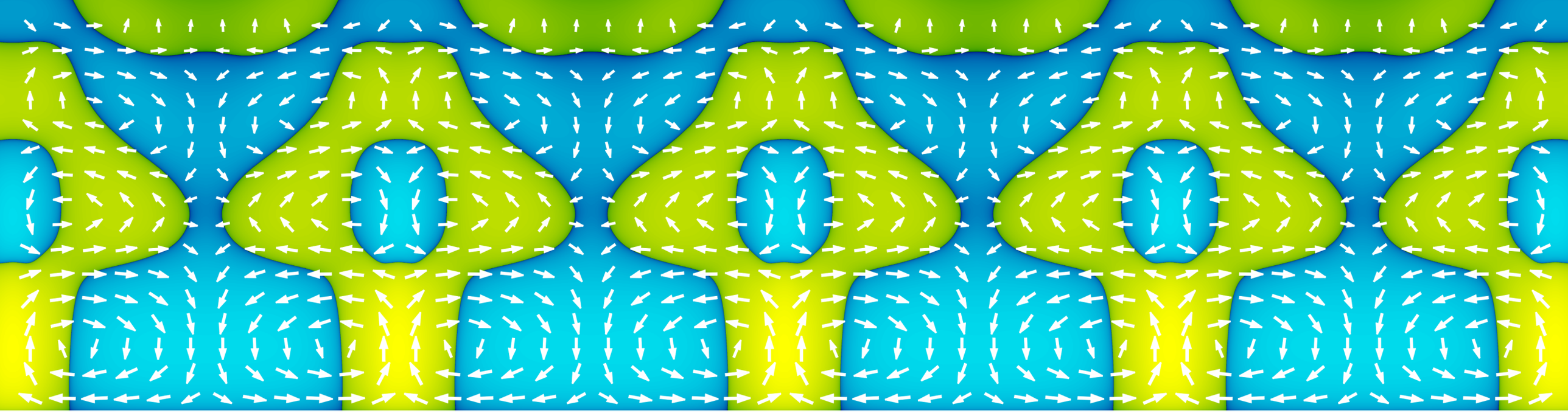
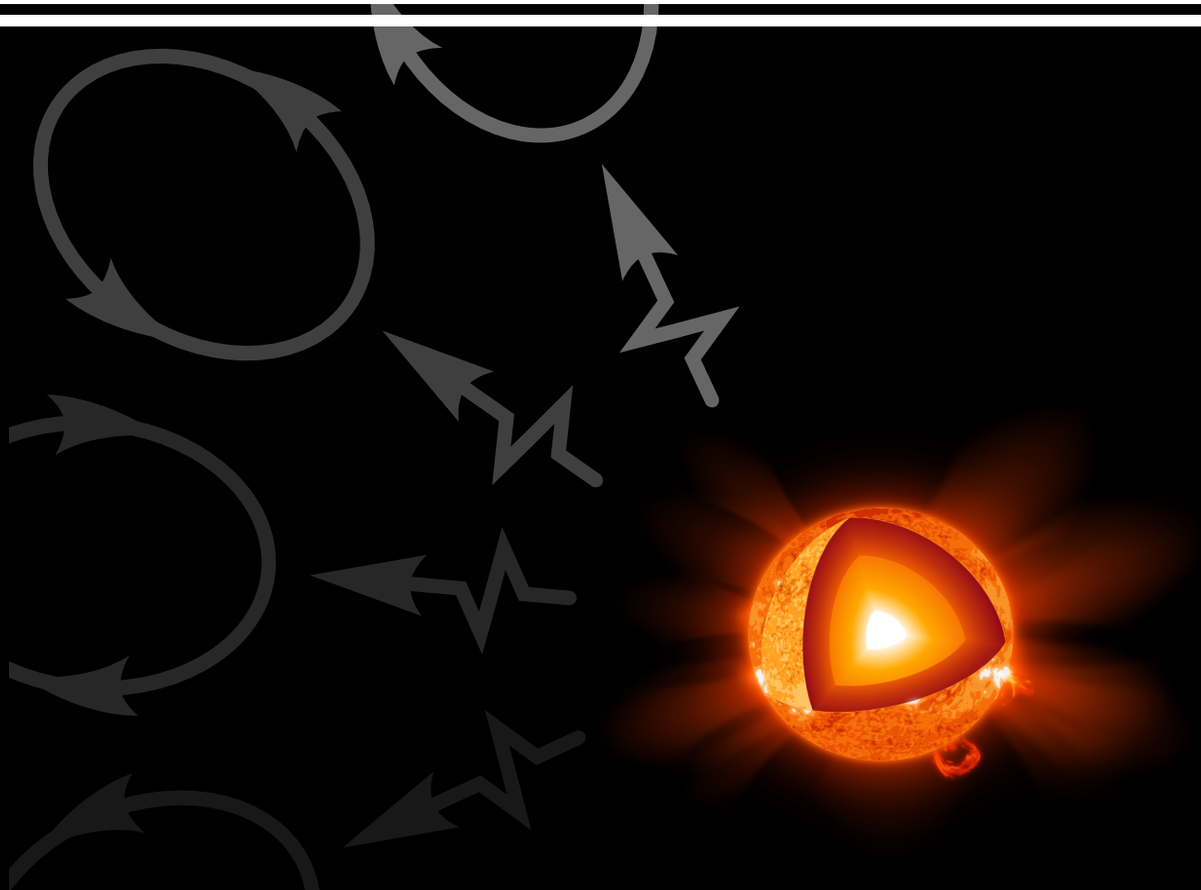


Convective overshooting in stars

Róbert Andrássy



Róbert Andrássy | Convective overshooting in stars

Convective overshooting in stars

Cover

Top: Structure of the diffusion-dominated, differential-heating flow. The dataset and its visualisation are identical to those in the right panel of Fig. 4.7.

Bottom left: Sketch showing the two main modes of heat transport in the stellar interior: radiative diffusion (jagged arrows) and convection (oriented ellipses).

Bottom right: Structure of the Sun. From the centre outwards are the core (white), the radiative zone (bright orange), the convection zone (dark orange), the photosphere (yellow section of the surface), the chromosphere (orange section of the surface) and the corona (orange "rays"). The picture is the work "Sun poster"[1] by Kelvin Song[2] (CC BY-SA[3]).

[1] http://commons.wikimedia.org/wiki/File:Sun_poster.svg
[2] http://commons.wikimedia.org/wiki/User:Kelvinsong
[3] http://creativecommons.org/licenses/by-sa/3.0/deed.en

Convective overshooting in stars

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. D. C. van den Boom
ten overstaan van een door het College voor Promoties ingestelde
commissie, in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 19 juni 2015, te 14:00 uur

door

Róbert Andrássy

geboren te Košice, Slowakije

Promotiecommissie

Promotor:	prof. dr. H. C. Spruit, Universiteit van Amsterdam
Overige leden:	prof. dr. L. Kaper, Universiteit van Amsterdam prof. dr. M. B. M. van der Klis, Universiteit van Amsterdam dr. O. R. Pols, Radboud Universiteit Nijmegen dr. A. L. Watts, Universiteit van Amsterdam prof. dr. R. A. M. J. Wijers, Universiteit van Amsterdam
Faculteit:	Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Contents

1 Introduction

Numerous observations provide evidence that the standard picture, in which convective mixing is limited to the unstable layers of a star, is incomplete. The mixing layers in real stars are significantly more extended than what the standard models predict. Some of the observations require changing the thermal structure of the stellar model, e.g. the poorly-fitting evolutionary tracks of intermediate-mass and massive stars, or the "inflated" cores of OB stars seen by asteroseismic techniques. On the other hand, there are effects that can be explained by some form of weak mixing with no influence the structure of the star, e.g. the Li and Be depletion in low-mass stars, or the abundance anomalies observed in red giants. Both kinds of mixing have traditionally been covered be the umbrella term "convective overshooting," although the physics involved is clearly different. There is a need to look at the overshooting problem from a more fundamental perspective. In this thesis, we investigate two mixing process, which could explain the Li and Be depletion in low-mass stars ("convective settling," Chaps. 2 and 3) and contribute to slow mixing at the boundaries of internal convection zones in a whole range of different stars ("differential heating," Chap. 4).

1.1 Stellar convection

Whatever their mass, stars contain various convection zones during their early formation phases, on the main sequence, and also in the late stages of their evolution. This represents a wide range of conditions from the fluid-dynamics point of view. Nevertheless, it is useful to distinguish two fundamental kinds of convection zone:

1. A *convective envelope* extends from some point in the star's interior all the way to its surface. The density ratio between the bottom and the top of such a convection zone is typically huge (10^6 in the Sun), the range of relevant length scales is wide (from 10^2 km at the surface to 10^5 km at the bottom in the Sun) as is the range of relevant time scales (from minutes at the surface to a month at the bottom in the Sun). The flow can become supersonic

close to the surface, but its Mach number is very low at the bottom. Envelope convection is driven from the top. The strong radiative cooling in the photosphere creates low-entropy downdrafts, which rapidly sink back to the interior. The upflow occurs due exclusively to mass conservation and is almost isentropic.

2. A *convective core or shell* is typically much less stratified than a convective envelope. Consequently, the range of relevant length and time scales is rather narrow. The driving of convection is provided by energy overproduction, which is usually concentrated to the hottest layers at the bottom. All entropy fluctuations are several orders of magnitude smaller that those at the top of a convective envelope, because a convective core or shell is by definition fully embedded in the star. The convective flow is characterised by a low Mach number.

Current stellar-evolution models do not make this distinction. Convection is generally described by the mixing-length theory (MLT) of Böhm-Vitense (1958) or a variant thereof. The MLT is an order-of-magnitude estimate of the convective flux. It is based on a picture, in which a large "blob" of fluid is accelerated by its buoyancy in the local mean stratification. The blob disintegrates after having travelled a distance of l_m, the so-called mixing length. This distance is taken to be a constant multiple of the local pressure scale height, $l_\mathrm{m} = \alpha_\mathrm{MLT} H_p$, where α_MLT is a tuning parameter. The value of α_MLT is usually obtained by the calibration of the standard solar model to the current solar radius and luminosity,[1] which leads to $1.5 \lesssim \alpha_\mathrm{MLT} \lesssim 2.0$. There is no physical reason for α_MLT to be constant. Indeed, numerical simulations of stellar surface convection show a dependence of α_MLT on the local physical conditions in the convection zone (Trampedach et al. 2014; Magic et al. 2015).

One of the main deficiencies of the MLT is its being local, i.e. that all its predictions are based on local quantities only. Consequently, it predicts zero velocities outside the region of linear instability against convection. Convection, however, is non-local by definition. Several physical processes can extend the mixing effect of convective motions to the neighbouring stable stratification, see Sect. 1.3. There are two wide-spread methods of treating this so-called *overshooting* in stellar-evolution codes. In the first, a layer of thickness

$$l_\mathrm{ov} = \alpha_\mathrm{ov} H_p \tag{1.1}$$

beyond the convection zone is assumed to be isentropic and fully mixed with the convection zone. This parametrisation is motivated by the phenomenon of convective penetration, see Sect. 1.3. The other popular parametrisation represents

[1] The other free parameter in the calibration is the initial He abundance.

overshooting as a diffusive process with a diffusion coefficient

$$D(z) = D_0 \exp\left(\frac{-2z}{f_{\rm ov} H_p}\right), \tag{1.2}$$

where D_0 is the diffusion coefficient at the convective boundary (estimated using the MLT just inside the convection zone), z the distance from the convective boundary, and $f_{\rm ov}$ a tuning parameter. This formulation is motivated by the results of numerical simulations of Freytag et al. (1996). Both $\alpha_{\rm ov}$ and $f_{\rm ov}$ are typically calibrated so as to reproduce a pre-selected set of observations and assumed to be constant for all stars. Neither of the parametrisations takes into account that the physical mechanism responsible for overshooting may depend on the type of convection zone and on the properties of its boundary.

There is a whole class of models (e.g. Kuhfuss 1986; Xiong & Deng 2001; Kupka & Montgomery 2002; Marik & Petrovay 2002; Deng & Xiong 2008), which employ the Reynolds-averaged Navier-Stokes equations in an attempt to describe the non-locality of convection. The averaging procedure leads to a set of equations containing higher-order correlations in the turbulent field. These correlations have to be expressed in terms of known quantities by the so-called closure conditions, which involve further assumptions about the transport and mixing processes. Models of this kind usually predict substantial overshoot. This comes about because they describe the convection zone and the overshoot region in the same way, but the mixing effect of overturning convection is quite different from that of internal gravity waves.

1.2 Observational evidence of overshooting

Colour–magnitude diagrams of open clusters The shape of stellar isochrones depends on the amount and type of core overshooting included in the calculation. The turn-off region is influenced the most. Numerous studies of open-cluster stars have shown that a certain amount of overshooting is needed for the theoretical isochrones to fit the colour-magnitude diagrams of the clusters (Gim et al. 1998; Rosvick & Vandenberg 1998; Pietrinferni et al. 2004; VandenBerg & Stetson 2004). The best-fit models with overshooting, however, generally give a different age estimate for the cluster compared with the standard models without overshooting.

Double-lined eclipsing binaries If spectral lines of both components of an eclipsing binary are detected it is possible to derive the masses and radii of both stars. A subset of these binaries happen to be composed of a star just beyond

the terminal-age main sequence and a main-sequence companion. Such systems provide tight constraints on convective core overshooting, because the position of the evolved component in the $T_{\rm eff} - \log g$ plane (Kiel diagram) at a given age varies significantly with the amount of core overshooting whereas the position of its companion does not. Various authors typically derive $0.1 \lesssim \alpha_{\rm ov} \lesssim 0.5$, but it is unclear whether $\alpha_{\rm ov}$ depends on the stellar mass or not (Schroder et al. 1997; Ribas et al. 2000; Guinan et al. 2000; Claret 2003, 2007; Lacy et al. 2008; Tkachenko et al. 2014; Stancliffe et al. 2015).

Li and Be depletion The abundances of Li and, to a much lesser extent, also of Be in the photospheres of low-mass stars decrease continually during their main-sequence lives (Baumann et al. 2010; Meléndez et al. 2010; Delgado Mena et al. 2012; Monroe et al. 2013; Meléndez et al. 2014). The standard solar model (SSM) and the standard models of other low-mass stars cannot reproduce these observations, because the convection zones in them do not reach the Li- and Be-burning layers (see e.g. Schlattl & Weiss 1999). Thus, there must be a process that provides a slow mass exchange between these layers and the convection zone. Observations of open clusters suggest that there may actually be two distinct mixing processes at work, one that is responsible for the smooth, global decrease in the Li abundance towards low effective temperatures and another one that is responsible for the sudden drop in Li abundances in stars of $T_{\rm eff} \approx 6700\,{\rm K}$ that is known as the lithium dip (Boesgaard & Tripicco 1986; Balachandran 1991).

Helioseismology The large number of oscillation modes observed on the solar surface allows us to reconstruct the Sun's internal structure with unprecedented accuracy. There is a highly significant difference between the sound-speed profile derived by the helioseismic inversion techniques and the prediction of the standard solar model (SSM), which reaches a maximum just below the convection zone. This may have to do with the absence of any overshooting physics in the SSM. Another piece of evidence is encoded in the oscillatory signal in the mode frequencies that arises due to the rapid change in the stratification at the bottom of the convection zone. The observed signal is weaker than what the SSM predicts (Christensen-Dalsgaard et al. 2011), hence there must be a process that smooths the transition at the convection zone's bottom. This constraint also rules out simple models of overshooting, which have an abrupt transition to the radiative stratification at the bottom of the overshoot layer.

Asteroseismology The number of oscillation modes observed in massive stars is much smaller than that in the Sun, but it has recently become possible to measure the size of their convective cores in some cases. Aerts (2013) analysed a

sample of 68 OB pulsators. She derived the value of the overshooting parameter α_{ov} in six cases and placed an upper limit on α_{ov} in further five cases. The typical values are $0.1 \lesssim \alpha_{ov} \lesssim 0.25$ with individual uncertainties of 0.05, but there is one case with $\alpha_{ov} = 0.44 \pm 0.07$. The data does not indicate any dependence on the stellar mass, rotation period, or core hydrogen mass fraction, but the sample is small. Montalbán et al. (2013) recently proposed a method to constrain the extent of the convective core in the H- and He-burning phases in intermediate-mass stars using period spacing of the solar-like oscillations observed in red giants.

1.3 Physics of overshooting

Ballistic overshooting A convective element approaching the boundary of a convection zone carries a certain amount of momentum. Its inertia allows it to enter the stable stratification and to travel some distance in it before it is stopped by the buoyancy force. Convective flows in the stellar interior typically occur at low Mach numbers ($\approx 10^{-4}$), hence this kind of overshooting can influence at most the first few per cent of the pressure scale heigh beyond the boundary of the convection zone. This estimate, first made by Roxburgh (1965) and Saslaw & Schwarzschild (1965), may be somewhat optimistic since it ignores the deceleration of the convective element by the adverse pressure gradient inside the convection zone. It is this very gradient that causes the large-scale convective flows to turn over at the boundary and sink back to the convection zone. Moreover, there is often a jump in the mean molecular weight that makes the boundary extremely "stiff" and the overshooting distance becomes even shorter.

Shear instabilities The large-scale convective eddies' turning around at the boundary of the convection zone produces shear that may be prone to shear instabilities. Meakin & Arnett (2007) used numerical simulations to show that the mixing at the convective boundary is indeed dominated by shear instabilities and "wave breaking" phenomena. The thickness of the mixing layer decreases as the "stiffness" of the boundary increases. Nevertheless, the material from the outside of the convection zone can be entrained into it even if the boundary is so stiff that it is stable against the shear as shown by Woodward et al. (2015). Their simulations reveal that the entrainment occurs in a thin boundary-separation layer just inside the formal convection zone. This layer abounds in Kelvin-Helmhotz instabilities, which slowly "peel off" material from the outside of the convection zone. It is quite possible that this process is responsible for a slow entrainment of fuel-rich material into a convective core or shell, in which this fuel is burnt.

Convective penetration It was realised early on (Shaviv & Salpeter 1973; van Ballegooijen 1982; Zahn 1991) that the vigorous mixing in a thin layer beyond a convection zone's boundary may decrease the stability of the thermal stratification in that layer and it may allow the layer to grow on a longer time scale. This is most easily seen in the entropy representation, in which the convection zone is characterised by a single value of entropy and the adjacent stable stratification by an entropy gradient. Heat diffusion is inefficient for large-enough convective eddies, so they will overshoot adiabatically, mixing their entropy value into the entropy gradient outside the convection zone. This process flattens the gradient in the overshoot layer and makes further penetration easier. The growth of the layer is ultimately stopped by thermal diffusion. Simplified models (van Ballegooijen 1982; Schmitt et al. 1984; Rempel 2004) mostly predict the whole penetration layer to be a few tenths of the pressure scale height thick. It is terminated by a very thin thermal-adjustment layer, in which the temperature gradient decreases to the radiative value (van Ballegooijen 1982; Schmitt et al. 1984, but see Rempel 2004 for the opposite conclusion).

Internal gravity waves The perpetual rippling of the convective/stable interface radiates internal gravity waves into the stable medium. The rippling, however, predominantly occurs on the dynamical time scale of convection, which is typically much longer than the buoyancy time scale deep in the stable stratification. By the physics of internal waves (Lighthill 1978), their surfaces of constant phase as well as the direction of their propagation become almost horizontal. As soon as this happens, radiative diffusion becomes efficient at damping their amplitude. Damped waves, unlike the non-damped ones, can produce a mean flow (Lighthill 1978). A superposition of many damped waves with random amplitudes and phases will redistribute a passive scalar in a diffusive way (Schatzman 1996; Montalbán 1994). In addition to that, internal waves of finite amplitude produce shear flows that may lead to shear instabilities and boost the diffusion rate (Press 1981; Garcia Lopez & Spruit 1991). The exact structure and spectrum of the fluctuations that generate the waves is not well constrained and constitutes the main source of uncertainty in both kinds of mixing models.

Convective settling The envelope convection zone of a low-mass star is exposed to vacuum at the top. The rapid radiative cooling in the photosphere creates downflows of very low entropy, which rapidly sink back to the convection zone. On their way down, they merge, entrain mass and get heated by radiative diffusion on small scales. Both numerical simulations (Kerr 1996; Viallet et al. 2013; Woodward et al. 2015) and laboratory experiments (Heslot et al. 1987; Libchaber et al. 1990; Zocchi et al. 1990) show that such downflows can span the

whole convection zone even at extreme values of the Rayleigh number. If some of them retain a substantial fraction of their initial entropy contrast until they reach the convection zone's bottom, their negative buoyancy will make them sink further down irrespectively of their velocity. Each downflow will ultimately settle at the point where the entropy the surrounding stratification equals that of the downflow, i.e. at the point of neutral buoyancy. An upflow due to mass conservation will act as a "conveyor belt" bringing the mass of the settling layer back to the convection zone. The entropy the downflows start with in the photosphere is low enough for them to reach the Li-burning layer, hence convective settling has a potential to explain the Li depletion observed in low-mass stars. This idea was first suggested by Spruit (1997) and later elaborated by Andrássy & Spruit (2013) and Andrássy & Spruit (2015b), see Chaps. 2 and 3.

Differential heating Flows in the convection zone create large-scale, time-dependent temperature fluctuations at the convection zone's boundary. Thermal diffusion transports these fluctuations into the neighbouring stably-stratified medium, where they perturb the hydrostatic equilibrium and set off a flow. The random changes in the heating pattern due to convection will turn this flow into a diffusion process, which may provide mixing at some distance from the boundary. This idea was first proposed and elaborated by Andrássy & Spruit (2015a), see Chap. 4.

1.4 Main results of the thesis

- The process of "convective settling," so far neglected in stellar-evolution calculations, may contribute significantly to the mixing below the convective envelopes of low-mass stars.

- Convective settling cannot smooth the sound-speed profile at the bottom of the solar convection zone, at least not in the parametrisation employed in this work. Inclusion of He sedimentation in the model might change this result, though.

- A simple model of convective settling developed in this work can well approximate the trends in the depletion of both Li and Be that are observed in low-mass, main sequence stars.

- The process of "differential heating" is identified as an additional, potential contributor to the mixing close to the boundary of a stellar convection zone.

- A simplified differential-heating problem is formulated as a fundamental, non-linear flow problem of low-Péclet-number hydrodynamics.
- An analytical model of the stationary, high-Reynolds-number, differential-heating flow is derived. The model shows that the decrease in the flow amplitude with increasing distance from the differentially-heated boundary is faster than exponential. The flow essentially vanishes at a finite distance.
- The mixing layer due to the differential-heating flow is estimated to extend about 4% of the pressure scale height above the core of a $10\,M_\odot$ zero-age main sequence star.

2 Overshooting by convective settling

R. Andrássy, H. C. Spruit

A&A 559, A122 (2013)

Abstract: We study a process of slow mixing in stars with convective envelopes, which is driven by the settling of cool downward plumes below the base of the convection zone. If a small fraction (of order 10^{-7}) of the material cooled at the surface retains a significant entropy deficit while descending in plumes, it can reach the depth where lithium burning takes place. The model calculates the thermal response and mixing below the convection zone due to the settling process, assuming that the plumes arrive at the base of the convection zone with a broad range of entropy contrasts. We obtain a good fit to the observed lithium depletion in the Sun by assuming that the settling mass flux is distributed with respect to the entropy contrast as a power law with a slope around -2. We find convective settling to have a negligible influence on the stratification below the convection zone, although mixing induced by it could modify the gradient of helium concentration.

2.1 Introduction

In stellar evolution, the term 'overshooting' is used to denote any process that can extend the mixing effect of a convective flow beyond the region of linear instability of the stratification. A number of conceptually different processes have been considered, operating on a range of length scales and time scales.

The first overshooting mechanism one can think of is the extension of the convection zone as caused by the inertia and entropy contrast of the convective flows, which they developed before leaving the convection zone proper. Roxburgh (1965) and Saslaw & Schwarzschild (1965) both estimated the expected extent of overshooting due to this process using characteristic values of the speed and entropy contrast predicted by the mixing-length theory (MLT). They obtained a negligible penetration distance as a consequence of the steep entropy gradient in the radiative zone and low convective velocities connected to the efficiency of deep convection in stars.

This process can be called 'overshooting' in the literal (ballistic) sense, operating on a short time scale and over a negligible distance. It was recognised early on (e.g. Shaviv & Salpeter 1973) that the convective cells at the boundary with a stable zone could also have a slower but systematic effect that extends the region of convectively overturning motion. Such effects would be much more relevant on the longer time scales of stellar evolution. To distinguish it from the ballistic process, this kind of process can be called 'convective penetration' (Zahn 1991). Models have been developed by several authors. Van Ballegooijen (1982) studied the long-term response of the sub-adiabatic interior to a stationary convective flow at the base of the convection zone. The model predicted a mildly sub-adiabatic overshoot region terminated by a thin boundary layer. Related ideas were developed by Schmitt et al. (1984) and Rempel (2004) using models for the interaction of downward plumes with their environment.

For the discussion of overshooting/penetration it is useful to make a distinction between the case of a convective envelope like the Sun and Sun-like stars and the conditions in internal convective zones in stars. Owing to the very low gas density at the stellar surface, where the flows are driven, the flows in a convective envelope are far more 'plume-like' than in a convective core, where the density stratification is much less extreme. This has major consequences for the overshooting problem. Whereas in core convection the order of magnitude of velocities and temperature fluctuations can be plausibly estimated from a mixing-length formula, the downward plumes in a convective envelope have much stronger entropy contrasts than mixing-length estimates based on some average of the stratification. To the (as yet poorly known) extent to which these plumes survive mixing, entrainment and merging (see e.g. Viallet 2012), they will arrive at the base of the convection zone with a broad range of entropy contrasts. This mixture will settle in the stable layers below at a range of depths, which are determined by the distribution of entropy contrasts (see Nordlund & Stein 1995 for an example of a simulation showing the effect qualitatively). We call this form of gentle overshooting 'convective settling' (shortened to 'settling' hereinafter). The explicit inclusion of the cold plumes makes it a rather different contribution to overshooting compared with the processes mentioned above. On the long time scales relevant to stellar evolution, it has the potential of producing a weak mixing extending to deeper layers.

Numerical simulations are naturally restricted to limited time scales that cover the ballistic process much better than the slower processes of penetration and convective settling. Extrapolations have to be made to translate simulation results from a numerically accessible regime to astrophysically relevant conditions. Since different contributing processes act on different time scales and depths, such extrapolation would require disentangling them from a numerical simulation. This

is not a straightforward task. Early two-dimensional computations by Hurlburt et al. (1986) showed strong downward-directed plumes penetrating deep into the stable stratification and generating gravity waves there. Later studies (e.g. Hurlburt et al. 1994; Singh et al. 1995; Brummell et al. 2002; Rogers & Glatzmaier 2005) focussed on the dependence of penetration depth on the stiffness of the interface separating the stable and the unstable stratification. With 3D simulations of a convective envelope model, Nordlund & Stein (1995) showed qualitatively how the transition from convection to the stable interior is softened by the spread in entropy of the downward plumes, i.e. the 'settling' process above. Realistic simulations of a convective envelope are still quite outside the accessible domain. This is due to the daunting range of length and time scales needed to cover the driving at the stellar surface, the mixing and entrainment in the plumes, and their compression over the enormous density range between the surface and the base (cf. Rogers & Glatzmaier 2005; Rogers et al. 2006).

A further phenomenon that could lead to mixing at the bottom of such a convection zone is a meridional circulation in the radiative zone, driven by the latitudinal differential rotation of the convective envelope. This was studied by Spiegel & Zahn (1992), among others, who calculated how viscous stress causes the pattern of differential rotation to penetrate the stable stratification on a long time scale.

McIntyre (2007) points out that the problem is identical to the so-called 'gyroscopic pumping' process (Haynes et al. 1991), in which thermal diffusion plays a dominant role. Garaud & Acevedo Arreguin (2009) show how this process is capable of producing meridional flows penetrating deep into the radiative interior, provided that there is a source of stress in it.

Stochastic convective motions generate a whole spectrum of internal gravity waves propagating into the stable stratification. These can in principle also produce a modest amount of mixing at some distance away from a convection zone, see e.g. Press (1981); Garcia Lopez & Spruit (1991); Schatzman (1996).

An important clue about the deeper, slow mixing processes is the long-term lithium depletion observed in low-mass main-sequence stars (Herbig 1965; Boesgaard 1976; Pinsonneault 1994, and references therein). From the complex overshooting physics described above, we decided to isolate the settling process since it could be the most relevant for lithium depletion as suggested before in Spruit (1997). We present the physics of settling and a simplified model of it in Sec. 2.2.1, then we formulate the model mathematically in Sec. 2.2.2 and describe our approach to estimating the extent of lithium depletion in Sec. 2.2.3. We focus on the solar case, because we can also apply helioseismic constraints there (Christensen-Dalsgaard et al. 1996; Bahcall et al. 1998; Schlattl & Weiss 1999). Results of our study are summarised in Sec. 2.3 and discussed in Sec. 2.4.

2.2 Model

As discussed above, several processes are present at the same time at the boundary of a stellar convection zone and act on a range of depths and time scales. We focus on lithium depletion here, so we need to consider mixing processes working deep in the stable stratification ($T \gtrsim 2.5$ MK) on nuclear-burning time scales. We envisage a mechanism based on the idea that, apart from the MLT-like flows, there might be a tiny fraction of substantially colder material at the lower boundary of an envelope convection zone. It would be a (partially mixed) remnant of the photospheric downflows. Its low entropy would make it sink until it has reached neutral buoyancy. Because of the large entropy deficit, with which the plumes start, this can in principle happen rather deep in the stable stratification. One would expect, however, that the larger the entropy deficit, the smaller the fraction of the downflowing material involved. We expect the settling rate at lithium-burning depths to be so low that settling negligibly affects the stratification there, but still high enough to ensure significant mass exchange with the convection zone on very long time scales. On the other hand, the settling rate in the vicinity of the convection zone could be much greater, possibly leaving a footprint on the sound-speed profile.

Settling itself involves a vast range of time scales. The downflows sink towards the settling point on their dynamical time scale. Since only a small fraction of the plumes will survive with a significant entropy deficit, the filling factor of this fraction is small. This implies that the upflow enforced by mass conservation is very slow. This circulation disturbs the radiative equilibrium, which in a steady state is balanced by radiative diffusion. It is a certain time average of this equilibration process that represents the influence of settling on the stratification on the time scale of the star's lifetime.

What fraction of the photospheric downflows should keep their low entropy? To get an order-of-magnitude estimate from the observed value of the Sun's lithium depletion, assume that a fraction ϵ of the photospheric downward mass flux $\mathcal{F}_{\rm ph}$ sinks deep enough to reach lithium-burning conditions, burning its lithium content instantaneously there. To burn all the observable lithium, the whole mass of the convection zone $M_{\rm cz}$ (per unit area) must be replaced by the flux $\epsilon\mathcal{F}_{\rm ph}$. If this is to happen on the time scale τ, $\epsilon = M_{\rm cz}/(\mathcal{F}_{\rm ph}\tau)$. This way we obtain $\epsilon \approx 10^{-7}$ for solar values ($M_{\rm cz} \approx 7\times10^{8}\,\mathrm{g\,cm^{-2}}$, $\mathcal{F}_{\rm ph} \approx 3\times10^{-2}\,\mathrm{g\,cm^{-2}\,s^{-1}}$ and $\tau \approx 5\times10^{9}\,\mathrm{yr}$). Such a minute amount would hardly be traceable by current ab initio simulations.

