

# 400A - Convection & MLT

Mathieu Renzo

February 14, 2025

**Materials:** Kippenhahn's book Ch. 6 and 7, Cox & Giuli vol. 1 Ch. 14, Schwarzschild Ch. 7, [Jermyn et al. 2022](#), [Anders et al. 2022](#), [Matteo Cantiello's talk at KITP in 2017](#), review by [Joyce & Tayar 2023](#).

## Energy transport in stars 2/2

In the [previous lecture on energy transport](#) we have dealt with the situation where energy diffusion (carried by photons, i.e., radiative energy transport, or carried by electrons, i.e., conductive energy transport) is sufficient to sustain the star and carry all the flux to maintain *local* energy conservation. However, this is *not* always the case! For example, we know this is not the case in the outer layers of the Sun, we can [directly see that](#):

We can also see this is not the case in the envelopes of red supergiants, as for example Betelgeuse, VY Canis Majoris, etc. Less directly, we can also infer this to happen in the cores of massive stars: *some stellar layers are not stably stratified and energy transport is not only diffusive*.

## Convection

In this lecture we will deal with convection, which allows for *energy transport through macroscopic motion of matter* resulting in a *non-zero energy flux with a zero matter flux*.

Convection involves turbulence (the "last" open problem of classical physics), and is because of this inherently multi-dimensional. Because of this, convection is usually one of the most approximate ingredients in stellar calculations, and often the root of many open problems.

Convection is a *thermal* instability (as we will see it kicks in if the temperature gradient is steeper than some threshold condition), although it results in *local* motion of gas/plasma: as we will see the velocities are very small

compared to thermal velocities. Moreover, it is *not* unique to stars: the monsoon clouds in Tucson's summer are also manifestation of convection in the Earth atmosphere!

Before discussing the details of the physics of convection by using oversimplified pictures dating back to [Prandtl](#), have a look at the following animations of multi-dimensional simulations of convection, these are probably/hopefully closer to reality than many of the oversimplifications we will use later in this lecture.

**N.B.:** It is important to keep in mind that we adopt modeling simplifications to make the simulation of stars tractable, but these introduce systematic uncertainties which are active topic of research. Approximations are present also in the multidimensional simulations shown below, and therefore they should *not* be taken as the truth! However, these multi-dimensional simulations do not need to make the same assumptions we will discuss in this lecture, which makes them informative on how rough these approximations are.

What all these simulations of convection in different settings show is that the morphology of the convective flow is *more complicated* than what we will assume. You can imagine "thermal flux tubes" carried by the gas that transport energy in all these situations, but the way we describe them in spherically symmetric models of stars *cannot* (and does not attempt to) capture all the details we can see in the Sun and or a few nearby red supergiants, and we can simulate in restricted time and spatial domains.

### Core convection in a massive star ([Anders et al. 2022b](#))

This is a simulation of the temperature fluctuations (right) and vertical velocities (left) in "code units" using the code [Dedalus](#).

### Envelope convection in a red super-giant ([Goldberg et al. 2022](#))

This is a 3D radiation-hydrodynamics simulation of large portions of the envelope of a red supergiant (not that different from Betelgeuse), computed with the [ATHENA++](#) code.

### Thermonuclear runaway during a Nova explosion ([Casanova et al. 2018](#))

This is a 2D simulation of convection developing during a nova explosion using the [FLASH](#) code. The movie shows the ( $\log_{10}$ ) mass fraction of  $^{20}\text{Ne}$ ,

and is taken from [J. Jordi's personal webpage](#) (one of the co-authors).

## The timescale for convection

We can define a timescale associated to the macroscopic motion of matter with net zero flux that is carrying energy  $\tau_{\text{conv}} = \Delta r / v_{\text{conv}}$ . This is the "convective turnover timescale", where  $\Delta r$  is the spatial extent of a convective region, and  $v_{\text{conv}}$  is the velocity of the blobs of matter moving around.

In the following we will see how to find  $\Delta r$  and get an estimate of  $v_{\text{conv}}$ .

- **Q:** We have already encountered two *global* timescales relevant to stars, the free fall timescale  $\tau_{\text{ff}}$  and the Kelvin-Helmholtz timescale  $\tau_{\text{KH}}$ .  $\tau_{\text{conv}}$  is instead a *local* timescale, relevant only to the convective layers. How do you think it compares to the two global timescales we have defined?

Since convection is an *instability* we expect that it kicks in when the stability of the matter stratification and transport of energy by the other mechanisms we have already discussed breaks down: *convection occurs if energy diffusion is insufficient to carry the flux.*

How convection exactly turns on/off and especially how to treat this in stellar evolution models is still a debated issue - but we can construct a **stability criterion**: when this is violated, convection develops. [Later on](#) we will develop a theoretical model for the *time- and spatially- averaged steady state* which we expect convection to reach when the instability saturates, glossing over the growth phase of the instability. Hopefully, stellar *evolution* timescales are long enough that describing the averaged steady state resulting from convection is sufficient for our purposes (but there are exceptions, for example during explosions or when looking at phenomena on timescales shorter than  $\tau_{\text{conv}}$ ).

