

Semiconvection: numerical simulations[★]

F. Zaussinger^{1,2,3} and H. C. Spruit¹

¹ Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85748 Garching, Germany

² Faculty of Mathematics, University of Vienna, Nordbergstraße 15, 1090 Vienna, Austria

³ Dept. Aerodynamics and Fluid Mechanics, BTU Cottbus, Siemens-Halske-Ring 14, 03046 Cottbus, Germany
e-mail: florian.zaussinger@tu-cottbus.de

Received 16 October 2012 / Accepted 18 March 2013

ABSTRACT

A grid of numerical simulations of double-diffusive convection is presented for the astrophysical case where viscosity (Prandtl number Pr) and solute diffusivity (Lewis number Le) are much lower than the thermal diffusivity. As in laboratory and geophysical cases, convection takes place in a layered form. The proper translation of subsonic flows in a stellar interior and an incompressible (Boussinesq) fluid is given, and the validity of the Boussinesq approximation for the semiconvection problem is checked by comparison with fully compressible simulations. The predictions of a simplified theory of mixing in semiconvection given in a companion paper are tested against the numerical results, and used to extrapolate these to astrophysical conditions. The predicted effective He-diffusion coefficient is nearly independent of the double-diffusive layering thickness d . For a fiducial main sequence model ($15 M_{\odot}$) the inferred mixing time scale is of the order of 10^{10} yr. An estimate for the secular increase in d during the semiconvective phase is given. It can potentially reach a significant fraction of the pressure scale height.

Key words. convection – methods: numerical – diffusion – hydrodynamics – stars: interiors

1. Introduction

In models of stellar structure, situations are found where the heavier products of nuclear burning provide stability to a zone that otherwise would be unstable to convective overturning. Such a zone, or part of it, would become convective if something managed to mix its composition (Tayler 1953). The question whether such a zone should be treated as if it were mixed or not has become known as the semiconvection problem. Answers to this question differ substantially. In practice, recipes are used containing a free parameter that allows the degree of mixing to be varied. Calculations in which such a parameter is adjusted to match observations are then called “with semiconvection”. Commonly used prescriptions are those of Langer (1985) and Maeder (1997).

The presence of a semiconvective zone only has a minor effect on the thermal structure of the star. The assumed amount of mixing of composition is critical, however, because the evolution of the star is sensitive to the precise distribution of products of nuclear burning with the depth in the star. The main goal of a theory for semiconvection is thus a good determination of the rate of mixing. From the perspective of the stellar evolutionist, the theory would ideally provide a formula for the rates of mixing and transport (the effective diffusivities), as functions of local thermodynamic state and composition, and their gradients.

These formulas were derived in Spruit (1992, hereafter S92) by adapting the known physics of double-diffusive convection (Turner 1979, 1985; Proctor 1981; Huppert & Turner 1981; Schmitt 1994) to the case of a stellar interior (discussed first in this context by Spiegel 1969, 1972). The expression developed in S92 makes use of simplifications valid in the limiting case of

very large Rayleigh number and very low solute diffusivity. In a companion paper (hereafter S13), the analysis in S92 is extended to cover the more moderate conditions accessible with numerical simulations. In the following, mixing rates and their dependence on the intrinsic parameters governing the problem are measured with such simulations. The predictions of S13 are tested against these results, and then used to estimate the expected mixing rates in semiconvective zones of stars.

One of the predictions in S13 is the existence of a maximum density ratio (the ratio R_{ρ} of stabilizing solute to destabilizing thermal stratification) for which a steady layered state is possible. In a slightly different guise, this limit also figures prominently in Proctor’s (1981) analytic analysis. In this analysis, he proved that in the limit of vanishing solute diffusivity, a layered state exists at any Rayleigh number above the critical value for convection in the absence of a stabilizing solute gradient, provided the density ratio is below this critical value. The model in S13 does not predict what happens for density ratios just above or below this maximum value. The numerical results presented in Sect. 5.2 clarify how the system behaves in this case.

The development of a linear initial gradient into the final state of overturning layers separated by diffusive steps is studied with a few examples. This transient process shows the “Kato-oscillations” expected from linear theory (see Sect. 5 for an example). It is not very relevant for astrophysical application, however, since the transition to the layered state happens on a time scale (a finite number of buoyancy oscillations) that is negligible compared with the time scales of interest, and is found to depend on the details of the initial state. Instead, the focus here is on the transport of heat and solute in a double-diffusive staircase with layers of given thickness, as a function of the intrinsic parameters of the problem.

[★] Movies associated to Figs. 5 and 7 are available in electronic form at <http://www.aanda.org>

In Sect. 2 the known physics of double-diffusive convection of the semiconvective type is reviewed, in general and from an astrophysical perspective. Section 3 describes the transport properties of a single double-diffusive layer in terms of the model of S13. The numerical methods are given in Sect. 4. In Sect. 5 the results of our parameter study are shown and compared with the predictions of this model. Application to the case of semiconvection in a $15 M_{\odot}$ main sequence star is discussed in Sect. 6.

2. Semiconvection and double-diffusive convection

Situations where a fluid is stabilized by the density gradient due to a dissolved heavy constituent occur in nature. An example is convection in arctic oceans (fresh meltwater cooled from above, stabilized by the salts dissolved in the sea water, see the review in Schmitt 1994). Intensively studied are East-African rift lakes (Lakes Kivu, Nyos, and Mounon, cf. Schmid et al. 2010). These are heated from below by volcanic activity, which is also a source of dissolved gases (carbon dioxide and methane, hereafter the “solute”). Their density stratification is stabilized against convection by the stable gradient resulting from the weight of the carbon dioxide. Efforts to prevent catastrophic release of carbon dioxide (Lake Nyos, e.g. Sigvaldason 1989) or to enable safe commercial exploitation of methane (Lake Kivu, Nayar 2009) have led to extensive study of the fluid flows, heat flux, and mixing rates in these natural double-diffusive systems.

The gradients in temperature and solute in the East-African lakes and the arctic are observed to be “stepped”: consisting of a stack (called “staircase”) of thin layers (decimetres to decametres). Inside a layer, overturning convection keeps the composition nearly uniform, with stable gradients in temperature and composition separating the layers. The physics involved is easily reproduced under controlled laboratory conditions (Turner & Stommel 1964; Turner 1985)¹. The layers are very long-lived: of the order of months or more in the geophysical examples mentioned, which are orders of magnitude longer than the convective turnover times inside the layers.

In the stable gradients between the overturning layers the transport of the stabilizing solute takes place by diffusion instead of convection. This strongly limits the effective transport of solute through the double-diffusive staircase. Residence times of the order of 1000 yrs are inferred for the solutes in Lake Kivu, for example (Schmid et al. 2010). This is eight orders of magnitude longer than the convective turnover times in these layers. The transport of heat is also strongly reduced and this is exploited for heat storage in solar ponds (cf. Lu & Swift 2001). Similarly low fluxes have been measured in thermohaline staircases in the arctic and antarctic oceans (e.g. Padman & Dillon 1987).

Theoretically, the observed layered nature of double-diffusive systems is now understood well. Central to this understanding is the fact that linear stability analysis does not provide relevant clues to their behaviour, because the double-diffusive case of thermal convection stabilized by a slowly-diffusing solute is *subcritical*. That is, a nonlinear form of the instability, in

the form of a stable overturning flow, already exists below the critical temperature gradient for onset of linear instability. The linear stability condition is thus not relevant for the behaviour of the system (cf. Schladow et al. 1987).

Linear instability predicts that internal gravity modes set in above some critical value of the temperature gradient, growing in amplitude by the effect of thermal diffusion: the so-called Kato oscillations (Kato 1966). Such oscillations (cf. movie in Fig. 5) transport a negligible amount of heat or solute, compared with overturning motions of the same amplitude. For the reason alone, linear stability arguments cannot be used for estimates of the mixing rate in semiconvective zones. More important, however, is the subcritical nature of double-diffusive convection. Proctor (1981) shows analytically that, within the limit of vanishing diffusivity of the solute, the layered form of convection exists whenever the Rayleigh number exceeds the critical value for ordinary convection, irrespective of the strength of the stabilizing component. This assumption of vanishing solute diffusivity is eminently satisfied in the astrophysical case and holds reasonably well in geophysical and laboratory experiments.

In Stevenson (1979), it was assumed that the nonlinear development of the overstable oscillations would lead to saturation of the wave at a finite amplitude. This has been the rationale for some prescriptions used in astrophysics (e.g. Langer et al. 1985). The assumption of saturation at finite amplitude is appropriate for the more common supercritical forms of oscillatory instability, but not for the subcritical case, where a finite amplitude state exists for parameters where the system is still linearly stable. The assumption also conflicted with observations: the laboratory and geophysical systems all showed the same characteristic form of convection in a system of overturning layers (e.g. Turner & Stommel 1964), but none that settled into finite amplitude oscillations.

In the literature on semiconvection, it is often argued that the Prandtl number in astrophysics is much lower than in the geophysical and laboratory cases. The implied assumption that the understanding of double-diffusive convection developed in these contexts can be set aside, however, is not necessary. Proctor’s analysis, for example, is largely independent of Prandtl number. It is valid as long as Pr is not larger than of order unity, and viscosity is higher than the solute diffusivity. This is also satisfied in the astrophysical case.

In physical terms, the reason for the subcritical behaviour can be understood with an energy consideration (S92). The amount of energy it takes to overturn a layer of thickness d against a stable gradient scales as d^2 (as in a harmonic oscillator at the buoyancy frequency of the stratification). The expense in initial energy needed per unit mass to put the system into its finite-amplitude, layered state thus vanishes as d , down to some value where damping losses stabilize the system. A small initial perturbation or an initial Kato oscillation is sufficient to provide the energy for overturning into thin layers. Once established, this layered state is a stable form of convection. In fact, reproduction of the oscillatory phase in the laboratory requires a very careful setup of the initial stratification (Shirtcliffe 1967, see also the discussion in Huppert & Turner 1981).