2.2.1 Model physics and simplifications

We have constructed a simplified one-dimensional model of settling (Sec. 2.2.2) based on the ideas described above. It includes three assumptions. First, we ignore any exchange of heat or mass between the downflows and the upflow within the region bounded by the Schwarzschild boundary at the top and the point where the given downflow settles at the bottom (i.e. all downflows are adiabatic in our model). Secondly, we do not consider the kinetic energy of the downflows and its dissipation. Third, we regard the overshoot region as chemically homogeneous (with the exception of lithium). The last assumption has a significant influence on the sound-speed profile and is justified a posteriori in Sec. 2.3.

Since we are only interested in very long time scales, the model is formulated as stationary — we base it on the equality of a local cooling rate due to advection (i.e. settling) and a local heating rate due to a negative radiative flux divergence. The very nature of settling, in which cold material sinks downwards and pushes hot material upwards, causes the convective flux to be positive, which is unusual for overshooting models. Conservation of the total flux then makes the radiative flux decrease, which is achieved by a decrease in the temperature gradient ∇. This immediately tells us that a model with settling will be warmer than the one without it. The decrease in ∇ will also shift the formal boundary of convective instability in the settling model higher up into the stratification and slightly *reduce* the depth of the convection zone. The settling process adds mass from the convection zone to the stratification below, which responds by a slow upward flow, eventually returning mass to the convection zone. The location where the mass settles depends on its entropy since it is the depth at which it matches that of the stratification. The model calculates how settling changes the steady state stratification under the combined effects of thermal diffusion and the assumed settling mass flux.

The model thus needs a description for the flux of settling mass as a function of its entropy. Since current numerical simulations still cannot reach the parameter values needed to get this distribution from first principles, it needs to be parametrised. We describe it with a power law. We define the mass flux (either upflow or downflow) in terms of the quantity $\eta\rho v$, where η is the filling factor (the relative geometrical area covered by the flow, $0 \leq \eta \leq 1$), ρ is the density, and v the velocity. We avoid any explicit usage of the unknown values of v and η this way. (The model only requires $\eta \ll 1$ for the downflow, see above.) The range of entropies in the mass flux distribution is bounded between the entropy of a typical downflow in the photosphere and the MLT estimate of the entropy contrast in the lower part of the convection zone.

2.2.2 Mathematical formulation

To describe the settling process mathematically, we model the settling region as a plane parallel layer (instead of a spherical shell) of ideal gas in a homogeneous gravitational field. These assumptions are not essential but make the mathematics more transparent and are not likely to change the outcomes substantially.

We describe the thermodynamic state by dimensionless variables $\hat{p}$, $\hat{\rho}$, $\hat{T}$, and $\hat{s}$, which stand for the pressure, density, temperature, and specific entropy, respectively:

$$\hat{p} = \frac{p}{p_0}, \tag{2.1}$$

$$\hat{\rho} = \frac{\rho}{\rho_0}, \tag{2.2}$$

$$\hat{T} = \frac{T}{T_0}, \tag{2.3}$$

$$\hat{s} = \frac{s - s_0}{\mathbb{R}}, \tag{2.4}$$

where p, ρ, T, and s are the physical state variables, the index zero corresponds to a reference point, and $\mathbb{R}$ is the gas constant. We use $\mathbb{R} \equiv k_\mathrm{B}/(\mu\, m_\mathrm{p})$, so that it includes the mean molecular weight $\mu = \text{const.}$ (see Sec. 2.2.1). The reference point, which must not be influenced by settling, can be put anywhere beneath the settling layer. The thermodynamic state we refer to is a horizontal average set by the upflow state, because the downflow filling factor is assumed to be small. We describe the thermodynamics of settling in terms of this average state.

We introduce a dimensionless independent variable

$$\psi = -\ln(p/p_0), \tag{2.5}$$

which can be related to the geometrical height z by using the condition of hydrostatic equilibrium,

$$dp = -\rho\, g\, dz, \tag{2.6}$$

where $g = \text{const.}$ is the gravitational acceleration, and we set $z = 0$ at the reference point. If we differentiate Eq. 2.5, combine it with Eq. 2.6, and use the ideal gas law $p = \rho\, \mathbb{R}\, T$, we obtain a recipe for converting derivatives to the dimensionless form,

$$\frac{\mathrm{d}}{\mathrm{d}\psi} = \frac{\mathbb{R}T}{g}\frac{\mathrm{d}}{\mathrm{dz}}. \tag{2.7}$$

The local pressure scale height is $H = \mathbb{R}T/g$. Elementary thermodynamics provide

us with the following useful relations

$$\hat{\rho} = \hat{p}^{1-\nabla_{\rm ad}} \exp\left(-\nabla_{\rm ad}\hat{s}\right), \tag{2.8}$$

$$\hat{T} = \hat{p}^{\nabla_{\rm ad}} \exp\left(\nabla_{\rm ad}\hat{s}\right), \tag{2.9}$$

$$\frac{\mathrm{d}\hat{s}}{\mathrm{d}\psi} = 1 - \frac{\nabla}{\nabla_{\rm ad}}, \tag{2.10}$$

where $\nabla_{\rm ad}$ is the usual adiabatic temperature gradient.

Let $\mathrm{d}\dot{m}(\hat{s})$ be the differential settling mass flux per unit of entropy $\hat{s}$ at the base of the convection zone. Its functional form is assumed to be a power law,

$$\mathrm{d}\dot{m} = -\dot{M} f(\hat{s})\,\mathrm{d}\hat{s}, \tag{2.11}$$

$$f(\hat{s}) = N \left(\frac{\hat{s}_{\rm sb} - \hat{s}}{(\delta\hat{s})_{\rm min}}\right)^{-\beta}, \tag{2.12}$$

where $\dot{M} > 0$ is the total settling mass flux, $f(\hat{s})$ is a distribution function, N a normalisation factor, $\hat{s}_{\rm sb}$ the entropy of the stratification at the Schwarzschild boundary (where $\nabla = \nabla_{\rm ad}$ and $\mathrm{d}\hat{s}/\mathrm{d}\psi = 0$, see Eq. 2.10), $(\delta\hat{s})_{\rm min} > 0$ is the lowest entropy contrast of a downflow with respect to $\hat{s}_{\rm sb}$, and $\beta > 0$ describes the steepness of the distribution. We also define a maximum entropy contrast $(\delta\hat{s})_{\rm max}$, which is determined by the entropy of the downflows at the surface (see Sec. 2.2.1), and set $f(\hat{s}) = 0$ for $\hat{s} < \hat{s}_{\rm sb} - (\delta\hat{s})_{\rm max}$ and $\hat{s} > \hat{s}_{\rm sb} - (\delta\hat{s})_{\rm min}$. We require that the area under $f(\hat{s})$ be unity, so that (for $\beta \neq 1$)

$$N = \frac{\beta - 1}{(\delta\hat{s})_{\rm min}} \left[1 - \left(\frac{(\delta\hat{s})_{\rm max}}{(\delta\hat{s})_{\rm min}}\right)^{-(\beta-1)}\right]^{-1}. \tag{2.13}$$

We define a cumulative distribution function

$$F(\hat{s}) = 1 - \int_{\hat{s}}^{\infty} f(\hat{s}')\,\mathrm{d}\hat{s}', \tag{2.14}$$

which describes the relative amount of downflows that settle *below* the level where the entropy in the stratification equals $\hat{s}$. All downflows of entropy $\hat{s}' > \hat{s}$ have already reached neutral buoyancy and settled higher up ($\mathrm{d}\hat{s}/\mathrm{d}z > 0$ for $\nabla < \nabla_{\rm ad}$, see Eqs. 2.7 and 2.10). Therefore $F(\hat{s}) = 0$ for $\hat{s} \leq \hat{s}_{\rm sb} - (\delta\hat{s})_{\rm max}$, $0 < F(\hat{s}) < 1$ for $\hat{s}_{\rm sb} - (\delta\hat{s})_{\rm max} < \hat{s} < \hat{s}_{\rm sb} - (\delta\hat{s})_{\rm min}$, and $F(\hat{s}) = 1$ for $\hat{s} \geq \hat{s}_{\rm sb} - (\delta\hat{s})_{\rm min}$.

We write the energy equation in the upflow in terms of entropy, in the Lagrangian form

$$\rho T \frac{\mathrm{D}s}{\mathrm{D}t} = -\frac{\mathrm{d}\mathcal{F}_{\rm rad}}{\mathrm{d}z}, \tag{2.15}$$

where t is the time, $\mathcal{F}_{\rm rad}$ the radiative flux, $\mathrm{D}/\mathrm{D}t = \partial/\partial t + v\,\mathrm{d}/\mathrm{d}z$ is the Lagrangian time derivative, and v the upflow velocity.

Global mass conservation requires the total amount of mass being transported downward to be equal to the total amount of mass being transported upward through any surface $z = \text{const}$. With our approximation that the filling factor of the upflow is close to unity, the upward mass flux $\dot{M}\,F(\hat{s})$ is given by the product of the upflow density and velocity,

$$\dot{M}\,F(\hat{s}) = \rho\,v. \tag{2.16}$$

This allows us to formulate the model without the knowledge of filling factors or using a momentum equation.

The stationary nature of our model eliminates the $\partial/\partial t$ term in Eq. 2.15, and using Eq. 2.16, we can write the upflow energy balance in the form

$$\dot{M}\,F(\hat{s})\,T\frac{\mathrm{d}s}{\mathrm{d}z} = -\frac{\mathrm{d}\mathcal{F}_{\rm rad}}{\mathrm{d}z}. \tag{2.17}$$

One can obtain a dimensionless form of this equation by using Eqs. 2.3, 2.4, and 2.7 and introducing a dimensionless radiative flux $\hat{\mathcal{F}}_{\rm rad} = \mathcal{F}_{\rm rad}/\mathcal{F}_{\rm tot}$, where $\mathcal{F}_{\rm tot}$ is the total flux (being equal to the radiative one at the reference point). The state variables can be related to the entropy by Eqs. 2.8 and 2.9. The diffusive approximation of the radiative flux is

$$\mathcal{F}_{\rm rad} = \frac{16}{3}\frac{g\,\sigma T^4}{p\,\kappa}\nabla. \tag{2.18}$$

With Eq. 2.10, Eq. 2.17 then yields

$$C\,\hat{T}(\psi,\hat{s})\frac{\mathrm{d}\hat{s}}{\mathrm{d}\psi}F(\hat{s}) = -\frac{\mathrm{d}\hat{\mathcal{F}}_{\rm rad}(\psi,\hat{s},\mathrm{d}\hat{s}/\mathrm{d}\psi)}{\mathrm{d}\psi}, \tag{2.19}$$

where

$$C = \frac{\dot{M}\,\mathbb{R}T_0}{\mathcal{F}_{\rm tot}} \tag{2.20}$$

is a ratio of a characteristic convective flux to the star's net energy flux $\mathcal{F}_{\rm tot}$ at the base of the convection zone. It is a dimensionless measure of the settling mass flux $\dot{M}$. The explicit form of the derivative $\mathrm{d}\hat{\mathcal{F}}_{\rm rad}/\mathrm{d}\psi$ on the right-hand side of

Eq. 2.19 is

$$\begin{aligned}\frac{\mathrm{d}\hat{\mathcal{F}}_{\mathrm{rad}}}{\mathrm{d}\psi} =& \frac{\nabla_{\mathrm{ad}}}{\nabla_0}\hat{\kappa}^{-1}\exp\left[(1-4\nabla_{\mathrm{ad}})\psi + 4\nabla_{\mathrm{ad}}\hat{s}\right]\times \\ & \left[1-4\nabla_{\mathrm{ad}} - \frac{\partial \ln\hat{\kappa}}{\partial\psi} + \left(8\nabla_{\mathrm{ad}} - 1 + \frac{\partial\ln\hat{\kappa}}{\partial\psi} - \frac{\partial\ln\hat{\kappa}}{\partial\hat{s}}\right)\frac{\mathrm{d}\hat{s}}{\mathrm{d}\psi} - \right. \\ & \left.\left(4\nabla_{\mathrm{ad}} - \frac{\partial\ln\hat{\kappa}}{\partial\hat{s}}\right)\left(\frac{\mathrm{d}\hat{s}}{\mathrm{d}\psi}\right)^2 - \frac{\mathrm{d}^2\hat{s}}{\mathrm{d}\psi^2}\right], \end{aligned} \tag{2.21}$$

where $\hat{\kappa} = \kappa/\kappa_0$ is the dimensionless opacity function and κ_0 the opacity value at the reference point.

Equation 2.19 governs the whole settling process in our model. It is a non-linear, second-order ordinary differential equation for the entropy profile $\hat{s}(\psi)$. The opacity function $\hat{\kappa}(\psi, \hat{s})$ in Eq. 2.21 could by specified by standard opacity tables, but that is not necessary in such a simplified model. In our sample calculations presented in Sec. 2.3, we used the opacity law

$$\ln\hat{\kappa} = \alpha\,\psi, \tag{2.22}$$

where the exponent α is a fitting parameter. In this simplification the opacity depends only on the pressure (via ψ) and not on the full thermodynamic state. This prescription is sufficient for our purposes if it reasonably fits the opacity profile of the stratification without settling, and if the state change due to settling is small.

Equation 2.19 implicitly contains an a priori unknown value of the entropy at the Schwarzschild boundary, $\hat{s}_{\mathrm{sb}}$, as an input parameter of the distribution function $F(\hat{s})$, see Eqs. 2.12 and 2.14. Therefore any solution procedure must involve iterations. One could pick an initial guess $\hat{s}^0_{\mathrm{sb}}$ and integrate Eq. 2.19 from the reference point (where $\psi = 0$, see Eq. 2.5) upwards. The initial conditions would be $\hat{s} = 0$, $\mathrm{d}\hat{s}/\mathrm{d}\psi = 1 - \nabla_0/\nabla_{\mathrm{ad}}$ (see Eqs. 2.4 and 2.10), where ∇_0 is the known temperature gradient at the reference point. The integration would then be stopped at the point ψ^0_{sb}, where the entropy reaches a maximum (i.e. the Schwarzschild boundary). The solution value at this point, $\hat{s}^0\left(\psi^0_{\mathrm{sb}}\right) \neq \hat{s}^0_{\mathrm{sb}}$ (in general), could be used as a new estimate $\hat{s}^1_{\mathrm{sb}}$, and the whole process could be repeated until convergence. In reality, the steep profile of $F(\hat{s})$ renders this method highly unstable.

Our sample calculations shown in Sec. 2.3 were computed by modifying this method. We started the i-th iteration by integrating Eq. 2.19 from an estimated position of the Schwarzschild boundary $\left(\psi^i_{\mathrm{sb}},\ s^i_{\mathrm{sb}}\right)$ downwards to the reference point. The initial conditions were $\hat{s} = s^i_{\mathrm{sb}}$, $\mathrm{d}\hat{s}/\mathrm{d}\psi = 0$ (see Eq. 2.10). We stopped

the integration at ψ_0^i such that $\hat{s}^i\left(\psi_0^i\right) = 0$. This, in general, leads to $\psi_0^i \neq 0$ and $\mathrm{d}\hat{s}/\mathrm{d}\psi\left|\left(\psi = \psi_0^i\right) \neq 1 - \nabla_0/\nabla_{\mathrm{ad}}\right.$; i.e., the solution curve misses the reference point. Therefore we make a new estimate $\left(\psi_{\mathrm{sb}}^{i+1},\ s_{\mathrm{sb}}^{i+1}\right)$ and repeat the procedure until the solution passes close enough to the reference point, and the entropy gradient gets close enough to $1 - \nabla_0/\nabla_{\mathrm{ad}}$ there. This method converges smoothly to the desired solution.

2.2.3 Lithium burning

The extent of lithium depletion in the convection zone is given by two time scales, which are both strong functions of depth. Taking a horizontal layer of thickness $\mathrm{d}z$, we introduce a 'recycling' time scale τ_{r} by defining the mass exchange rate (per unit area) in this layer as $\rho\,\mathrm{d}z/\tau_{\mathrm{r}}$. The height interval $\mathrm{d}z$ corresponds to an entropy interval $\mathrm{d}\hat{s}$ in the stratification, which implies that the mass exchange rate due to settling is $\dot{M} f(\hat{s})\mathrm{d}\hat{s}$ (see Eq. 2.11). Equating the last two expressions, we obtain

$$1/\tau_{\mathrm{r}} = \frac{\dot{M}}{\rho} f(\hat{s}) \frac{\mathrm{d}\hat{s}}{\mathrm{d}z} = \frac{g}{p} \dot{M} f(\hat{s}) \frac{\mathrm{d}\hat{s}}{\mathrm{d}\psi}, \tag{2.23}$$

where we have also used Eq. 2.7 and the equation of state.

The second time scale describes the speed of the burning itself. Lithium is burned by the reaction $^7\mathrm{Li}(\mathrm{p}, \alpha)\alpha$, which causes a decrease in its abundance $A \equiv N_{\mathrm{Li}}/N_{\mathrm{H}}$ on the time scale $\tau_{\mathrm{b}} \equiv -(\mathrm{d}\ln A/\mathrm{d}t)^{-1}$. With the reaction's astrophysical S-factor $S_{\mathrm{b}}(0) = 55\,\mathrm{keV\,barns}$ from Lattuada et al. (2001), the burning time scale is (see e.g. Hansen & Kawaler 1994)

$$\tau_{\mathrm{b}} = \left[9.02 \times 10^6\, X\rho\, \xi^2 \exp(-\xi)\right]^{-1}\ \mathrm{yr}, \tag{2.24}$$

where X is the hydrogen mass fraction, the density ρ is in $\mathrm{g\,cm^{-3}}$, and $\xi = 84.5\, T_6^{-1/3}$ with the temperature T_6 in MK.

We discretise the settling layer into a grid of n sub-layers ordered by height. The bottom of the i-th sub-layer is located at a height of z_i, $i = 1, 2, \ldots, n$. The resolution of the grid is chosen by setting a maximum to the relative changes in $F[\hat{s}(z)]$, $\tau_{\mathrm{r}}(z)$ and $\tau_{\mathrm{b}}(z)$ between grid points. We put the first grid point to the maximal depth settling can reach, i.e. $z_1 = \max\{z\colon F[\hat{s}(z)] = 0\}$. The topmost point represents the convection zone itself including a well-mixed upper part of the overshoot region, where the burning rate is negligible. Lithium is assumed to be a trace element.

We model the burning process by the set of equations

$$\frac{\mathrm{d}N_i}{\mathrm{d}t} = R_{\mathrm{b},i} + R_{\mathrm{s},i} + R_{\mathrm{a},i}, \tag{2.25}$$

where N_i is the number density of lithium atoms per unit area in the i-th sub-layer, t is the time, $R_{\mathrm{b},i}$ a burning rate, $R_{\mathrm{s},i}$ a settling rate of 'fresh' lithium atoms from the convection zone, and $R_{\mathrm{a},i}$ a rate of lithium transport by advection. Nuclear burning is an exponential decay process, hence

$$R_{\mathrm{b},i} = -\frac{N_i}{\tau_{\mathrm{b},i}}, \tag{2.26}$$

where $\tau_{\mathrm{b},i}$ can be defined e.g. as $\tau_{\mathrm{b},i} = [\tau_{\mathrm{b}}(z_i) + \tau_{\mathrm{b}}(z_{i+1})]/2$ with a special case $\tau_{\mathrm{b},n} = 0$ (see above). We introduce the mass settling rate $\dot{m}_i$ in the i-th sub-layer (cf. Eq. 2.14),

$$\dot{m}_i = \dot{M}[F(z_{i+1}) - F(z_i)] \tag{2.27}$$

and set $\dot{m}_n = 0$. We model settling as a process that extracts mass from the convection zone and deposits it over a range of depths without any mixing in between (see Sec. 2.2.1). Therefore the whole lithium content of $\dot{m}_i$ gets into the i-th sub-layer, and the lithium settling rate there is

$$R_{\mathrm{s},i} = A_n \frac{X\dot{m}_i}{m_{\mathrm{p}}}, \tag{2.28}$$

where the hydrogen mass fraction X and proton mass m_{p} are used to obtain the settling rate of hydrogen atoms, $X\dot{m}_i/m_{\mathrm{p}}$. The lithium abundance in the convection zone, A_n, then converts the hydrogen settling rate to the lithium settling rate. The advection part of Eq. 2.25 refers to the transport of lithium by the upflow, which causes a mass flux of $\sigma_i = \dot{M}F(z_i)$ through the bottom of the i-th sub-layer. This corresponds to a flux of lithium atoms of $A_{i-1}X\sigma_i/m_{\mathrm{p}}$ (coming from the $(i-1)$-st sub-layer). The rate of advective transport $R_{a,i}$ is then the difference between the inflow into and the outflow from the i-th sub-layer,

$$R_{\mathrm{a},i} = A_{i-1}\frac{X\sigma_i}{m_{\mathrm{p}}} - A_i\frac{X\sigma_{i+1}}{m_{\mathrm{p}}}, \tag{2.29}$$

which is negative.

Equation 2.25 can be expressed in terms of abundances in the following way. First, insert Eqs. 2.26, 2.28 and 2.29 into Eq. 2.25 and divide each resulting equation by the corresponding hydrogen number density $N_{\mathrm{H},i}$. Second, notice that $\sigma_i = \sum_{k=1}^{i-1}\dot{m}_k$ and define a discrete version of τ_{r} as $\tau_{\mathrm{r},i} = M_i/\dot{m}_i$, where the mass of the i-th sub-layer (per unit area) is $M_i = N_{\mathrm{H},i}\,m_{\mathrm{p}}/X$. Third, rearrange terms to obtain

$$\frac{\mathrm{d}\ln A_i}{\mathrm{d}t} = -\frac{1}{\tau_{\mathrm{b},i}} + \frac{A_n/A_i - 1}{\tau_{\mathrm{r},i}} - \left(1 - \frac{A_{i-1}}{A_i}\right)\sum_{k=1}^{i-1}\frac{M_k}{M_i}\frac{1}{\tau_{\mathrm{r},k}}. \tag{2.30}$$

Model id.	$\hat{\mathcal{F}}_{\mathrm{conv,\,sb}}$	β	$\dot{M}$ [g cm^{-2} s^{-1}]	$\max\left(\delta c_{\mathrm{s}}^2/c_{\mathrm{s}}^2\right)$	$\log_{10}[A_{\mathrm{cz}}(t_\odot)]$
A1.5	10^{-2}	1.5	4.5×10^{-3}	8.7×10^{-3}	<-10
A2.0	10^{-2}	2.0	4.4×10^{-1}	2.2×10^{-3}	-4.8
A2.5	10^{-2}	2.5	2.2×10^{0}	4.0×10^{-5}	-0.091
B1.5	10^{-1}	1.5	4.9×10^{-2}	8.7×10^{-2}	<-10
B2.0	10^{-1}	2.0	4.5×10^{0}	2.2×10^{-2}	<-10
B2.5	10^{-1}	2.5	2.2×10^{1}	3.7×10^{-4}	-0.72

Table 2.1: Properties of the six settling models.

The physical effect of each of the three terms on the right-hand side of Eq. 2.30 can now be seen easily. The first one exponentially destroys lithium on the local burning time scale $\tau_{\mathrm{b},i}$. The second term strives to equalise the local lithium abundance A_i with that of the convection zone. This, as a direct effect of settling, happens on the local recycling time scale $\tau_{\mathrm{r},i}$. The last term describes how A_i tends to approach A_{i-1}, i.e. the slow rising of the lithium stratification due to the upflow induced by settling. Its strength depends on the total 'speed' ($\sim 1/\tau_{\mathrm{r}}$) of settling beneath the i-th sub-layer.

2.3 Results

Numerical simulations of the solar photosphere show that the entropy contrast between the upflow and a typical downflow is $\delta s = 1.8{\times}10^8$ erg K^{-1} g^{-1} (see Fig. 29 in Stein & Nordlund 1998). We set the maximal entropy contrast in the mass flux distribution to this value, which corresponds to $(\delta\hat{s})_{\max} = 1.3$ in our dimensionless units. We use $(\delta\hat{s})_{\min} = 1.0\times10^{-6}$ for the minimal entropy contrast in the mass flux distribution. We put the reference point to $r = 0.50R_\odot$ in the standard solar model (SSM), which is slightly deeper than the bottom of the settling layer for our choice of $(\delta\hat{s})_{\max}$. The gravitational acceleration is set such that the mass of the region with $T > 2.5\times10^6$ K in our model (with $\dot{M} = 0$) is close to the corresponding value from the SSM.[1] We also adjust the opacity parameter α such that the entropy difference between the Schwarzschild boundary and the reference point matches the value from the SSM (again with $\dot{M} = 0$).

We first show the typical behaviour of settling on a set of six models (see Table 2.1), roughly sampling the corner of the parameter space that is likely to be relevant. We use three different values of β and, instead of setting $\dot{M}$ to any particular value, we adjust it iteratively in order to reach certain values of the

[1] This of some importance for the lithium-burning calculation. Equation 2.23 shows that $\tau_{\mathrm{r}} \propto 1/g$.

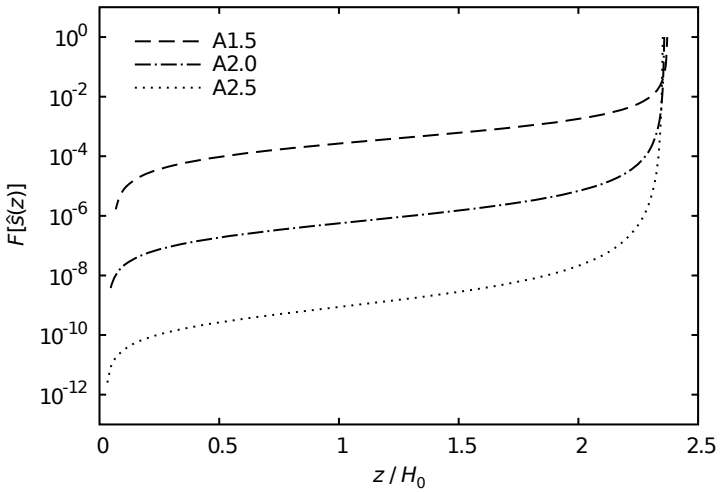


Figure 2.1: The cumulative mass flux distributions in the A-type settling models plotted as functions of height. $F\,[\hat{s}(z)]$ is the fraction of the mass flux $\dot{M}$ that settles below the height level z. The height z is counted upwards from a point beneath the settling layer.

convective flux at the Schwarzschild boundary, $\hat{\mathcal{F}}_{\mathrm{conv,\,sb}} \equiv 1 - \hat{\mathcal{F}}_{\mathrm{rad,\,sb}}$ (see Table 2.1). This is motivated by the fact that our mass flux distributions (see Fig. 2.1) are to represent a low-entropy tail *appended* to a distribution of MLT-like flows, which do not appear in our model. Since the $F(\hat{s})$ distributions have a sharp peak close to the entropy values predicted by the MLT, one can expect the convective flux induced by such set of downflows to be a non-negligible (although not precisely known) fraction of the total flux. Therefore we use $\hat{\mathcal{F}}_{\mathrm{conv,\,sb}} = 0.01$ and $\hat{\mathcal{F}}_{\mathrm{conv,\,sb}} = 0.1$ in A-type and B-type models, respectively (see Table 2.1). The cumulative mass flux distributions (Fig. 2.1) show that in all cases only a tiny fraction of the input mass flux reaches substantial depths. We do not show the distributions of the B-type models, because their profiles are almost the same as the ones shown, but shifted along the z axis.

We compare the properties of our settling models with a reference one having $\dot{M} = 0$ (i.e. without settling). The comparison is made at the same geometrical height z, normalised to the pressure scale height at the reference point, $H_0 = \mathbb{R}T_0/g$. The convective part of the reference model extends a bit deeper than in the settling models; i.e., the settling process *reduces* the depth of the convection zone (cf. discussion in Sec. 2.4). We set $\nabla = \nabla_{\mathrm{ad}} = 0.4$ in the convectively unstable part of the reference model to make the comparison of thermodynamic quantities

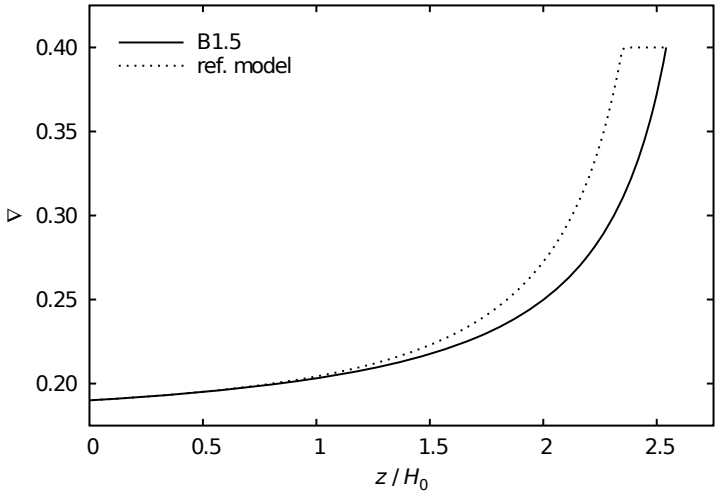


Figure 2.2: The temperature gradient in the model B1.5 (solid line) compared with the reference model (dotted line). In the convectively unstable part of the reference model $\nabla = \nabla_{\rm ad} = 0.4$ (see Sec. 2.3).

possible.

Figure 2.2 compares the temperature gradient in the model B1.5 (the strongest settling) to the reference one. It clearly shows that the temperature gradient in the settling models has to decrease in order to reduce the radiative flux, as expected in Sec. 2.2.1. We can see that models with settling indeed reach convective instability higher up in the stratification. It is also evident from Fig. 2.2 that settling preserves the discontinuity in the slope of ∇ (hence in the second derivative of the sound speed) at the Schwarzschild boundary, which has consequences for helioseismology.

Figure 2.3 shows the relative *increase* in the squared sound speed due to settling,

$$\frac{\delta c_{\rm s}^2}{c_{\rm s}^2} = \frac{c_{\rm s}^2 - c_{\rm s,r}^2}{c_{\rm s,r}^2}, \tag{2.31}$$

where $c_{\rm s}$ is the sound speed in the settling model and $c_{\rm s,r}$ is the sound speed in the reference one. The maximum values of $\delta c_{\rm s}^2/c_{\rm s}^2$ reached by our models are also listed in Table 2.1.

Plots of the recycling time scale in Fig. 2.4 shows that a considerable fraction of the settling layer gets mixed with the convection zone on the time scale of the Sun's lifetime. Such mixing could alter the chemical gradients caused by slow

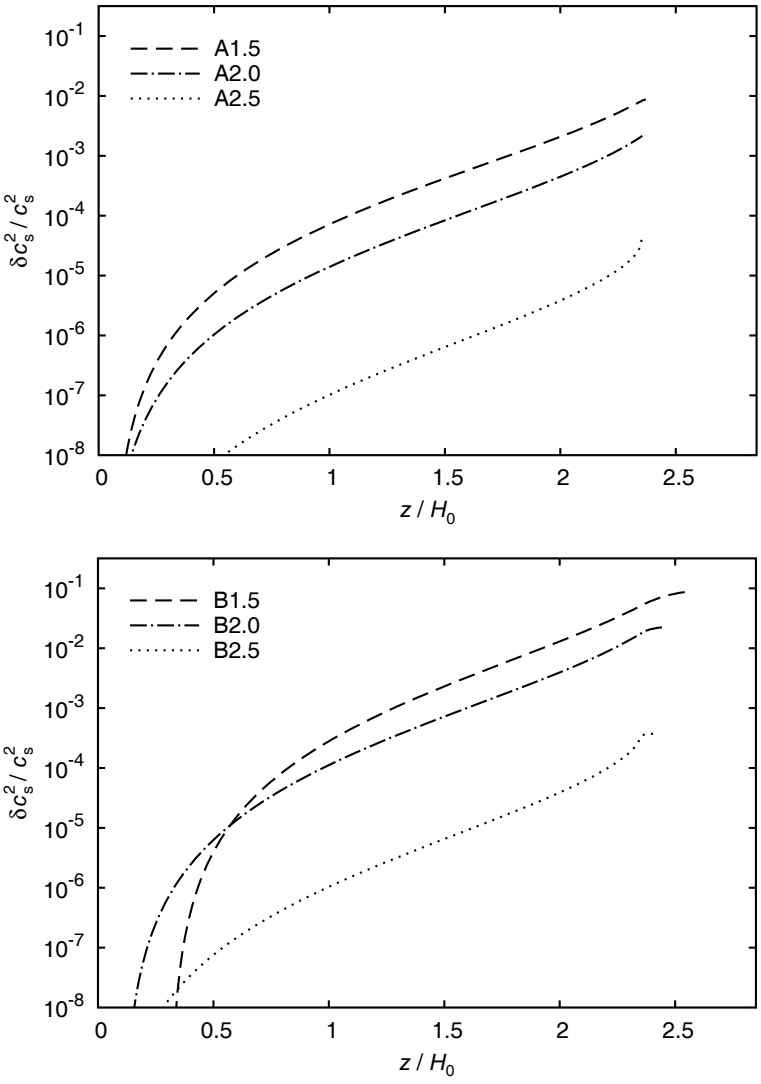


Figure 2.3: The differences in the squared sound speed between the settling models and the reference model. The right end of each curve marks the position of the Schwarzschild boundary.