## Convective stability criterion: Schwarzschild & Ledoux criterion

To derive a stability criterion, let's assume to start from a stable situation, where the temperature gradient is determined by the (radiative) diffusion of energy:  $dT/dr \propto \kappa L / (4\pi r^2)$ .

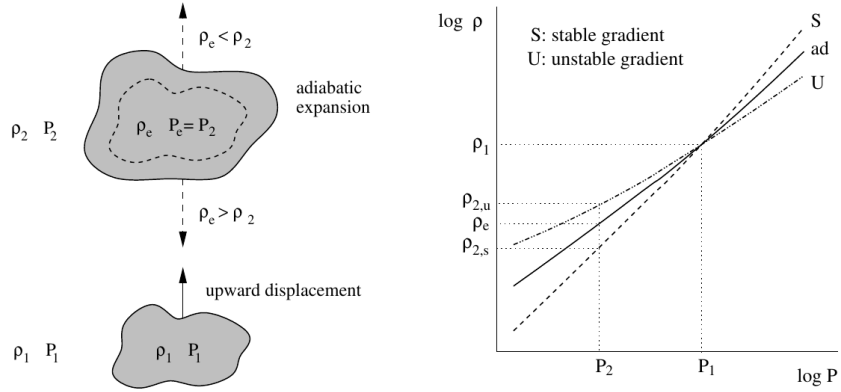


Figure 1: Schematic picture of the convective stability based on Prandtl over-simplified "bubble picture". The right panel shows schematically the  $\rho(P)$  track, in the left panel the bubble is moved upwards (i.e. towards lower pressure), so we expect adiabatic expansion of the bubble to maintain pressure equilibrium with the environment. This is Fig. 5.3 in Onno Pols' lecture notes.

Let's consider a parcel of gas initially in equilibrium with its surroundings at pressure  $P_1$  and density  $\rho_1$ . To determine a stability criterion, let's perturb such parcel by displacing it by a certain (small) amount  $\Delta$  and discuss how the parcel reacts: if things act to move the parcel of gas back towards its original position we have a stable situation, if instead a seed initial displacement  $\Delta$  result in more displacement, we have an unstable situation.

As the gas parcel moves, we can assume it maintains hydrostatic equilibrium with the surrounding: we are looking for a thermal instability that acts on a longer timescale than dynamical timescale. Moreover, if this were not the case, any pressure imbalance would be equalized through sound-waves. Therefore, throughout the path  $\Delta$  and at the final position we have  $P_{\text{bubble}} = P_{\text{environment}}$ .

Let's calculate the density. Since we assume  $\Delta$  to be "small" (w.r.t. the relevant spatial scales in the star), a first order approximation is sufficient:  $\rho_{\text{bubble}} = \rho_1 + (d\rho/dr)_{\text{ad}} \Delta$ . The relevant way to calculate  $d\rho/dr$  here is to assume that the gas parcel moves *adiabatically*: there is no time for heat exchange,  $dq=0$ , and the specific entropy of the bubble remains constant  $ds = 0$  (recall thermodynamics!).

We want to compare  $\rho_{\text{bubble}}$  after the displacement by  $\Delta$  to the environ-

ment density. Once again we can use a first order approximation, but for the environment we have assumed an initially stable stratification, meaning  $d\rho/dr$  is not adiabatic, but the radiative gradient. Therefore  $\rho_2 = \rho_1 + (d\rho/dr)_{\text{env}} \Delta \equiv \rho_1 + (d\rho/dr)_{\text{rad}} \Delta$ , because we are assuming the surrounding environment to be in radiative equilibrium (and assessing the stability of that equilibrium).

**N.B.:** we have derived an equation for  $dT/dr$  in radiative equilibrium (i.e., when the energy is transported by the diffusion of photons), which combined with the EOS can be turned into  $(d\rho/dr)_{\text{rad}}$ .

**N.B.:** In reality, the "bubble" picture is a gross oversimplification. In a convective layer what really moves around are "flux tubes" of thermal energy carried by gas, but there is not a true "bubbling". A common misconception is that water boiling is exhibiting convection: that is not exactly correct. Water boiling is *by definition* a phase transition from liquid to gas, however, shortly before the phase transition occurs, it is true that conduction in the water is typically insufficient to carry the energy release at the bottom, and convective motion can be spotted as a small simmering of the water breaking down into small waves at the surface.