This agrees with the observation in laboratory experiments and geophysical systems like Lake Kivu mentioned above, i.e. that the layering first sets in at a small thickness (cf. footnote above). In this context, the formation of layers found in the simulations by Rosenblum et al. (2011) is in line with expectations. Much less well-defined is the evolution on longer time scales, in particular the question of how and on which time scales the layer thickness evolves (see Sect. 6.6 and S13).

¹ Also on a coffee table. Latte macchiato in a tall glass often shows the effect nicely. After the coffee is added to the milk, a stably stratified gradient of milk/coffee mix develops (showing internal gravity waves in the form of a sloshing motion with a period of a few seconds). After about a minute, the initially smooth gradient starts dividing into thin (a few mm) layers, visible at low contrast. In the course of several minutes these merge into fewer, more clearly defined layers.

2.1. Double-diffusive convection in stars

Double-diffusive convection in stars has traditionally been regarded as a part of physics to be treated separately from the geophysical examples, since the numerical values of controlling parameters such as the Prandtl number are quite different. Apart from the difference in equation of state, however, the hydrodynamic equations are identical. Differences in physics that might be present between the two cases are in fact not apparent in the elementary recipes for semiconvection used in stellar evolution codes. The failure of these recipes when applied to the geophysical case is traditionally not considered an argument against their application in stars. Such recipes include (a) to assume that no mixing takes place at all (the ‘‘Ledoux’’ recipe); another (b) to ignore the stabilizing effect of the solute gradient (‘‘Schwarzschild’’ recipe, yielding a very high mixing rate); (c) some interpolation between these recipes; (d) to assume that the amount of mixing is such that the layer becomes marginally stable to overturning; and (e) the somewhat more physically motivated oscillation-based recipe of Langer et al. above.

Define the Prandtl number $Pr = \nu/\kappa_T$ and the Lewis number $Le = \kappa_S/\kappa_T$, where κ_T is the thermal diffusivity and κ_S the diffusivity of the stabilizing ‘‘solute’’ (He diffusing in H, say). Because of the high thermal diffusivity mediated by photons, Pr and Le are very small numbers, some eight to ten orders of magnitude less than in geophysical cases.

Such low parameter values cannot be covered realistically in numerical simulations. If τ_c is a typical convective time scale (as estimated from the superadiabaticity and pressure scale height in the semiconvective zone), very small length scales, of the order of $(\tau_c \kappa_S)^{1/2}$ would have to be resolved to represent the interaction between flow and diffusion. This is not computationally possible at present, even in two-dimensional simulations.

Translation from the numerically accessible parameter values to an astrophysically relevant parameter range therefore requires scaling of the results over the orders of magnitude in between. A valid extrapolation cannot be found by mere intuitive inspection of numerical results, since these are too far from the target regime. On their own, the applicability of numerical results to the astrophysical case is bound to remain diffuse, or applicability as an explicit goal has to be given up altogether.

Some physical understanding that includes this asymptotic regime is necessary. A theory that accomplishes this is given in the companion paper (S13). It is made possible by making explicit use of the observed separation into layers of convective overturning between stable diffusive interfaces.

An important difference between the astrophysical and geophysical cases is the equation of state. For the thin layering expected, sound travel times are very short compared with convective time scales. As a consequence, the fluid in a semiconvective zone behaves as if it were nearly incompressible. A Boussinesq approximation can then be used for the calculations, provided a small complication is properly taken into account. Whereas in the incompressible case the convective and diffusive heat fluxes are both governed by the temperature gradient, convection and radiative diffusion are governed by different quantities in the compressible case (entropy gradient and temperature gradient). This affects, in particular, the definition of the Nusselt number as a measure of the efficiency of heat transport. An exact translation between these cases is possible (Massaguer & Zahn 1980, see below in 6.1).

A second difference concerns the status of the heat flux in formulating the problem. The effect of semiconvective mixing on the star’s structure in evolution calculations is found to be

small during the semiconvective phase itself, and is somewhat independent of the way semiconvection is approximated. (This is in part because of the limited extent of a semiconvective zone.) Its effect on the radial profile of elemental composition, however, is of lasting importance for later evolutionary stages (see Langer et al. 1989).

The consequence is that, in contrast to laboratory and geophysical situations, the heat flux F , rather than the temperature gradient in a stellar model, can be considered as known. Since the radiative contribution F_r to the heat flux is known to good approximation from the thermal structure of the star, the convective heat flux $F_c = F - F_r$ transported by semiconvection is also known. The efficiency of convection: i.e. how close the mean thermal gradient is to the adiabatic gradient, follows from the imposed heat flux (instead of the other way around as in a laboratory experiment). We return to this distinction in Sect. 6.4.

2.2. Size of the parameter space

The fluid is described by the thermodynamic variables defining its local state (e.g. temperature and density) and the material functions (e.g. pressure, diffusivities, viscosity). In addition, the gradients of the thermodynamic variables with depth and the acceleration of gravity are relevant for the properties of the flow. Taken together, these quantities form a large parameter space, and it might be concluded that realistic numerical simulations of semiconvection would have to be done for individual zones in individual stellar models.

The equations of fluid dynamics have symmetries, however, so that the independent degrees of freedom are far fewer. They can be represented by five dimensionless parameters: a Rayleigh number Ra , the layer thickness d in units of the pressure scale height H , the Prandtl number, the Lewis number, and a density ratio R_ρ that measures the ratio of the stabilizing (solute) gradient to the destabilizing thermal gradient. The behaviour of semiconvection at some point in a star can be defined in terms of these parameters.

The Boussinesq approximation corresponds to the limit $\epsilon = d/H \ll 1$. The pressure scale height disappears as a parameter in this limit, and all dependence on d is subsumed in Ra . This reduces the number of parameters of the problem to four.

By a fortunate coincidence, it turns out that as long as the Prandtl number is less than unity, the results are effectively independent of Pr . This further reduces the number of independent parameters to only three. Since measurement of the mixing rate in each individual case does not require a very expensive simulation, this allows a significant volume of parameter space to be covered, and a comparison with the model predictions in S13 to be made.

3. The layered state

3.1. Layer formation, layer thickness

The formation of a layer from an initially smooth and static gradient starts with a well-known oscillatory instability, with oscillation periods close to the buoyancy period of the stratification. This is followed by nonlinear development into an overturning flow. The transition can be observed in numerical simulations (Merryfield 1995; Rosenblum et al. 2011, Sect. 5.5 below), and in very carefully designed laboratory experiments, but is not seen in geophysical cases. An unperturbed smooth initial gradient in solute and temperature is an artificial case that is unlikely to be realized in nature. In addition, the intermediate oscillatory state

lasts for only a finite number of oscillation periods (5–10 in the results below), a very short time scale compared with the life time of the double-diffusive layers. In the case of Lake Kivu, for example, buoyancy periods are in the range 5–30 min, the life time for individual layers of a few months.

In laboratory and kitchen table experiments the layers are initially very thin. This is understood from the energy argument above. The layers slowly merge, either by fading of contrast between neighbouring layers, or the vertical migration of interfaces towards adjacent layers. The details of this process have not been studied much (but see McDougall 1981; Young & Rosner 2000; Ross & Lavery 2009). A plausible estimate of the rate of growth of the layer thickness can be given in terms of the effective solute diffusivity of the system, however (S13, and Sect. 6.5).

In the simulations reported here, the evolution of layer thickness is not included. The thickness (d) is therefore treated as a free parameter of the problem. It enters through the Rayleigh number (in the Boussinesq limit, see 2.2 above). In the astrophysical application of imposed heat flux, however, the resulting mixing rate is effectively independent of the layer thickness (S13 and Sect. 6.5 below). This is a fortunate circumstance, since following the evolution of a stack of layers through its merging processes would require much lengthier simulations (for an example see Young & Rosner 2000).

3.2. Structure of a layer

An individual layer in the double-diffusive staircase consists of a zone of overturning convection, separated from its neighbours by stably stratified stagnant zones. In a stagnant zone the transport both of heat and solute takes place by diffusion, so that the profiles of S and T are steep and approximately linear with depth in this zone. In the overturning zone the gradients are shallow, except in the thin boundary layers at its interface with the stagnant zone. The thickness d_s of the stagnant zone is usually only a small fraction of the layer thickness d , but can be as high as 20% close to marginal conditions for layer formation (for a discussion of its dependence on parameters see S13).

Apart from deformations by internal gravity waves in the stagnant zone, both the vertical and the horizontal velocity components vanish at its interface with the overturning zone. Apart from the small amount of solute carried by the flow, the overturning zone thus behaves essentially like laboratory convection in a box, the interfaces with the stagnant zones acting like the top and bottom plates, with viscosity enforcing no-slip conditions. Because of the periodic boundary conditions in the horizontal direction, free slip conditions would allow the stagnant zones as a whole to start moving sideways. To the overturning flow, the stagnant zones would still present internal no-slip interfaces, however. We have chosen the no-slip conditions since the freedom of such large-scale flows is somewhat of an artefact of the periodic boundary conditions in the horizontal direction. In layers of realistically large horizontal extent they would probably not happen. This picture has already been put forward early on in the interpretation of laboratory and geophysical observations (Shirtcliffe 1967; Linden & Shirtcliffe 1978) and of the results from numerical simulations, such as those reported below. It has become the standard interpretation used for deducing heat and solute fluxes in thermohaline staircases in geophysics (e.g. Padman & Dillon 1987; Turner & Stommel 1964; Schmid et al. 2010).

3.3. The overturning zone

At the boundary with the stagnant zone, the overturning flow has three nested boundary layers: for temperature, solute, and flow speed, there are the thermal, solute, and viscous boundary layers, respectively. Their thicknesses are determined by the thermal diffusivity κ_T , solute diffusivity κ_S , and (kinematic) viscosity ν . The highest diffusivity (thermal) has the widest boundary layer. For astrophysical conditions, $\kappa_T \gg \kappa_S, \nu$, so the solute and viscous sublayers are thinner than the thermal boundary layer.