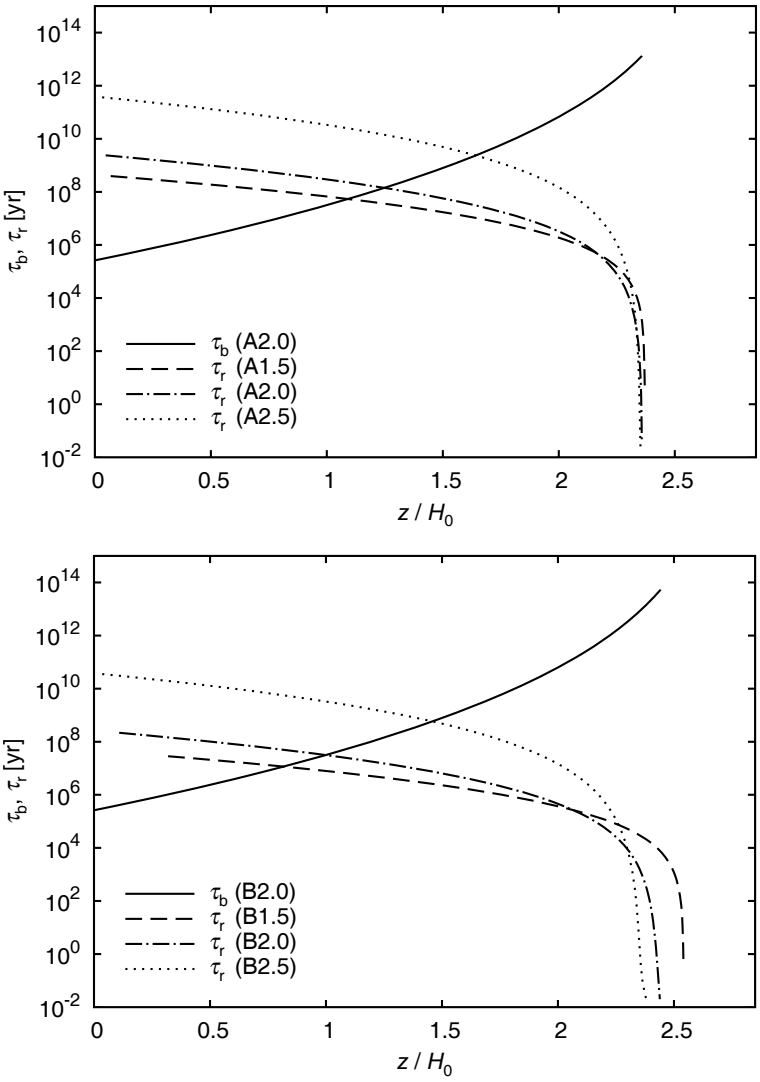


Figure 2.4: The lithium-burning time scale τ_b and the recycling time scale τ_r in the settling models. τ_b is only plotted for the models with $\beta = 2.0$, because it is very similar in the other ones. The right end of each curve marks the position of the Schwarzschild boundary.

element diffusion and change the sound speed profile with respect to the SSM.

The conditions for lithium burning can be qualitatively judged using Fig. 2.4, which shows the recycling time scale τ_r (Eq. 2.23) and the lithium-burning time scale τ_b (Eq. 2.24) as functions of height. Since they both change by many orders of magnitude, we estimate the overall extent of lithium depletion in the convection zone by integrating Eq. 2.30 with the initial condition $A_i = 1, i = 1, 2, \ldots, n$, so that our calculations show the *relative* change in the lithium abundance with respect to the initial one. The integration is stopped at $t = t_\odot = 4.5 \times 10^9$ yr. The strongly varying extent of lithium depletion in the convection zone $A_{cz} \equiv A_n$, listed in Table 2.1, shows the extreme sensitivity of the lithium depletion rate to τ_r, τ_b and thus to β. We only present models for one value of $(\delta\hat{s})_{max}$, but the consequences of changing this parameter can be judged using Fig. 2.4.

First, consider a model with $\tau_r \gg t_\odot$ at the bottom of the settling layer. Under this condition, the convection zone does not „feel" the bottom of the settling layer, because the material exchange between them over the Sun's lifetime is negligible. The lithium depletion cannot depend much on $(\delta\hat{s})_{max}$ in this case. If, on the other hand, we take a model with $\tau_r \ll t_\odot$ at the bottom of the settling layer, the exchange of mass with the convection zone is efficient. The rate of lithium depletion in the convection zone is sensitive to the maximal depletion rate in the settling layer, i.e. the value at its bottom. Such a model must therefore be sensitive to $(\delta\hat{s})_{max}$ since this parameter determines the depth of the settling layer.

Observations tell us that the Sun has depleted 2.21 ± 0.11 dex of its initial lithium content (Asplund et al. 2009). We have calibrated a set of models (not shown in Table 2.1) by adjusting $\dot{M}$ at any given value of β until they all predicted the observed lithium depletion. These models use the same values of $(\delta\hat{s})_{min}$ and $(\delta\hat{s})_{max}$ as the ones listed in Table 2.1. The derived values of $\dot{M}$ and $\hat{\mathcal{F}}_{conv,\,sb}$ are plotted in Fig. 2.5 as functions of β. The recycling time scale at the bottom of the settling layer in this set of models is $\sim 6.9 \times 10^9$ yr with only 16% variation over the range of β plotted in Fig. 2.5. This is more than the age of the Sun, so that the burning-calibrated models are only mildly sensitive to the assumed value of $(\delta\hat{s})_{max}$. The peak values of $\delta c_s^2/c_s^2$ in these models range from 6×10^{-4} at $\beta = 1.5$ to 2×10^{-3} at $\beta = 2.5$.

2.4 Summary and discussion

We have shown that there are several processes contributing to overshooting under a convective stellar envelope, which cover a wide range of depth and time scales. The 'ballistic' form of overshooting acts on the dynamical time scale in a shallow boundary layer. The process of convective penetration needs more

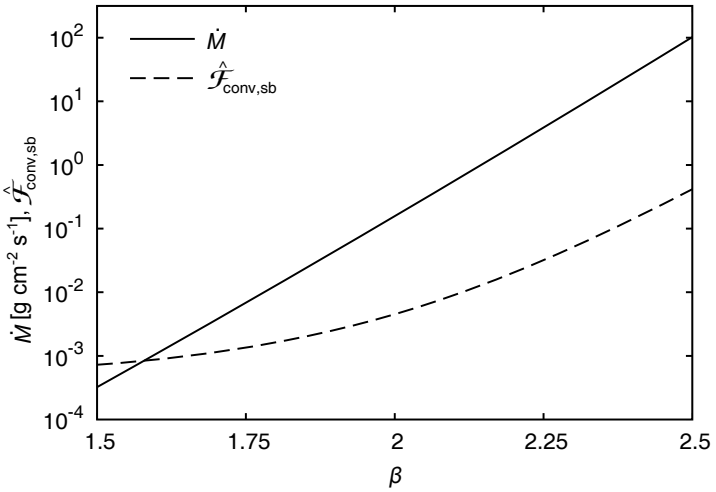


Figure 2.5: The settling mass flux $\dot{M}$ and the convective flux at the Schwarzschild boundary $\hat{\mathcal{F}}_{\rm conv,\,sb}$ plotted as functions of β in a set of models calibrated to produce the observed lithium depletion in the Sun.

time to spread but reaches deeper. Convective settling, potentially the deepest-reaching overshooting process, only plays a role on the time scale of the star's lifetime. Finally, gyroscopic pumping and internal gravity waves may also contribute to slow mixing.

We have isolated the process of convective settling, in which a small fraction of the cold photospheric material is assumed to maintain some of its low entropy until it has settled at its neutral-buoyancy level in the stable stratification beneath the convection zone. The typical entropy contrast observed in photospheres of solar-type stars is strong enough so that settling can in priniciple penetrate the lithium-burning layers (Spruit 1997). However, an order-of-magnitude estimate shows that the mass flux of this material required to explain the observed lithium depletion in the Sun on the nuclear time scale is tiny (a fraction of order 10^{-7} of the mass flux in the downflows at the surface). Effects as weak as this cannot be measured from direct hydrodynamic simulations.

On its way down, the cold material mixes by entrainment with its surroundings, a process that cannot be captured realistically with existing means. The likely outcome at the base of the convection zone is a distribution of entropy contrast in the downflows, with most of the mass flux near the value corresponding to a mixing-length estimate, but with a tail of unknown shape extending to much

lower entropy values. In the absence of further information, we have parametrised this tail by a power law.

We use a one-dimensional model that treats the settling material as a depth-dependent source of mass in the layers below the base of the convection zone. The response of this region is a slow upward flow. Accompanying it, there is a thermal adjustment by radiative diffusion. The result is a slow, depth-dependent circulation of material. The free parameters of the entropy distribution model are adjusted such that lithium depletion takes place on the observed time scale.

The results show that the radiative flux and temperature gradient decrease to compensate for a positive (upward) convective flux caused by settling. This decrease in ∇ leads to a slight reduction of the convection zone's depth. The settling process is calculated only as a perturbation of a precalculated solar model, however; i.e., we do not model the evolution of the Sun with settling included. In a self-consistent and properly calibrated stellar-evolution model the change in the depth of the convection zone might actually have the opposite sign.

The resulting change in the sound speed profile due to settling is rather small and is concentrated in a thin layer below the Schwarzschild boundary. The calculation, however, only includes the direct effect of the settling process. It does not include the changes in the Sun's structure during its evolution on the main sequence. It is possible that this will redistribute the structural changes due to settling over a larger portion of its radius.

Another factor not included in our calculations is the stratification of helium concentration below the convection zone. It is caused by gravitational settling (also called 'diffusion') and known to significantly influence the solar sound-speed profile (see Christensen-Dalsgaard & Di Mauro 2007, and references therein). This helium-concentration gradient could be modified by the mixing induced by convective settling.

As Fig. 2.2 shows, settling preserves the discontinuity in the second derivative of the sound speed. That the helioseismic observations favour smoother sound-speed profiles (Christensen-Dalsgaard et al. 2011) is evidence for the existence of an additional overshooting mechanism acting closer to the boundary of convective instability than the settling process studied here.

The predicted lithium depletion changes by many orders of magnitude, as a function of the two model parameters (see Table 2.1). This is a natural consequence of the high temperature sensitivity of the burning reaction, combined with rapid changes of the mass flux in the lithium-burning layers as we change the slope β of the mass flux distribution. Therefore we use the observed lithium depletion in the Sun to constrain our model. This constraint yields a dependence of the total input mass flux $\dot{M}$ on slope β, reducing the number of free parameters to one (see Fig. 2.5). The linearity of $\log_{10} \dot{M}(\beta)$ comes as no surprise if we inspect

Fig. 2.1 in detail. We see that the $F[\hat{s}(z)]$ distributions at significant depths are self-similar and they apparently shift in proportion to β (in the logarithmic space of Fig. 2.1); i.e., a change in β can be directly translated to an equivalent change in $\dot{M}$ if the overall extent of lithium depletion is fixed. Figure 2.5 also shows that the most relevant values of β lie somewhere in the interval (1.5, 2.5) or even (2.0, 2.5). At higher values we could not meet the lithium depletion constraint because such models would require negative radiative flux at the top of the settling layer. At the low end, the convective flux becomes a negligible fraction of the total flux, and it gets difficult to interpret our power law as a tail of some more general mass flux distribution.

One might ask how our strongly simplified model compares to other models of overshooting. The most striking difference is that usually some form of the MLT is used to provide estimates of the velocities and entropy fluctuations at the boundary of the convection zone (e.g. Roxburgh 1965; Saslaw & Schwarzschild 1965; Shaviv & Salpeter 1973; van Ballegooijen 1982; Pidatella & Stix 1986; Zahn 1991), whereas we explicitly add the hypothesised low-entropy flows from the photosphere. Perhaps the closest to our ideas are the works of Rempel (2004) and Schmitt et al. (1984), who model the non-local convection by plumes.

The approach of modelling higher-order correlations in a turbulent field (e.g. Kuhfuss 1986; Xiong & Deng 2001; Marik & Petrovay 2002; Deng & Xiong 2008) tends to produce much deeper overshooting zones than the models above (depending on the value of the free parameters of the models). This can be traced to the fact that these models lack an essential aspect of the transition between convection and the stable interior. While velocity *amplitudes* vary rather smoothly across the boundary, their mixing effect varies strongly. In the convective region, the flows are of the efficiently mixing, overturning kind. In the stable part, however, the flows take the form of internal waves, which have a very weak mixing effect. These models also cannot capture the process of convective settling since the rare low-entropy downflows are not present in this picture.

In summary, we have shown that the convective settling process studied here can in principle explain long-term lithium depletion in the Sun and solar-type stars. This can be tested further by applying the model, calibrated to the lithium depletion observed in the Sun, to stars of different masses and ages.

3 Convective settling in main sequence stars: Li and Be depletion

R. Andrássy, H. C. Spruit

to be submitted to A&A

Abstract: The process of convective settling is based on the assumption that a small fraction of the low-entropy downflows sink from the photosphere down to the bottom of the star's envelope convection zone retaining a substantial entropy contrast. We have previously shown that this process could explain the slow Li depletion observed in the Sun. We construct a parametric model of convective settling to investigate the dependence of Li and Be depletion on stellar mass and age. Our model is generally in good agreement with the Li abundances measured in open clusters and solar twins, although it seems to underestimate the Li depletion in the first $\sim 1\,\mathrm{Gyr}$. The model is also compatible with the Be abundances measured in a sample of field stars.

3.1 Introduction

Low-mass, main sequence stars, with their deep convective envelopes, are astrophysical laboratories that allow us to investigate weak mixing processes under stellar conditions. The small distance between the convection zone proper and the Li-burning layer in these stars makes the surface abundance of Li a sensitive indicator of any mixing process that may be at work just below the convection zone. The surface abundance of Be, which is burnt at a somewhat higher temperature, provides an additional constraint on the extent of this mixing.

The standard solar model predicts essentially zero Li depletion on the main sequence (see e.g. Schlattl & Weiss 1999). Measurements of the Li abundances in solar analogues and twins, however, show indisputable evidence for Li depletion (Baumann et al. 2010; Meléndez et al. 2010; Monroe et al. 2013; Meléndez et al. 2014). The abundances of Be in the Sun, solar twins 16 Cyg A, B (6.6 Gyr), and solar analogues α Cen A, B (5–6 Gyr) show little depletion, if any (Asplund et al. 2009; Deliyannis et al. 2000; King et al. 1997). On the other hand, cool ($T_{\mathrm{eff}} \lesssim$

5500 K), main-sequence field stars show significant Be depletion (Delgado Mena et al. 2012, and references therein). All these observations imply that there is a mixing process operating below the convective envelopes of low-mass, main sequence stars, which works on a time scale of ~ 1 Gyr at Li-burning depths, but it is much less effective in the deeper layers where Be is burnt.

Several mechanisms have been proposed that could provide mixing beyond the formal boundary of a convection zone. Among the most well-know are „ballistic" overshooting (Roxburgh 1965; Saslaw & Schwarzschild 1965), convective penetration (Shaviv & Salpeter 1973; van Ballegooijen 1982; Zahn 1991), internal waves (Press 1981; Garcia Lopez & Spruit 1991; Montalbán 1994; Schatzman 1996), or shear-induced instabilities (Meakin & Arnett 2007). In this paper, we focus on the process of "convective settling", which was first proposed by Spruit (1997) and later elaborated in Andrássy & Spruit (2013, Paper I hereinafter).

Envelope convection zones are dominated by large-scale downflows spanning the whole convection zone (cf. Nordlund & Stein 1997; Trampedach et al. 2014). They are generated by the strong radiative cooling in the photosphere, which makes their initial entropy much lower than that of the nearly isentropic upflow. As they sink, they merge, entrain mass from the hot upflow and are heated by radiative diffusion on small scales. If a small fraction ($\approx 10^{-7}$, see Paper I) of the photospheric downflows retains a substantial entropy deficit until they have arrived at the base of the convection zone, they will continue sinking until each of them has settled on its level of neutral buoyancy, hence the name "convective settling." The entropy the downflows start with is low enough for them to reach the Li-burning layer. In general, there will be a broad distribution of entropy contrasts at the base of the convection zone, spanning from the low values predicted by the mixing-length theory (MLT) up to the highest values the downflows start with in the photosphere. They will settle at a range of depths, and mass conservation will enforce an upflow carrying the Li- and Be-depleted material back to the convection zone, reducing the surface abundances.

The envelope convection problem is too difficult even for the state-of-the-art numerical simulations. Hence, the distribution mentioned above can only be parametrised. In Paper I, we show that a power-law parametrisation leads to a model that can explain the slow, main-sequence Li depletion in the Sun without changing the thermal stratification so much as to come into conflict with the results of helioseismology. The simplified model presented in Paper I, however, is based on an approximate, non-evolving model of the Sun and it does not utilise the Be constraint. In this paper, we construct a model of the convective settling process that takes stellar evolution into account and applies both the Li and Be constraints. We also extend the computation to a range of stellar masses (from $0.8\,M_\odot$ to $1.2\,M_\odot$) in order to compare predictions of the model with abundance

measurements in open clusters, solar twins, and field stars.

3.2 Model

3.2.1 Overview

We construct a two-component, kinematic model of the convective settling process. The first component is an ensemble of downflows leaving the convection zone and sinking towards the deeper layers. They have been created by the rapid cooling in the photosphere and the entrainment and heating processes have changed their distribution on their way through the convection zone. We do not model these processes. Instead, we parametrise their output by a distribution of a mass flow rate[1] in the downflows with respect to their entropy contrast and model how they settle below the convection zone. Each downflow settles at the point where its entropy equals that of its surroundings, i.e. when it becomes neutrally buoyant. The other component of the model is an upflow due to mass conservation, the strength of which at any given depth depends on the total settling rate below that depth. The upflow advects the Li- and Be-depleted gas back to the convection zone.

The distribution of the mass flow rate is a central element in our model. The downflows at the base of an envelope convection zone are characterised by a wide range of entropy contrasts δs with respect to the isentropic upflow. This range can be split into two parts. One part represents the well-mixed material and the fluctuations created *locally* by turbulent convection. Their typical amplitude $(\delta s)_{\mathrm{typ}}$ is very small due to the high efficiency of convection deep in the stellar interior. Therefore, they settle in a very thin layer below the convection zone. Downflows with an entropy contrast significantly greater than $(\delta s)_{\mathrm{typ}}$ correspond to the incompletely-mixed remnants of the photospheric downflows. They span several orders of magnitude in δs, up to $(\delta s)_{\mathrm{max}}$, which is the maximum entropy contrast reached by the downflows just below the photosphere. The coolest of them settle deep below the convection zone. We ignore the well-mixed component in our model and parametrise the distribution of the mass flow rate carried by the incompletely-mixed downflows. We do not model the entrainment and heating processes explicitly; the mass flow rate as a function of the contrast value as parametrised in the model refers to the point where the downflow ultimately settles. We call this distribution *the settling rate distribution*.

[1] We use the term mass flux (measured in $\mathrm{g\,cm^{-2}\,s^{-1}}$ in cgs units) in Paper I. In this paper, we work in spherical geometry and the mass flow rate (measured in $\mathrm{g\,s^{-1}}$ in cgs units) becomes a more convenient quantity.

We assume that the settling rate distribution is a power law defined over the range of downflow entropy contrasts $(\delta s)_{\mathrm{min}} \leq \delta s \leq (\delta s)_{\mathrm{max}}$. The minimal entropy contrast considered, $(\delta s)_{\mathrm{min}}$, corresponds to the transition point to the range of the well-mixed downflows, thus we set $(\delta s)_{\mathrm{min}} = (\delta s)_{\mathrm{typ}}$. We estimate $(\delta s)_{\mathrm{typ}}$ using the MLT and extract $(\delta s)_{\mathrm{max}}$ from radiation-hydrodynamic simulations of stellar photospheres, which are readily available today (see Sect. 3.2.3).

In the settling paradigm, the slope β of the settling rate distribution should be a characteristic value resulting from the physics of the entrainment and heating processes in the convection zone and in the settling region. Therefore, we keep this value, albeit unknown, constant for all stars and all evolutionary stages of theirs and investigate what influence β has on the results.

The total mass flow rate $\dot{M}$ of the distribution is given by the mass flow rate that is leaving the photosphere, but it is considerably modified by the entrainment and heating processes in the convection zone and in the settling region. One could argue that the "destruction rate" of the cold downflows should somehow depend on the „strength" of the radiation effects on the downflows' way from the photosphere to their settling point, because it is radiative diffusion that provides heat exchange between the core of the downflow and the entrained material on small scales. The relative importance of radiation in the convection zone is parametrised through the value of $(\delta s)_{\mathrm{typ}}$. The more important the radiation, the less efficient the convection and the higher the value of $(\delta s)_{\mathrm{typ}}$. On the other hand, the more important the radiation, the smaller the fraction of the cold downflows that reach the settling region. To take this into account we scale $\dot{M}$ both in proportion to the mass downflow rate in the photosphere and in inverse proportion to $(\delta s)_{\mathrm{typ}}$. We use this scaling such as to include the qualitative effect of the heating process, although we realise that more parameters are likely to play a role and the dependence is much more complicated in reality. The constant of proportionality in the scaling is adjusted until the model reproduces the observed Li depletion in the Sun for a given value of β.

The convective settling process, depending on its strength, can change the thermal stratification below the convection zone. In Paper I, we have quantified this effect and shown it to be negligible in a solar model calibrated to reproduce the Li depletion observed in the solar photosphere if the input settling rate distribution is not too steep ($\beta \lesssim 2.5$). In this work, we neglect the influence of convective settling on the thermal stratification and compute Li and Be depletion using a few pre-computed stellar-evolution models.

3.2.2 Stellar models

We consider the main-sequence evolution of five solar-composition stars with masses of 0.8 $M_\odot$, 0.9 $M_\odot$, 1.0 $M_\odot$, 1.1 $M_\odot$, and 1.2 $M_\odot$. The models have been computed with the stellar-evolution code GARSTEC (Weiss & Schlattl 2008), neglecting the processes of convective overshooting and gravitational settling (also called sedimentation or diffusion). Since the code does not output the stratification of the specific entropy s in the star, we compute and approximation to it from the stratification of the temperature T and pressure p using the ideal-gas expression

$$s = \mathbb{R} \ln\left(\frac{T^{5/2}}{p}\right), \tag{3.1}$$

where $\mathbb{R} = k_\mathrm{B}/(\mu\, m_\mathrm{u})$ is the gas constant including the mean molecular weight μ, k_B the Boltzmann constant, and m_u the atomic mass unit. The mean molecular weight is constant in the region we are interested in. Equation 3.1 assumes constant level of ionisation, which is a good assumption at the temperatures and densities prevailing below the convection zones of the stars considered.

3.2.3 Mathematical formulation

Parametric model

The settling rate $\dot{m}$ in the downflows is distributed as

$$\mathrm{d}\dot{m} = \dot{M} f(\delta s)\, \mathrm{d}\delta s \tag{3.2}$$

where $\dot{M} \geq 0$ is the total settling rate, $\delta s = s_0 - s_\mathrm{d}$ the entropy contrast with s_0 the entropy of the stratification at the bottom of the convection zone and s_d the entropy of the downflow. As discussed in Sect. 3.2.1, the distribution function $f(\delta s)$ is assumed to be a power-law,

$$f(\delta s) = \begin{cases} N \left(\frac{\delta s}{(\delta s)_\mathrm{min}}\right)^{-\beta} & \text{for } (\delta s)_\mathrm{min} \leq \delta s \leq (\delta s)_\mathrm{max}, \\ 0 & \text{otherwise,} \end{cases} \tag{3.3}$$

where N is a normalisation factor, $(\delta s)_\mathrm{min} > 0$ and $(\delta s)_\mathrm{max} > 0$ are the bounds, and $\beta > 0$ is the slope of the distribution. We require $\int_{-\infty}^{\infty} f(\delta s)\, \mathrm{d}\delta s = 1$, so that

$$N = \begin{cases} \frac{\beta-1}{(\delta s)_\mathrm{min}} \left[1 - \left(\frac{(\delta s)_\mathrm{max}}{(\delta s)_\mathrm{min}}\right)^{-(\beta-1)}\right]^{-1} & \text{for } \beta \neq 1, \\ \left[(\delta s)_\mathrm{min} \ln\left(\frac{(\delta s)_\mathrm{max}}{(\delta s)_\mathrm{min}}\right)\right]^{-1} & \text{for } \beta = 1. \end{cases} \tag{3.4}$$

A downflow of entropy $s_\mathrm{d} = s_0 - \delta s$ reaches neutral buoyancy and settles down at the point where the entropy of the surrounding stratification s equals s_d. Mass conservation requires the upward mass flow rate at this point to be $\dot{M}F(s)$, where $F(s)$ is the fraction of downflows that settle below[2] this point and is given by the cumulative distribution function

$$F(s) = \int_{-\infty}^{s} f(s_0 - s')\,\mathrm{d}s'. \tag{3.5}$$

Parameter scaling

As described above, there are four parameters to be specified: $(\delta s)_\mathrm{min}$, $(\delta s)_\mathrm{max}$, $\dot{M}$, and β. In Paper I, we use fixed values because we only model one star, the evolution of which is also neglected. Now, we intend to model a range of different stars and follow their evolution, hence we have to adapt the parameter values to the changing physical conditions in the convection zone. The only exception is β, which is held constant (see Sect. 3.2.1).

The entropy contrast of the coldest downflow in the ensemble, $(\delta s)_\mathrm{max}$, is given by the maximum entropy contrast that the cooling process in the photosphere can create. This value reaches a well-defined maximum (in a time-averaged sense) just below the photosphere and can be extracted from radiation-hydrodynamic simulations of stellar photospheres. We use the grid of models computed by Magic et al. (2013) and approximate the dependence of $(\delta s)_\mathrm{max}$ on the effective temperature T_eff of the star and on its surface gravity $\log g$ by the fitting function

$$\log(\delta s)_\mathrm{max} = a_0 + a_1 x + a_2 y, \tag{3.6}$$

where $x = (T_\mathrm{eff} - 5777)/1000$, $y = \log g - 4.44$, $a_0 = 8.164$, $a_1 = 0.491$, and $a_2 = -0.461$ with cgs units assumed throughout.

We set the lowest entropy contrast considered, $(\delta s)_\mathrm{min}$, equal to a typical entropy contrast $(\delta s)_\mathrm{typ}$ predicted by the MLT (see Sect. 3.2.1). We use the MLT formulation of Kippenhahn et al. (2012) with $\alpha_\mathrm{MLT} = 1.65$ to estimate the superadiabatic temperature gradient $\Delta\nabla = \nabla - \nabla_\mathrm{ad}$ at the point where the pressure $p = p_0\,\mathrm{e}^{-1/2}$, where p_0 is the pressure at the bottom of the convection zone. We then estimate the entropy contrast an adiabatic convective element would reach after having overcome approximately one pressure scale height in such environment,

$$(\delta s)_\mathrm{min} \equiv (\delta s)_\mathrm{typ} = \frac{\mathbb{R}}{\nabla_\mathrm{ad}}\Delta\nabla. \tag{3.7}$$

[2]Since specific entropy decreases with increasing pressure in a stable thermal stratification.

The total mass settling rate $\dot{M}$ is scaled as (see Sect. 3.2.1)

$$\dot{M} = \dot{M}_0 \frac{\dot{M}_{\rm phot}}{2.04 \times 10^{21}\,{\rm g\,s^{-1}}} \left(\frac{(\delta s)_{\rm typ}}{7.96 \times 10^{1}\,{\rm erg\,g^{-1}\,K^{-1}}} \right)^{-1}, \tag{3.8}$$

where $\dot{M}_{\rm phot}$ is the mass downflow rate at the point (close to the photosphere) where the downflows reach the maximum entropy contrast $(\delta s)_{\rm max}$, and the constant $\dot{M}_0$ is adjusted until the solar model reproduces the observed Li abundance in the Sun. The numbers in the denominators in Eq. 3.8 correspond to the current solar values. The downflow mass flux in the photosphere, $\mathcal{F}_{\rm phot}$, is computed in the same way as $(\delta s)_{\rm max}$ (Eq. 3.6),

$$\log \mathcal{F}_{\rm phot} = b_0 + b_1 x + b_2 y, \tag{3.9}$$

where $b_0 = -1.475$, $b_1 = -0.239$, and $b_2 = 0.511$ in cgs units. No extrapolation is needed when using Eqs. 3.6 or 3.9 with the stellar models considered in this work. The mass downflow rate $\dot{M}_{\rm phot}$ is then

$$\dot{M}_{\rm phot} = 4\pi R_*^2 \mathcal{F}_{\rm phot}, \tag{3.10}$$

where R_* is the radius of the star.

Computational approach

We neglect the thermodynamic response of the star to the convective settling process. We do, however, estimate the convective flux that would be caused by this process to further constrain our model and check the plausibility of our assumptions (Sect. 3.3.4). All downflows considered pass through the bottom of the convection zone (as defined by the Schwarzschild criterion) and their entropy contrast with respect to their surroundings is the highest at that point, hence the convective flux reaches a maximum there. Its value relative to the total flux of energy is estimated to be

$$\hat{\mathcal{F}}_{\rm conv} = \frac{c_p T \overline{\frac{\Delta T}{T}} \dot{M}}{L}, \tag{3.11}$$

where c_p is the heat capacity at constant pressure, $\overline{\Delta T/T}$ the mean temperature contrast in the distribution (weighted by the mass flow rate), and L the luminosity of the star, all evaluated at the bottom of the convection zone. As explained in Sect. 3.2.1, the mass flow rate for every downflow in our distribution corresponds to the point where the downflow settles, because we do not model mass entrainment explicitly. Most of the convective flux, however, is carried by the downflows that settle very close to the convection zone's bottom, i.e. to the place where we

compute $\hat{\mathcal{F}}_{\rm conv}$. Hence, we regard Eq. 3.11 as a reasonable order-of-magnitude estimate.

