$ and setting $\beta = 1$ because radiation pressure is not relevant in the laboratory)

$$r = \left(\frac{\kappa_s}{\kappa_t} \right)^{1/2}. \quad (46)$$

This depends only on the ratio of the diffusivities; in particular, it is independent of the strength of the stabilizing solute gradient.

This is just what is observed (cf. Turner 1985, p. 15). The square root dependence is also as observed (Turner et al. 1970), and even the coefficient in front is of the order unity in the experiments.

The simple result (46) is a consequence of the marginal buoyancy assumption (33). The reason for assuming this was that the density contrast due to the solute regulates itself to such a value that it roughly balances the thermal density contrast. Parts of the solute boundary layer with a smaller density contrast are carried with the flow as well, but contribute less because they contain little solute, while parts with larger density contrast than the thermal value are too buoyant to be carried with the flow.

The square root dependence of $\kappa_{s, \text{eff}}$ on the microscopic diffusivity κ_s is a very general result, and not specific to the present model. It occurs also in the extensively studied cases of convective expulsion and passive diffusion of contaminants in prescribed (oscillatory or steady) flows. It is the result of the balance between advection of the contaminant into a boundary layer, and the diffusive broadening of this boundary layer. Calculations of this process have been made for steady overturning flows (Weiss 1966; Shraiman 1987; Rosenbluth et al. 1987), and oscillatory flows (Rhines & Young 1983; Merryfield & Knobloch 1991). Details of the time-dependent problem of the formation of the boundary layers from an initially uniform gradient have been clarified by Rhines & Young (1983) for fairly general flows.

5. Corrections when convection is not efficient

In Sect. 4, I have assumed that the thickness d of the overturning layers is large compared to the thickness of the thermal boundary layers at the interfaces. Since we do not know d very well, it is useful to explore how the mixing rate depends on d if this assumption is not made. In this case, the turnover time is not short compared to the thermal diffusion time across the layer, so that the thermal boundary layer thickness is not small compared to the layer thickness. At the same time the buoyancy driving the flow is reduced by thermal diffusion across the cell.

In this section I set $\varepsilon = \eta = \zeta = 1$ for simplicity and because elaborations to derive more precise values are probably not warranted at the present level of the theory. Instead of the boundary layer thickness (15), its maximum value must be used,

$$\delta_t = d/2. \quad (47)$$

If $\tau_t = d^2/\kappa_t$ is the thermal diffusion time, and $\tau = d/v$ the turnover time, the typical value of the entropy fluctuation driving the flow is reduced by a factor of the order τ_t/τ . Hence,

$$\delta S = \frac{dc_p}{2H} (\nabla - \nabla_a) \frac{\tau_t}{\tau} \quad (\tau_t < \tau) \quad (48)$$

instead of (9). The corresponding temperature fluctuation is $\delta \ln T = \delta S/c_p$. With (11) the convective velocity becomes

$$v = gd^3 \frac{1}{\kappa_t H} \left(\frac{4}{\beta} - 3 \right) (\nabla - \nabla_a). \quad (49)$$

With (12) and (13) the convective heat flux is

$$F_c = \frac{1}{4} g^2 d^8 H^{-3} \kappa_t^{-3} \rho c_p T \left(\frac{4}{\beta} - 3 \right)^2 (\nabla - \nabla_a)^3. \quad (50)$$

With $(F_c + F_r)/F_r = \nabla_r/\nabla$, and F_r given by (18), this yields

$$\nabla_r = \frac{1}{4} u^2 (\nabla - \nabla_a)^3 + \nabla, \quad (51)$$

[with u given by (21)] from which ∇ can be solved. Since we are interested in the limit where the radiative flux dominates, we can expand in $u \ll 1$, and find

$$\nabla - \nabla_a = (\nabla_r - \nabla_a) \left[1 - \frac{1}{4} (\nabla_r - \nabla_a)^2 u^2 \right] \quad [(\nabla_r - \nabla_a) u \ll 1]. \quad (52)$$