The horizontal velocity vanishes in the stagnant zone (apart from internal wave modes), so it imposes no-slip conditions on the overturning flow. Since the viscous boundary layer is thin compared with the thermal boundary layer; however, this boundary condition has little effect on turnover times and the resulting heat flux, and it becomes noticeable only at low Rayleigh numbers.

The need to properly resolve the narrowest of the boundary layers in a numerical simulation determines the number of grid points needed in the vertical direction. In the astrophysical case, viscosity is typically similar to or somewhat larger than solute diffusivity, so the required resolution is set by the solute boundary layer (cf. Sect. 4.4).

3.4. Heat transport

The flow in the overturning zone is driven entirely by boundary layers at the top and bottom steps of the layer, much like in laboratory convection in a box (Niemela et al. 2000). Except for these thin boundary layers, the (horizontal averages of) entropy and composition are almost uniform inside the layers. Under the conditions in a stellar interior, the thermal diffusivity is much higher than the diffusivity of the solute (helium in hydrogen, say). The thickness of the solute boundary layers is then much smaller than the thermal boundary layers, and the amount of solute in transit across the layer vanishes in this limit. The convective flux is then almost the same as in the absence of a stabilizing solute (cf. Schmitt 1994). Under these asymptotic conditions, the dependence of heat flux on Rayleigh number can be taken from a simple estimate, as was done in S92, or else a result from laboratory measurements of convection can be used (e.g. Niemela et al. 2000). With the convective flow thus known, the flux of solute can be calculated as well.

Convection experiments in gaseous Helium at temperatures just above the critical point by Castaing et al. (1989) yielded the fit

$$\text{Nu}_T = 0.23 \text{ Ra}^{0.282} \quad (1)$$

over the range $10^7 < \text{Ra} < 10^{14}$. More recent measurements by Niemela et al. (2000) using superfluid ^3He , over the remarkable range $10^6 \leq \text{Ra} \leq 10^{17}$ gave a marginally different result:

$$\text{Nu}_T = 0.124 \text{ Ra}^{0.309} \quad (2)$$

This can be compared with the estimate based on a two-dimensional argument in S92:

$$\text{Nu}_T \approx 0.5 \text{ Ra}_*^{0.25} \quad (3)$$

Since the Prandtl number in the laboratory experiments (of order 0.7) is not far from unity, Ra in (1, 2) can be identified with Ra_* in (3). Expression (3) then agrees within a factor 2 with (2), up to $\text{Ra}_* = 10^{15}$.

The expressions above take the form $\text{Nu}_T = a \text{Ra}^b$. They are fits for large Ra ; to better cover lower Rayleigh numbers as well, we modify the expression slightly:

$$\text{Nu}_T - 1 = a (\text{Ra}_* - \text{Ra}_{*c})^b, \quad (4)$$

where $\text{Ra}_{*c} = \text{Pr} \text{Ra}_c$, and $\text{Ra}_c \approx 1400$ the critical Rayleigh number for convection with no-slip boundary conditions. With this change the Nusselt number approaches the diffusive value of 1 as Ra approaches the critical value.

3.5. Solute transport

The transport of solute is governed by the diffusion from the overturning flow into the steep gradient in the stagnant zone. This happens in a narrow boundary layer of the flow. Its thickness δ_s is of the order of $\delta_s = (\kappa_S \tau)^{1/2}$, where τ is the time the flow travels in contact with the boundary before descending/ascending back into the interior. In the same way, the thermal boundary layer has thickness $\delta_T = (\kappa_T \tau)^{1/2}$. The fluxes of solute and heat carried in these boundary layers also determines the flux across the overturning zone as a whole. This predicts that the solute and heat fluxes F_S , F_T are related by

$$F_S/F_T \sim (\kappa_S/\kappa_T)^{1/2}. \quad (5)$$

This is the well-known relation derived in various ways for double-diffusive convection (cf. Turner 1985). In a more accurate form it appears in the model of S13 and in the numerical experiments reported below.

The amount of solute transported across the overturning zone is limited by the fact that convective plumes only develop from material with a net buoyancy of the unstable sign. This limits the density contrast of the stable solute that can be carried to a fraction $1/R_\rho$ of the density contrast across the layer as a whole (Schmitt 1994, S92, S13). Solute with a higher density contrast is not mixed into the convective flow, but remains in the stagnant zone. The solute flux carried by convection in the overturning zone has to match the diffusive flux in the stagnant zone, and idem for the heat flux. In S13 it was shown that these conditions, together with expression (4), form a complete model, determining the effective transport coefficients, as well as the thickness of the stagnant zone, as functions of the three parameters of the Boussinesq problem, R_ρ , Ra_* , Le .

In the estimates above the amount of solute that is transported is given by the amount flowing through the solute boundary layer of thickness δ_s . This assumes that the interaction with the stagnant zone is determined entirely by diffusion of the solute. This is a minimum: the stagnant zone as a whole is stable in the sense of the Richardson condition (one verifies that this is equivalent to the fact that $R_\rho > 1$), but the density gradient is lower near its boundary with the overturning zone, and a certain degree of mixing by shear instabilities is to be expected. This will increase the amount of solute that has low enough buoyancy to be carried with the flow. As argued in S13, this additional amount scales with δ_s , so that its net effect is equivalent to an increase in δ_s by a factor q , of order unity, increasing the solute flux by the same factor. The predicted relation between solute and thermal Nusselt numbers was derived in S13, and is

$$\text{Nu}_S - 1 = \frac{q}{\text{Le}^{1/2} R_\rho} (\text{Nu}_T - 1) \quad (R_\rho < \text{Le}^{-1/2}). \quad (6)$$

The value of q will be a fitting parameter in the quantitative comparison with the numerical results. (The value $R_\rho = \text{Le}^{-1/2}$ is the highest for which the model predicts existence of a layered state at any Rayleigh number, see also Sect. 5.2 below.)

3.6. Model summary

The model used for comparison with the numerical simulations is defined uniquely by the ingredients described above. It makes use of only minimal assumptions about a double-diffusive layer: (a) convection in the overturning zone of a layer is described by a fit to laboratory measurements, (b) transport in the stagnant zone is by diffusion only, and (c) the fluxes of solute and heat are related by the buoyancy limit in the overturning zone (S13 Eq. (17)). These together determine the thickness of the stagnant zone and the effective transport coefficients of heat and solute through the layer as functions of the parameters of the problem, Le , Ra , and R_ρ . The behaviour of the model is discussed in more detail in S13.

4. Numerical simulations

All equations were calculated on a two-dimensional rectangular Cartesian grid in terms of finite differences. The horizontal coordinate is labelled with x , the vertical coordinate with z , where $z = 0$ is the bottom and $z = 1$ the top boundary. The set of equations are implemented in the ANTARES software framework (Muthsam et al. 2010). Advective currents are treated by a weighted essentially nonoscillatory (WENO) finite volume scheme at fifth order (Shu & Osher 1988), and the physical diffusion is handled by a fourth-order finite difference discretization. A second-order total variation diminishing scheme is chosen as time integrator. To avoid odd-even decoupling, a MAC grid, which locates vector variables at cell faces and scalar variables at cell centres, is used.

For a detailed description of the numerical solution of the binary mixture equations presented here, see Zaussinger (2011).

4.1. Boussinesq approximation for a binary mixture

In the Boussinesq approximation the continuity equation is reduced to that of an incompressible fluid,

$$\frac{d\rho}{dt} = 0, \quad (7)$$

such that in the equation of motion the density is taken to be a constant ρ_0 ,

$$\frac{d\mathbf{u}}{dt} = -\frac{1}{\rho_0} \nabla P + (-\alpha_T \Theta + \alpha_S S) \mathbf{g} + \nabla \cdot (\nu \nabla \mathbf{u}), \quad (8)$$

where the ‘‘expansion coefficients’’ α_T , α_S describe the density effects of variations Θ and S in temperature and solute, which is assumed to be low compared with the mean temperature $\bar{\Theta}$ and salinity \bar{S} . They are given by advection-diffusion equations:

$$\frac{d\Theta}{dt} = -\mathbf{u} \cdot \nabla \bar{\Theta} + \nabla \cdot (\kappa_T \nabla \Theta), \quad (9)$$

$$\frac{dS}{dt} = -\mathbf{u} \cdot \nabla \bar{S} + \nabla \cdot (\kappa_S \nabla S), \quad (10)$$

where ν is the kinematic viscosity. The mean gradients $\nabla \bar{\Theta}$ and $\nabla \bar{S}$ can be expressed more usefully in terms of the buoyancy frequencies:

$$N_T^2 = \alpha_T \mathbf{g} \cdot \nabla \bar{\Theta}, \quad (11)$$

$$N_S^2 = \alpha_S \mathbf{g} \cdot \nabla \bar{S}. \quad (12)$$

The density ratio R_ρ is then defined as in the compressible case, Eq. (32). The numerical algorithm solving this set of equations

is based on a semi-implicit scheme. Intermediate values for the velocity field \mathbf{u}^* are calculated explicitly from the equations of motion. By the nature of the incompressible equations, the pressure update is done implicitly, by solving a Poisson equation:

$$\Delta P = \frac{\rho_0}{\Delta t} (\nabla \cdot \mathbf{u}^*). \quad (13)$$

The resulting pressure P leads to the required divergence free velocity field at the new time step $n + 1$.

$$\mathbf{u}^{n+1} = \mathbf{u}^* - \frac{\Delta t}{\rho_0} \nabla P. \quad (14)$$

4.2. Compressible fluid equations for a binary mixture

Verification of the Boussinesq results has been done with simulations of the fully explicit compressible fluid equations. The fluid is assumed to be an ideal gas, which is a good approximation to a binary gas mixture of our interest. These are the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (15)$$

the partial density equation

$$\frac{\partial (\rho c)}{\partial t} + \nabla \cdot (\rho c \mathbf{u}) = \nabla \cdot (\rho \kappa_c \nabla c), \quad (16)$$

the momentum equation

$$\frac{\partial (\rho \mathbf{u})}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u} + P \mathbf{I}) = \rho \mathbf{g} + \nabla \cdot \boldsymbol{\tau}, \quad (17)$$

the total energy equation

$$\frac{\partial \rho E}{\partial t} + \nabla \cdot [\mathbf{u} (\rho E + P)] = \nabla \cdot (K_h \nabla T) + \nabla \cdot (\mathbf{u} \boldsymbol{\tau}) + \rho \mathbf{g} \cdot \mathbf{u}, \quad (18)$$

and the equation of state

$$P = \frac{\mathcal{R} \rho T}{\mu(c)}, \quad (19)$$

where c is the solute mass fraction, $E = \frac{1}{2} |\mathbf{u}|^2 + e$ is the total specific energy in units of energy per mass, e the specific internal energy, K_h the heat conduction coefficient, κ_c the solute diffusion coefficient, $\boldsymbol{\tau}$ the viscous stress tensor, \mathbf{g} the gravitational acceleration, \mathbf{I} the unit tensor, $\mu = \mu(c)$ the mean molecular weight, and \mathcal{R} the gas constant.