To compute the Li and Be burning, we map the GARSTEC models (Sect. 3.2.2) on a grid equidistant in the mass fraction q.[3] We only include the outermost 10 – 15% of the stellar mass, which are relevant for the convective settling process. Interpolation of the models in time is done via the nearest-neighbour algorithm. We model the burning and transport of Li and Be using the set of equations

$$\frac{\mathrm{d}A_i}{\mathrm{d}t} = R_{\mathrm{b},i} + R_{\mathrm{m},i} + R_{\mathrm{s},i} + R_{\mathrm{a},i}, \tag{3.12}$$

where $A_i = N_i/N_H$ is the abundance of Li/Be in the i-the grid cell, N_i the number of Li/Be nuclei in the i-th grid cell, N_H the number of hydrogen nuclei per grid cell,[4] i increases with radius, t is the time, $R_{\mathrm{b},i}$ the burning rate, $R_{\mathrm{m},i}$ a mixing rate, $R_{\mathrm{s},i}$ the settling rate, and $R_{\mathrm{a},i}$ the advection rate due to the upflow. The burning rate

$$R_{\mathrm{b},i} = -\frac{A_i}{\tau_{\mathrm{b},i}} \tag{3.13}$$

is related to the nuclear-burning time scale $\tau_{\mathrm{b},i}$, which we compute using the standard expressions for low-energy nuclear reaction rates in an ideal gas that can be found e.g. in Hansen & Kawaler (1994). We consider the burning of ^{7}Li by the reaction ^{7}Li(p, α)α and the burning of ^{9}Be by the reactions ^{9}Be(p, α)^{6}Li and ^{9}Be(p, d)^{8}Be.[5] The products of the latter two reactions are quickly transformed to ^{3}He and ^{4}He nuclei, respectively, and are of no interest for this work. The low-energy astrophysical S-factors of the three reactions are taken from the NACRE-II database (Xu et al. 2013). Electron screening is neglected. The burning time scale below the settling layer is set equal to that at the bottom of the layer for numerical reasons. The rate at which the downflows settle in the i-th grid cell is

$$\dot{m}_i = \dot{M}\left[F(s_{i+1/2}) - F(s_{i-1/2})\right], \tag{3.14}$$

where the cumulative distribution function $F(s)$ is given by Eq. 3.5, s is the local entropy of the stratification, the index $i+1/2$ refers to the top and the index $i-1/2$ to the bottom of the i-th grid cell. Every single cell of our equidistant grid contains a mass of Δm, so we can define a "recycling" time scale,

$$\tau_{\mathrm{r},i} = \frac{\Delta m}{\dot{m}_i}, \tag{3.15}$$

[3] Mass loss is negligibly small for the stars considered.

[4] There is no gradient in the hydrogen mass fraction, because we neglect the gravitational settling of He (see Sect. 3.2.2) and because convective settling does not reach the core of the star.

[5] We omit the atomic mass numbers in the rest of the text.

which is the time it takes the convective settling process to completely replace the content of the i-th grid cell by "fresh" material from the convection zone. Most of the downflows settle just below the convection zone and $\tau_{\mathrm{r},i}$ can become very short there. To avoid severe time-step restrictions, we define a "well-mixed zone," which is composed of the convection zone and the region in that $\tau_{\mathrm{r}} < 10^5$ yr and $\tau_{\mathrm{b}} > 10^8$ yr. The fast recycling together with the slow burning prevent the formation of any significant gradients in the abundances of Li and Be in the well-mixed zone. We homogenise this zone using the artificial mixing term

$$R_{\mathrm{m},i} = \frac{\overline{A}_{\mathrm{wmz}} - A_i}{\Delta t}, \tag{3.16}$$

where $\overline{A}_{\mathrm{wmz}}$ is the average abundance of Li/Be in the well-mixed zone and Δt is the length of the current time step. We set $R_{\mathrm{m},i} = 0$ outside the well-mixed zone. The downflows in our model bring the Li- and Be-rich material from the convection zone and deposit it in the settling layer. The settling rate of Li/Be nuclei is

$$\dot{N}_{\mathrm{s},i} = A_{\mathrm{cz}} \frac{X \dot{m}_i}{m_{\mathrm{p}}}, \tag{3.17}$$

where A_{cz} is the abundance of Li/Be in the convection zone, X the hydrogen mass fraction and m_p the proton mass. The rate of change of the abundance due to settling is then

$$\dot{R}_{\mathrm{s},i} = \frac{\dot{N}_{\mathrm{s},i}}{N_H} = A_{\mathrm{cz}} \frac{\dot{m}_i}{\Delta m} = \frac{A_{\mathrm{cz}}}{\tau_{\mathrm{r},i}}, \tag{3.18}$$

Finally, there is an upflow due to mass conservation, which we model by the advection rate of Li/Be nuclei,

$$\dot{N}_{\mathrm{a},i} = A_{i-1} \frac{X \sigma_{i-1/2}}{m_{\mathrm{p}}} - A_i \frac{X \sigma_{i+1/2}}{m_{\mathrm{p}}}, \tag{3.19}$$

where $\sigma_{i-1/2} = \sum_{k=0}^{i-1} \dot{m}_k$ is the mass inflow rate to the i-th grid cell at its bottom boundary and $\sigma_{i+1/2} = \sum_{k=0}^{i} \dot{m}_k$ is the mass outflow rate from the i-th grid cell at its top boundary. The rate of change of the abundance due to advection is then

$$R_{\mathrm{a},i} = A_{i-1} \sum_{k=0}^{i-1} \frac{1}{\tau_{\mathrm{r,i}}} - A_i \sum_{k=0}^{i} \frac{1}{\tau_{\mathrm{r,i}}}. \tag{3.20}$$

Equation 3.12 is independent of the absolute abundance scale (see also Eqs. 3.13, 3.16, 3.18, and 3.20). Therefore, we start all our calculations with $A_i = 1$ and integrate Eq. 3.12 using the standard, 4th-order Runge-Kutta method. The initial condition that we want to impose (see Sect. 3.3.1) is then taken care of simply by rescaling the results accordingly. We use the usual astrophysical notation $\log \epsilon = 12 + \log(A_{\mathrm{cz}})$ in the rest of the text.

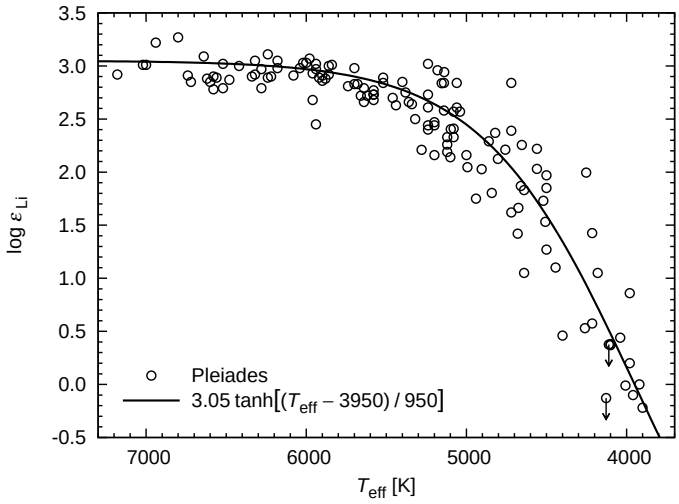


Figure 3.1: Lithium distribution in the Pleiades (Sestito & Randich 2005). The mean trend is approximated by a smooth function, which provides an initial condition for our model.

3.3 Results

3.3.1 Preliminaries

Cool stars deplete significant amounts of their Li during their pre-main-sequence (PMS) evolution. Changes in the structure of a PMS star, however, are rather dramatic and show very large scatter even in a single cluster (Baraffe et al. 2009). Applying our simple scaling relations to such a wide range of conditions would be questionable. Instead, we start all our computations at the age of the Pleiades, for which we adopt a value of 100 Myr. We approximate the observed Li distribution in this cluster by a smooth function (see Fig. 3.1) and use it as an initial condition for the Li abundance. The depletion of Be is much lower and difficult to measure at low $T_{\rm eff}$, so we use the meteoritic value $\log \epsilon_{\rm Be} = 1.30$ (Asplund et al. 2009) as an initial condition for the Be abundance. We assume that both Li and Be are homogeneously distributed in the interior of the star at the start of the computation.

The dependences of $(\delta s)_{\rm max}$, $(\delta s)_{\rm min}$, and $\dot{M}$ on stellar mass and age, as defined by Eqs. 3.6, 3.7, and 3.8, are shown in Figs. 3.2, 3.3, and 3.4. The values of $\dot{M}_{\rm phot}$ in Eq. 3.8 turn out to lie within ~ 30% of one another for all of the stars considered,

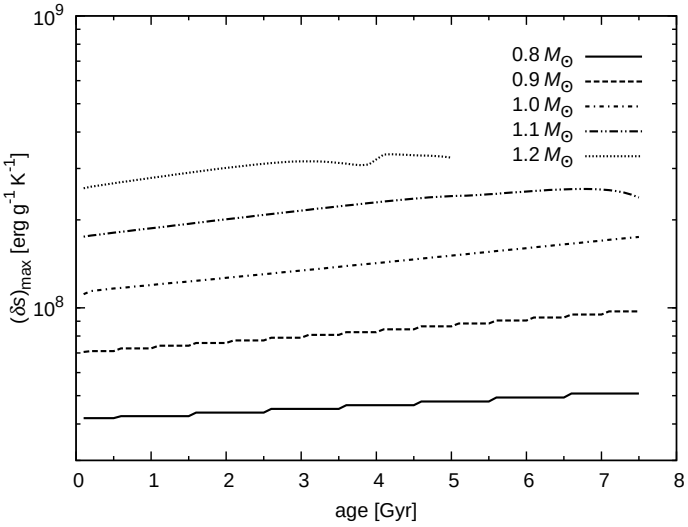


Figure 3.2: Age dependence of $(\delta s)_{\rm max}$ in the stellar models used in this work. The discontinuities are caused by the nearest-neighbour interpolation that we use.

although the spread in the values of $\mathcal{F}_{\rm phot}$ in Eq. 3.10 is an order of magnitude larger. The scaling of $\dot{M}$ in Eq. 3.8 is thus dominated by the $(\delta s)_{\rm typ}$ factor.

3.3.2 Li and Be depletion in the Sun

The influence of stellar evolution on the Li depletion in the Sun is illustrated in Fig. 3.5. The depletion rate becomes quasi-stationary after ~ 200 Myr and slowly decreases as the Sun ages. The depletion rate is hardly influenced by the assumed value of β. The observational data over-plotted in Fig. 3.5 suggest a somewhat more pronounced slowdown in the depletion rate, although the error bars are quite large. We also show a non-evolving model, in which the stratification is given by the solar-structure model at an age of 4.6 Gyr and is not allowed to evolve during the computation. In this case, the depletion rate becomes constant after the initial transition, as could be expected.

The abundance of Be predicted by the evolving, 1.0 $M_\odot$ model at an age of 4.6 Gyr ranges from 1.15 at $\beta = 1.5$ to 1.17 at $\beta = 2.5$, which deviates from the observed value of 1.38 ± 0.09 (Asplund et al. 2009) by -2.5σ. Note that the meteoritic value is only 1.30 ± 0.03 (Asplund et al. 2009).

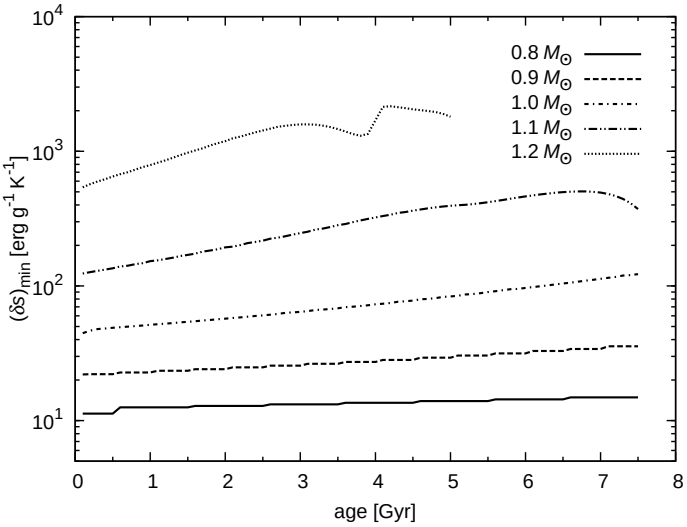


Figure 3.3: As Fig. 3.2, but $(\delta s)_{\rm min}$ is plotted.

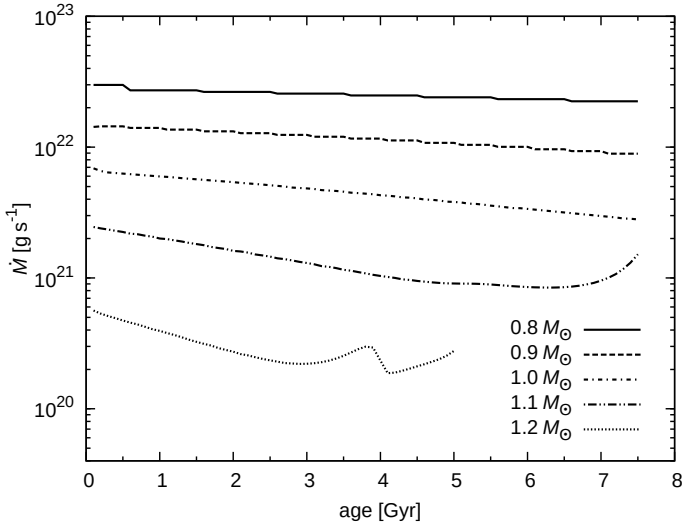


Figure 3.4: As Fig. 3.3, but $\dot{M}$ is plotted. This figure assumes $\beta = 2.0$. The curves would be shifted downwards by 2.8 dex at $\beta = 1.5$ and upwards by 2.9 dex at $\beta = 2.5$.

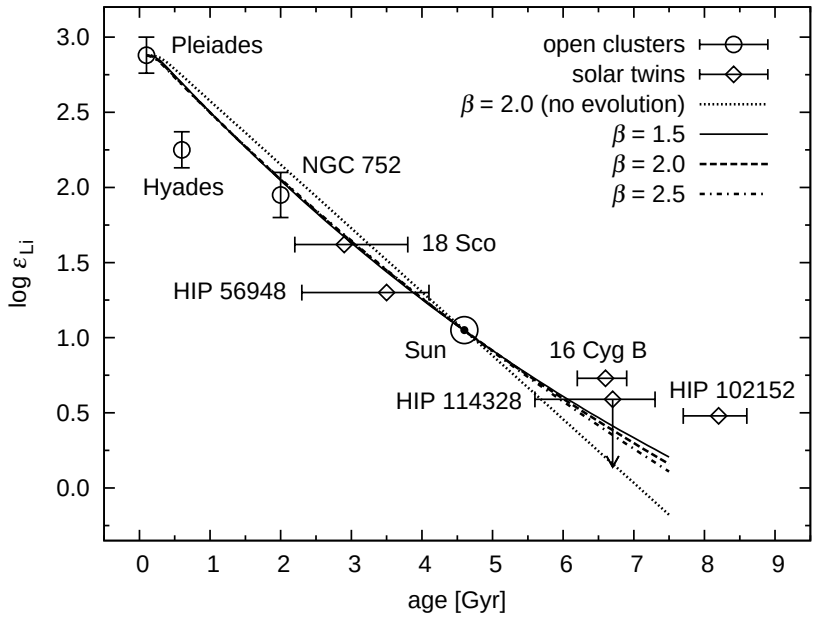


Figure 3.5: Age dependence of the Li depletion in the Sun as predicted by the convective settling model compared with the Li abundances in solar twins and in solar-type stars of a few open clusters. The thermal stratification in the model plotted by the dotted line is not allowed to evolve in time. The abundances in the open clusters are from Sestito & Randich (2005) and correspond to the solar effective temperature at the age of the cluster; the error bars show the typical scatter in the data. The measurements in 18 Sco and HIP 102152 are from Monroe et al. (2013), HIP 56948 and HIP 114328 from Meléndez et al. (2012, 2014), respectively, and 16 Cyg B from Ramírez et al. (2011); the age of 16 Cyg B is from Tucci Maia et al. (2014).

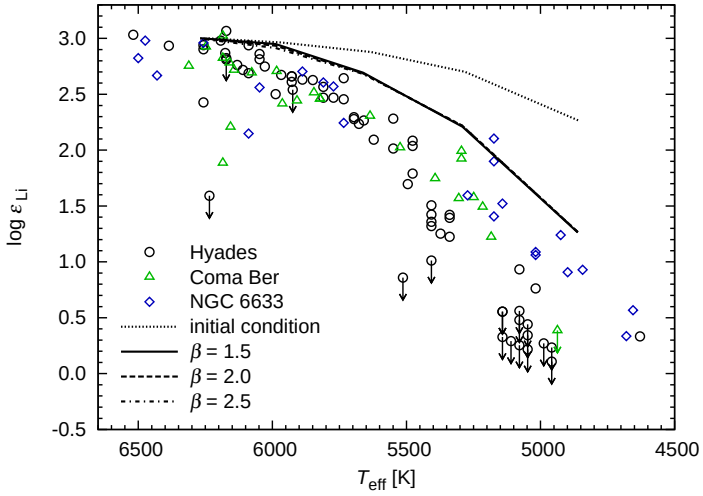


Figure 3.6: Dependence on the effective temperature of the Li depletion predicted by the convective settling model compared with the Li abundances in three 600-Myr-old open clusters as determined by Sestito & Randich (2005). Almost no dependence on β can be seen at this age.

3.3.3 Mass dependence of Li and Be depletion

Figures 3.6 and 3.7 show the dependence of the Li depletion on the effective temperature of the star (hence on its mass) at a fixed age as compared with Li abundances observed in open clusters. The metallicity of the stars may influence the extent of Li depletion, because the higher the metallicity, the higher the opacity and the deeper the convection zone. The metallicities of the clusters used in this work are summarised in Table 3.1. Our model underestimates the Li depletion at an age of 600 Myr (Fig. 3.6) independently of the value of β. This may stem from our assumption of a homogeneous Li distribution in the stellar interior at the start of the computation. The fit to the Hyades data (see Figs. 3.5 and 3.6) would also improve if the age of this cluster was not 600 Myr as we assume, but 950 Myr as Brandt & Huang (2015) recently suggested. The data are scarce at an age of 2 Gyr (Fig. 3.7), but the overall trend fits better. A weak dependence on β can be seen at high effective temperatures.

The depletion of Be is much smaller and much more difficult to measure than the depletion of Li. Therefore, we resort to a comparison with field stars in Fig. 3.8. We only show models with $\beta = 2.0$, because their dependence on β is

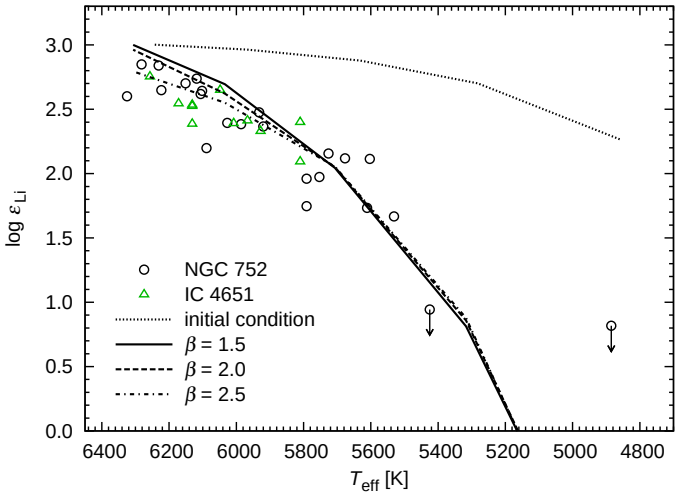


Figure 3.7: As Fig. 3.6, but the comparison is made with two 2-Gyr-old open clusters; the measurements are from Sestito & Randich (2005). A slight dependence on β becomes visible at this age.

Cluster	Age [Gyr]	[Fe/H]
Pleiades	0.1	−0.03
Coma Ber	0.6	−0.05
Hyades	0.6	+0.13
NGC 6633	0.6	−0.10
IC 4651	2	+0.10
NGC 752	2	+0.01

Table 3.1: Ages and metallicities of the open clusters used in this work. Adapted from Sestito & Randich (2005); the age of the Pleiades rounded to 0.1 Gyr.

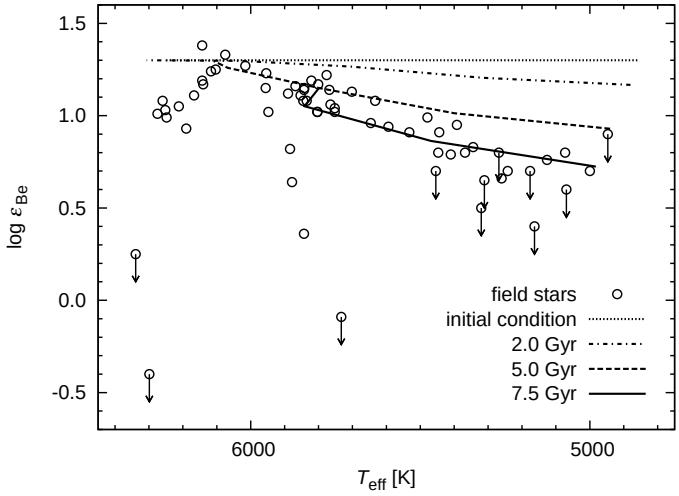


Figure 3.8: Dependence on the effective temperature of the Be depletion predicted by the convective settling model in a range of ages compared with the Be abundances measured in a sample of field stars from Santos et al. (2004). The fastest-evolving, $1.2\,M_\odot$ star is computed up to 5 Gyr only.

rather weak. The trend in the Be depletion is well reproduced assuming that the stars with $T_{\rm eff} \lesssim 5500\,$K are older than $\sim 5\,$Gyr, which is a reasonable assumption for cool field stars.

3.3.4 Heat flux due to convective settling

The convective flux at the bottom of the convection zone, $\hat{\mathcal{F}}_{\rm conv}$ (Eq. 3.11), unlike the depletion of Li and Be, is very sensitive to the assumed value of β, see Fig. 3.9. This comes about because the settling rate distribution spans several orders of magnitude in the downflow entropy contrast δs and the slope β of the distribution is constant over the whole range (see Sect. 3.2.1). The calibration of the $1\,M_\odot$ model to the observed solar Li depletion sets the total amount of material that has to settle in the Li-burning layer. Thus, the calibration fixes the tail of the distribution where the entropy contrast is high, of the order of $(\delta s)_{\rm max}$. If we increase the value of β and recalibrate the model, the settling rate integrated over the Li-burning layer will not change much, but that just below the convection zone, where $\delta s \approx (\delta s)_{\rm min} \ll (\delta s)_{\rm max}$, will increase considerably. The main

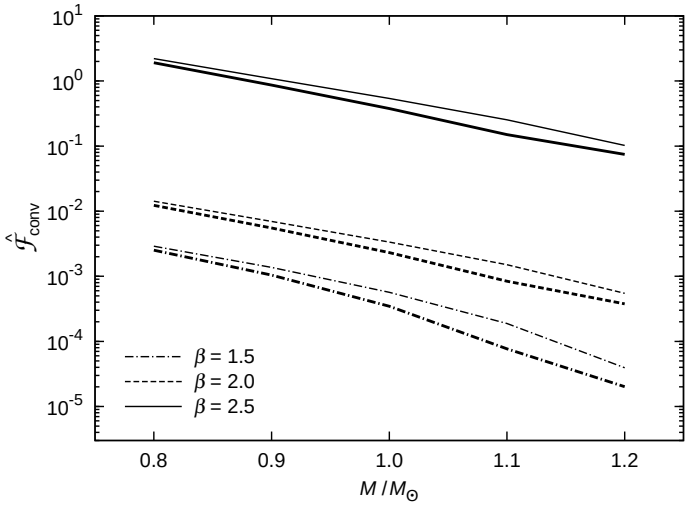


Figure 3.9: Dependence on the stellar mass of the convective flux due to the convective settling process at the bottom of the convection zone. Three sets of models with $\beta \in \{1.5, 2.0, 2.5\}$ are show at ages of 1 Gyr (thin lines) and 5 Gyr (thick lines).

cause of the mass dependence of $\hat{\mathcal{F}}_{\rm conv}$ seen in Fig. 3.9 is the variation of $(\delta s)_{\rm min}$ between different stars. This parameter changes $\hat{\mathcal{F}}_{\rm conv}$ both directly, by changing the normalisation factor N of the settling rate distribution (Eq. 3.4), and indirectly, via the scaling relation for $\dot{M}$, see Eq. 3.8 and the related discussion in Sect. 3.2.3. Since the values of $(\delta s)_{\rm max}$ span a much narrower range than those of $(\delta s)_{\rm min}$, the influence of this parameter on $\hat{\mathcal{F}}_{\rm conv}$ is correspondingly weaker. The differences in the thermal stratification between the stars considered are significantly larger than those due to their main-sequence evolution. The dependence of $\hat{\mathcal{F}}_{\rm conv}$ on the age of the star is therefore much weaker than on its mass (see Fig. 3.9). The models with $\hat{\mathcal{F}}_{\rm conv}$ approaching or even exceeding unity are in conflict with our assumption of no thermal feedback of convective settling on the star (see the last paragraph of Sect. 3.2.1), so their predictions should be taken with a grain of salt.

3.4 Summary and discussion

The process of convective settling is based on the idea that the envelope convection zones of low-mass, main sequence stars are dominated by large-scale downflows spanning the whole convection zone. If a tiny fraction of the low-entropy, photospheric downflows crosses the whole convection zone without having experienced much heating, their strongly negative buoyancy will make them sink and settle as deep as in the Li- and Be-burning layers. Mass conservation implies an upflow of the Li- and Be-depleted material back into the convection zone, reducing the photospheric abundances.

Building on the results of Paper I, we have explored the dependence of the Li and Be depletion on stellar mass in the range from $0.8\,M_\odot$ to $1.2\,M_\odot$. We assume that the mass settling rate is distributed with respect to the entropy contrast of the downflow as a power law. It spans a wide range of entropy contrasts from $(\delta s)_{\rm min}$, given by the MLT at the bottom of the convection zone, to $(\delta s)_{\rm max}$, given by the maximum entropy contrast the downflows reach just below the photosphere. The slope β of the distribution, assumed to be a constant, parametrises the physics of entrainment and heating processes on the downflows' way to their settling points. In the absence of information on the dependence of the total mass settling rate $\dot{M}$ on the structure of the star, we scale this parameter in proportion to the mass downflow rate in the photosphere and in inverse proportion to $(\delta s)_{\rm min}$, so that we qualitatively capture the effect of radiative diffusion on the heating of downflows. The mass downflow rate in the photosphere turns out to be essentially constant for the stars considered. We calibrate the scaling of $\dot{M}$ such as to reproduce the solar Li depletion. We allow for stellar evolution, but neglect the thermal feedback of convective settling on the star because we have shown in Paper I that the feedback is negligible provided that $\beta \leq 2.5$. The computation is started at 100 Myr, using the observed Li distribution in the Pleiades and the meteoritic Be abundance as initial conditions for Li and Be burning, respectively.

Changes in the solar structure cause a slowdown in the Li-depletion rate as the Sun ages. This slowdown seems to be somewhat milder in the model compared to the observed Li evolution in open clusters and solar twins. The main discrepancy occurs in the first ~ 1 Gyr when real stars deplete Li faster than those in our model. This may be a consequence of our assumption that both Li and Be are homogeneously distributed in the star when the computation is started. Despite this issue, it is encouraging that the model can well reproduce the observed dependence of the Be depletion on the effective temperature in a sample of field stars. The current abundance of Be in the Sun predicted by the model is also compatible with the observed value. These conclusions are essentially independent of the assumed value of β, which is likely caused by the similarity of the internal

structures of the stars considered.

We show that the convective flux at the bottom of the convection zone is very sensitive to β. Our model includes the low-entropy tail of the mass-flow-rate distribution only. We imagine it as an extension of the mass-flow-rate distribution in the "MLT-like" downflows, which carry a significant portion of the total flux even close to the bottom of the convection zone. Thus, one would expect the convective flux in our model to fall roughly into the interval $10^{-2} \lesssim \hat{\mathcal{F}}_{\mathrm{conv}} \lesssim 10^{0}$ to be compatible with our conceptual picture of convective settling. This corresponds to $2.0 \lesssim \beta \lesssim 2.5$ in the solar model, in agreement with the conclusions of Paper I.

The main caveat of our present analysis is the qualitative nature of the scaling of the total mass settling rate with the properties of the star. It determines the sensitivity of the predicted Li- and Be-depletion rates and of the convective flux on the stellar mass and age. We would need to model the details of the downflows' mass entrainment and heating to shed light on this issue as well as on the question whether the mass-settling-rate distribution is a power law in the first place.

4 Overshooting by differential heating

R. Andrássy, H. C. Spruit

accepted for publication in A&A

Abstract: On the long nuclear time scale of stellar main-sequence evolution, even weak mixing processes can become relevant for redistributing chemical species in a star. We investigate a process of "differential heating," which occurs when a temperature fluctuation propagates by radiative diffusion from the boundary of a convection zone into the adjacent radiative zone. The resulting perturbation of the hydrostatic equilibrium causes a flow that extends some distance from the convection zone. We study a simplified differential-heating problem with a static temperature fluctuation imposed on a solid boundary. The astrophysically relevant limit of a high Reynolds number and a low Péclet number (high thermal diffusivity) turns out to be interestingly non-intuitive. We derive a set of scaling relations for the stationary differential heating flow. A numerical method adapted to a high dynamic range in flow amplitude needed to detect weak flows is presented. Our two-dimensional simulations show that the flow reaches a stationary state and confirm the analytic scaling relations. These imply that the flow speed drops abruptly to a negligible value at a finite height above the source of heating. We approximate the mixing rate due to the differential heating flow in a star by a height-dependent diffusion coefficient and show that this mixing extends about 4% of the pressure scale height above the convective core of a 10 $M_\odot$ zero-age main sequence star.

4.1 Introduction

Our lack of understanding of (magneto)hydrodynamic transport processes in stars has hampered progress in developing the stellar evolution theory since its earliest beginnings. One particular aspect of the problem is the mixing in the boundary layers between convection and radiative zones in stellar interiors, which is also known as the problem of convective overshooting. Despite the indisputable advance in numerical simulations, the problem remains extremely challenging owing to the extreme range of the length and time scales involved in it.

The set of physical mechanisms that provide mixing at a convective/stable interface very likely depends on the type of convection zone involved. Because it is exposed to outer space at the top, a convective envelope is driven by the cold plumes originating in the photosphere. It is quite possible that the plumes span the whole convection zone and even provide mixing at its bottom boundary (cf. Andrássy & Spruit 2013, and references therein). A convective core or shell is, on the other hand, fully embedded in the star, its stratification is much weaker, and the temperature fluctuations within it are much smaller. Consequently, a different set of physical mechanisms may dominate mixing at its boundary.

It has long been known that the kinetic energy of the low-Mach-number flow in a convective core (or shell) is so low that the convective motions are stopped within about one per cent of the pressure scale height once they enter the steep entropy gradient of the radiative zone (Roxburgh 1965; Saslaw & Schwarzschild 1965). The motions can reach much farther, though, if they are vigorous enough to flatten the radiative entropy gradient above the convective core. In this case, known as the process convective penetration, the motions gradually "erode" the radiative stratification on the thermal time scale until radiative diffusion stops any further advance of the erosion front (Shaviv & Salpeter 1973; van Ballegooijen 1982; Zahn 1991). Finally, the fluid parcels hitting the stable stratification always generates a spectrum of internal gravity waves, which may also provide a certain amount of mixing (Press 1981; Garcia Lopez & Spruit 1991; Schatzman 1996).

Several of the processes mentioned may operate at the boundary of a convection zone at the same time. Their effects on long time scales and at long distances from the boundary are very different. In full numerical hydrodynamic simulations, the restrictions on time scales that can be covered makes it difficult to disentangle these effects. Physical insight developed by different means is needed to extrapolate them to longer time scales and distances.

We take a closer look at one specific process operating at a convective/stable interface in the interior of a star. Thermal diffusion causes temperature fluctuations from the convection zone's boundary to spread into the stable stratification. Temperature differences on surfaces of constant pressure set up a flow even in the absence of momentum transport by hydrodynamic stress. We call this process "differential heating", explore the physics of it in an idealised set-up, and estimate what amount of mixing it could cause in the stellar interior.