To calculate the mixing, assume that the solute boundary layer is still very thin, as before. With (34)–(36) and (48) this yields

$$F_s = \frac{1}{2} d^{3/2} \kappa_s^{1/2} v^{3/2} \frac{1}{H} \frac{c}{\chi} \left(\frac{4}{\beta} - 3 \right) (\nabla - \nabla_a) \kappa_t^{-1}. \quad (53)$$

Hence, with (37)

$$F_s = \frac{1}{2} d^6 \kappa_s^{1/2} g^{3/2} H^{-5/2} \left(\frac{4}{\beta} - 3 \right)^{5/2} (\nabla - \nabla_a)^{5/2} \kappa_t^{-5/2} c/\chi. \quad (54)$$

With (34)–(36) and (48) this yields

$$\kappa_{s, \text{eff}} = \frac{1}{2} u^{3/2} (\kappa_s \kappa_t)^{1/2} \left(\frac{4}{\beta} - 3 \right) (\nabla_r - \nabla_a)^{5/2} / \nabla_\mu. \quad (55)$$

5.1. Summary of results

The two limiting cases of efficient and inefficient convection, expressions (23) and (52), can be combined into

$$\frac{\nabla - \nabla_a}{\nabla_r - \nabla_a} = \min[1, 2^{4/5} q^{-1/5}], \quad (56)$$

where

$$q = u (\nabla_r - \nabla_a) = \frac{gd^4}{H \kappa_t^2} \left(\frac{4}{\beta} - 3 \right) (\nabla_r - \nabla_a), \quad (57)$$

and I have used $f = 1/2$ as before [Eq. (26)]. A smoother interpolation between the two cases could be considered but would probably not be more accurate. For the effective diffusivity we get from (45) and (55)

$$\kappa_{s, \text{eff}} = (\kappa_s \kappa_t)^{1/2} \left(\frac{4}{\beta} - 3 \right) \frac{\nabla_r - \nabla_a}{\nabla_\mu} \min \left[1, \frac{1}{2} q^{3/2} \right]. \quad (58)$$

Since q depends on the assumed layer thickness, the mixing depends on it as well if $q \lesssim 1$. In practice one finds that $q < 1$ only for very small layer thickness d . For example in a semiconvective zone above the helium burning core in an evolved $20M_\odot$ star ($g = 5 \cdot 10^5$, $H = 10^{10}$, $\kappa_t = 10^8$, $\nabla_r - \nabla_a = 0.07$, $\beta = 0.9$) this would be the case for layers thinner than 1 km. In laboratory experiments and in geophysical cases, it is observed that the layer thickness grows in time by merging of the layers with each other. In the astrophysical case the lifetime of a semiconvective zone, in units of the convective turnover time in the layers, is probably very long compared to these experiments, so that this merging process is probably important. Thus, it may be justified to assume that d is large enough so that $q > 1$. The results (23) and (44) then hold, i.e.

$$\nabla - \nabla_a \ll \nabla_r - \nabla_a, \quad (59a)$$

and

$$\kappa_{s, \text{eff}} = (\kappa_s \kappa_t)^{1/2} \left(\frac{4}{\beta} - 3 \right) \frac{\nabla_r - \nabla_a}{\nabla_\mu}. \quad (59b)$$

In front of the right-hand side an adjustable numerical factor of order unity may be added [see Eq. (23)] in view of the uncertainties in the derivation. Too much freedom in the choice of

its value does not exist, however, if we want the theory to apply also to the experimental results (Sect. 4.2), for which this factor is unity within a factor of two or so.