4.3. Units, boundary and initial conditions

As unit of length we use the layer thickness d for the Boussinesq cases and the pressure scale height H for the compressible calculations. A nominal convective turnover time is used as unit of time. As units of temperature (potential temperature Θ in the compressible case) we use $1/\alpha_T$ and $1/\alpha_S$ for solute concentration. The density ratio then becomes $R_\rho = \nabla \bar{S} / \nabla \bar{T}$. The Boussinesq equations are invariant to arbitrary additive constants in temperature and solute. We set these such that T and S are zero at the top boundary.

Because of the symmetries of the problem, there are fewer independent parameters than physical variables describing it. So some of the physical quantities appearing in the problem can be set to unity. For these we choose the temperature difference

between top and bottom of the layer, the density at the bottom of the layer, and the acceleration of gravity. The Rayleigh number Ra_* is then controlled through the thermal diffusivity κ_T , the solute difference between top and bottom through the density ratio R_ρ , and the solute diffusivity through the Lewis number Le .

Most of the calculations were done in a box simulating a single layer from the double-diffusive staircase, so the top and bottom boundaries coincide with the steps between layers. This ignores the distortions of the interfaces by surface waves, but since the essence of the double-layering phenomenon is that the transport across the interface is by diffusion, this is not expected to make a big difference. To check that this is indeed the case, a smaller set of simulations was done in which a step is present inside the volume (Sect. 5.7). The vertical boundary conditions are thus taken to be impermeable and stress-free. In the horizontal direction periodic conditions are used:

$$u_z = 0 \quad (z = 0, 1) \quad (20)$$

$$\frac{\partial u_x}{\partial z} = 0 \quad (z = 0, 1) \quad (21)$$

$$S = R_\rho \quad (z = 0) \quad (22)$$

$$T = 1 \quad (z = 0) \quad (23)$$

$$S = 0 \quad (z = 1) \quad (24)$$

$$T = 0 \quad (z = 1). \quad (25)$$

As initial condition the stratification of temperature and solute was taken to be horizontally uniform with either a linear gradient between the values at the top and bottom (the ‘‘linear’’ case below), or something approximating the boundary layer structure expected of the final state (the ‘‘step’’ case). Small initial random perturbations are applied on the solute field.

The numerical algorithm for setting up the initial conditions for the compressible fluid equations is an extension of the procedure presented in Muthsam et al. (1995, 1999).

4.4. Numerical setup

As the thinnest boundary layer in the problem, δ_S determines the numerical resolution needed near the boundaries. Since the fine structure in the interior of the layer largely consists of boundary layers ‘‘peeled off’’ from the boundaries, the same resolution is needed in the interior of the layer as well, and there is no need or justification for using non-uniform grids.

The expected solute boundary thickness for $Ra_* = 1.6 \times 10^5$ and a Lewis number of $Le = 0.1$ is 1.6% of the layer thickness d . With the high-order spatial discretization used, a nominal resolution of $n_{BL} = 3$ points is needed across the critical structures that must be captured. The transport of solute across the layer is determined by the vertical structure of the boundary layers. Resolving these with $n_{BL} = 3$ translates to the $n_z = 200$ points needed in the vertical direction for this case.

To test the actual convergence of the results with respect to resolution, a series of simulations with $n_z = 100, 200, 300$ and 600 has been run for $Ra_* = 1.6 \times 10^5$, $Le = 0.1$, and $R_\rho = 1.15$, see Table 1. Convergence to 15% is already reached at $n_z = 100$. For the lowest Lewis number used in the results reported below in Table 2, $Le = 0.01$, the number of grid points needed would be $\sqrt{10}$ times higher. The number of points across the height of the box used in these simulations, $n_z = 300$, is deemed sufficient for the level of accuracy we are aiming for.

Most of the calculations were done with a horizontal-to-vertical aspect ratio of 2:1. A few tests with different ratios

Table 1. Test series showing convergence of the Nusselt numbers with respect to resolution (number of grid points n_z) across the height of the box.

n_z	Nu_S	Nu_T
100	26	9.5
200	27	10
300	30	10
600	30	10

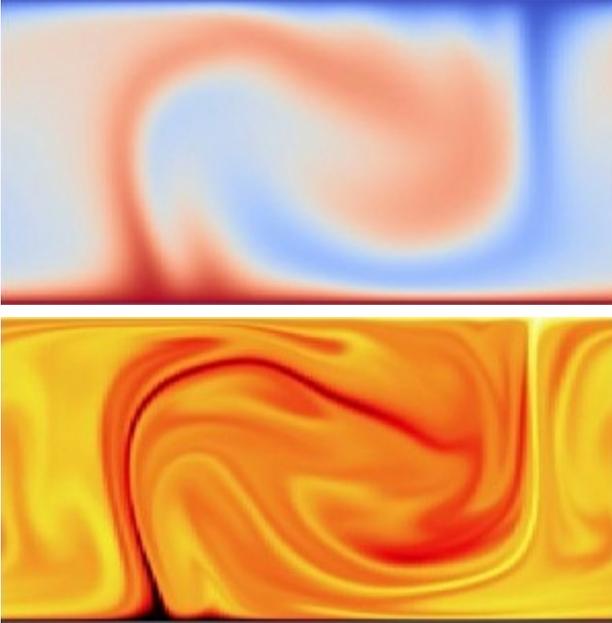


Fig. 1. Flow structure in a double-diffusive layer. The temperature field (*top*) is more diffuse than the solute (“helium”) concentration, as a result of the high thermal diffusivity.

showed that this choice does not affect the measured Nusselt numbers significantly (Sect. 5.4).

5. Results of numerical simulations

Over a wide range in parameter values, about 100 numerical simulations were done in the Boussinesq approximation and about 20 simulations were performed with the fully compressible code. Only the cases closest to the parameter range of astrophysical interest (about 60) are reported here (Figs. 2, 4, Table 1).

An example of the flow structure is shown in Fig. 1, with $Pr = 0.1$, $Ra_* = 5 \times 10^5$, $Le = 0.01$, and $R_\rho = 1.15$. It shows the key characteristics of double-diffusive convection: the boundary layers and the plumes of the solute are narrower than those of temperature, on account of the low solute diffusivity.

5.1. Dependence on RA_* and R_ρ

Figure 2 shows the dependence of the measured Nusselt numbers Nu_T and Nu_S on the parameters Ra_* and R_ρ , for the case $Le = 0.01$ and $Pr = 0.1$. The results show some fluctuation, as expected from the limited number of overturning times for which the simulations were run. This results in uncertainties in the Nusselt numbers of the a factor 1.5. For comparison with the theoretical model predictions (dashed lines) the adjustable erosion factor q in Eq. (6) has been set to 1.5. Within the scatter,

the predictions appear consistent with the numerical results to a factor of 2 or less. In fact, leaving out the erosion factor (i.e. setting $q = 1$) does not make the fit much worse.

The upper right-hand panel shows Nu_S vs. Nu_T , where each curve represents the results for a fixed density ratio R_ρ and concurrently increasing Ra_* . The approximately linear relation between the two shows that the ratio of thermal-to-solute transport depends primarily on Lewis number and density ratio, but not much on Ra , as predicted by (6). The actual transport efficiency of both depends, of course, on the Rayleigh number Ra_* as well (upper left panel).

5.2. Behaviour near the maximum density ratio

The data in Fig. 2 show how the Nusselt number Nu_T declines with increasing density ratio. The theoretical analysis in S13 predicts a maximum density ratio $R_{\rho \max}$ for the existence of a (statistically) stationary layered state. It is a slowly increasing function of Ra_* , asymptotically approaching the value $Le^{-1/2}$. To test this a series of simulations was done at $Ra_* = 10^6$, $Le = 0.1$. The predicted value of $R_{\rho \max}$ for this combination is 1.2 (the asymptotic value would be 1.3). Figure 3 shows the resulting Nusselt numbers as functions of time.

For $R_\rho = 1.2$ and 1.3, the heat flux in the simulations quickly settles to a statistically steady value after a transient due to the initial conditions used. For the three highest values of R_ρ , the initial state smoothly settles in to the purely diffusive state $Nu_T = 1$; the faster and smoother it is, the higher the density ratio. Values in between 1.5 and 1.8 are characterized by large fluctuations superposed on a slow general decline, as if the system vacillates between an overturning or a static. A precise boundary for deciding when an overturning state can exist, is therefore somewhat hard to define: it appears to depend on the length of time over which the flow is followed.

The time dependences of the Nusselt number shown in Fig. 3 suggest a value $R_{\rho \max} \approx 1.4$, somewhat higher than the theoretical value of 1.2. For density ratios slightly higher than 1.2, it looks at first as if a statistically steady state has been reached, but on a longer time scale, a continuing decline of Nu_T is observed. The behaviour with increasing R_ρ is therefore somewhat gradual, at least for the finite length of time over which the simulations have followed the development of the flow. Interestingly, the distinction between diffusive and overturning final states becomes sharper with increasing length of time over which the simulation is followed. This also affects the results in Fig. 2. At the highest density ratios shown there, the numerically measured Nusselt numbers are systematically higher than predicted. This can be attributed to the fact that at these R_ρ the simulations were not run as long as those of Fig. 3.