4.2 The differential heating problem

4.2.1 Problem formulation and simplification

Consider a horizontal, solid surface with a stably-stratified fluid overlying it.[1] A temperature fluctuation imposed at the surface propagates into the fluid by a diffusive process and upsets the hydrostatic equilibrium. We investigate what the properties of the resulting flow are.

By replacing the convective/stable interface by a solid wall, we eliminate all the phenomena related to the inertia of the convective flows and the shear induced by them. This allows us to study the physics of differential heating in isolation. The upper boundary is taken far enough not to influence the flow. Next we introduce further assumptions to facilitate the mathematical description and the subsequent analysis of the problem:

(1) The flow is confined to a layer that is significantly thinner than the pressure scale height.

(2) The fluid is a chemically homogeneous, ideal gas.

(3) The Brunt-Väisälä (buoyancy) frequency of the stratification is constant.

(4) Thermal diffusivity is constant.

(5) The gravitational field is homogeneous.

(6) The differentially heated surface is flat and horizontal.

(7) The flow is constrained to two spatial dimensions.

Assumption (1) allows us to use the Bussinesq approximation and turns out to be justified. The chemical homogeneity that we assume in (2) is, at least for the nuclear-burning layers in a star, only realistic at the onset of the burning. The differential heating process above a convective core would weaken as the nuclear burning progresses owing to the increase in the mean molecular weight in the core. We focus on the chemically homogeneous case to keep the number of parameters tractable. We introduce (3) and (4) for the same reason. The Brunt-Väisälä frequency depends on the distance from the convective/stable interface in a real star. The constant frequency in our analysis can be thought of as a typical value for the layer influenced by differential heating. Finally, we add the last

[1] Equivalently, the stably stratified fluid could be placed under the differentially heated surface. The role of hot and cold spots on the surface would be reversed in this case. We discuss only one case for the sake of concreteness.

three assumptions to make our analysis more transparent and to reduce the computational costs of the numerical solutions. We have to keep in mind, however, that the constraint (7) might influence the stability properties of the flow, and thus some of our conclusions may not apply to the three-dimensional case.

The Bussinesq equations are (Spiegel & Veronis 1960)

$$\nabla \cdot \boldsymbol{u} = 0, \tag{4.1}$$

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\frac{1}{\rho_\mathrm{m}}\nabla p' + \frac{T'}{T_\mathrm{m}} g\boldsymbol{k} + \nu\,\nabla^2\boldsymbol{u}, \tag{4.2}$$

$$\frac{\mathrm{D}T'}{\mathrm{D}t} = -\frac{T_\mathrm{m}N^2}{g} w + \varkappa\,\nabla^2 T', \tag{4.3}$$

where $\boldsymbol{u}$ is the fluid velocity, $\mathrm{D}/\mathrm{D}t = \partial/\partial t + \boldsymbol{u}\cdot\nabla$ the Lagrangian time derivative, ρ_m and T_m are the mean density and temperature, respectively, p' and T' the pressure and temperature perturbations, respectively, g is the gravitational acceleration, $\boldsymbol{k}$ a unit vector pointed in the vertical direction, ν the kinematic viscosity, N the Brunt-Väisälä frequency, w the vertical velocity component, and $\varkappa$ the thermal diffusivity.

Equations 4.1, 4.2, and 4.3 still contain several dimensional parameters. It is crucial to realise that there is a natural system of units for the differential heating problem that makes the equations dimensionless. The flow in this problem is set off by thermal diffusion in a stably stratified medium, therefore the inverse of the Brunt-Väisälä frequency, $1/N$ (or a multiple of it), is a natural unit of time. Having made this choice, we can define a natural unit of distance as $\sqrt{\varkappa/N}$, which is a typical thermal-diffusion length scale on the time scale $1/N$. The dimensionless Bussinesq equations are then

$$\nabla \cdot \boldsymbol{u} = 0, \tag{4.4}$$

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\nabla p + \vartheta\boldsymbol{k} + \mathrm{Pr}\;\nabla^2\boldsymbol{u}, \tag{4.5}$$

$$\frac{\mathrm{D}\vartheta}{\mathrm{D}t} = -w + \nabla^2\vartheta, \tag{4.6}$$

where we omit any symbol to indicate the new units. We have also introduced a new pressure-like variable $p = p'/\rho_\mathrm{m}$ and the buoyancy acceleration $\vartheta = gT'/T_\mathrm{m}$, which we continue to call the "temperature fluctuation" in the rest of the paper, because that is the central concept in the differential heating process. The Prandtl number $\mathrm{Pr} = \nu/\varkappa$ now becomes a measure of kinematic viscosity, because the new unit of diffusivity is $\varkappa$. Equations 4.4, 4.5, and 4.6 are particularly well suited to theoretical studies since their solution is fully determined by the Prandtl number, the initial, and the boundary conditions.

The distance unit $\sqrt{\varkappa/N}$ is rather short in stellar interiors, and it only weakly depends on the stratification. To see this, we express the Brunt-Väisälä frequency in terms of the more common stellar-structure parameters,

$$N^2 = \frac{g}{H_\mathrm{p}}(\nabla_\mathrm{ad} - \nabla), \tag{4.7}$$

where H_p is the pressure scale height, ∇_ad the adiabatic temperature gradient, and ∇ the actual temperature gradient. Close to a convection zone's boundary, we can write

$$\nabla_\mathrm{ad} - \nabla = \alpha \frac{z}{H_\mathrm{p}}, \tag{4.8}$$

where $\alpha \approx 10^{-1}$ is a coefficient of proportionality and z the distance from the boundary ($z > 0$ in the stable stratification). When using Eqs. 4.7 and 4.8, the unit of distance can be expressed as

$$\sqrt{\frac{\varkappa}{N}} = \varkappa^{1/2} \left(\alpha \frac{g}{H_\mathrm{p}} \frac{z}{H_\mathrm{p}} \right)^{-1/4}, \tag{4.9}$$

which is about 10^7 cm for values typical of a point close to the convective/stable interface ($z \approx 10^{-2} H_\mathrm{p}$) in the core of a massive ($10\,M_\odot$), main-sequence star ($\varkappa \approx 10^{10}\,\mathrm{cm^2\,s^{-1}}$, $\alpha \approx 10^{-1}$, $g \approx 10^5\,\mathrm{cm\,s^{-2}}$, $H_\mathrm{p} \approx 10^{10}$ cm).

Two distinct regimes of differential heating can be expected, depending on the amplitude and the spatial scale of the temperature fluctuation imposed on the differentially heated surface. If the heating is strong enough, the heat transport is advection-dominated (i.e. the flow's Péclet number is high), and the flow is generally unsteady. A similar phenomenon takes place right at the point where the convective flow leaves the unstable stratification, still retaining some positive temperature fluctuation. It quickly cools down as it rises in the stable medium, its temperature fluctuation turns negative, and the flow is brought to a halt. This is the place where we can impose a lower boundary condition for a much weaker kind of differential-heating-induced flow, in which diffusive heat transport plays a major role (i.e. the flow's Péclet number is low). The latter case is the main focus of this paper. We show in Sect. 4.3 that such a flow is generally smooth and reaches a stationary state (to be specified in Sect. 4.3.1) even at rather high values of the Reynolds number, up to $\mathrm{Re} = 4 \times 10^3$. This allows us to gain some insight into the problem by exploring the scaling properties of the stationary differential-heating equations, which we do in the next section.

4.2.2 Analytical considerations

The stationary differential-heating problem is described in two dimensions by the set of equations (cf. Eqs. 4.4, 4.5, 4.6)

$$\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} = 0, \tag{4.10}$$

$$\frac{\partial (uu)}{\partial x} + \frac{\partial (uw)}{\partial z} = -\frac{\partial p}{\partial x} + \mathrm{Pr}\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2}\right), \tag{4.11}$$

$$\frac{\partial (uw)}{\partial x} + \frac{\partial (ww)}{\partial z} = -\frac{\partial p}{\partial z} + \vartheta + \mathrm{Pr}\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial z^2}\right), \tag{4.12}$$

$$\frac{\partial (u\vartheta)}{\partial x} + \frac{\partial (w\vartheta)}{\partial z} = -w + \frac{\partial^2 \vartheta}{\partial x^2} + \frac{\partial^2 \vartheta}{\partial z^2}, \tag{4.13}$$

where x and z are the horizontal and vertical coordinates, respectively, with the z axis pointed against the gravitational acceleration vector, u is the horizontal velocity component, and w the vertical one. In what follows, we show how the characteristic properties of the stationary flow depend on the typical amplitude Θ and the typical horizontal length scale L of the heating applied.

Assume that there is a well-defined vertical length scale H in the differential heating flow pattern. Let us denote the typical horizontal and vertical velocities by U and W, respectively, and the typical pressure fluctuation by P. We then introduce a new set of variables $\hat{x}$, $\hat{z}$, $\hat{u}$, $\hat{w}$, $\hat{p}$, and $\hat{\vartheta}$, which all reach values of the order of unity close to the differentially heated surface, and

$$x = L\hat{x}, \tag{4.14}$$

$$z = H\hat{z}, \tag{4.15}$$

$$u = U\hat{u}, \tag{4.16}$$

$$w = W\hat{w}, \tag{4.17}$$

$$p = P\hat{p}, \tag{4.18}$$

$$\vartheta = \Theta\hat{\vartheta}. \tag{4.19}$$

Upon making these substitutions in Eq. 4.10, we obtain

$$\frac{U}{L}\frac{\partial \hat{u}}{\partial \hat{x}} + \frac{W}{H}\frac{\partial \hat{w}}{\partial \hat{z}} = 0, \tag{4.20}$$

which implies the approximate relation

$$\frac{U}{L} \approx \frac{W}{H}. \tag{4.21}$$

The horizontal momentum equation (Eq. 4.11) attains the form

$$\frac{\partial(\hat{u}\hat{u})}{\partial\hat{x}} + \frac{\partial(\hat{u}\hat{w})}{\partial\hat{z}} \approx -\frac{P}{U^2}\frac{\partial\hat{p}}{\partial\hat{x}} + \frac{\mathrm{Pr}}{UL}\frac{\partial^2\hat{u}}{\partial\hat{x}^2} + \frac{\mathrm{Pr}}{WH}\frac{\partial^2\hat{u}}{\partial\hat{z}^2}, \tag{4.22}$$

where Eq. 4.21 has been used, so the equality is only approximate. Nonetheless, we can see that the viscous terms are of the order of $1/\mathrm{Re}_x \equiv \mathrm{Pr}/(UL)$ and $1/\mathrm{Re}_z \equiv \mathrm{Pr}/(WH)$, where Re_x and Re_z are Reynolds-like numbers associated with horizontal and vertical motions, respectively. We introduce this unusual notation to characterise the relative contributions of the two viscous terms in the case of $L \gg H$. We focus on this limit because it turns out to be the relevant one in stellar interiors (see Sect. 4.5). From now on, we assume $\mathrm{Re}_x \gg 1$ and $\mathrm{Re}_z \gg 1$. Equation 4.22 shows that pressure fluctuations are of the order of U^2 in this high-Reynolds-number limit, so that we can estimate

$$P \approx U^2. \tag{4.23}$$

The vertical momentum equation (Eq. 4.12), with the substitutions defined above and Eqs. 4.21 and 4.23, becomes

$$\frac{\partial(\hat{u}\hat{w})}{\partial\hat{x}} + \frac{\partial(\hat{w}\hat{w})}{\partial\hat{z}} \approx \frac{L}{H}\left(-\frac{\partial\hat{p}}{\partial\hat{z}} + \frac{H\Theta}{U^2}\hat{\vartheta}\right) + \frac{1}{\mathrm{Re}_x}\frac{\partial^2\hat{w}}{\partial\hat{x}^2} + \frac{1}{\mathrm{Re}_z}\frac{\partial^2\hat{w}}{\partial\hat{z}^2} \tag{4.24}$$

and implies a close balance between the vertical component of the pressure gradient and the buoyancy-acceleration term provided that $L \gg H$ in addition to $\mathrm{Re}_x \gg 1$ and $\mathrm{Re}_z \gg 1$. This allows us to estimate

$$U^2 \approx H\Theta, \tag{4.25}$$

which is a plain, order-of-magnitude equality of the characteristic kinetic and potential energies. Finally, the energy equation (Eq. 4.13) becomes

$$\frac{\partial(\hat{u}\hat{\vartheta})}{\partial\hat{x}} + \frac{\partial(\hat{w}\hat{\vartheta})}{\partial\hat{z}} \approx -\frac{H}{\Theta}\hat{w} + \frac{1}{UL}\frac{\partial^2\hat{\vartheta}}{\partial\hat{x}^2} + \frac{1}{WH}\frac{\partial^2\hat{\vartheta}}{\partial\hat{z}^2}. \tag{4.26}$$

The diffusion terms in Eq. 4.26 are of the order of $1/\mathrm{Pe}_x \equiv 1/(UL)$ and $1/\mathrm{Pe}_z \equiv 1/(WH)$, where Pe_x and Pe_z are Péclet-like numbers associated with horizontal and vertical motions, respectively. We introduce them for the very same reason as we did in the case of Re_x and Re_z. Making use of Eqs. 4.21 and 4.25, we can put Eq. 4.26 into the form

$$\frac{\partial(\hat{u}\hat{\vartheta})}{\partial\hat{x}} + \frac{\partial(\hat{w}\hat{\vartheta})}{\partial\hat{z}} \approx \frac{1}{\mathrm{Pe}_x}\frac{\partial^2\hat{\vartheta}}{\partial\hat{x}^2} + \frac{1}{\mathrm{Pe}_z}\left(-\frac{H^{7/2}}{L\Theta^{1/2}}\hat{w} + \frac{\partial^2\hat{\vartheta}}{\partial\hat{z}^2}\right), \tag{4.27}$$

which can be greatly simplified in the double limit of $\mathrm{Pe}_x \gg \mathrm{Pe}_z$ and $\mathrm{Pe}_z \ll 1$. In that case, the two terms in the parentheses on the right-hand side have to closely balance one another, so that we can estimate

$$H \approx \Theta^{1/7} L^{2/7}, \tag{4.28}$$

and Eq. 4.27 becomes linear,

$$\frac{\partial^2 \hat{\vartheta}}{\partial \hat{z}^2} = \hat{w}. \tag{4.29}$$

Equation 4.29 is a special case of the energy equation in the low-Péclet-number approximation of Lignières (1999).

Using Eq. 4.28, we eliminate H from Eq. 4.25 to get an estimate of $U(\Theta, L)$ and, with Eq. 4.21, also an estimate of $W(\Theta, L)$. The resulting relations also enable us to express Re_x, Re_z, Pe_x, and Pe_z as functions of Θ, L, and Pr. This way we obtain

$$U \approx \Theta^{4/7} L^{1/7}, \tag{4.30}$$

$$W \approx \Theta^{5/7} L^{-4/7}, \tag{4.31}$$

$$\mathrm{Re}_x \approx \Theta^{4/7} L^{8/7} \mathrm{Pr}^{-1}, \tag{4.32}$$

$$\mathrm{Re}_z \approx \Theta^{6/7} L^{-2/7} \mathrm{Pr}^{-1}, \tag{4.33}$$

$$\mathrm{Pe}_x \approx \Theta^{4/7} L^{8/7}, \tag{4.34}$$

$$\mathrm{Pe}_z \approx \Theta^{6/7} L^{-2/7}. \tag{4.35}$$

One might be tempted to estimate the time scale τ of flow acceleration towards the stationary state directly from the buoyancy acceleration Θ provided by the temperature fluctuation imposed on the bottom boundary. It is crucial to realise that, as Eq. 4.24 shows, the buoyancy acceleration is almost completely compensated for by the vertical component of the pressure gradient in the case $L \gg H$. It is only their difference that contributes to the vertical acceleration. We can, however, consider the horizontal acceleration provided by the horizontal component of the pressure gradient and write $U/\tau \approx P/L \approx U^2/L$ (see Eq. 4.23). Using Eq. 4.30 we obtain

$$\tau \approx \Theta^{-4/7} L^{6/7}. \tag{4.36}$$

Finally, we would like to point out that the characteristic thermal-diffusion length scale corresponding to the time scale τ is $\tau^{1/2} \approx \Theta^{-2/7} L^{3/7}$, which scales with Θ and L in quite a different way than H does (see Eq. 4.28). This comes about because our estimates take the back reaction of the flow on the temperature distribution into account.

4.2.3 Numerical solutions

The order-of-magnitude estimates derived in the preceding section assume that the flow is stationary and that there is a well-defined vertical length scale in the flow pattern. We performed a series of time-dependent, numerical simulations of the differential heating problem to confirm these assumptions and to determine how the solutions depend on the Reynolds number and how they decrease with height.

We have developed a specialised code dedicated to the study of the differential-heating problem, because the problem places rather high demands on the numerical scheme. For instance, it has to tackle the highly diffusive nature of the flow and its high aspect ratio and resolve a wide dynamic range within a single simulation box. The code is of the finite-difference type, and it solves the differential-heating equations on a collocated grid using a variant of the MacCormack integration scheme. The Poisson equation for pressure, which can be derived from Eqs. 4.4 and 4.5 (or 4.37, see below), is solved by a spectral method. Heat-diffusion terms are treated implicitly, again by a spectral method. In what follows, we discuss a few selected issues related to the numerical solution of the differential-heating equations that need to be borne in mind when interpreting our results. The reader interested in the details of the numerical scheme is referred to App. A.

We use periodic boundaries in the horizontal direction and force the shear stress and the vertical velocity component to vanish at the lower and upper boundaries of the computational domain. One could also use non-slip boundaries, but these are hardly more akin to the physical reality that motivated this study in the first place, so we omit this case. We impose a temperature fluctuation in the form $\vartheta(x,0) = \Theta \sin(\pi x/L)$ at the bottom boundary and force the temperature fluctuation to vanish at the upper boundary. The parameters Θ and L can be identified with the same symbols as introduced in Sect. 4.2.2.

The high thermal diffusivity in the differential-heating problem forces us to use long implicit time steps for the heat-diffusion terms, which might have an adverse effect on the accuracy of the results. To show that this is not the case, we re-computed the simulations sr03, sr30, and Re1024 (see Tables 4.1 and 4.2 and Sect. 4.3), decreasing the time step by a factor of ten. This brings about a change in the velocity field, which is of the order of 0.1% in the cases sr03 and sr30 and of the order of 1% in the case of Re1024 (measured well away from the field's zeroes). The reason for this insensitivity to the time step is the low Péclet number of the flow. Lignières (1999) shows that in the low-Pe regime, the energy equation can be approximated by a Poisson equation for the temperature fluctuation with w as a source term (see also our Eq. 4.29). We do not use this approximation to

Id.	Θ	L	Pe_x	Pe_z
sr00	10^{0}	10^{1}	8.5×10^{0}	2.5×10^{0}
sr01	10^{0}	10^{2}	1.4×10^{2}	1.3×10^{0}
sr02	10^{0}	10^{3}	2.0×10^{3}	5.5×10^{-1}
sr03	10^{0}	10^{4}	2.9×10^{4}	2.7×10^{-1}
sr10	10^{-1}	10^{1}	2.5×10^{0}	3.1×10^{-1}
sr11	10^{-1}	10^{2}	4.0×10^{1}	1.4×10^{-1}
sr12	10^{-1}	10^{3}	5.5×10^{2}	6.8×10^{-2}
sr13	10^{-1}	10^{4}	7.7×10^{3}	3.5×10^{-2}
sr20	10^{-2}	10^{1}	7.2×10^{-1}	4.1×10^{-2}
sr21	10^{-2}	10^{2}	1.1×10^{1}	1.8×10^{-2}
sr22	10^{-2}	10^{3}	1.5×10^{2}	9.3×10^{-3}
sr23	10^{-2}	10^{4}	2.1×10^{3}	4.8×10^{-3}
sr30	10^{-3}	10^{1}	2.0×10^{-1}	5.4×10^{-3}
sr31	10^{-3}	10^{2}	2.9×10^{0}	2.5×10^{-3}
sr32	10^{-3}	10^{3}	4.0×10^{1}	1.3×10^{-3}
sr33	10^{-3}	10^{4}	5.5×10^{2}	6.7×10^{-4}

Table 4.1: Parameters of the series of simulations sampling a patch of the parameter space $\{\Theta, L\}$ at the constant value of $\mathrm{Re} = 2.6\times10^{2}$.

make our code more versatile; instead, we naturally obtain a close equilibrium between the terms $\nabla^2\vartheta$ and w in Eq. 4.6 when the Péclet number is low. This equilibrium is reached so quickly that details of the evolution of ϑ towards the equilibrium become irrelevant.

It is a well-known fact that any numerical advection scheme either requires adding a so-called artificial-viscosity term to guarantee stability or it involves some viscous behaviour implicitly. In either case, the effective Reynolds number does not even come close to the astrophysical regime with current computing facilities. The artificial viscosity (be it explicit or implicit) thus exceeds the physical one by a wide margin, so it demands special attention.

Suppose we include an explicit viscous term as in Eq. 4.5 to model the artificial viscosity. Equations 4.32 and 4.33 show that for $L \approx 10^3$ (equivalent to $\sim H_p$ in the astrophysical case mentioned in Sect. 4.2.1) we have $\mathrm{Re}_z \approx 10^{-4}\,\mathrm{Re}_x$ as a consequence of $H \ll L$. Using equidistant grids with up to 10^3 grid points in each direction, we can achieve $\mathrm{Re}_x \approx 10^3$. It follows that $\mathrm{Re}_z \lesssim 10^{-1}$ and the vertical momentum transport is dominated by the artificial-viscosity term. A value $\mathrm{Re}_z \gg 1$ is, however, expected in stellar interiors. We use a simple workaround,

Id.	Resolution	Re
Re32	32×32	3.2×10^1
Re64	64×64	6.4×10^1
Re128	128×128	1.3×10^2
Re256	256×256	2.6×10^2
Re512	512×512	5.1×10^2
Re1024	1024×1024	1.0×10^3
Re2048	2048×2048	2.0×10^3
Re4096	4096×4096	4.1×10^3

Table 4.2: Parameters of the series of simulations with Re increasing at the fixed values of $\Theta = 10^{-3}$ and $L = 10^0$.

replacing the viscous term $\mathrm{Pr}\nabla^2\boldsymbol{u}$ by the anisotropic form $\mathrm{Pr}_x\,\partial^2\boldsymbol{u}/\partial x^2 + \mathrm{Pr}_z\,\partial^2\boldsymbol{u}/\partial z^2$. The coefficients Pr_x and Pr_z are re-computed at each time step from the relations $\mathrm{Pr}_x = h_x \max|u|\,\mathrm{Re}_{\mathrm{grid}}^{-1}$ and $\mathrm{Pr}_z = h_z \max|w|\,\mathrm{Re}_{\mathrm{grid}}^{-1}$, where h_x and h_z are the horizontal and vertical grid spacings, respectively, and $\mathrm{Re}_{\mathrm{grid}}$ is the Reynolds number on the grid scale. We performed a few numerical tests of the code on a convection problem to determine that the value $\mathrm{Re}_{\mathrm{grid}} = 4$ is a conservative trade-off between the amount of viscous dissipation and the code's stability, so we use this value in all the simulations presented here.

The anisotropic form of artificial viscosity enables us to reach $\mathrm{Re}_x \gg 1$ and $\mathrm{Re}_z \gg 1$ at the same time on a grid of reasonable size. We show in Sect. 4.3 that the solutions with constant values of Pr_x and Pr_z decay exponentially with height. The effective, local Reynolds number decreases in proportion to the flow speed, and the solutions quickly become dominated by the artificial viscosity. This would also happen in an (otherwise idealised) stellar interior at some point but that point would be much farther from the convection zone's boundary. Therefore, we generalise the artificial-viscosity terms, and the momentum equation (Eq. 4.5) in 2D becomes

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} = -\nabla p + \vartheta\boldsymbol{k} + \frac{\partial}{\partial x}\left[\mathrm{Pr}_x(z)\frac{\partial \boldsymbol{u}}{\partial x}\right] + \frac{\partial}{\partial z}\left[\mathrm{Pr}_z(z)\frac{\partial \boldsymbol{u}}{\partial z}\right], \tag{4.37}$$

where are $\mathrm{Pr}_x(z)$ and $\mathrm{Pr}_z(z)$ are proportional to $\mathrm{e}^{-\eta z}$ with η being an adjustable parameter. We set $\eta = 0$ when we are not interested in the precise vertical profiles and use $\eta > 0$ to suppress the viscous terms when examining how the solutions decrease with height. The latter case, $\eta > 0$, is a rather touchy problem because one has to increase η in a few steps, always using the (almost) stationary flow

from a previous run as an initial condition for the next run. Overestimating the value of η can lead to a lack of viscous dissipation in some parts of the computational domain and a numerical instability ensues. Finally, the very goal that we want to achieve by this treatment, i.e. the flow dynamics' being dominated by inertial terms at great heights, becomes an issue since such a flow evolves on the extremely long time scale corresponding to its low speed.

4.3 Results

4.3.1 The stationarity and structure of the flow

Our numerical investigation of the diffusion-dominated differential heating problem reveals that the flow reaches a stationary state at all values of the Reynolds number that we have been able to achieve (up to Re = 4×10^3). We use the rate of change of the quantity $u_{\max} = \max|u|$ (taken over the whole simulation box) as a convergence monitor and a basis of our criterion for deciding the flow's stationarity. We show in Sect. 4.2.2 that the relevant dynamical time scale near the differentially heated boundary should be close to τ given by Eq. 4.36 (confirmed a posteriori, see below). We pronounce the flow stationary and stop the simulation once the condition

$$\left| \frac{1}{u_{max}} \frac{\partial u_{\max}}{\partial t} \tau \right| < 10^{-3} \tag{4.38}$$

has been met at least for one time scale τ. A direct implementation of this condition would involve extrapolation from the time scale of one time step, Δt, to a much longer time scale τ, which would amplify the round-off noise by a factor of $\tau/\Delta t \gg 1$. Instead, we approximate Eq. 4.38 by

$$\left| \frac{1}{u_{max}} \frac{u_{\max} - \overline{u_{\max}}}{\tau} \tau \right| < 10^{-3}, \tag{4.39}$$

where $\overline{u_{\max}}$ is the Euler-time-stepped solution of the equation

$$\frac{\partial \overline{u_{\max}}}{\partial t} = \frac{u_{\max} - \overline{u_{\max}}}{\tau}. \tag{4.40}$$

Thus, $\overline{u_{\max}}$ is a smoothed version of $u_{\max}$, lagging behind it approximately by τ in time.

The flow in all of the simulations presented in this paper is composed of several layers of overturning cells with flow speed rapidly decreasing from one layer to the next (see Figs. 4.1 or 4.7). We characterise the flow properties close to the differentially heated surface by a vertical length scale H, defined as the height

above the hottest spot at which the flow first turns over (i.e. $w(L/2, H) = 0$), and the typical horizontal and vertical velocity components $U = \frac{1}{2}\max(u)$ and $W = \frac{1}{2}\max(w)$, respectively, where the maxima are taken over the whole simulation box. The symbols H, U, and W can be identified with the same symbols as used in Sect. 4.2.2. The flow always reaches its maximal horizontal speed at the bottom boundary and the maximal vertical speed in the first overturning cell above the hot spot. The flow pattern is asymmetric, with the maximum downward flow speed (reached above the cold spot), $\max(-w)$, always lower than the maximal upward flow speed, $\max(w)$. We define the characteristic numbers Re_x, Re_z, Pe_x, and Pe_z in an analogous way to what is used in Sect. 4.2.2 with the difference that now we have two Prandtl-like numbers Pr_x and Pr_z instead of one Prandtl number Pr.

4.3.2 Scaling relations

We computed a grid of 16 simulations to verify the analytical relations derived in Sect. 4.2.2. All of these simulations, summarised in Table 4.1, have a resolution of 256×512, and the vertical grid spacing was adjusted so as to obtain $\mathrm{Re}_x = \mathrm{Re}_z \equiv \mathrm{Re} = 2.6 \times 10^2$. The decision to fix the value of Re is motivated by the fact that the flow pattern turns out to be scalable over a large part of the parameter space provided that Re = const.. In other words, while changing the heating parameters Θ and L at Re = const. does change the amplitude and the vertical scale of the flow, the structure of the flow, as seen in a system of normalised coordinates x/L and z/H, remains unchanged (see Fig. 4.1). We show in Figs. 4.2 and 4.3 our numerical results as compared with the scaling relations fitted to all but the four data points at $L = 10^1$. The excluded data points do not comply well with the premise $L \gg H$ and are thus expected not to follow the scaling relations. Allowing only the constants of proportionality to change in the fitting process, we obtain

$$H = 1.3\,\Theta^{1/7} L^{2/7}, \tag{4.41}$$

$$U = 0.77\,\Theta^{4/7} L^{1/7}, \tag{4.42}$$

$$W = 2.7\,\Theta^{5/7} L^{-4/7}, \tag{4.43}$$

$$\tau = 0.76\,\Theta^{-4/7} L^{6/7}. \tag{4.44}$$

The unexpectedly good fit is a result of the flow's scalability.

The constants of proportionality in Eqs. 4.41 – 4.44, as well as the structure of the flow, depend on Re. We illustrate this dependence by computing a series of simulations with resolution ranging from 32^2 to 4096^2. This way, we cover about two orders of magnitude in the Reynolds number from 3×10^1 to 4×10^3

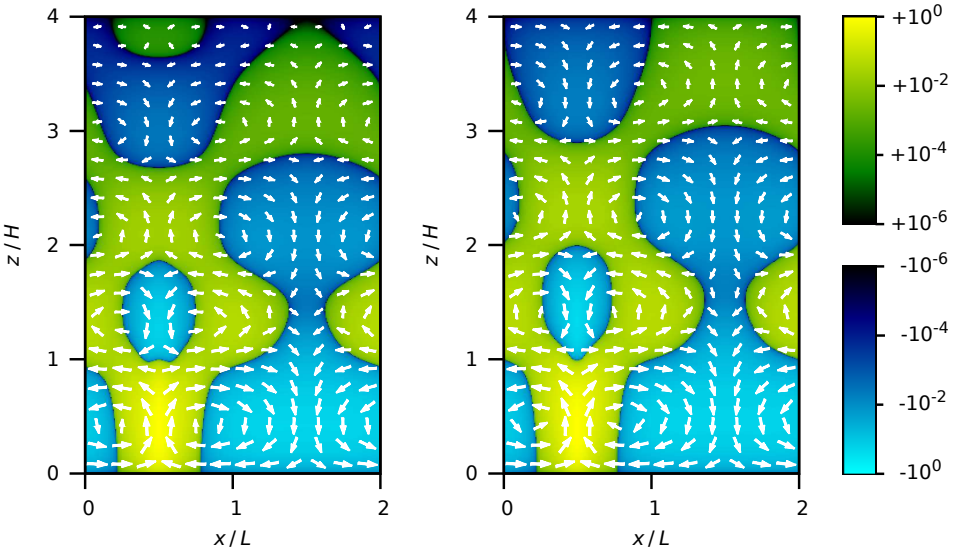


Figure 4.1: Comparison of the flow structure in two simulations with very disparate heating parameters. The left panel shows simulation sr30 ($\Theta = 10^{-3}$, $L = 10^{1}$). The right panel shows simulation sr03 ($\Theta = 10^{0}$, $L = 10^{4}$). In both cases, the vertical velocity component w, normalised to its maximal absolute value, is plotted on a split logarithmic colour scale. The length of the velocity vectors (arrows) is scaled in a non-linear way to aid visualisation. The spatial coordinates are normalised by the characteristic length scales defined in Sects. 4.2.3 and 4.3.