5.2. The microscopic diffusion coefficient

For convenience I summarize here some standard results. Let the mixture consist of two ion species of charge Z_1 and Z_2 (in units of the electron charge e) and masses m_1 and m_2 . Let the concentration c of species 2 be defined as $c = n_2/n_1$, where n_1, n_2 are the number densities per unit volume. If the gradient of c is the main cause of diffusion, the drift velocity of species 2 with respect to 1 is

$$v = -\kappa_s \nabla \ln c. \quad (60)$$

In addition to this, there are contributions to the drift velocity from the temperature gradient, the pressure gradient, gravity, and the gradient in the radiation density (for a review see Vauclair & Vauclair 1982). In the present case of a narrow gradient in concentration all these effects can be ignored. Further complications arise if more than two species (in addition to electrons) are present. For a discussion see Montmerle & Michaud (1976). If $c \ll 1$, the diffusion coefficient κ_s is given by its Chapman value (Chapman & Cowling 1960; Aller & Chapman 1960; see also Spitzer 1962; Urpin and Muslimov 1982):

$$\kappa_s = D_{1(2)} = \frac{3}{16n} \left(\frac{2kT}{\pi m} \right)^{1/2} \left(\frac{2kT}{Z_1 Z_2 e^2} \right)^2 \frac{1}{\ln \Lambda}, \quad (61)$$

where

$$m = m_1 m_2 / (m_1 + m_2), \quad n = n_1 + n_2,$$

and $\ln \Lambda$ is the Coulomb factor:

$$\Lambda = 1 + \left(\frac{4d_D kT}{Z_1 Z_2 e^2} \right)^2,$$

where d_D is the Debye length:

$$d_D = \left(\frac{kT}{4\pi n_e e^2} \right)^{1/2},$$

and n_e the electron density. If c is not small, corrections must be applied which can be as large as a factor of two. Expressions for these corrections are given in Montmerle & Michaud (1976). The viscosity is, again for $c \ll 1$ so that it is due mainly to species 1 (Chapman & Cowling 1960):

$$\nu = \frac{5}{16n} \left(\frac{2kT}{\pi m_1} \right)^{1/2} \left(\frac{2kT}{Z_1^2 e^2} \right)^2 (\ln \Lambda' - 1 + 1/\Lambda')^{-1} + \nu_r, \quad (62)$$

where

$$\Lambda' = 1 + \left(\frac{4d_D kT}{Z_1^2 e^2} \right)^2,$$

and ν_r is the radiative viscosity (Mihalas & Mihalas 1984):

$$\nu_r = \frac{16}{15} \frac{\sigma T^4}{\kappa_R \rho^2 c^2}. \quad (63)$$

If the radiative viscosity is negligible and $\Lambda \gg 1$, the ratio of viscosity to diffusivity is therefore

$$\nu/\kappa_s \approx \frac{5}{3} \left(\frac{m}{m_1} \right)^{1/2} \left(\frac{Z_2}{Z_1} \right)^2,$$

which is about 13 for helium diffusing in hydrogen. In semi-convective zones, ν_r is not negligible, and ν/κ_s can be much larger. If the radiation pressure is a substantial fraction of the total pressure, the diffusion of both momentum and heat is mainly due to photons. In this case, the Prandtl number ν/κ_i is:

$$\nu_r/\kappa_i = \frac{1}{5} \frac{T c_p}{c^2}, \quad (64)$$

which is still small for a nonrelativistic gas.

6. Discussion

I have developed a theory for mixing in semiconvective zones under the assumption that the layered convection observed in similar laboratory situations takes place. Arguments are given that this nonlinear form of motion is also preferred in the astrophysical case. The theory takes into account that boundary layers with steep gradients in entropy and mean weight develop at the interface between adjacent layers. A part of the theory describes the efficiency of heat transport (in the form of an expression similar to that for the temperature gradient in mixing length theory). This part is tested by comparison with recent measurements of laboratory convection at very high Rayleigh numbers; the expression for the heat flux agrees within a factor of two. The mixing efficiency, in the form of an effective diffusion coefficient is then derived by similar arguments and tested against the main experimental results. It reproduces, quantitatively and in functional form, the mixing rate deduced from these experiments (though it must be recognized that these cover only a rather small range in parameters).