The numerical results thus confirm the predicted maximum density ratio for the overturning layered state, but with the added twist of very long settling times when R_ρ is near the maximum value. This also answers a question that was left undecided in S13: the theory does not predict what happens close to the maximum density ratio. The results in Fig. 3 indicate that the system vacillates between overturning and diffusive states for increasingly long periods of time as the maximum is approached.

The fast settlement towards a diffusive state for high values of R_ρ is not surprising given that the linear stability criterion predicts exactly such a result for $R_\rho > (1+Pr)/(Le+Pr)$ (e.g Veronis 1968; Stevenson 1979). For the parameter values of Fig. 3, this yields $R_\rho \gtrsim 1.8$. The difference between 1.8, where $Nu_T = 1$ drops only initially, but then only mildly so if at all, and a value

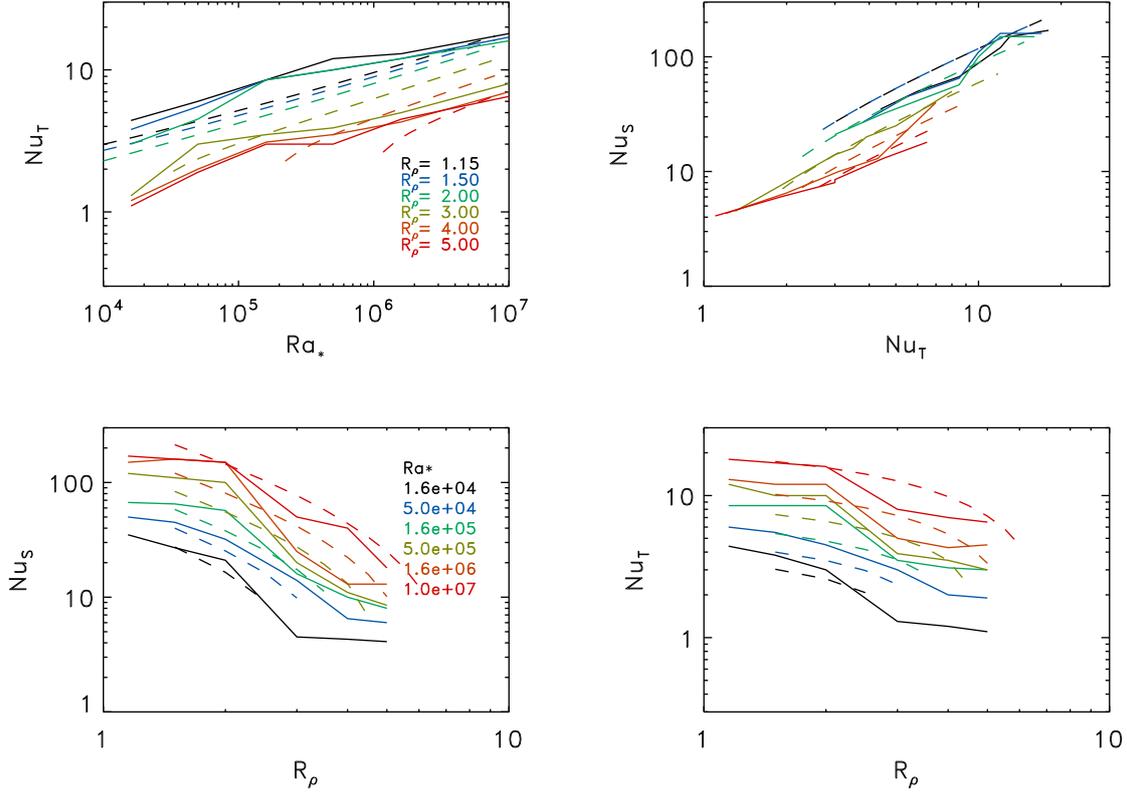


Fig. 2. Dependence of the Nusselt numbers on density ratio R_ρ and Ra_* , for $Le = 0.01$ and $Pr = 0.1$. Solid: numerical results; dashed: model predictions for erosion factor $q = 1.5$.

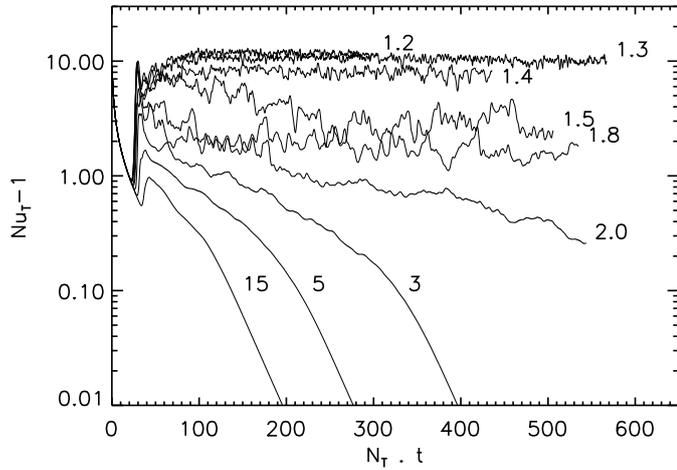


Fig. 3. Time dependence of the thermal Nusselt number Nu_T in simulations with $Pr = 1$, $Le = 0.1$, and $Ra_* = 10^6$, for density ratios (labeled) near the maximum value for which a layered state is predicted by the theoretical model. Time t in units of the thermal buoyancy time scale N_T^{-1} .

of 2.0, which shows a continuous decay seems consistent with this classical result.

The results also provide an interesting link to Proctor's (1981) analysis. In this theory the maximum value $R_\rho = Le^{-1/2}$ appears as a limit for its applicability. Proctor's analysis therefore does not answer what happens near this critical number, any more than the model in S13 does, but it is gratifying that it appears to have a consistent place in both, as well as in the numerical results.

5.3. Dependence on Pr

Figure 4 shows the results of a set of simulations testing the dependence on Prandtl number. The Lewis number is fixed at $Le = 0.01$. The density ratio is varied over the range $1.15 < R_\rho < 5$, Prandtl number from $Pr = 0.01$ to $Pr = 1$. Within the limited numerical sensitivity due to the stochastic nature of the flow, little systematic dependence on Pr is detectable. As argued above, a low dependence on Pr was to be expected in the limit $Pr \downarrow 0$. It appears that this also extends to a Prandtl number of order unity.

5.4. Dependence on aspect ratio

The aspect ratio of 2:1 used in the results reported above was tested against 5:1 and 10:1 at the same spatial resolution. For the reference simulation ($Pr = 0.1$, $Le = 0.01$, $Ra_* = 10^5$, $R_\rho = 1.15$), we find $Nu_S = 90$ and $Nu_T = 8.5$. By comparison the simulation with 5:1 results in $Nu_S = 90$ and $Nu_T = 9.0$. The most extended box with an aspect ratio of 10:1 and a spatial resolution of 1500×300 has Nusselt numbers of $Nu_S = 80$ and $Nu_T = 8.75$. The aspect ratio thus has no significant influence on the dependences of the fluxes on input parameters in our simulations, within the fluctuations due to the stochastic nature of the flow.

5.5. Dependence on initial stratification

As initial state we used either a constant linear gradient of temperature and solute between top and bottom values ("linear"), or a profile approximating the expected steplike combination of a stagnant and an overturning zone. The linear case shows how the oscillatory phase due to the Kato instability develops into

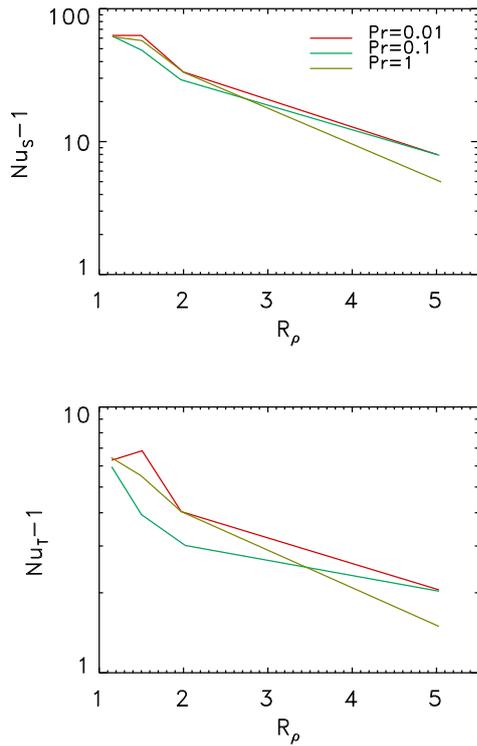


Fig. 4. Dependence on Prandtl number and density ratio. Solute Nusselt number (*top*) and thermal Nusselt number (*bottom*) for $Pr = 1, 10^{-1}$, and 10^{-2} , with $Ra_* = 1.6 \times 10^5$ and Le fixed at 0.01.

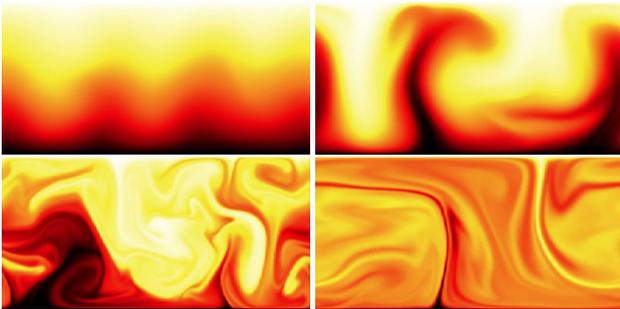


Fig. 5. Example of the development of an overturning flow from Kato oscillations for $Pr = 1$, $Le = 10^{-1}$, $R_\rho = 1.15$, and $Ra_* = 1.6 \times 10^5$ starting from a linear stratification. Time from left to right and top to bottom. Wave braking occurs after 5 oscillations (Fig. 5.2). The layer is fully evolved after 10 oscillations (Fig. 5.4). See also the movie online.

overturning flow, see Figs. 5 and 6. The duration of the initial formation process is mainly determined by R_ρ and Ra_* . It is in the range from five to ten oscillation periods of the initial stratification. The end state in both cases is statistically the same. The “step” as initial condition also covers cases that are stable in linear theory because of the subcriticality of the system, and would not develop from a linear initial profile. With this initial state, computing time can be saved for low values in Ra_* and high R_ρ . It was used in most of the results reported above.