(again with $\mathrm{Re}_x = \mathrm{Re}_z \equiv \mathrm{Re}$). We perform this experiment at $L = \text{const.}$ and $\Theta = \text{const.}$, so any change in Re reflects a change in Pr_x and Pr_z. Nevertheless, we present the dependence on Re, because the scalability of the flow shows that the absolute values of Pr_x and Pr_z do not matter. Ideally, we should choose the heating parameters so as to have $\Theta \ll 1$ and $L \gg 1$ as the scaling relations hold true in this limit (see Sect. 4.2.2).

Equation 4.44 shows, however, that the flow's dynamical time scale becomes extremely long in the same limit, thus making any high-resolution computation unfeasible. Therefore we use $\Theta = 10^{-3}$ and $L = 10^{0}$, which still keeps the energy equation approximately linear (since $\mathrm{Pe}_x \approx \mathrm{Pe}_x \approx 10^{-2}$), but we forgo having $L \gg H$. Nevertheless, we expect the changes in the flow with increasing Re in this case to be similar to those that would be seen in a simulation with $L \gg H$ because of the energy equation's being linear in both cases. All of this series of simulations, summarised in Table 4.2, reach the stationary state as defined by Eq. 4.39.

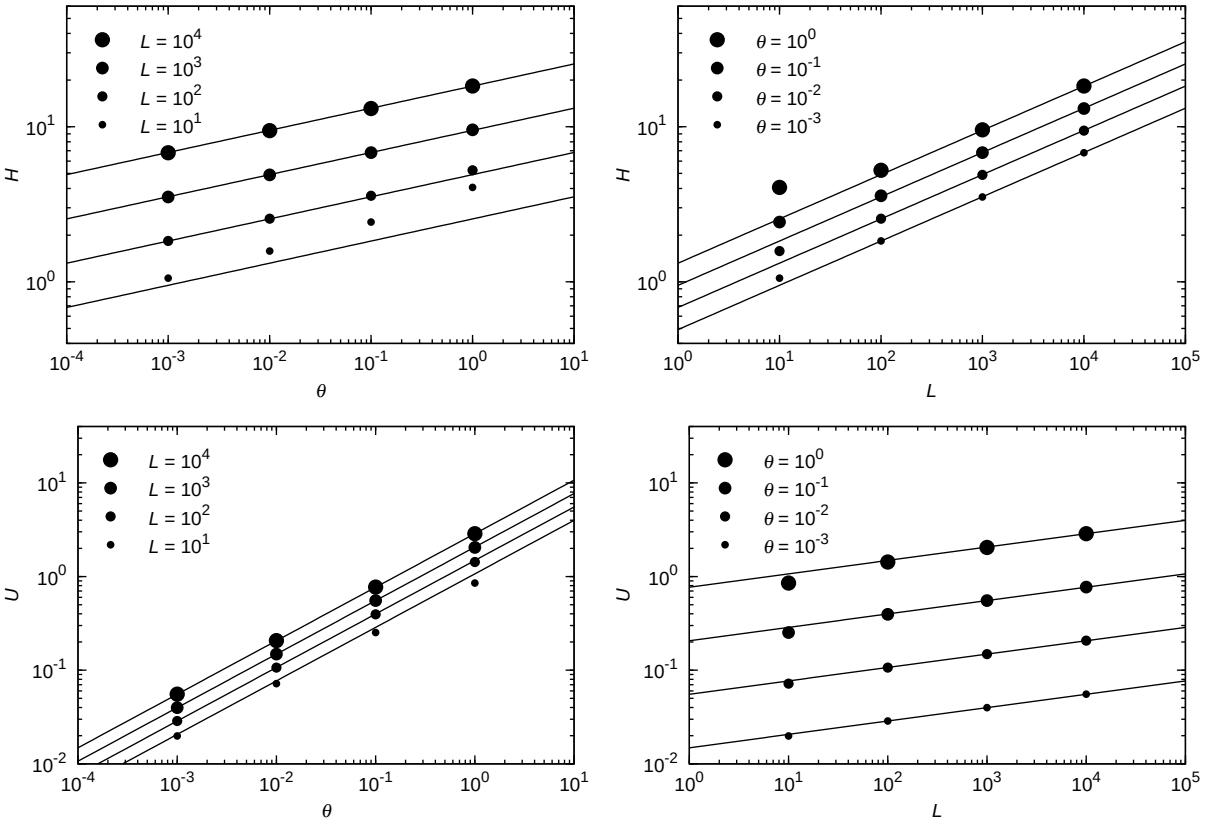


Figure 4.2: Dependence of the global flow characteristics on the heating amplitude Θ and length scale L. Circles show the values derived from numerical simulations (Table 4.1). Solid lines show the scaling relations (Eqs. 4.28 and 4.30), normalised to fit all but the four simulations at $L = 10^1$, which are expected to deviate from the scaling relations.

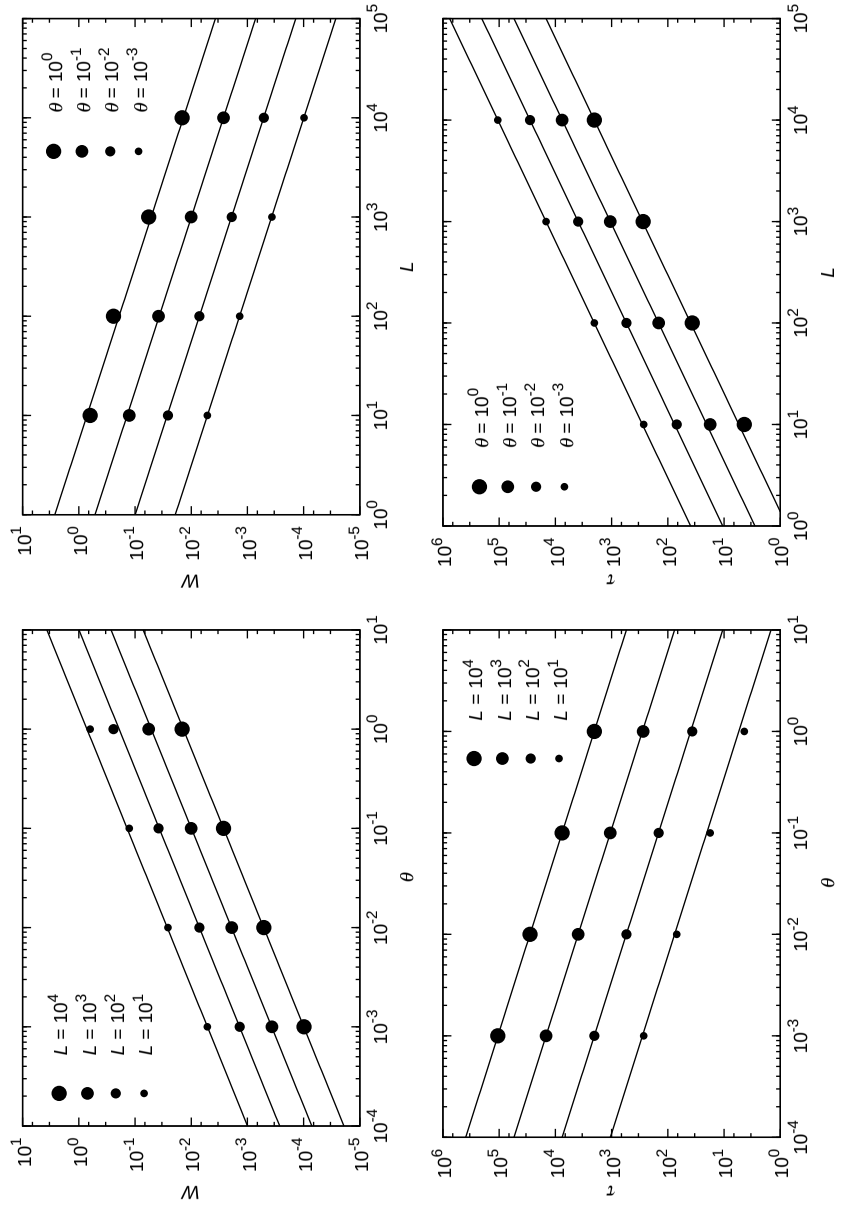


Figure 4.3: Dependence of the global flow characteristics on the heating amplitude Θ and length scale L. Circles show the values derived from numerical simulations (Table 4.1). Solid lines show the scaling relations (Eqs. 4.31 and 4.36), normalised to fit all but the four simulations at $L = 10^1$, which are expected to deviate from the scaling relations.

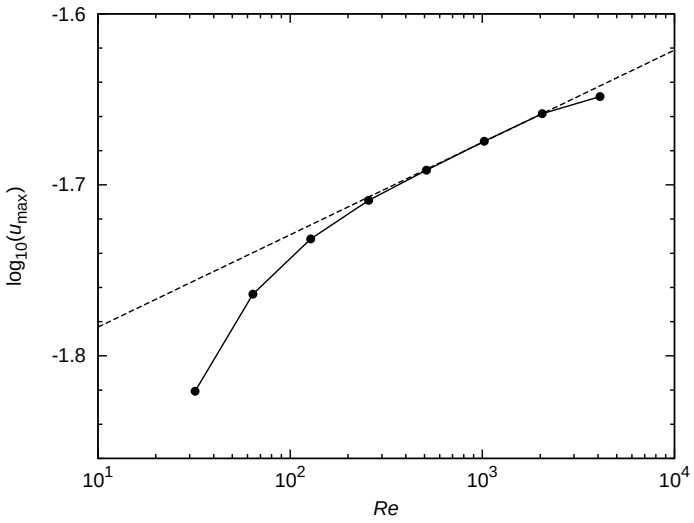


Figure 4.4: Dependence of the maximal horizontal velocity on the Reynolds number. Simulation data (circles) are connected by the solid line to guide the eye. The scaling law $u_{max} \propto Re^{0.054}$ is shown by the dashed line for comparison.

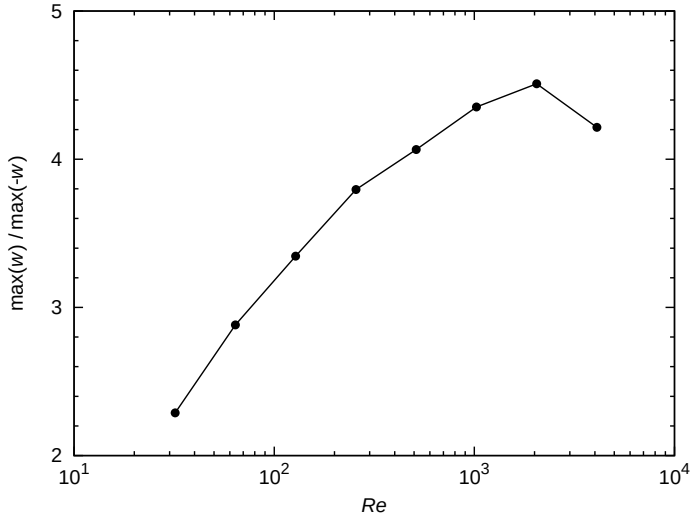


Figure 4.5: Dependence on the Reynold number of the flow's asymmetry, characterised by the ratio of the maximum upward and downward flow speeds.

Figure 4.4 shows that the maximum horizontal velocity in the computational domain slowly increases in proportion to $\mathrm{Re}^{0.054}$ in the high-Re regime. The flow also becomes increasingly asymmetric, as shown by the ratio of the maximum upward and downward flow speeds plotted as a function of Re in Fig. 4.5. The seemingly asymptotic trend changes at the highest Reynolds number considered, but we do not know the reason for this change.

4.3.3 Flow at great heights

The flow speed in all our simulations quickly decreases with height. Figure 4.6 compares the vertical profiles of the root-mean-square (rms; computed in the x direction) vertical velocity component, $w_{\mathrm{rms}}(z)$, in four simulations with widely disparate heating parameters (sr00, sr03, sr30, and sr33). We find that w_{rms} decreases approximately as $e^{-\beta_w z/H}$ in a global sense with $\beta_w \doteq 1.5$ almost independently of Θ and L. We use the values $H(\Theta, L)$ given by Eq. 4.41 instead of those measured in the simulations to normalise the z coordinate, because this brings the slopes much closer to one another. We have to keep in mind, though, that these flows are reasonably close to a stationary state only up to $z/H \doteq 2.5$, because our convergence criterion (Eq. 4.39) is ignorant of the weak flow in the upper part of the simulation box, and consequently, that part of the flow is still slowly evolving when the computation is stopped.

The simulations discussed so far use constant artificial-viscosity parameters Pr_x and Pr_z, which leads to a rapid decrease in the local Reynolds number with height (in proportion to the decreasing flow speed). We computed another two simulations, this time with $\Theta = 10^{-4}$ and $L = 10^1$. In the first one, we set $\mathrm{Pr}_x = \text{const.}$ and $\mathrm{Pr}_z = \text{const.}$ (the constant-Pr case hereinafter), just as we have done so far. In the other one, we set $\mathrm{Pr}_x \propto e^{-\eta z}$ and $\mathrm{Pr}_z \propto e^{-\eta z}$ as described in Sect. 4.2.3 to keep a local version of the Reynolds number approximately constant (the constant-Re case hereinafter). We increased the slope η from 0 in a few steps in order to make the ratio of the rms advection terms to the rms viscous terms, i.e. the local Reynolds number, as independent of height as possible; $\eta = 2$ turns out to be a good compromise in this case. There is a large-scale, residual variation by about a factor of four in the local Reynolds number, because the simple exponential profile of the artificial viscosity is not flexible enough to compensate for it. Using Fig. 4.4 we estimate that this variation can change the velocities by ~ 0.1 dex at most. Since our usual stopping condition, Eq. 4.39, cannot "sense" the weak flow at great heights, we judge the stationarity of the flow by comparing the rms values of the $\partial/\partial t$ terms to the rms values of all other terms that appear in Eq. 4.37 and require the former to be significantly smaller than the latter. This way, we obtain the results summarised in Figs. 4.7, 4.8, and 4.9. The constant-

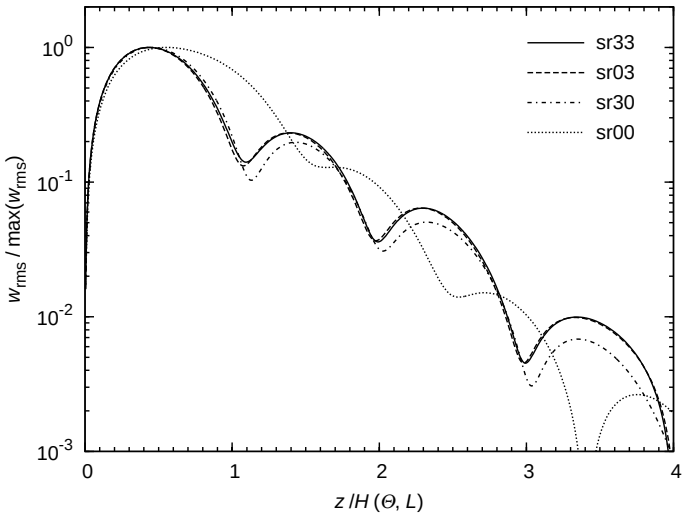


Figure 4.6: Decline with height of the relative rms vertical velocity in four simulations with very different heating parameters. The flows are reasonably close to a stationary state only up to $z/H \doteq 2.5$.

Pr flow can be considered stationary over the whole range shown, whereas the constant-Re flow is only stationary for $z/H \lesssim 3.8$, because the topmost part of that flow evolves so slowly that a global oscillation develops before it has reached equilibrium (see Sect. 4.3.4 for details).

Figure 4.8 illustrates that the flow is somewhat faster at $z/H > 1$ in the constant-Re case, as could be expected from the massive increase in the local Reynolds number by as much as two orders of magnitude at $z \doteq 3$. Much more interesting is, however, that the overturning cells in the constant-Re case become apparently thinner with increasing height, hence with decreasing local temperature fluctuation. This observation suggests that the scaling relations derived in Sect. 4.2.2 could be used locally (see the dependence of H on Θ in Eq. 4.28). Another piece of evidence for this hypothesis is shown in Fig. 4.9, in which we compare the relative rates of decrease in $\vartheta_{\rm rms}(z)$, $u_{\rm rms}(z)$, and $w_{\rm rms}(z)$. The envelope of $\vartheta_{\rm rms}(z)$ can be approximated well by the function $\vartheta_{\rm e}(z) \propto e^{-\beta_\vartheta z/H}$ with $\beta_\vartheta = 1.7$ for $z/H \lesssim 3$. We then regard $\vartheta_{\rm e}(z)$ as an estimate of the local temperature fluctuation and rewrite the scaling relations for the velocity components, Eqs. 4.30 and 4.31, to obtain

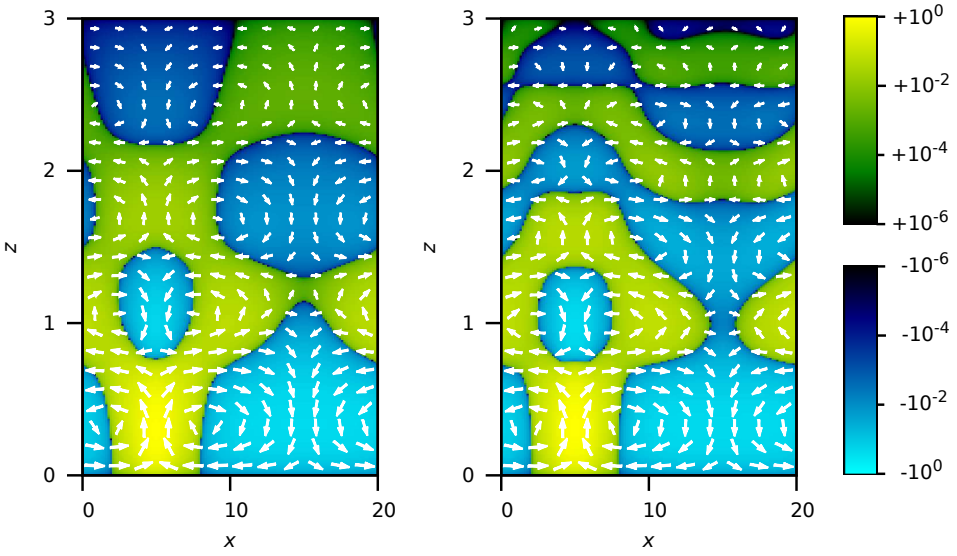


Figure 4.7: Effect of two different artificial-viscosity prescriptions on the flow structure. The constant-Pr flow (Pr_x, Pr_z = const.; left panel) is compared with the constant-Re flow ($Pr_x, Pr_z \propto \mathrm{e}^{-2z}$; right panel). The vertical velocity component w is in both cases plotted on a split-logarithmic colour scale. The length of the velocity vectors (arrows) is scaled in a non-linear way to aid visualisation.

their local versions,

$$u_\mathrm{e}(z) \approx \vartheta_\mathrm{e}(z)^{4/7} L^{1/7}, \tag{4.45}$$

$$w_\mathrm{e}(z) \approx \vartheta_\mathrm{e}(z)^{5/7} L^{-4/7}, \tag{4.46}$$

where $u_\mathrm{e}(z)$ and $w_\mathrm{e}(z)$ are expected to be good envelope models of $u_\mathrm{rms}(z)$ and $w_\mathrm{rms}(z)$. In other words, we expect $u_\mathrm{e}(z) \propto e^{-\beta_u z/H}$ and $w_\mathrm{e}(z) \propto e^{-\beta_w z/H}$ with $\beta_u = \frac{4}{7}\beta_\vartheta$ and $\beta_w = \frac{5}{7}\beta_\vartheta$. Indeed, these scalings turn out to be correct, as shown in Fig. 4.9. Similarly, we can produce a local version of Eq. 4.28,

$$h(z) \approx \vartheta_\mathrm{e}(z)^{1/7} L^{2/7}, \tag{4.47}$$

where $h(z)$ is a local, height-dependent estimate of a vertical length scale analogous to H. As a result, we expect $h(z) \propto e^{-\beta_h z/H}$ with $\beta_h = \frac{1}{7}\beta_\vartheta$, i.e. a slow thinning of the overturning cells with increasing height, similar to what we observe in Figs. 4.7, 4.8, and 4.9. This seemingly innocuous phenomenon has very grave consequences for the flow at great heights. Instead of fading out exponentially, it

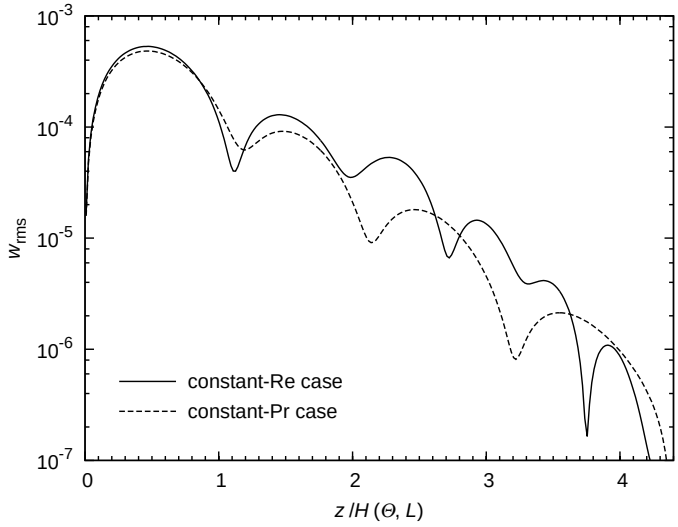


Figure 4.8: Effect of two different artificial-viscosity prescriptions on the rms vertical velocities. Note that the constant-Pr flow can be considered stationary over the whole range shown whereas the constant-Re flow is stationary only up to $z/H \doteq 3.8$.

decreases even faster (see Fig. 4.9). We expand on this in Sect. 4.4.1 and derive a better model for the flow's decline with height to show that the flow speed drops dramatically above a certain point.

4.3.4 Late-time evolution of the flow

Having continued some of our simulations for as much as $10^4\tau$, we discover an intriguing phenomenon. At first, a horizontal mean shear flow develops on top of the differential-heating flow. Its amplitude grows, and the shear flow begins to oscillate at some point. Finally, the oscillation saturates at an amplitude ranging from $\sim 10^{-3}$ to $\sim 10^{0}$ of the differential heating flow's amplitude, depending on the parameters of the simulation. The oscillation's period and development time strongly decrease with increasing Reynolds number. They do not seem to have an upper limit but approach 10τ at $\mathrm{Re} \approx 10^3$. This phenomenon most likely has a physical origin because decreasing the time step by a factor of ten does not affect the shear flow or its behaviour significantly. Any detailed study of this phenomenon is certainly beyond the scope of this paper, but our preliminary research suggests that it is unlikely to be a cumulative effect induced by in-

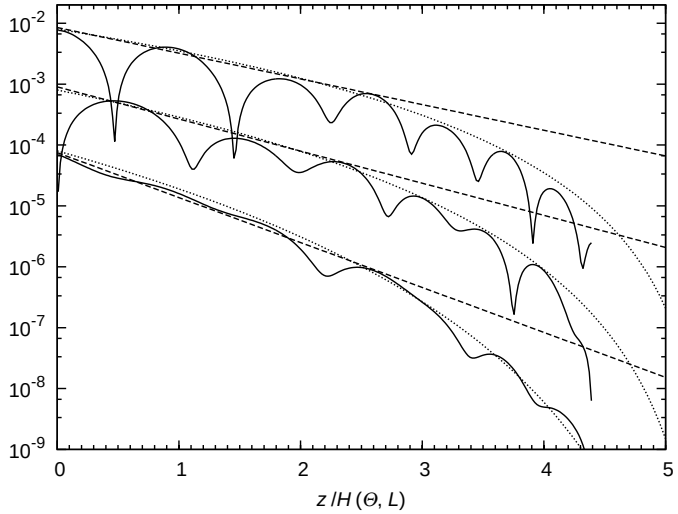


Figure 4.9: Comparison of the rms velocities and temperature fluctuations in the constant-Re case with two models approximating their global behaviour. Solid lines show $u_{\rm rms}$ (top), $w_{\rm rms}$ (middle), and $\vartheta_{\rm rms}$ (bottom). Dashed lines show the model, in which $u_{\rm e}(z) \propto e^{-\beta_u z/H}$, $w_{\rm e}(z) \propto e^{-\beta_w z/H}$, and $\vartheta_{\rm e}(z) \propto e^{-\beta_\vartheta z/H}$ with $\beta_u = \frac{4}{7}\beta_\vartheta$, $\beta_w = \frac{5}{7}\beta_\vartheta$, and $\beta_\vartheta = 1.7$. Dotted lines show the improved model given by Eqs. 4.53, 4.54, and 4.55 with $\gamma = 1.3$. The coefficients of proportionality have been adjusted for each variable independently.

ternal gravity waves since it (1) also occurs in very small computational boxes, in which all internal-wave modes are over-damped by radiative diffusion, and (2) the temporal spectra of the average horizontal velocity are featureless at periods significantly shorter than that of the shear flow oscillation.

4.4 Interpretation of the results

4.4.1 Improving the model at great heights

After picking up the threads of Sect. 4.3.3, we presently find that the diffusion-dominated, high-Re differential-heating flow actually decreases faster than exponentially with height. To see this, we make use of two results from Sect. 4.3.3. First, that the scaling relations derived in Sect. 4.2.2 have their local analogues, which hold within the flow (compare Eqs. 4.28, 4.30, and 4.31 with Eqs. 4.47, 4.45, and 4.46, respectively). Second, that the envelope of $\vartheta_{\rm rms}(z)$ can be approximated

well by the function $\vartheta_e(z) \propto e^{-\beta_\vartheta z/H}$ at low heights, where β_ϑ is independent of Θ and L. This allows us to write

$$\frac{\mathrm{d}\ln\vartheta_e}{\mathrm{d}z} = -\frac{\beta_\vartheta}{H}. \tag{4.48}$$

The characteristic vertical scale H is linked to the heating amplitude Θ by Eq. 4.28 and is thus relevant close to the differentially heated surface, where the typical temperature fluctuations are of the order of Θ. A straightforward generalisation of Eq. 4.48 is obtained by replacing H by the local, height-dependent estimate $h(z)$ given by Eq. 4.47. Upon doing so, we have

$$\frac{\mathrm{d}\ln\vartheta'_e}{\mathrm{d}z'} = -\frac{\beta_\vartheta}{h'(z')}, \tag{4.49}$$

where we have introduced the new variables $\vartheta'_e(z) = \vartheta_e(z)/\Theta$, $z' = z/H$, and $h'(z) = h(z)/H$. By Eqs. 4.28 and 4.47 we have

$$h'(z') \approx \vartheta'_e(z')^{1/7}, \tag{4.50}$$

and Eq. 4.49 becomes

$$\frac{\mathrm{d}\ln\vartheta'_e}{\mathrm{d}z'} = -\beta\vartheta'^{-1/7}_e, \tag{4.51}$$

where β may differ slightly from β_ϑ, because we have used an order-of-magnitude relation in the last step. Equation 4.51 shows that $\ln\vartheta'_e(z)$ decreases with a fairly constant slope over a few orders of magnitude, but the slope starts to change as soon as a wider dynamic range is considered. Since the slope is proportional to $\vartheta'^{-1/7}_e$, Eq. 4.51 describes a runaway process. Indeed, the solution is

$$\vartheta'_e(z') = \left[\vartheta'_e(0)^{1/7} - \frac{\beta}{7}z'\right]^7 \tag{4.52}$$

and vanishes at a finite height of $z'_0 = 7/\beta$. The constant $\vartheta'_e(0)^{1/7}$ must be very close to unity as $\vartheta'_e(0) = \vartheta_e(0)/\Theta \approx 1$, and we can simplify Eq. 4.52 to obtain

$$\vartheta_e(z) \propto \left(1 - \frac{\gamma}{7}\frac{z}{H}\right)^7, \tag{4.53}$$

where we have returned to the non-primed variables and introduced a new constant $\gamma = \beta\vartheta'_e(0)^{-1/7}$, which is a parameter to be adjusted to fit the numerical data. Using the local scaling relations, Eqs. 4.45, 4.46, and 4.47, we derive the functional

dependencies

$$u_e(z) \propto \left(1 - \frac{\gamma}{7}\frac{z}{H}\right)^4, \tag{4.54}$$

$$w_e(z) \propto \left(1 - \frac{\gamma}{7}\frac{z}{H}\right)^5, \tag{4.55}$$

$$h(z) \propto 1 - \frac{\gamma}{7}\frac{z}{H}. \tag{4.56}$$

The functions $\vartheta_e(z)$, $u_e(z)$, and $w_e(z)$ are shown in Fig. 4.9. The constants of proportionality in Eqs. 4.53, 4.54, and 4.55 have been adjusted independently, but all three functions share the value $\gamma = 1.3$. The good fit indicates that our line of reasoning is probably correct.

Can we conclude that the flow stops at the finite height we have just derived? No, since the scaling relations only work in the high-Re regime. Provided that Re is high close to $z = 0$, the flow speed quickly decreases according to Eqs. 4.54 and 4.55 until $\mathrm{Re} \approx 1$ is achieved at some height $z_1 < 7H/\gamma$. The weak flow above this point is supported by viscosity and gradually vanishes as $z \to \infty$.

4.4.2 Allowing for a buoyancy-frequency gradient

So far, we have assumed that the flow occurs in a particularly simple type of thermal stratification — one characterised by a typical buoyancy frequency $N_{\mathrm{typ}} = \mathrm{const}$. Nevertheless, we aim to apply our results to the immediate vicinity of a convection zone, i.e. to a medium, in that the buoyancy frequency rises continuously from zero to a finite value. In this section, we first show how to estimate the value of N_{typ} in such a setting and then reapply the techniques developed in Sect. 4.4.1 to demonstrate how the varying buoyancy frequency affects the global flow field.

To do this, we have to recover the dependence of all the relevant flow properties on N_{typ} by returning to a system of physical units. We recall that we use $1/N_{\mathrm{typ}}$ as a unit of time and $(\varkappa/N_{\mathrm{typ}})^{1/2}$ as a unit of distance, which implies that the unit of velocity is $(\varkappa N_{\mathrm{typ}})^{1/2}$ and the unit of acceleration (hence of ϑ) is $(\varkappa N_{\mathrm{typ}}^3)^{1/2}$. We use these conversion factors throughout this section without mentioning them further. The height of the bottommost overturning cell is by Eq. 4.41

$$H_{\mathrm{ph}} = 1.3\left(\frac{\varkappa}{N_{\mathrm{typ}}^2}\right)^{2/7} \Theta_{\mathrm{ph}}^{1/7} L_{\mathrm{ph}}^{2/7}, \tag{4.57}$$

where we have introduced the index „ph" to indicate the use of physical units for quantities that are dimensionless in the rest of our analysis. The buoyancy

frequency N is now an increasing function of z_{ph} and can be approximated by Eqs. 4.7 and 4.8,

$$N(z_{\mathrm{ph}}) = \left(\frac{\alpha g}{H_p^2}\right)^{1/2} z_{\mathrm{ph}}^{1/2}. \tag{4.58}$$

The overall scale of the flow pattern is given by the bottommost overturning cell, which is thus the most important. Therefore we estimate $N_{\mathrm{typ}} = N(H_{\mathrm{ph}}/2)$, i.e.

$$N_{\mathrm{typ}} = \left(\frac{\alpha g}{H_p^2}\right)^{1/2} \left(\frac{H_{\mathrm{ph}}}{2}\right)^{1/2}, \tag{4.59}$$

and combine Eqs. 4.57 and 4.59 to obtain

$$H_{\mathrm{ph}} \doteq 1.4 \left(\frac{\varkappa^2 H_p^6}{\alpha^2 g}\right)^{1/9} \left(\frac{\Delta T}{T_{\mathrm{m}}}\right)^{1/9} \left(\frac{L_{\mathrm{ph}}}{H_p}\right)^{2/9}, \tag{4.60}$$

where we have also expanded $\Theta_{\mathrm{ph}} = g\Delta T/T_{\mathrm{m}}$ to emphasise the dependence on the imposed temperature fluctuation $\Delta T/T_{\mathrm{m}}$. We use the sign $\doteq$ in Eq. 4.60 and also in Eqs. 4.61, 4.62, and 4.63 below to indicate that we do not expect these estimates to be off by more than a few tens of percent. The dependence of H_{ph} on the heating amplitude and length scale is somewhat weaker in Eq. 4.60 compared with Eq. 4.57, because Eq. 4.60 takes into account that any gain in the flow's vertical extent brings about an increase in the typical buoyancy frequency, which in turn makes further penetration harder. This effect can also be seen when we express the characteristic velocity components and the flow's dynamical time scale in physical units,

$$U_{\mathrm{ph}} \doteq 0.8 \left(\frac{\varkappa g^4 H_p^3}{\alpha}\right)^{1/9} \left(\frac{\Delta T}{T_{\mathrm{m}}}\right)^{5/9} \left(\frac{L_{\mathrm{ph}}}{H_p}\right)^{1/9}, \tag{4.61}$$

$$W_{\mathrm{ph}} \doteq 3 \left(\frac{\varkappa g}{\alpha}\right)^{1/3} \left(\frac{\Delta T}{T_{\mathrm{m}}}\right)^{2/3} \left(\frac{L_{\mathrm{ph}}}{H_p}\right)^{-2/3}, \tag{4.62}$$

$$\tau_{\mathrm{ph}} \doteq 0.7 \left(\frac{\alpha H_p^6}{\varkappa g^4}\right)^{1/9} \left(\frac{\Delta T}{T_{\mathrm{m}}}\right)^{-5/9} \left(\frac{L_{\mathrm{ph}}}{H_p}\right)^{8/9}, \tag{4.63}$$

where the exponents have slightly changed compared with Eqs. 4.42, 4.43, and 4.44.