The most important property of the derived mixing rate is that it is proportional to the square root of the microscopic diffusion coefficient. This reflects the fact that both diffusion and advection by overturning cells play an essential role in the transport mechanism. In its simplest form (applicable when the thickness of the layers is larger than a rather small minimum), the diffusion coefficient depends only on the energy flux, the thermodynamic variables and the chemical composition. In particular, it is independent of the somewhat uncertain thickness of the double-diffusive layers assumed.

The process of layer formation, assumed here as the basis for the calculations in this paper, is reminiscent of a curious result in experimental stellar evolution. Langer et al. (1985), and Langer (1991), using the Langer et al. (1983) mixing rate in semiconvective zones, observed that semiconvective zones typically develop in the form of a sequence of semiconvective layers sandwiched between normal convective layers (the number of such layers growing with time). Since the physics in the stellar evolution models does not include the hydrodynamics that is responsible for the actual layer formation process as seen, for example, in the laboratory, the reason for this must be somewhat different. The effective mixing across the stack of convective and semiconvective layers in these models is also different from that derived above. Yet it may reflect, at a more basic level, the same tendency for double-diffusive systems to break up into a sequence of layers.

Acknowledgements. I thank Dr. Mounib El Eid for the stimulating discussions which provided the incentive to write this text, Dr. Norbert Langer for useful comments and communication of recent results, and Dr. Edgar Knobloch for comments and for pointing out relations to other work.

References

- Aller L.H., Chapman S., 1960, ApJ 132, 461
Batchelor G.K., 1956, J. Fluid Mech. 1, 177
Chapman S., Cowling T.G., 1960, The Mathematical Theory of Non-Uniform Gases. Cambridge Univ. Press, Cambridge, 2nd ed. p. 179
Castaing B., Gunaratne G., Heslot F., et al., 1989, J. Fluid Mech. 204, 1
Cox J.P., Giuli R.T., 1968, Principles of stellar structure, Gordon and Breach, New York
Gough D.O., Toomre J., 1982, J. Fluid Mech. 125, 75
Huppert H., Turner J.S., 1981, J. Fluid Mech. 106, 299
Kato S., 1966, PASJ 18, 374
Kippenhahn R., Ruschenplatt G., Thomas H.-C., 1982, A&A 91, 175
Langer N., Sugimoto D., Fricke K.J., 1983, A&A 126, 207
Langer N., El Eid M.F., Fricke K.J., 1985, A&A 145, 179
Langer N., 1990 (preprint)
Libchaber A., Maurer J., 1982, in: Nonlinear phenomena at phase transitions and instabilities, ed. T. Riste, Plenum, New York, p. 259
Merryfield W.J., Knobloch E., 1991, in: Challenges to Theories of the Structure of Moderate Mass Stars, eds. D.O. Gough, J. Toomre, Springer, New York Berlin Heidelberg (in press)
Mihalas D., Mihalas B.W., 1984, Foundations of Radiation Hydrodynamics, Oxford University Press, New York
Montmerle T., Michaud G., 1976, ApJS 31, 489
Parker E.N., 1963, ApJ 138, 552
Proctor M.R.E., 1981, J. Fluid Mech. 105, 507
Rhines P.B., Young W.R., 1983, J. Fluid Mech. 133, 133
Spiegel E.A., 1972, ARA&A 10, 261
Spitzer L., 1962, The Physics of fully ionized Gases, Interscience, New York, 2nd Edition
Shraiman B.I., 1987, Phys. Rev. A 36, 261
Turner J.S., 1985, Ann. Rev. Fluid Mech. 17, 11
Turner J.S., Shirtcliffe T.G.L., Brewer P.G., 1970, Nat 228, 1083
Ulrich R.K., 1972, ApJ 172, 165
Urpin V., Muslimov A.G., 1982, Astron. Zh 59, 318 (translation in Soviet Astronomy 26, 193 (1982))
Vauclair S., Vauclair G., 1982, ARA&A 20, 37
Weiss N.O., 1966, Proc. Roy. Soc. London A, 293, 310