5.6. Comparison with compressible results

The compressible simulations are based on a fifth-order weighted ENO scheme. Compressible fluids lead to restrictions in time stepping (due to the need to resolve sound waves). The compressible simulations take up to 100 times longer for the

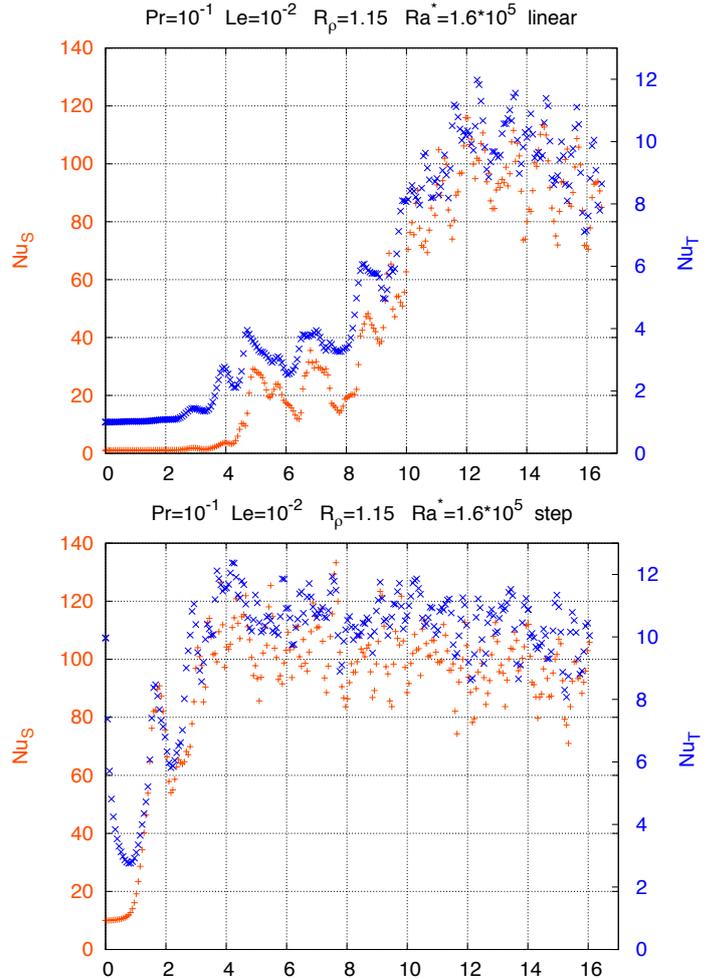


Fig. 6. *Top panel:* evolution of the Nusselt number for the linear initial stratification. Convective cells get established from Kato oscillations (cf. Fig. 5) after about 10 turnover times. Starting the simulation from a step (*bottom*) saves computing time. At the end of the runs the Nusselt numbers of both simulations are the same within the statistical variations.

same resolution compared to the incompressible solver. In both regimes a resolution of 300×300 points has been set. Therefore only a few tests have been done with the compressible code. The degree of compressibility is governed by the $\epsilon = d/H$ of layer thickness to pressure scale height. In the limit $\epsilon \rightarrow 0$, there is a direct translation between the compressible and the Boussinesq case (Sect. 6.1). The results of a numerical comparison with $\epsilon = 0.1$ is shown in Table 1. The resulting Nusselt numbers do not differ significantly.

Simulations done with $\epsilon = 1.0$ behave quite similar to those done with $\epsilon = 0.1$. The mixing processes do not significantly differ, at least for $Pr \leq 0.1$ and as long as the Rayleigh number is high enough, $Ra > 5.0 \times 10^5$. Differences of around 50% in the Nusselt numbers were found, but considering the present level of accuracy (a factor of 2, say), we conclude that the Boussinesq approximation qualitatively gives the right results, even for layer thicknesses approaching a scale height. At Pr of order unity or larger, the effect of viscous damping might become more significant in more strongly stratified cases, however.

Table 2. Comparison of compressible and incompressible simulations.

Pr	Le	R_ρ	Ra_*	Nu_S^B	Nu_T^B	Nu_S^c	Nu_T^c
10^{-1}	10^{-2}	2.0	1.6×10^5	60	8	55	7
10^{-1}	10^{-2}	1.2	1.6×10^5	110	12	110	11
10^{-1}	10^{-2}	2.0	1.6×10^6	150	12	130	10
10^{-1}	10^{-2}	1.2	1.6×10^6	200	16	200	14
1.0	10^{-1}	2.0	1.6×10^5	3.5	2	11	1.5
1.0	10^{-1}	1.2	1.6×10^5	45	15	17	5
1.0	10^{-1}	2.0	1.6×10^6	4	3.5	26	10
1.0	10^{-1}	1.2	1.6×10^6	33	11	26	10

Notes. Thermal (T) and solute (S) Nusselt numbers from the Boussinesq (B) and compressible (c) results. Layer thickness d is 0.1 pressure scale height.

5.7. Multilayer simulations

In all of the above we have assumed that the interfaces between the double-diffusive layers can be approximated as solid boundaries. To test the reliability of this assumption, a few cases were run where the initial state consisted of two steps instead of a single one. In some, though not all of these runs, the division into two layers remained until the end. An example is shown in Fig. 7 for a case with $Pr = 1$, $Le = 10^{-2}$, $R_\rho = 1.15$, and $Ra_* = 6 \times 10^5$ and a resolution of 500×500 . We note the approximate (anti-)symmetry of the plumes near the interface in the middle, a phenomenon known from laboratory experiments (e.g. Fig. 1 in Turner 1985). It is caused by the continuity of the horizontal velocity across the interface enforced by viscosity. This symmetry is less marked in simulations at lower Prandtl numbers, when the thickness of the viscous boundary layer becomes smaller than the thickness of the stagnant zone.

Figure 8 shows horizontal and temporal average profiles of temperature and solute with height. The transition in the middle is broad compared with the boundary layers at top and bottom. Inspection of the time-dependent flow shows that this is due to two separate effects. One is the displacements of the interface by surface waves, which smoothen the average gradient without changing the actual fluxes across the interface. In addition there is possibly some real mixing associated with breaking of the surface waves, but it remains localized around the interface between overturning and stagnant zone.

6. Application to stars

6.1. Boussinesq limit: thin layers

If the double-diffusive layering is thin, convective overturning times are longer than the sound crossing time of the layer thickness, and a Boussinesq (or other “low Mach number” approximation) can be used for a compressible fluid. By the presence of a stratification in density and temperature, however, a heat flux is present even in a convectively neutral stratification. This requires some care in translating Boussinesq results to the astrophysical case.

6.2. Heat flux

We let $F_{r,c}$ be the radiative and hydrodynamic (semiconvective) contributions to the heat flux in the star. The radiative heat flux is proportional to the logarithmic temperature gradient $\nabla = d \ln T / d \ln P$:

$$F_r = k \nabla = k \nabla_a + k(\nabla - \nabla_a) \equiv F_{ra} + F_{rs}, \quad (26)$$

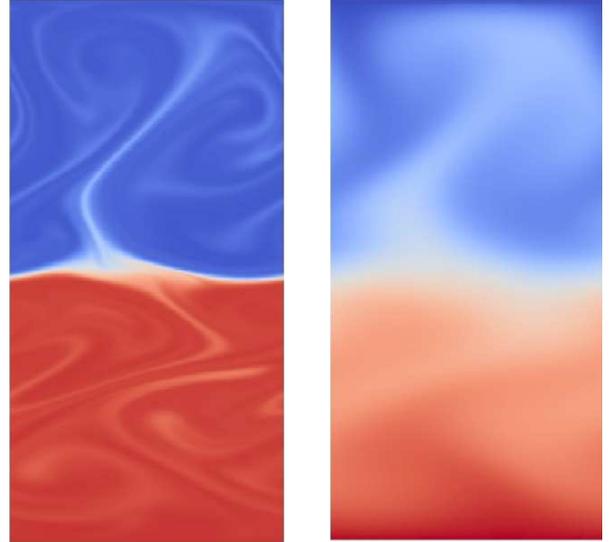


Fig. 7. Snapshot of a simulation including the free interface between two layers. *Left:* solute, *Right:* temperature. $Pr = 1.0$, $Le = 0.01$, $R_\rho = 1.15$, $Ra_* = 6 \times 10^5$. See also the movie online.