The spatial variation of N brings on a first-order effect, too; that is, the stratification offers less resistance to overturning in the bottom part of the flow field compared with the rest of it. We mimic this effect by using the flow's excellent

scaling properties under the assumption that the flow behaves locally as if N was constant. Our goal is to improve upon the envelope models of $\vartheta_{\mathrm{rms}}(z)$, $u_{\mathrm{rms}}(z)$, and $w_{\mathrm{rms}}(z)$ derived in Sect. 4.4.1 by taking the dependence of N on height into account.

Our starting point is Eq. 4.49 with the difference that now we define $\vartheta'_{\mathrm{e}} = \vartheta_{\mathrm{e,ph}}/\Theta_{\mathrm{ph}}$, $z' = z_{\mathrm{ph}}/H_{\mathrm{ph}}$ and $h' = h_{\mathrm{ph}}/H_{\mathrm{ph}}$. We caution the reader that $\vartheta_{\mathrm{e,ph}}$ refers to a model with $N = N(z)$ and not to a direct translation of ϑ_{e} that appears in Eq. 4.49 to physical units. The local vertical length scale of the flow, $h(z)$ given by Eq. 4.47, can be translated to physical units directly,

$$h_{\mathrm{ph}} \approx \varkappa^{2/7} N^{-4/7} \vartheta_{\mathrm{e,ph}}^{1/7} L_{\mathrm{ph}}^{2/7}. \tag{4.64}$$

This equation, together with Eq. 4.57, implies

$$h' \approx N'^{-4/7} \vartheta_{\mathrm{e}}'^{1/7}, \tag{4.65}$$

where $N' = N/N_{\mathrm{typ}} = (2z')^{1/2}$ (see Eqs. 4.58 and 4.59). It is evident that h' diverges for $N' \to 0^+$, i.e. $z' \to 0^+$. This effect is purely artificial because the divergence occurs within the bottommost overturning cell of the flow, and the large-scale model we are developing here cannot capture such local phenomena. We ignore the divergence for now because only h'^{-1} appears in Eq. 4.49 and use the same procedure as in Sect. 4.4.1 to derive a generalised version of Eq. 4.51,

$$\frac{\mathrm{d} \ln \vartheta'_{\mathrm{e}}}{\mathrm{d}z'} = -\beta z'^{2/7} \vartheta_{\mathrm{e}}'^{-1/7}, \tag{4.66}$$

where the parameter β has absorbed all coefficients of the order of unity. Its value should still be of the order of unity, but it may be different in this model compared with the model developed in Sect. 4.4.1. By analogy to the derivation in Sect. 4.4.1, we can write the solution to Eq. 4.66 in the form

$$\vartheta_{\mathrm{e,ph}}(z_{\mathrm{ph}}) \propto \left[1 - \frac{\gamma}{9}\left(\frac{z_{\mathrm{ph}}}{H_{\mathrm{ph}}}\right)^{9/7}\right]^7, \tag{4.67}$$

where we have also returned to the non-primed quantities, and $\gamma = \beta \vartheta'_{\mathrm{e}}(0)^{-1/7}$ is a parameter of the order of unity. The typical velocity components and the typical vertical vertical length scale can be estimated using the local scaling relations,

Eqs. 4.45, 4.46, 4.47, and 4.67. We obtain

$$u_{\mathrm{e,ph}}(z_{\mathrm{ph}}) \propto \left[\frac{N(z_{\mathrm{ph}})}{N_{\mathrm{typ}}}\right]^{-2/7}\left[1-\frac{\gamma}{9}\left(\frac{z_{\mathrm{ph}}}{H_{\mathrm{ph}}}\right)^{9/7}\right]^{4}, \tag{4.68}$$

$$w_{\mathrm{e,ph}}(z_{\mathrm{ph}}) \propto \left[\frac{N(z_{\mathrm{ph}})}{N_{\mathrm{typ}}}\right]^{-6/7}\left[1-\frac{\gamma}{9}\left(\frac{z_{\mathrm{ph}}}{H_{\mathrm{ph}}}\right)^{9/7}\right]^{5}, \tag{4.69}$$

$$h_{\mathrm{ph}}(z_{\mathrm{ph}}) \propto \left[\frac{N(z_{\mathrm{ph}})}{N_{\mathrm{typ}}}\right]^{-4/7}\left[1-\frac{\gamma}{9}\left(\frac{z_{\mathrm{ph}}}{H_{\mathrm{ph}}}\right)^{9/7}\right], \tag{4.70}$$

where an explicit dependence on N appears after the transition to physical units. These expressions diverge for $z \to 0^+$ where $N \to 0^+$ (see Eq. 4.58), which is just another illustration of the envelope models' inability to capture local phenomena (see also the discussion above). The bottommost part of the flow should in reality behave approximately as if it was in a medium with $N = N_{\mathrm{typ}} = \text{const.}$, so we can cut off the problematic part of the $N(z)$ profile and use, for example, the function

$$\widetilde{N}(z) = \begin{cases} N_{\mathrm{typ}} & \text{for } 0 \le z_{\mathrm{ph}} \le \frac{1}{2}H_{\mathrm{ph}} \\ N_{\mathrm{typ}}\left(\frac{2z_{\mathrm{ph}}}{H_{\mathrm{ph}}}\right)^{1/2} & \text{for } z_{\mathrm{ph}} > \frac{1}{2}H_{\mathrm{ph}} \end{cases} \tag{4.71}$$

instead of $N(z)$ in practical calculations. Doing so makes the right-hand sides of Eqs. 4.68, 4.69, and 4.70 converge to unity as $z_{\mathrm{ph}} \to 0^+$.

Just as the results of Sect. 4.4.1 do not mean that the flow vanishes at a finite height, neither the results of this section mean that. Again, the sudden drop in the typical velocities predicted by Eqs. 4.68 and 4.69 only signifies that the flow undergoes a transition to the low-Re regime at a relatively low height. Eqs. 4.68 and 4.69 cease to be usable from that point on and the weak flow supported by viscosity gradually vanishes as $z_{\mathrm{ph}} \to \infty$.

4.5 Application to stellar conditions

The flow in a layer of thickness $h_{\mathrm{ph}}(z_{\mathrm{ph}})$ and vertical velocity $w_{\mathrm{e,ph}}(z_{\mathrm{ph}})$ at distance z_{ph} from the boundary overturns a passive tracer in it on a time scale $\tau_{\mathrm{m}} = h_{\mathrm{ph}}/w_{\mathrm{e,ph}}$. This suggests an effective diffusion coefficient $D_{\mathrm{eff}} \approx h_{\mathrm{ph}} w_{\mathrm{e,ph}}$. For the first layer above the boundary, this is

$$D_{\mathrm{eff}}(0) = W_{\mathrm{ph}} H_{\mathrm{ph}}. \tag{4.72}$$

At distance $z_{\rm ph}$, Eqs. 4.69 and 4.70 give

$$D_{\rm eff}(z_{\rm ph}) = D_{\rm eff}(0)\left[\frac{\widetilde{N}(z_{\rm ph})}{N_{\rm typ}}\right]^{-10/7}\left[1-\frac{\gamma}{9}\left(\frac{z_{\rm ph}}{H_{\rm ph}}\right)^{9/7}\right]^{6}, \tag{4.73}$$

where $N_{\rm typ}$ is given by Eq. 4.59 and we have replaced $N(z_{\rm ph})$ in Eqs. 4.69 and 4.70 by $\widetilde{N}(z_{\rm ph})$ given by Eq. 4.71 as discussed in Sect. 4.4.2. The constant γ is of the order of unity but cannot be constrained further by our present analysis. It determines the maximum height $z_{\rm max,ph}$ that the mixing process can reach, $z_{\rm max,ph} = (9/\gamma)^{7/9}H_{\rm ph}$.

For a specific example, consider the boundary of the core convection zone in a $10\,M_\odot$ zero age main sequence star. This environment is characterised by a thermal diffusivity $\varkappa = 5.9\times10^{10}\,{\rm cm^2\,s^{-1}}$, $\alpha = {\rm d}(\nabla_{\rm ad}-\nabla)/{\rm d}(z_{\rm ph}/H_p) = 0.14$, a gravitational acceleration $g = 1.1\times10^5\,{\rm cm\,s^{-2}}$ and a pressure scale height $H_{\rm p} = 2.9\times10^{10}$ cm. A mixing-length estimate for convection in the core produces temperature fluctuations $\Delta T/T_{\rm m}\approx10^{-6}$ on a horizontal length scale $L_{\rm ph}\approx H_{\rm p}$. Equation 4.60 then predicts that the typical vertical length scale is $H_{\rm ph}\approx 2\times10^8\,{\rm cm} = 7\times10^{-3}H_{\rm p}$. The typical vertical velocity (Eq. 4.62) is $W_{\rm ph}\approx 5{\times}10^1\,{\rm cm\,s^{-1}}$. These numbers imply ${\rm Pe}_z = (W_{\rm ph}H_{\rm ph})/\varkappa\approx 2$; i.e., the bottom part of the flow is located right at the transition between the regions of advection-dominated and diffusion-dominated heat transport. This is not a coincidence, because we are modelling the region where heat leaks from the convective eddies, allowing them to turn over and sink back to the convection zone. Such a flow has to have ${\rm Pe}_z\approx 1$. Therefore, the effective diffusivity close to the convection zone, $D_{\rm eff}(0)$ in Eq. 4.73, is of the same order as the diffusivity of heat $\varkappa$. Diffusivities that are several orders of magnitude smaller than $\varkappa$ can be important on the long nuclear time scale. The maximum height reached by the differential heating process on this time scale can thus be approximated by $z_{\rm max,ph}$. Assuming $\gamma = 1$ we obtain $z_{\rm max,ph}\approx 4\times10^{-2}H_{\rm p}$.

Equation (4.73) is likely to be somewhat of an overestimate of the actual mixing rate of the differential-heating process. The layers mix on the hydrodynamic time scale in their interiors, but as long as they are stationary, transport of the tracer between layers takes place by diffusion. As in the case of semiconvective layering (cf. Spruit 2013), this reduces the effective mixing rate to the geometric mean of the microscopic diffusion coefficient $\kappa_{\rm t}$ of the tracer and the estimate (4.73).

More significantly, the picture is complicated by the time dependence of the convective heat source. For the $10M_\odot$ example, only the bottommost part of the flow can approach the stationary flow speed before the heating pattern changes because the dynamical time scale $\tau_{\rm ph}\approx 5\times10^6\,{\rm s}$ (Eq. 4.63 with the parameter values stated above) is of the same order as the convective overturning time scale in the core. This is likely to lead to some form of averaging, reducing the effective

amplitude of the source. The level of this effect can probably be investigated with a time-dependent simulation.

4.6 Summary

Various observations show that there is a need for some additional mixing at the interfaces between the convective and radiative layers of stars. Even processes that are too weak to be detectable in numerical hydrodynamic simulations need to be considered as candidate sources of this mixing, because the nuclear time scale on the main sequence is so much longer than the dynamical time scale of convection, and cumulative effects are likely to play an important role.

In this work, we have investigated one such weak process, which we call "differential heating". The differential heating process occurs when radiative diffusion transports a temperature fluctuation from the boundary of a convection zone into the neighbouring stable stratification. The resulting perturbation of hydrostatic equilibrium triggers a weak flow, which may provide mixing up to some distance from the convection zone. We investigated the flow that is driven by a static temperature fluctuation varying sinusoidally along the solid horizontal boundary of a stably stratified, thin layer of gas. This low-Péclet number problem (i.e. a slow flow dominated by thermal diffusion) turns out to be intrinsically nonlinear, in the sense that the horizontal structure of the flow is asymmetric. Even for symmetric boundary conditions, the upflow is narrower than the downflowing part for the flow, and the shape of the flow pattern is nearly independent of the amplitude of the driving temperature perturbation.

A few additional assumptions (Sect. 4.2) allow us to describe the problem by a set of dimensionless equations, the solution to which depends (apart from the boundary and initial conditions) only on the Prandtl number. We analysed these differential-heating equations for their scaling properties under the assumption that the flow is stationary (Sect. 4.2.2). An astrophysically interesting corner of the parameter space is characterised by $\mathrm{Re}_x \gg 1$, $\mathrm{Re}_z \gg 1$, $\mathrm{Pe}_x \gg \mathrm{Pe}_z$, and $\mathrm{Pe}_z \ll 1$. (The x and z directions have to be distinguished because such flow has a high aspect ratio.) In this limit we derive a set of simple relations (Eqs. 4.28 and 4.30–4.36) to describe how the global flow properties depend on the heating amplitude Θ and length scale L. We find, in particular, that the characteristic vertical length scale H depends only weakly on the heating parameters (Eq. 4.28).

We developed a dedicated numerical code to solve the equations. The main difficulties are related to the highly diffusive nature of the flow, its high aspect ratio, and the need to resolve a wide dynamic range in the flow amplitude within the computational box (as much as five orders of magnitude). The flow in our

two-dimensional, time-dependent simulations reaches a stationary state at all values of the Reynolds number that we have been able to achieve (up to $\mathrm{Re} \equiv \mathrm{Re}_x = \mathrm{Re}_z = 4 \times 10^3$). The flow is always composed of several layers of overturning cells, the shape of which depends only on the Reynolds number and not on the heating length scale L and amplitude Θ. This property makes the flow scaleable in the sense that the flow field corresponding to some heating parameters L_1, Θ_1 can be stretched in space and scaled in amplitude to get a good approximation of the flow field corresponding to a different set of heating parameters L_2, Θ_2 provided that Re is in both cases the same. This is also the reason the scaling relations derived in Eq. 4.2.2 fit the simulation data remarkably well at $\mathrm{Re} = \mathrm{const.}$ (see Figs. 4.2 and 4.3). Increasing the Reynolds number has little influence on the flow speed, but it makes the flow pattern increasingly asymmetric.

We decrease the artificial-viscosity coefficients in the code with height in order to keep the Reynolds number approximately the same in every layer of flow cells. The numerical data show that the global scaling relations derived in Sect. 4.2.2 have their local analogues, which can be used within the flow. The flow speed's decrease with height, being locally exponential, steepens with the decreasing flow amplitude according to the local scaling relations. Based on this we derive a model of the flow's dependence on height, which closely fits the numerical data over the whole dynamic range that we have been able to cover (as much as five orders of magnitude, see Fig. 4.9). The model shows that the flow speed drops abruptly to a negligible value at a finite height. The local scaling relations also allow us to generalise our results to the more realistic case, in which the buoyancy frequency N increases with height (see Sect. 4.4.1).

We illustrated the typical scales associated with the stellar differential-heating process with the example of the convective core of a $10\,M_\odot$ zero-age main sequence star (see Sect. 4.5). We approximate the mixing due to the differential-heating flow by an "effective" diffusion coefficient D_{eff}, which is of the order of the diffusivity of heat near the convection zone and decreases with height according to Eq. 4.73. The mixing relevant for stellar evolution extends about 4% of the pressure scale height above the convection zone.

5 Samenvatting

5.1 Kontext

Het proefschrift houdt zich bezig met konvektie in sterren, zowel konvektie in de buitenlagen van sterren zoals de zon, als in konvektieve zones in het inwendige van sterren. Doel hierbij niet zozeer theorien van konvektie 'an sich', als wel om toepassingen in de sterevolutietheorie. De standaardbeschrijving in de leerboeken zegt dat je alleen maar hoeft te weten waar de temperatuurgradient steiler is dan 'Schwarzschild'. Zo ja, dan zet je de temperatuurgradient gelijk aan de adiabatische gradient. Afgezien van konvektie dicht bij het steroppervlak verwacht je dat dit een uitstekende benadering is. Stermodellen die op deze manier berekend worden vertonen echter diskrepanties met de waarnemingen. Naarmate de waarnemingen precieser worden, blijken de diskrepanties ook hardnekkiger. In vrijwel alle gevallen lijkt het erop alsof konvektieve lagen uitgebreider zijn dan ze volgens 'Schwarzschild' zouden moeten zijn (*dunnere* konvektieve lagen dan verwacht schijnen niet voor te komen[1]).

Alvorens in te gaan op voorgestelde oorzaken en mogelijke technische 'fixes': het beeld is niet erg eenduidig. Om de berekende ontwikkelingsweg in het HRD in overeenstemming met waarnemingen te brengen is het gebruikelijk een 'overshoot' parameter te implementeren die de konvektielaag een fraktie van een drukschaalhoogte uitbreidt. Populair in sterevolutie berekeningen is het gebruik van een 'turbulente diffusiecoefficient', die kunstmatig de samenstelling van de ster mengt. De waarde ervan wordt aangepast tot optimale overeenstemming, en het resultaat heet evolutie met overshooting. De aldus bepaalde overshoot parameter blijkt minder stabiel dan meestal aangenomen: afhankelijk niet alleen van het gebruikte evolutieprogramma, maar ook van de waarnemingen waaraan aangepast wordt (zie de voorbeelden in hoofdstuk 1).

Het is nuttig een onderscheid te maken tussen verschillende soorten diskrepanties. Fouten in een evolutiespoor duiden op een duidelijk verschil in de struktuur van de ster. Daar zijn niet-triviale aanpassingen van de diepte van een kon-

[1] Daarbij even aangenomen dat dit niet een eenvoudig 'file drawer effect' is: het effekt dat niet passende resultaten als onfysisch in de onderste lade van het bureau verdwijnen.

vektieve laag voor nodig (aangenomen dat daar dus de oorzaak in ligt). Veel minder dramatisch zijn anomalien in aan het oppervlak gemeten abondanties van de 'metalen'. In de eerste plaats gaat het daarbij om waarnemingen van de relatieve abondanties van de CNO elementen, zowel in rode als asymptotische reuzensterren. Deze tonen aan dat in sommige evolutiefasen materiaal met elementverhoudingen van de CNO-cyclus vanuit het energie-producerende inwendige naar het oppervlak getransporteerd wordt (tijdens zgn. 'dredgup' fasen). Het transport moet door gebieden plaatsgegrepen hebben die volgens standaard evolutieberekeningen stabiel gelaagd zijn. Het is niet duidelijk hoe dit gebeurt[2], maar het kan in principe een zwak proces zijn dat maar een enkele keer mengt en daar een paar millioen jaar de tijd voor heeft. Dit in tegenstelling tot processen die de thermische struktuur veranderen: om die tegen de thermische relaxatie in in stand te houden is een kontinue bron van energie nodig.

Overwegingen als deze doen vermoeden dat een enkele turbulente diffusie parameter misschien een aantal verschillende fysische processen op een hoop gooit. Voor een betere behandeling van het thema 'overshooting' in de sterevolutie is een wat nauwkeuriger onderzoek van mogelijkerwijs relevante fysische processen noodzakelijk. Daar zijn er een aantal van. Afgezien van direkte manifestaties van de konvektieve stroming zelf (meer daarover in het vervolg) kun je aan afschuifinstabiliteiten tengevolge van differentiele rotatie denken. Verders interne zwaartekrachtsgolven (die grote afstanden kunnen overbruggen), en aan het oppervlak van hete sterren kunnen gravitationele bezinking en differentiele opwaartse krachten in het stralingsveld de abondanties lokaal sterk veranderen ('die französische Diffusion'[3]).

De term 'overshooting' suggereert een dynamisch process: zoiets als konvektieve cellen die met hun kinetische energie in stabiele lagen onder of boven de konvektieve laag dringen. Dit is in de praktijk een verwaarloosbaar effekt. Voor de omstandigheden aan de bodem van de konvektielaag van de Zon bijvoorbeeld levert dit energieargument een overshoot-diepte in de orde van 100 km, verwaarloosbaar in vergelijking met de drukschaalhoogte van 50 000 km. Dit ligt aan de enorme stabiliteit van het stratifikatie van het stralingsinwendige: reeds bij een fraktie van een drukschaalhoogte onder de bodem is de Brunt-Väisälä frekwentie van de orde van 10^{-4} s^{-1}; een faktor 300 sneller dan de tijdschaal van konvektieve bewegingen aan de bodem van de konvektielaag ($\approx 3\,10^6$ s volgens het menglengte formalisme). De energie per massa-eenheid die nodig is om een vloeistofelement in radiele richting over een signifikante fraktie van een schaalhoogte in de stabiele laag te verplaatsen gaat met het kwadraat van dit getal, dwz. is 10^5 maal hoger dan de energiedichtheid van vloeistofbeweginging in de

[2]Een aardige term hiervoor is 'dark mixing' (naar het voorbeeld van de kosmologie).
[3]R. Kippenhahn, 1981.

konvektiezone.

Dit heeft dramatische konsekwenties voor modellen die overshooting beschrijven als een verlenging van dezelfde turbulente diffusie die deze modellen ook voor de konvektiezone zelf gebruiken. Zulke modellen, ook beschreven met termen als 'nonlocal convection' of 'Reynolds averages' zijn gebaseerd op een frappant misverstand. Het is weliswaar juist dat de konvektieve storingen aan de bodem van een konvektieve laag vloeistofbewegingen eronder induceren, en de amplituden daarvan zijn ook vergelijkbaar met die van de konvektie erboven. Maar de aard van deze bewegingen is essentieel anders: het zijn golven, oscillaties die het medium niet mengen maar vloeistofelementen op de zelfde plaats terugbrengen waar ze vandaan komen. Het mengeffekt van deze zwaartekrachtsgolven is subtieler en in de praktijk verwaarloosbaar, in overeenstemming met het energie-argument hierboven.

Er is dus reden het begrip overshooting vanuit een wat fundamentelere hoek te benaderen. Dit inzicht begint geleidelijk sociaal acceptabel te worden, hoewel vaak nog geformuleerd met de weinig zinvolle beschrijving '3-dimensionale sterevolutie'. Door schaarste aan relevant theoretisch werk over de afgelopen tientallen jaren biedt dit tevens een ideale gelegenheid: met relatief bescheiden middelen zijn heel nieuwe, relevante bijdragen tot de theorie van sterevolutie mogelijk. Dit is de motivatie voor het werk in het proefschrift, dat twee verschillende processen onder de loep neemt. Beide passen onder het algemene hoofdje 'konvektieve overshooting' maar behandelen zeer verschillende fysika. Ze hebben ook gemeen dat het om zwakke, langzame processen gaat, bedoeld als mogelijke verklaringen voor Lithium depletie in hoofdreekssterren en bijvoorbeeld de afwijkende C-isotoopverhoudingen in reuzen.

5.2 Konvektieve buitenlagen

Een konvektielaag in het inwendige van een ster heeft een diepte van de orde van een drukschaalhoogte; de omstandigheden aan boven- en onderkant ervan zijn dan hydrodynamisch enigzins vergelijkbaar. In de buitenlagen van sterren is de situatie dramatisch anders. In het geval van de Zon bijvoorbeeld is de dichtheid van het gas aan het oppervlak een miljoen maal lager dan aan de bodem van de konvektiezone. De temperatuurfluktuatie geassocieerd met de konvektieve energieflux is aan de bodem van de konvektiezone van de Zon een paar Kelvin, aan het oppervlak duizend keer zo hoog. De konvektieve stroming in dit geval is extreem asymmetrisch: opstijgende stroming heeft een temperatuurkontrast van de orde 10^{-6}, de dalende intergranulatiestroming een kontrast van 30%. De bijbehorende kontrasten in gasdichtheid zijn ongeveer hetzelfde. Sinds de nume-

rieke simulaties van Nordlund omstreeks 1980 is duidelijk dat konvektie in dit geval niet zinvol met het traditionele menglengte beeld beschreven wordt. Veel relevanter is een twee-komponentenbeeld: de dalende stroming vindt plaats in de vorm van 'pluimen' die met zo'n sterk dichtheidsdeficiet beginnen dat ze ook nog na menging ('entrainment') met 100 - 1000 maal meer omgevingsmateriaal nog zo zwaar zijn dat ze omlaag blijven zinken, tot aan de bodem van de konvektiezone[4]. Alleen opstijgend gas dat nooit met een pluim in aanraking gekomen is heeft een kans het zonsoppervlak te bereiken, en heeft daar dan nog steeds zijn adiabatische temperatuurkontrast van 10^{-6}.

Het werk in hoofdstuk 2 beschrijft een model gebaseerd op dit 2-komponentenbeeld. De konsekwentie van het pluimbeeld is dat dalend gas aan de bodem van de konvektielaag aangekomen een inhomogeen mengsel is, met een verdeling van zeer verschillende temperatuurkontrasten, afhankelijk van individuele verschillen in de entrainment-geschiedenis van de vloeistofelementen. Het is zelfs denkbaar dat een (zeer) kleine fraktie (van de orde 10^{-7}) per stochastisch toeval met niet meer dan een faktor 10 verdund wordt, en daardoor nog een dichtheidskontrast heeft waarmee het tot een fraktie van een drukschaalhoogte onder de nominale bodem van de konvektiezone kan zinken. Voor een ster als de Zon is het met de huidige numerieke middelen volstrekt onmogelijk deze verdeling te voorspellen (extreme lengte- en tijdschaalproblemen). Bij gebrek aan een kwantitatief steekhoudend pluimmodel moet dus geparametriseerd worden. Daartoe wordt in hoofdstuk 2 een gedetailleerd model ontwikkeld. Het laat zien dat de Lithium-depletie van de Zon plausibel met het pluimplaatje te verklaren is. Toegepast op andere hoofdreekssterren met konvektieve buitenlagen maakt het model voorspellingen voor de Li- en Be-depletie. Dit wordt beschreven in hoofdstuk 3. Hoewel natuurlijk afhankelijk van een astrofysische parametrisering, levert het een overtuigende verklaring van het geleidelijk verloop van de Li-depletie met spektraaltyp and leeftijd, en de veel zwakkere Be-depletie.

5.3 Differentiele verwarming

Op de grens tussen een konvektiezone en een stabiele laag treden drukfluktuaties op benevens afschuifstromingen en temperatuurfluktuaties; alle oefenen op hun eigen manier tot op een zekere afstand van de grens hun invloed uit. Ons interesseren vooral de zwakke effekten die op tijdschalen van miljoenen jaren werken. Hydrodynamika direkt aan de grens, die op konvektieve tijdschalen van maanden werkt ligt tegenwoordig goed in het bereik van algemene numerieke hydro-

[4]Een meer symmetrisch maar afgezien daarvan identiek gedrag treedt op in laboratoriumkonvektie en numerieke simulaties ervan.

dynamische simulaties. Maar juist de langzame effekten die op grotere afstand nuttig worden blijven daarmee volstrekt onbereikbaar. Daar zijn gespecialiseerde methoden voor nodig, gebaseerd op diepergaand fysisch onderzoek[5].

Afschuifinstabiliteiten en zwaartekrachtgolven blijken moeilijk. Er is echter nog een op het eerste gezicht eenvoudig proces: de stroming die in een stabiele laag geinduceerd wordt door een konvektieve temperatuurvariatie aan de grens eronder. Hoofdstuk 4 beschrijft dit proces, zowel met analytisch werk als een speciaal voor dit doel ontwikkelde numerieke methode. Het geidealiseerde model neemt een stationaire tempertuurvariatie aan de onderrand van een stabiele laag aan, in de vorm van een sinusvormige afhankelijkheid van de horizontale koordinaten. De intuitieve verwachting is dat dit voor voldoende zwakke amplitude een lineaire respons levert, o.a. dus met dezelfde horizontale afhankelijkheid en iets als een exponentieel afvallen met de hoogte. Het probleem blijkt echter matematisch interessanter. Het is intrinsiek nietlinear, en een mooi voorbeeld van stromingen bij laag 'Péclet-getal' (dwz. gedomineerd door thermische diffusie). In Boussinesque benadering heeft het slechts twee onafhankelijke parameters en het gedrag ervan kan daardoor omvattend ontleed worden. De geinduceerde stroming heeft de vorm van een stapel platte cellen. De amplitude ervan neemt met de hoogte af, maar het stromingspatroon blijft hetzelfde. In plaats van exponentieel af te vallen bereikt de stapel een eindige hoogte. Vertaald van de dimensieloze matematische parameterruimte naar astrofysische omgeving is deze hoogte maximaal in de orde van een paar procent van een drukschaalhoogte, met slechts een zwakke afhankelijkheid van de preciese stellaire omgeving. Hoewel het model dus niet onmiddelijk een overtuigende verklaring voor 'dark mixing' levert is het een fraai stuk matematische fysika gebleken.

[5] Om Martin Schwarzschild te citeren (QJRAS, 1970): 'You may counter me with the question: What can be the problem; since clearly the numerical techniques are in hand? [..., ...] No, I feel we are once again forced to think'.