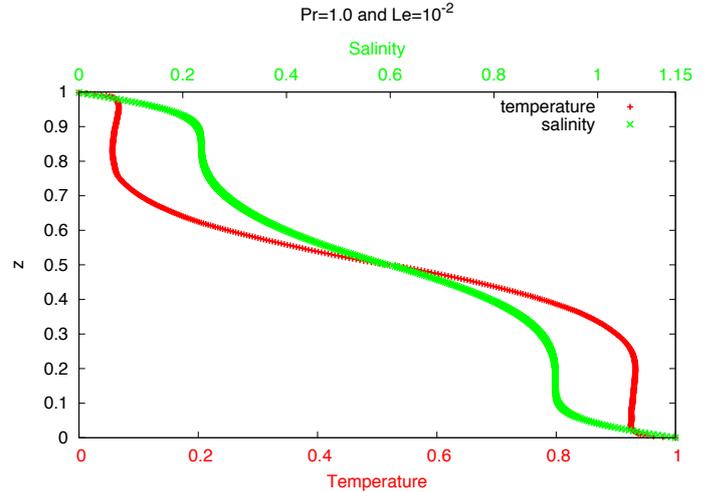


Fig. 8. Average temperature and solute profiles with height z in the double-layer simulation of Fig. 7.

where k is a constant depending on the local thermodynamic state, ∇_a the adiabatic gradient, and F_{ra} , F_{rs} are the contributions to the radiative heat flux of the adiabatic and the superadiabatic parts of the mean temperature gradient, respectively. In the Boussinesq model, the contribution F_{ra} is absent: the convective and radiative heat fluxes F_c and F_r are governed by the same temperature gradient. Related to this, the Boussinesq model has one parameter less: the pressure scale height H . Because of this difference, the heat flux in the stellar (compressible) model cannot be compared directly with the heat flux in an incompressible model. Instead, the ratio of convective to radiative heat flux in the Boussinesq model is to be identified with the ratio $f = F_c / F_{rs}$ of the convective flux to the superadiabatic component of the radiative flux in the star, in the limit $H \rightarrow \infty$ (cf. Massaguer & Zahn 1980). The Boussinesq model is thus the limit $\epsilon \downarrow 0$ (taken at fixed Ra_*). If the semiconvective layer thickness ϵ is sufficiently small, the compressible case can be compared directly with the Boussinesq model (as verified by the numerical tests with $\epsilon = 1$ and $\epsilon = 0.1$ in Sect. 5.6 above).

This makes the semiconvection problem more amenable to numerical simulation. In contrast to a convective stellar envelope, for example, with its many scale heights to be covered, the layered nature of double-diffusive convection puts it in a parameter range that is much closer to conditions accessible with realistic *ab initio* calculations.

With (26) for the radiative flux, the total heat flux can be written as

$$F = F_r + F_{\text{conv}} = k\nabla_a + \text{Nu} k(\nabla - \nabla_a), \quad (27)$$

where Nu is the Nusselt number. By taking out the radiative flux associated with the adiabatic gradient, the Nusselt number as defined in (27) can now be identified with the Boussinesq equivalent (see also Massaguer & Zahn 1980). The relevant Nusselt number is thus not simply the ratio of total to radiative heat flux².

6.3. Rayleigh number and density ratio

The modified Rayleigh number for a layer of thickness d is, in astrophysical notation:

$$\text{Ra}_* = \text{Pr} \text{Ra} = \frac{gd^4}{\kappa^2 H} (\nabla - \nabla_a), \quad (28)$$

where κ is the thermal diffusivity (cm^2/s) and g the acceleration of gravity. Ra_* is typically a very large number unless the layer thickness d is small (for estimates see the $15 M_\odot$ example below). Apart from Le, the problem depends on the relative strength of the stabilizing solute gradient relative to the destabilizing thermal gradient. This can be measured in terms of the thermal and solute buoyancy frequencies N_T , N_S :

$$N_T^2 = \frac{g}{H} (\nabla_a - \nabla), \quad (29)$$

$$N_S^2 = \frac{g}{H} \nabla_\mu, \quad (30)$$

with

$$\nabla_\mu \equiv d \ln \mu / d \ln P, \quad (31)$$

where μ is the mean atomic weight per particle, and the gradients are understood as mean values over a double-diffusive step. The density ratio R_ρ ,

$$R_\rho \equiv -N_S^2 / N_T^2 = \frac{\nabla_\mu}{\nabla - \nabla_a}, \quad (32)$$

is then the equivalent of the density ratio in the Boussinesq formulation, and is the dimensionless measure we use for the relative strength of the stabilizing helium gradient. Semiconvection, i.e. a stratification between Schwarzschild and Ledoux then corresponds to $R_\rho > 1$. $\nabla - \nabla_a$ is typically a small number in a stellar interior (even in the presence of the double-diffusive steps), and R_ρ a large number (see below).

6.4. Extrapolation

The connection between the numerical results and the astrophysical conditions requires extrapolation that cannot at present be covered by the simulations. It is covered by the model in S13, as validated by the comparison with our numerical results. In

² The implications of this distinction are not apparent in some of the published work on semiconvection.

contrast with the laboratory setup where the temperature difference is the imposed quantity, the heat flux is the fixed quantity under the conditions in a stellar semiconvective zone, since the structure of the star only depends a little on the behaviour of the semiconvective zone. The asymptotic conditions for large Ra_* then lead (S13 Sect. 6) to simple expressions for the superadiabaticity $\nabla - \nabla_a$ and the effective Helium diffusivity $\kappa_{s \text{ eff}}$:

$$\nabla - \nabla_a \approx \text{Le}^{1/2} \nabla_\mu, \quad (33)$$

$$\kappa_{s \text{ eff}} = (\kappa_s \kappa)^{1/2} \frac{(\nabla_r - \nabla_a)}{\nabla_\mu} \quad (34)$$

(S13 Eqs. (57), (59)). The analysis in S13 does not take a modification due to the effect of radiation pressure discussed in S92 into account. With this modification, Eq. (34) becomes

$$\kappa_{s \text{ eff}} = (\kappa_s \kappa)^{1/2} \left(\frac{4}{\beta} - 3 \right) \frac{(\nabla_r - \nabla_a)}{\nabla_\mu}, \quad (35)$$

where $\beta = P_g / (P_g + P_r)$ is the ratio of gas to total pressure. The density ratio approaches the value

$$R_\rho \approx \text{Le}^{-1/2}, \quad (36)$$

while the relative thickness of the stagnant zones in this limit is of order (setting $\beta = 1$)

$$\delta \approx 1/\text{Nu}_s = \text{Le}^{1/2} \nabla_\mu / (\nabla_r - \nabla_a). \quad (37)$$

The range of validity of the limit is

$$l_0 \ll d \ll H, \quad (38)$$

where

$$l_0 = (\kappa^2 H / g)^{1/4} \quad (39)$$

is the length scale on which the thermal diffusion time scale equals the free fall time over a pressure scale height.

The asymptotic value $R_\rho = \text{Le}^{-1/2}$ (Eq. (36)) is also the maximum density ratio for which the theory predicts the existence of the layered state. Semiconvection in a stellar model with heat flux fixed is thus predicted to settle close to this maximum, if the layer thickness is not small ($d \gg d_0$ or $\text{Ra}_* \gg 1$). How close it actually settles can also be predicted, but requires analysis to the next order in the small quantity $1/\text{Ra}_*$.

In a stellar evolution model, expressions (33, 35) give nearly identical results to the model of Zaussinger (2011), which is based on S92. They are easier to implement since they have been restricted to conditions $d \gg l_0$ that are the most interesting for stellar evolution anyway.

6.5. A $15 M_\odot$ star

We are now in the position to estimate the range of parameter values for semiconvection in a star. Consider the important case of massive stars around main sequence turnoff. We use the model kindly provided by A. Weiss (model “fzm15_151”). Characteristic values for the physical quantities in the semiconvective zone of this model are $g \approx 10^6 \text{ cm/s}^2$, $\kappa_S \approx 1 \text{ cm}^2/\text{s}$, $\kappa \approx 3 \times 10^8 \text{ cm}^2/\text{s}$, $H \approx 2 \times 10^{10} \text{ cm}$, $\nabla_a = 0.4$, $\nabla_r - \nabla_a \approx 0.02$, and $\nabla_\mu \approx 1$.

With Eqs. (28) and (33), the Rayleigh number can be found for this semiconvective zone, as a function of the layer thickness d . This yields $\text{Ra}_* \sim 10^{12}$ for $d/H = 0.1$, or $\text{Ra}_* \sim 10^8$ for $d/H = 0.01$, for example. For such layer thicknesses, conditions

are thus in the asymptotic regime for which the expressions in 6.4 are approximately valid.

The effective He-diffusivity from (35) is $\kappa_{s\text{eff}} \approx 10^3 \text{ cm}^2/\text{s}$, three orders of magnitude above the microscopic value. The mixing time scale over a pressure scale height is thus about 10^{10} yr . The value of l_0 is $2 \times 10^5 \text{ cm}$, or about 10^{-5} pressure scale heights. The limiting expressions (33–37) are thus applicable over a large range in the value of the (uncertain) layer thickness. The lower end of this range is not likely to be very relevant, since at $d \sim l_0$ the expected time scale for layer merging (see below) is very short, of the order of $\text{Le}^{-1/2}$ times the free fall time scale over a pressure scale height (from Eqs. (39), (42)).

The time scale for the initial layer formation from a Kato oscillation is similarly short, a few times the instability growth time, which is close to the convective overturning time in the absence of the stabilizing μ -gradient. For the $15 M_\odot$ star, this works out to about one day. The asymptotic expressions (33) and (35) should therefore be adequate for use in stellar evolution codes.

6.6. Layer thickness

While the layer thickness has little influence on the effective mixing rate, it is useful to check whether it could become significant compared with the pressure scale height, in which case the limit of thin layering assumed would not be valid. This requires a model for the layer thickness, for which no good theory is available. Observations in laboratory experiments and in geophysical cases show that layer thickness is not constant, but grows in time by processes in the merging of neighbouring layers (McDougall 1981; Ross & Lavery 2009, see also the numerical experiment in Young & Rosner 2000). Layer thickness can therefore *not be treated independently of the history of the system*.

An estimate can be made, however, from the effective solute diffusivity. Changes in layer thickness involve the exchange of solute between neighbouring layers. Over a distance D , the time scale τ for exchange is given by the effective diffusivity $\kappa_{s\text{eff}}$: $\tau \approx D^2/\kappa_{s\text{eff}}$. Setting D equal to the layer thickness d itself then gives an estimate of the rate of change of thickness by merging

$$d \ln d / dt = 1/\tau, \quad (40)$$

or

$$\frac{d}{dt} d = \frac{\kappa_{s\text{eff}}}{d}. \quad (41)$$

If $\kappa_{s\text{eff}}$ is constant in time, then

$$d = (2\kappa_{s\text{eff}}t)^{1/2}, \quad (42)$$

where t is time since the semiconvective condition started. In the $15 M_\odot$ star the duration of the semiconvective state is of the order of 10^7 yr , so the layer thickness to be expected at the end of the semiconvective phase is of the order of 10^9 cm , or $0.05 H$. Given the uncertainties involved, it cannot be ruled out that semiconvective layer thickness d can actually become macroscopic, of order H , in the course of the star's evolution.