Bibliography

Aerts, C. 2013, EAS Publications Series, 64, 323

Andrássy, R., & Spruit, H. C. 2013, A&A, 559, A122

Andrássy, R., & Spruit, H. C. 2015a, accepted for publication in A&A

Andrássy, R., & Spruit, H. C. 2015b, to be submitted to A&A

Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481

Bahcall, J. N., Basu, S., & Pinsonneault, M. H. 1998, Physics Letters B, 433, 1

Balachandran, S. 1991, Mem. Soc. Astron. Italiana, 62, 33

Baraffe, I., Chabrier, G., & Gallardo, J. 2009, ApJL, 702, L27

Baumann, P., Ramírez, I., Meléndez, J., Asplund, M., & Lind, K. 2010, A&A, 519, A87

Boesgaard, A. M. 1976, PASP, 88, 353

Boesgaard, A. M., & Tripicco, M. J. 1986, ApJL, 302, L49

Böhm-Vitense, E. 1958, Zeitschrift für Astrophysik, 46, 108

Brandt, T. D., & Huang, C. X. 2015, arXiv:1501.04404

Brummell, N. H., Clune, T. L., & Toomre, J. 2002, ApJ, 570, 825

Christensen-Dalsgaard, J., Dappen, W., Ajukov, S. V., et al. 1996, Science, 272, 1286

Christensen-Dalsgaard, J., & Di Mauro, M. P. 2007, EAS Publications Series, 26, 3

Christensen-Dalsgaard, J., Monteiro, M. J. P. F. G., Rempel, M., & Thompson, M. J. 2011, MNRAS, 414, 1158

Claret, A. 2003, A&A, 399, 1115

Claret, A. 2007, A&A, 475, 1019

Delgado Mena, E., Israelian, G., González Hernández, J. I., Santos, N. C., & Rebolo, R. 2012, ApJ, 746, 47

Deliyannis, C. P., Cunha, K., King, J. R.,& Boesgaard, A. M. 2000, AJ, 119, 2437

Deng, L., & Xiong, D. R. 2008, MNRAS, 386, 1979

Freytag, B., Ludwig, H.-G., & Steffen, M. 1996, A&A, 313, 497

Frigo, M., Johnson, S. G. 2005, in Proceedings of the IEEE 93 (2), 216

Garaud, P., & Acevedo Arreguin, L. 2009, ApJ, 704, 1

Garcia Lopez, R. J., & Spruit, H. C. 1991, ApJ, 377, 268

Gim, M., Vandenberg, D. A., Stetson, P. B., Hesser, J. E., & Zurek, D. R. 1998, PASP, 110, 1318

Guinan, E. F., Ribas, I., Fitzpatrick, E. L., et al. 2000, ApJ, 544, 409

Hansen, C. J., & Kawaler, S. D. 1994, Stellar Interiors. Physical Principles, Structure, and Evolution, 1st ed., Springer-Verlag Berlin Heidelberg New York

Haynes, P. H., McIntyre, M. E., Shepherd, T. G., Marks, C. J., & Shine, K. P. 1991, Journal of Atmospheric Sciences, 48, 651

Herbig, G. H. 1965, ApJ, 141, 588

Heslot, F., Castaing, B., & Libchaber, A. 1987, Phys. Rev. A, 36, 5870

Hurlburt, N. E., Toomre, J., & Massaguer, J. M. 1986, ApJ, 311, 563

Hurlburt, N. E., Toomre, J., Massaguer, J. M., & Zahn, J.-P. 1994, ApJ, 421, 245

Kerr, R. M. 1996, Journal of Fluid Mechanics, 310, 139

King, J. R., Deliyannis, C. P., & Boesgaard, M. 1997,ApJ, 478, 778

Kippenhahn, R., Weigert, A., & Weiss, A. 2012, Stellar Structure and Evolution, 2nd ed., Springer-Verlag Berlin Heidelberg

Kuhfuss, R. 1986, A&A, 160, 116

Kupka, F., & Montgomery, M. H. 2002, MNRAS, 330, L6

Lacy, C. H. S., Torres, G., & Claret, A. 2008, AJ, 135, 1757

Lattuada, M., Pizzone, R. G., Typel, S., et al. 2001, ApJ, 562, 1076

Libchaber, A., Sano, M., & Wu, X. 1990, Physica A Statistical Mechanics and its Applications, 163, 258

Lighthill, J. 1978, Waves in fluids, 1st ed., Cambridge University Press Cambridge

Lignières, F. 1999, A&A, 348, 933

Magic, Z., Collet, R., & Asplund, M. 2013, EAS Publications Series, 63, 367

Magic, Z., Weiss, A., & Asplund, M. 2015, A&A, 573, A89

Marik, D., & Petrovay, K. 2002, A&A, 396, 1011

McIntyre, M. E. 2007, The Solar Tachocline, 183

Meakin, C. A., & Arnett, D. 2007, ApJ, 667, 448

Meléndez, J., Ramírez, I., Casagrande, L., et al. 2010, Ap&SS, 328, 193

Meléndez, J., Bergemann, M., Cohen, J. G., et al. 2012, A&A, 543, AA29

Meléndez, J., Schirbel, L., Monroe, T. R., et al. 2014, A&A, 567, L3

Monroe, T. R., Meléndez, J., Ramírez, I., et al. 2013, ApJL, 774, L32

Montalbán, J. 1994, A&A, 281, 421

Montalbán, J., Miglio, A., Noels, A., et al. 2013, ApJ, 766, 118

Nordlund, A., & Stein, R. F. 1995, Liege International Astrophysical Colloquia, 32, 75

Nordlund, A., & Stein, R. 1997, SCORe'96 : Solar Convection and Oscillations and their Relationship, 225, 79

Pidatella, R. M., & Stix, M. 1986, A&A, 157, 338

Pietrinferni, A., Cassisi, S., Salaris, M., & Castelli, F. 2004, ApJ, 612, 168

Pinsonneault, M. H. 1994, Cool Stars, Stellar Systems, and the Sun, 64, 254

Press, W. H. 1981, ApJ, 245, 286

Ramírez, I., Meléndez, J., Cornejo, D., Roederer, I. U., & Fish, J. R. 2011, ApJ, 740, 76

Rempel, M. 2004, ApJ, 607, 1046

Ribas, I., Jordi, C., & Giménez, Á. 2000, MNRAS, 318, L55

Rogers, T. M., & Glatzmaier, G. A. 2005, ApJ, 620, 432

Rogers, T. M., Glatzmaier, G. A., & Jones, C. A. 2006, ApJ, 653, 765

Rosvick, J. M., & Vandenberg, D. A. 1998, AJ, 115, 1516

Roxburgh, I. W. 1965, MNRAS, 130, 223

Santos, N. C., Israelian, G., Randich, S., García López, R. J., & Rebolo, R. 2004, A&A, 425, 1013

Saslaw, W. C., & Schwarzschild, M. 1965, ApJ, 142, 1468

Schatzman, E. 1996, Journal of Fluid Mechanics, 322, 355

Schlattl, H., & Weiss, A. 1999, A&A, 347, 272

Schmitt, J. H. M. M., Rosner, R., & Bohn, H. U. 1984, ApJ, 282, 316

Schroder, K.-P., Pols, O. R., & Eggleton, P. P. 1997, MNRAS, 285, 696

Sestito, P., & Randich, S. 2005, A&A, 442, 615

Shaviv, G., & Salpeter, E. E. 1973, ApJ, 184, 191

Singh, H. P., Roxburgh, I. W., & Chan, K. L. 1995, A&A, 295, 703

Spiegel, E. A. & Veronis, G. 1960, ApJ, 131, 442

Spiegel, E. A., & Zahn, J.-P. 1992, A&A, 265, 106

Spruit, H. 1997, Mem. Soc. Astron. Italiana, 68, 397

Spruit, H. C. 2013, A&A, 552, A76

Stancliffe, R. J., Fossati, L., Passy, J.-C., & Schneider, F. R. N. 2015, arXiv:1501.05322

Stein, R. F., & Nordlund, A. 1998, ApJ, 499, 914

Tkachenko, A., Degroote, P., Aerts, C., et al. 2014, MNRAS, 438, 3093

Trampedach, R., Stein, R. F., Christensen-Dalsgaard, J., Nordlund, Å., & Asplund, M. 2014, MNRAS, 445, 4366

Tucci Maia, M., Meléndez, J., & Ramírez, I. 2014, ApJL, 790, L25

van Ballegooijen, A. A. 1982, A&A, 113, 99

VandenBerg, D. A.,& Stetson, P. B. 2004, PASP, 116, 997

Viallet, M. 2012, SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, 393

Viallet, M., Meakin, C., Arnett, D., & Mocák, M. 2013, ApJ, 769, 1

Weiss, A., & Schlattl, H. 2008, Ap&SS, 316, 99

Woodward, P. R., Herwig, F., & Lin, P.-H. 2015, ApJ, 798, 49

Xiong, D. R., & Deng, L. 2001, MNRAS, 327, 1137

Xu, Y., Takahashi, K., Goriely, S., et al. 2013, Nuclear Physics A, 918, 61

Zahn, J.-P. 1991, A&A, 252, 179

Zocchi, G., Moses, E., & Libchaber, A. 1990, Physica A Statistical Mechanics and its Applications, 166, 387

Appendix A
Numerical methods

A.1 Integration scheme

We have adapted the standard MacCormack method to suit the differential-heating problem (Chap. 4). In the simplest case of a one-dimensional vector $\boldsymbol{q}$ of conserved quantities being advected on and equidistant grid with a spacing of Δx, MacCormack's method can be written as

$$\boldsymbol{q}_k^{(1)} = \boldsymbol{q}_k^n - \Delta t \frac{\boldsymbol{f}\left(\boldsymbol{q}_{k+1}^n\right) - \boldsymbol{f}\left(\boldsymbol{q}_k^n\right)}{\Delta x}, \tag{A.1}$$

$$\boldsymbol{q}_k^{(2)} = \boldsymbol{q}_k^{(1)} - \Delta t \frac{\boldsymbol{f}\left(\boldsymbol{q}_k^{(1)}\right) - \boldsymbol{f}\left(\boldsymbol{q}_{k-1}^{(1)}\right)}{\Delta x}, \tag{A.2}$$

$$\boldsymbol{q}_k^{n+1} = \frac{\boldsymbol{q}_k^n + \boldsymbol{q}_k^{(2)}}{2}, \tag{A.3}$$

where $\boldsymbol{q}_k^n$ is the value of $\boldsymbol{q}$ at the k-th grid point and the n-th time step, Δt the time step, $\boldsymbol{f}(\boldsymbol{q})$ the flux function, and we use the convention that any parenthesised upper index refers to a sub-step of the method instead of a time-step index. The method is linearly stable provided that the CFL condition $\Delta t \leq \Delta x/\rho(A)$ is met, where A is the Jacobian matrix of the flux vector and $\rho(A)$ is the largest characteristic value of A. Nonlinear stability typically requires the addition of some form of artificial viscosity. MacCormack's method is second-order accurate both in space and time.

We discretise Eqs. 4.4, 4.6, and 4.37 on a collocated, two-dimensional grid of $M \times N$ cells with constant cell spacing (Δx, Δz). The two spatial dimensions and the presence of source terms in the equations forces us to significantly extend the basic MacCormack scheme. We begin by advecting the vector of variables

$\boldsymbol{q} = (u, w, \vartheta)$ in both spatial directions using Strang splitting,

$$\boldsymbol{q}_{k,l}^{(1a)} = \boldsymbol{q}_{k,l}^{n} - \frac{\Delta t}{2} \frac{u_{k+1,l}^{n} \boldsymbol{q}_{k+1,l}^{n} - u_{k,l}^{n} \boldsymbol{q}_{k,l}^{n}}{\Delta x}, \tag{A.4}$$

$$\boldsymbol{q}_{k,l}^{(1b)} = \boldsymbol{q}_{k,l}^{(1a)} - \Delta t \frac{w_{k,l+1}^{(1a)} \boldsymbol{q}_{k,l+1}^{(1a)} - w_{k,l}^{(1a)} \boldsymbol{q}_{k,l}^{(1a)}}{\Delta z}, \tag{A.5}$$

$$\boldsymbol{q}_{k,l}^{(1c)} = \boldsymbol{q}_{k,l}^{(1b)} - \frac{\Delta t}{2} \frac{u_{k+1,l}^{(1b)} \boldsymbol{q}_{k+1,l}^{(1b)} - u_{k,l}^{(1b)} \boldsymbol{q}_{k,l}^{(1b)}}{\Delta x}, \tag{A.6}$$

where we have written out the explicit form of the flux terms. The indices k and l refer to the position along the x and z axes, respectively. We proceed by adding the source terms to the momentum equations,

$$u_{k,l}^{(1d)} = u_{k,l}^{(1c)} + \Delta t \left[-\frac{p_{k+1,l}^{n} - p_{k-1,l}^{n}}{2\Delta x} + \nu_l \frac{u_{k-1,l}^{(1c)} - 2u_{k,l}^{(1c)} + u_{k+1,l}^{(1c)}}{(\Delta x)^2} + \frac{\mu_{l+1/2} \left(u_{k,l+1}^{(1c)} - u_{k,l}^{(1c)} \right) - \mu_{l-1/2} \left(u_{k,l}^{(1c)} - u_{k,l-1}^{(1c)} \right)}{(\Delta z)^2} \right], \tag{A.7}$$

$$w_{k,l}^{(1d)} = w_{k,l}^{(1c)} + \Delta t \left[-\frac{p_{k,l+1}^{n} - p_{k,l-1}^{n}}{2\Delta z} + \vartheta_{k,l}^{(1c)} + \nu_l \frac{w_{k-1,l}^{(1c)} - 2w_{k,l}^{(1c)} + w_{k+1,l}^{(1c)}}{(\Delta x)^2} + \frac{\mu_{l+1/2} \left(w_{k,l+1}^{(1c)} - w_{k,l}^{(1c)} \right) - \mu_{l-1/2} \left(w_{k,l}^{(1c)} - w_{k,l-1}^{(1c)} \right)}{(\Delta z)^2} \right], \tag{A.8}$$

where we use second-order-accurate central differences to keep up with the order of accuracy of the advection scheme, $\nu_l = \mathrm{Pr}_x(z_l)$ and $\mu_l = \mathrm{Pr}_z(z_l)$ are the coefficients of our anisotropic artificial-viscosity prescription (see Sect. 4.2.3), and $\mu_{l+1/2} = (\mu_l + \mu_{l+1})/2$. The new velocity field $\boldsymbol{u}^{(1d)} = \left(u^{(1d)}, w^{(1d)} \right)$ is, in general, slightly divergent. We correct for this divergence by subtracting the gradient of a pressure-correction field, $\boldsymbol{u}^{(1)} = \boldsymbol{u}^{(1d)} - \Delta t \nabla (\Delta p)^{(1)}$. The condition $\nabla \cdot \boldsymbol{u}^{(1)} = 0$ leads to a Poisson equation for the pressure correction,

$$\nabla^2 (\Delta p)^{(1)} = \frac{\nabla \cdot \boldsymbol{u}^{(1d)}}{\Delta t}. \tag{A.9}$$

Since we use central differences to compute partial derivatives, the discrete form of the Laplace operator in Eq. A.9 should be derived by applying the central differences twice. That would, however, lead to a sparse operator and cause odd-even-decoupling problems on our collocated grid. Therefore we use the standard

compact Laplacian and solve the approximate pressure-correction equation

$$\frac{(\Delta p)^{(1)}_{k-1,l} - 2(\Delta p)^{(1)}_{k,l} + (\Delta p)^{(1)}_{k+1,l}}{(\Delta x)^2} + \frac{(\Delta p)^{(1)}_{k,l-1} - 2(\Delta p)^{(1)}_{k,l} + (\Delta p)^{(1)}_{k,l+1}}{(\Delta z)^2} = \frac{1}{\Delta t}\left[\frac{u^{(1\mathrm{d})}_{k+1,l} - u^{(1\mathrm{d})}_{k-1,l}}{2\Delta x} + \frac{w^{(1\mathrm{d})}_{k,l+1} - w^{(1\mathrm{d})}_{k,l-1}}{2\Delta z}\right]. \quad \text{(A.10)}$$

Equation A.10 is solved by a spectral solver, see Sect. A.3. Having computed the pressure correction, we apply it to the velocity field,

$$u^{(1)}_{k,l} = u^{(1\mathrm{d})}_{k,l} - \Delta t \frac{(\Delta p)^{(1)}_{k+1,l} - (\Delta p)^{(1)}_{k-1,l}}{2\Delta x}, \quad \text{(A.11)}$$

$$w^{(1)}_{k,l} = w^{(1\mathrm{d})}_{k,l} - \Delta t \frac{(\Delta p)^{(1)}_{k,l+1} - (\Delta p)^{(1)}_{k,l-1}}{2\Delta z}. \quad \text{(A.12)}$$

The approximate nature of the pressure-correction equation (Eq. A.10) causes $\nabla \cdot \boldsymbol{u}^{(1)}$ to be small, but non-zero. Practical experience has shown that the residual divergence is negligibly small in the flows analysed in this paper provided that the boundary conditions are treated properly, see Sect. A.2. We should also write $p^{(1)}_{k,l} = p^n_{k,l} + (\Delta p)^{(1)}_{k,l}$ at this point, but our numerical tests have shown that the residual divergence in the velocity field becomes much smaller if we set $p^{(1)}_{k,l} = p^n_{k,l}$, so we use the latter form. The next step is to integrate the remaining two terms in the energy equation. We begin by adding the $-w$ term,

$$\vartheta^{(1\mathrm{d})}_{k,l} = \vartheta^{(1\mathrm{c})}_{k,l} - \Delta t\, w^{(1)}_{k,l}, \quad \text{(A.13)}$$

where its latest available value, $-w^{(1)}$, has been used. The diffusion sub-step is given by the implicit equation

$$\vartheta^{(1)}_{k,l} = \vartheta^{(1\mathrm{d})}_{k,l} + \Delta t\left[\frac{\vartheta^{(1)}_{k-1,l} - 2\vartheta^{(1)}_{k,l} + \vartheta^{(1)}_{k+1,l}}{(\Delta x)^2} + \frac{\vartheta^{(1)}_{k,l-1} - 2\vartheta^{(1)}_{k,l} + \vartheta^{(1)}_{k,l+1}}{(\Delta z)^2}\right], \quad \text{(A.14)}$$

which is also solved by a spectral solver, see Sect. A.3. We have thus completed the first step of the MacCormack scheme, analogous to Eq. A.1, and obtained the new variables $u^{(1)}$, $w^{(1)}$, $p^{(1)}$, and $\vartheta^{(1)}$. The second step, which we do not do not go into detail on, differs from the first one at two points. First, advection is done using backward-space flux differencing, as in Eq. A.2 (compare with Eq. A.1). Second, the pressure field is updated in this step, i.e. $p^{(2)}_{k,l} = p^{(1)}_{k,l} + (\Delta p)^{(2)}_{k,l}$. The final step of the MacCormack's scheme, Eq. A.3, is used in the same form, with $\boldsymbol{q}$ =

(u, w, ϑ). We also update the pressure field in the same way, $p_{k,l}^{(n+1)} = \frac{1}{2}\left(p_{k,l}^{n} + p_{k,l}^{(2)}\right)$, so that we obtain an estimate of the pressure field for the next time step.

Finally, there is a simple way of increasing the accuracy of the scheme at a given grid resolution, which we use. The MacCormack method contains a built-in asymmetry: Eqs. A.1 and A.2 show that it always starts with forward-space flux differencing and continues with backward-space flux differencing. The two flux-differencing methods can be reversed, obtaining a "reverse" MacCormack method, without decreasing the order of accuracy of the overall scheme. We compute every time step using both the "direct" and the "reverse" methods and use the arithmetic average of the estimates given by the two methods.

A.2 Boundary conditions

The treatment of boundaries is restricted by our decision to use spectral solvers, which do not allow changing the differentiation operators anywhere in the computation domain. We use the ghost-cell technique for this reason. The boundary conditions we impose on the differential-heating flow are summarised in Sect. 4.2.3. The periodic boundaries in the horizontal direction are trivial to implement. The solid boundaries on the top and bottom of the computational domain, however, require much more care. We implement them using reflective boundary conditions for the velocity vector,

$$u_{k,-1} = u_{k,0}, \tag{A.15}$$

$$u_{k,N} = u_{k,N-1}, \tag{A.16}$$

$$w_{k,-1} = -w_{k,0}, \tag{A.17}$$

$$w_{k,N} = -w_{k,N-1}, \tag{A.18}$$

so that the imaginary walls are located at $l = -1/2$ and at $l = N - 1/2$. The conditions imposed on u also eliminate any shear on the boundary. The pressure field is required to be symmetric with respect to the solid boundaries,

$$p_{k,-1} = p_{k,0}, \tag{A.19}$$

$$p_{k,N} = p_{k,N-1}. \tag{A.20}$$

The conditions imposed by Eqs. A.15–A.20 can easily be shown to be consistent with the pressure-correction equation (Eq. A.10; sum both sides over $k = 0, 1, \ldots, M$ and $l = 0, 1, \ldots, N$). They typically do, however, bring about a cusp in the pressure field along the normal to the walls. The resulting discontinuity in the vertical pressure gradient then propagates to the rest of the domain and can

be seen as a low-amplitude oscillatory field superimposed on the true pressure field (see the left panel of Fig. A.1). We tried to cure this problem by changing the discretisation of the vertical-gradient operator at the walls, so that the ghost cells would not be used when computing the pressure gradient. This solution has met with very little success, most likely because the abrupt change in the operator brings about an abrupt change in the discretisation error so the problem remains. Quite surprisingly, preceding the pressure-gradient computation by high-order pressure extrapolation to the ghost cells has turned out to be an effective solution, able to eliminate nearly all of the spurious oscillations (see the middle and right panels of Fig. A.1). We therefore use sixth-order extrapolation in the simulations with constant artificial viscosity and increase the extrapolation order to ten when we let the artificial viscosity decrease with height. This technique cannot be viewed, however, as an all-purpose solution, because it is likely to be too unstable to be useful when computing highly turbulent flows.

We require the temperature fluctuation ϑ to have a fixed sinusoidal profile at the bottom boundary and to vanish at the upper boundary, which translates into

$$\vartheta_{k,-1} = -\vartheta_{k,0} + 2\Theta \sin\left(\frac{\pi x_k}{L}\right), \tag{A.21}$$

$$\vartheta_{k,N} = -\vartheta_{k,N-1}. \tag{A.22}$$

A.3 Spectral solvers

We use spectral methods to solve the two equations involving the Laplace operator, the Poisson equation for the pressure-correction equation (Eq. A.10) and the implicit heat-diffusion equation (Eq. A.14). We express both the knowns and unknowns as linear combinations of the Laplacian's eigenfunctions that comply with the desired boundary conditions. The solution procedure is then much simplified and effective, provided that the transform to the eigenfunction basis can be computed efficiently.

In case of the pressure-correction equation (Eq. A.10), we use the linear transform

$$\hat{f}_{m,n} = \frac{1}{2MN} \sum_{k=0}^{M-1} \left[2 \sum_{l=0}^{N-1} f_{k,l} \cos\left(\frac{\pi n \left(l + \frac{1}{2}\right)}{N} \right) \right] \exp\left(-\frac{2\pi i m k}{M} \right) \tag{A.23}$$

and its inverse

$$f_{k,l} = \sum_{m=0}^{M-1} \left[\hat{f}_{m,0} + 2 \sum_{n=1}^{N-1} \hat{f}_{m,n} \cos\left(\frac{\pi n \left(l + \frac{1}{2}\right)}{N} \right) \right] \exp\left(\frac{2\pi i m k}{M} \right) \tag{A.24}$$

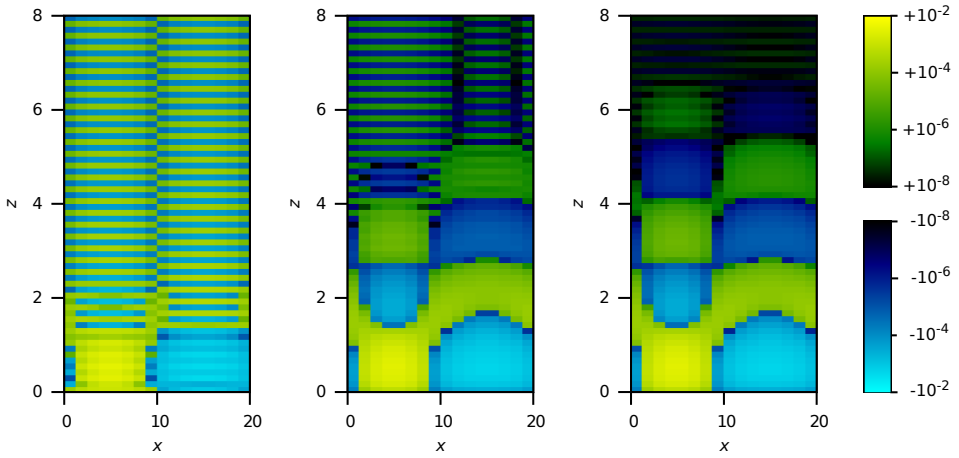


Figure A.1: Effect of three different methods of treating the pressure at the solid top and bottom boundaries. In all three panels, the vertical velocity component, w, is plotted on a split logarithmic colour scale. We use $\Theta = 10^{-3}$, $L = 10^{1}$, constant kinematic viscosity and set the resolution to only 16×64 to make the spurious oscillations visible. We obtain the result plotted in the left panel using the simple symmetry conditions for pressure (Eqs. A.19 and A.20). Preceding the pressure-gradient computation by third-order pressure extrapolation to the ghost cells reduces the oscillations' amplitude by a factor of ~ 100 (middle panel). Increasing the extrapolation order to six brings about another decrease by a factor of ~ 30 in the oscillations' amplitude (right panel). The pressure gradient is in all three cases computed by the second-order central differences in the whole computational domain.

to transform any field $f_{k,l}$ to an array of complex amplitudes $\hat{f}_{m,n}$ and back. We can see that the basis functions in Eq. A.24 are periodic in k and even around $l = -1/2$ and $l = N - 1/2$; i.e., they comply with our boundary conditions on the pressure field (see Sect. A.2). Upon using the spectral decomposition defined by Eq. A.24 on both sides of the pressure-correction equation (Eq. A.10), we readily obtain its solution in the wavenumber space,

$$(\Delta\hat{p})_{m,n} = \frac{\hat{S}_{m,n}}{\lambda_{m,n}}, \tag{A.25}$$

where we have omitted the upper indices because the expression applies to both steps of the MacCormack scheme, $\hat{S}_{k,l}$ is the transformed right-hand side

of Eq. A.10. The eigenvalues $\lambda_{m,n}$ of the Laplacian are

$$\lambda_{m,n} = -\frac{2 - 2\cos\left(\frac{2\pi m}{M}\right)}{(\Delta x)^2} - \frac{2 - 2\cos\left(\frac{\pi n}{N}\right)}{(\Delta z)^2} \tag{A.26}$$

and can be pre-computed. We set $\lambda_{0,0}$ to a large number to prevent division by zero and make the undetermined component $(\Delta\hat{p})_{0,0}$ vanish.

In case of the heat-diffusion equation (Eq. A.14), we use the linear transform

$$\hat{g}_{m,n} = \frac{1}{2MN} \sum_{k=0}^{M-1} \left[2 \sum_{l=0}^{N-1} g_{k,l} \sin\left(\frac{\pi\,(n+1)\left(l+\frac{1}{2}\right)}{N} \right) \right] \exp\left(-\frac{2\pi imk}{M} \right) \tag{A.27}$$

and its inverse

$$g_{k,l} = \sum_{m=0}^{M-1} \left[(-1)^l\, \hat{g}_{m,N-1} + 2 \sum_{n=0}^{N-2} \hat{g}_{m,n} \sin\left(\frac{\pi\,(n+1)\left(l+\frac{1}{2}\right)}{N} \right) \right] \exp\left(\frac{2\pi imk}{M} \right) \tag{A.28}$$

to transform any field $g_{k,l}$ to an array of complex amplitudes $\hat{g}_{m,n}$ and back. We can see that the basis functions in Eq. A.28 are periodic in k and odd around $l = -1/2$ and $l = N - 1/2$; i.e., they comply with our boundary conditions on the temperature field in case of a vanishing heating amplitude (see Sect. A.2). To allow for an arbitrary heating profile at the bottom boundary, we take out the known boundary term from the Laplacian on the right-hand side of Eq. A.14 and treat it as a source term. One can show that it is the same as replacing the diffusion equation $\partial\vartheta/\partial t = \nabla^2\vartheta$ by the equivalent equation $\partial(\vartheta - \zeta)/\partial t = \nabla^2(\vartheta - \zeta)$, where ζ is the static solution to the diffusion equation $\partial\zeta/\partial t = \nabla^2\zeta$ with the desired boundary conditions (ζ can be pre-computed for a fixed heating profile). The boundary conditions on the difference $\vartheta - \zeta$ are then identically zero, and the spectral decomposition defined by Eq. A.28 can be used. This way we obtain an explicit expression for the solution of the implicit Eq. A.14 in the wavenumber space,

$$\hat{\vartheta}^{(1)}_{m,n} = \frac{\hat{\vartheta}^{(1\mathrm{d})}_{m,n} - \hat{\zeta}_{m,n}}{1 - \Delta t \Lambda_{m,n}} + \hat{\zeta}_{m,n}, \tag{A.29}$$

where the eigenvalues $\Lambda_{m,n}$ of the Laplacian are

$$\Lambda_{m,n} = -\frac{2 - 2\cos\left(\frac{2\pi m}{M}\right)}{(\Delta x)^2} - \frac{2 - 2\cos\left(\frac{\pi(n+1)}{N}\right)}{(\Delta z)^2} \tag{A.30}$$

and can be pre-computed. An equation analogous to Eq. A.29 relates $\hat{\vartheta}^{(2)}$ to $\hat{\vartheta}^{(2\mathrm{d})}$. The typical distribution of the temperature fluctuation $\vartheta(x, z)$ in a stationary differential-heating flow computed by this algorithm is shown in Fig. A.2.

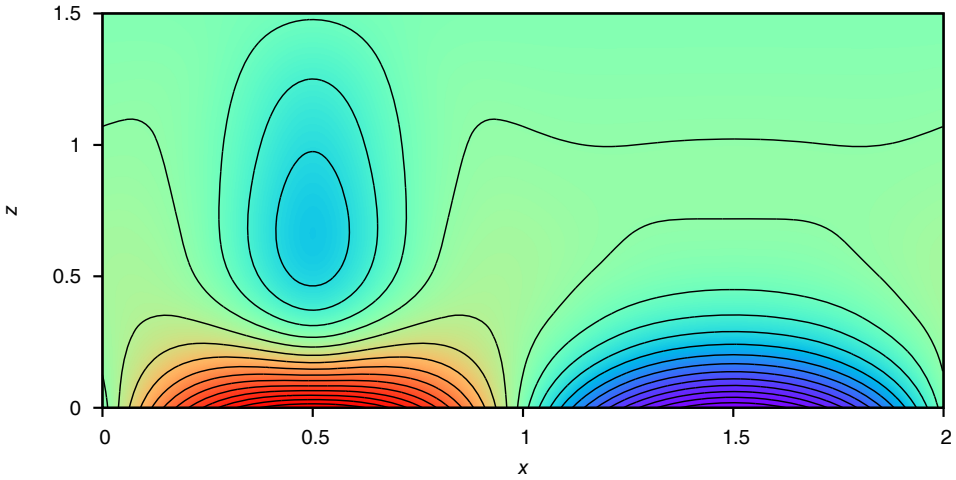


Figure A.2: Distribution of the temperature fluctuation ϑ in simulation Re4096 (Table 4.2). The colour scale encodes values from $\vartheta = -10^{-3}$ (blue) through $\vartheta = 0$ (green) to $\vartheta = +10^{-3}$ (red). The contours have a spacing of $\Delta\vartheta = 8 \times 10^{-5}$.

In the practical implementation, we use the FFTW library (Frigo & Johnson 2005) to compute the transforms in Eqs. A.23, A.24, A.27, and A.28. We combine standard, one-dimensional transforms of different kinds to obtain the non-standard, two-dimensional transforms that we need. Namely, Eq. A.23 is implemented as a series of DCT-II transforms over the rows of the input array, after which the columns of the resulting array are transformed by a series of DTF transforms. The backward transform (Eq. A.24) is then computed by a series of DFTs followed by a series of DCT-IIIs. The transforms for the diffusion equation (Eq. A.27 and A.28) are implemented in the same way, but simply replacing the DCT-IIs by DST-IIs and DCT-IIIs by DST-IIIs. The transforms from the FFTW library do not include the normalisation factor $(2MN)^{-1}$.

Acknowledgements

First of all, I would like to thank my supervisor Henk Spruit for everything he taught me, for the numerous discussions we had, for his encouragement when nothing seemed to work, and for his genuine interest in my future career. I also thank my co-advisors Achim Weiss and Ewald Müller for their expert advice. Many thanks go to all the former and current members of the SAGA/SESTAS research group, from whom I learnt so much in our Wednesday meetings, as well as to my long-term office mates Alexander Kolodzig, Chervin Laporte, and Hannes Grimm-Strele for the relaxed working environment and all the nice conversations we had. This study would not have been possible without my wonderful wife Lucka who rid me of any doubts whether to go study astrophysics many years ago and who has been making my life enjoyable ever since we first met. I am also grateful to the rest of my family for their support.

Róbert Andrássy
Garching, February 2015