7. Discussion

As in the case of ordinary convection, there are far fewer intrinsic parameters in semiconvection than the quantities defining the physical state in a stellar interior. This allows a significant volume of astrophysically useful parameter space to be covered by a

grid of numerical simulations. A physical model as applied here is still needed, however, to interpret the results and to extrapolate them meaningfully to the astrophysically relevant range.

The low value of the viscosity and solute diffusivity compared with the thermal diffusivity constitute a limiting case that actually simplifies the double-diffusive problem greatly. Among other things, the results become nearly independent of the Prandtl number in this limit (as suggested already by Proctor's 1981 analysis).

The simulations were all done in 2D, so tests in 3D will be needed for verification. It is unlikely, however, that the results will turn out very different, at least within an astrophysical factor of 2. The reason is that in the low-Pr, low-Le limit the flow in the layers is almost equivalent to ordinary convection between plates. Owing the low solute diffusivity, the amount of solute in the bulk of the layer is small. It can then be treated as a perturbation, as assumed in S92. Known laboratory results for convection at very high Rayleigh numbers can then be used to extrapolate the numerical results. As shown in S13, this makes predictions similar to the simple 2D model used in S92. In particular the effective mixing rate is very low, as also found (in a different region in parameter space) in geophysical examples, such as Lake Kivu and the (Ant)Arctic oceans (cf. Sect. 2 and S13).

Stochastic fluctuations in the flow produce scatter in the fluxes of heat and solute, which affect the measured averages (cf. Table 1). The accuracy of these averages could be improved with longer runs.

Most of the results presented are based on simulation of a single double-diffusive layer. As we found in Sect. 5.7, this does not fully reproduce broadening of the interfaces between layers by surface waves (see movie in Fig. 7).

The physics determining the thickness of the individual double-diffusive layers remains uncertain. In the equivalent geophysical examples semiconvective zones always consist of a large number of very long-lived layers. Observations in the East-African rift lakes (e.g. Schmid et al. 2010) show that layers first forming at the boundary of an expanding double-diffusive zone are always thin, subsequently growing slowly by a process of merging. This is understood in terms of an energy argument (cf. Sect. 2), and is likely to happen in a growing stellar semiconvective zone as well.

The estimate (Eq. (42)) suggests that layer thickness might approach a significant fraction of a pressure scale height. In this case, only a few layers would be left at the end of the semiconvective phase, and the location of their boundaries may well vary somewhat randomly between stars. As surmised in Zaussinger (2011), this might introduce a random element in the late stages of the evolution of massive stars.

Semiconvection as studied here addresses only one of two astrophysical meanings of the term. In addition to the effect of a stabilizing solute gradient, there is the effect of a composition-dependent opacity: the increase in the radiative gradient when mixing is assumed, with consequent onset of convection in a stratification that was radiative before mixing. Historically, this has actually been the main concern in stellar evolution computations. The physics of this kind of semiconvection is completely different from the double-diffusive kind (cf. Kippenhahn et al. 2013); it has not received the same theoretical attention.

8. Conclusions

The results show how the physics of double-diffusive convection, which is known from geophysical and laboratory studies, can be applied to the astrophysical case of low Prandtl and Lewis

numbers. As expected, the process takes place in the form of the characteristic double-diffusive layering known and theoretically understood since the 1980s.

In the asymptotic regime occupied by astrophysical conditions, the number of independent parameters determining the physics is effectively reduced to three, so that a meaningful range of parameter values can be covered with numerical simulations. We have compared a grid of simulations with the predictions of the model for a semiconvective layer in S13. In the parameter region where the estimate overlaps with the region covered by the simulations we found good agreement, even without significant tuning of the one fitting parameter used (the “erosion factor” q , Sect. 3.5).

The simulations also provide an answer to an interesting question raised by Proctor’s (1981) mathematical analysis and by the model in S13. Both predicted that a critical number $R_\rho = \text{Le}^{-1/2}$ plays a role in the steady layered state. The results presented in Sect. 5.2 clarify the behaviour of the layered state around this maximum density ratio. In a system set up in a layered state just above $R_{\rho \text{ max}}$, the Nusselt numbers show large fluctuations in amplitude, and the duration of the fluctuations lengthens as R_ρ approaches $R_{\rho \text{ max}}$. This explains why theories based on the assumption of stationarity fail near this limit.

Boussinesq and fully compressible simulations give equivalent results in the limit of small layer thickness. The Boussinesq calculations even reproduce the approximate Nusselt numbers for layers as thick as a scale height. The numerical results confirm the theoretical expectation that the dependence on viscosity is weak at low Prandtl number and extends to Pr of order unity.

The effective mixing rate in the semiconvective zone of a $15 M_\odot$ main sequence star is predicted to be low, though the semiconvective phase may possibly leave significant jumps in the profile of helium concentration.

Acknowledgements. We thank the referee for a perceptive and detailed review, which has led to substantial improvements. F. Zaussinger is grateful to F. Kupka and H.J. Muthsam for discussions of the numerical solution of binary mixture equations with high resolution methods. He was supported by the DFG within the project “Modelling of diffusive and double-diffusive convection” (projects KU 1954-3/1 and KU 1954-3/2 in SPP 1276/1 and SPP 1276/2, project leader F. Kupka) within the interdisciplinary Metstroem project. He was also supported by the Austrian Science Foundation (project P20973, Numerical Modelling of Semiconvection, project leader H.J. Muthsam). We also thank H. Grimm-Strele for implementing a parallel Poisson solver in the ANTARES code suite.

References

- Biello, J. A. 2001, Ph.D. Thesis, Univ. of Chicago
 Castaing, B., Gunaratne, G., Kadanoff, L., Libchaber, A., & Heslot, F. 1989, *J. Fluid Mech.*, 204, 1
 Huppert, H. E., & Turner, J. S. 1981, *J. Fluid Mech.*, 106, 299
 Kato, S. 1966, *PASJ* 18, 374
 Kippenhahn, R., Weigert, A., & Weiss, A., 2013, *Stellar Evolution* (Berlin: Springer)
 Langer, N., El Eid, M. F., & Fricke, K. J. 1985, *A&A*, 145, 179
 Langer, N., El Eid, M. F., & Baraffe, I. 1989, *A&A*, 224, L17
 Linden, P. F., & Shirtcliffe, T. G. L. 1978, *J. Fluid Mech.*, 87, 417
 Lu, H., Swift, H. P. 2001, in *Conference on desalination strategies*, Jerba 11-12 Sept. 2000, Elsevier (available at www.desline.com/articoli/4052.pdf)
 Maeder, A. 1997, *A&A*, 321, 134
 Massager, J. M., & Zahn, J.-P. 1980, *A&A* 87, 315
 McDougall, T. 1981, *Progr. Oceanogr.*, 10, 91
 Merryfield, W. J. 1995, *ApJ*, 444, 318
 Muthsam, H. J., Goeb, W., Kupka, F., Liebich, W., & Zochling, J. 1995, *A&A*, 293, 127
 Muthsam, H. J., Goeb, W., Kupka, F., & Liebich, W. 1999, *New Astron.*, 4, 405
 Muthsam, H. J., Kupka, F., Löw-Baselli, B., et al. 2010, *New Astron.*, 15, 460
 Nayar, A. 2009, *Nature*, 460, 321
 Niemela, J. J., Skrbek, L., Sreenivasan, K. R., & Donnelly R. J. 2000, *Nature*, 404, 837
 Padman, L. & Dillon, T. M., 1987, *J. Geophys. Res.*, 92, 10799
 Proctor, M. R. E. 1981, *J. Fluid Mech.*, 105, 507
 Rosenblum, E., Garaud, P., Traxler, A., & Stellmach, S. 2011, *ApJ*, 731, 66
 Ross, T., & Lavery, A. 2009, *Expe. Fluids*, 46, 355
 Schladow, S. G., & Imberger, J. 1987, *J. Geophys. Res.*, 92, 6501
 Schmid, M., Busbridge, M., & Wüst, A. 2010, *Limnol. Oceanogr.*, 55, 225
 Schmitt, R. 1987, *Deep Sea Research Part I: Oceanographic Research*, 34, 1655
 Schmitt, R. W. 1994, *Ann. Rev. Fluid Mech.*, 26, 255
 Shirtcliffe, T. G. L. 1967, *Nature* 213, 489
 Shu, C.-W., & Osher, S. 1988, *J. Comput. Phys.* 77, 439
 Sigvaldason, G. 1989, *J. Volcanol. Geoth. Res.*, 39, 97
 Spiegel, E. A. 1969, *Comments on Astrophysics and Space Physics*, 1, 57
 Spiegel, E. A. 1972, *ARA&A*, 10, 269
 Spruit, H. C. 1992, *A&A*, 253, 131 (S92)
 Spruit, H. C. 2013, *A&A*, 552, A76 (S13)
 Tayler, R. J. 1953, *Doctoral Dissertation*, Cambridge University
 Turner, J. S. 1979, *Buoyancy Effects in Fluids* (Cambridge University Press)
 Turner, J. S. 1985, *Ann. Rev. Fluid Mech.* 17, 11
 Turner, J. S., & Stommel, H. 1964, *Proc. Nat. Acad. Sci.*, 52, 49
 Stevenson, D. J. 1979, *MNRAS*, 187, 129
 Veronis, G. 1968, *J. Fluid Mech.*, 34, 315
 Young, Y., & Rosner, R. 2000, *Phys. Rev. E*, 61, 2676
 Zaussinger, F. 2011, Ph.D. Thesis, University of Vienna, <http://pubman.mpg.de/pubman/item/escidoc:1121652:2>