## Instability condition

Comparing the density of the displaced bubble  $\rho_{\text{bubble}}$  to the density of the environment  $\rho_2$  we can determine a condition for instability. If  $\rho_{\text{Bubble}} \leq \rho_2$  then there will be a buoyant force acting to displace it further up (Archimedes force, the bubble displaces a volume of fluid heavier than its own weight!):

$$\text{Instability if : } \rho_{\text{Bubble}} \leq \rho_2 \Rightarrow \left( \frac{d\rho}{dr} \right)_{\text{ad}} \leq \left( \frac{d\rho}{dr} \right)_{\text{env}} \equiv \left( \frac{d\rho}{dr} \right)_{\text{rad}} . \quad (1)$$

Therefore, the density stratification for radiative energy transport is *unstable w.r.t. buoyancy forces* **if** the density gradient  $(d\rho/dr)_{\text{rad}}$  is larger than the adiabatic gradient  $(d\rho/dr)_{\text{ad}}$ : if the gradient is *too steep* then radiative diffusion is not sufficient to carry the energy flux and convection kicks in, and the threshold defining what is *too steep* is the adiabatic gradient. This is the criterion derived by [Martin Schwarzschild](#), however, in stellar physics textbooks it is often re-written in a slightly different form.

## Schwarzschild criterion

For the same reason why we tend to use the Lagrangian enclosed mass as independent coordinate in stellar calculation, it is impractical to use the derivatives w.r.t. radius when trying to determine whether the stratification of gas is stable or not. For example, a red supergiant envelope (which is convective!) has a radial extent of  $\sim 100$ - $1000$ s of  $R_\odot$ . Moreover, since we are only dealing with properties of the stellar gas here, it is more practical to use as independent coordinate something else that is more directly related to the gas itself. The common choice is to use the pressure  $P$  itself.

**N.B.:**  $P(r)$  is also monotonically increasing outwards because of the hydrostatic equilibrium equation, much like  $m(r)$  is, so it is a valid variable to use as independent coordinate.

Thus, rewriting  $d\rho/dr = (d\rho/dP)/dP/dr$ , using the ideal gas EOS, and defining the pressure scale height as the e-folding length of the pressure:

$$H_p = -\frac{dr}{d\log(P)} \Leftrightarrow P(r) \simeq P_0 e^{-r/H_p} \quad , \quad (2)$$

we can rewrite:

$$\frac{d\rho}{dr} = -\frac{P}{H_p} \frac{d\rho}{dP} \quad . \quad (3)$$

**N.B.:** because of the minus sign in the definition of pressure scale height (which is there to make  $H_p$  a positive quantity), the signs change when going from  $(d\rho/dr)_i$  to  $\nabla_i$ .

Furthermore, it is helpful to rewrite this in terms of temperature gradients instead of density gradients. These changes of variables, assuming an ideal gas EOS with constant mean molecular weight  $\mu$  allow to re-write the instability condition 1 in the form most commonly called Schwarzschild criterion:

$$\text{Instability if : } \frac{\partial \log(T)}{\partial \log(P)}_{\text{rad}} = \nabla_{\text{rad}} > \frac{\partial \log(T)}{\partial \log(P)}_{\text{ad}} = \nabla_{\text{ad}} \quad , \quad (4)$$

with  $\mu = \text{constant}$ . Note that from the radiative transport equation we can directly obtain  $\nabla_{\text{rad}}$

$$\nabla_{\text{rad}} = \frac{3\kappa L P}{64\pi G m \sigma T^4} \propto \kappa L \quad , \quad (5)$$

So we see immediately that this is going to be large and prone to convective instability in regions where there is a large opacity  $\kappa \equiv \kappa(m)$  and/or regions with a large luminosity  $L \equiv L(m)$ .

### Ledoux criterion

In a star,  $\mu$  is *not* always constant: as we have already seen there can be regions of *partial ionization* where  $\mu$  changes as we move through them, and we already know that in the fully ionized inner regions of the star we have  $\mu \simeq 1/(2X+3Y/4+Z/2)$  so as we convert hydrogen into helium in the core (and later on helium into metals), we also expect  $\mu$  to change. We can thus rewrite the instability condition 1 without assuming  $\mu = \text{constant}$ .

To do this, it may be helpful to write the EOS functional dependence in a very general form  $P \equiv P(\rho, T, \{X_i\}) \equiv P(\rho, T, \mu)$ . By differentiating this we obtain

$$\frac{d\rho}{\rho} = \alpha \frac{dP}{P} - \delta \frac{dT}{T} + \varphi \frac{d\mu}{\mu} , \quad (6)$$

where  $\alpha$ ,  $\delta$ , and  $\varphi$  are coefficients that depend on the details of the EOS, but known as long as the EOS is known.

We can then rewrite 1 as

$$\text{Instability if : } \nabla_{\text{rad}} \geq \nabla_{\text{ad}} + \frac{\varphi}{\delta} \nabla_{\mu} , \quad (7)$$

with  $\nabla_{\mu} = \partial \log(\mu) / \partial \log(P)$ . Eq. 7 is usually referred to as the Ledoux criterion (Ledoux 1947).

### Secular mixing processes: semiconvection and thermohaline mixing

For stellar layers that are stable according to the Ledoux criterion but unstable according to the Schwarzschild criterion as we have defined them above, that is

$$\nabla_{\text{ad}} \leq \nabla_{\text{rad}} \leq \nabla_{\text{ad}} + \frac{\varphi}{\delta} \nabla_{\mu} , \quad (8)$$

the chemical potential gradient acts as a stabilizing force. There will *not* be a full blown instability, but rather, in the approximate toy model we have used to derive the instability criterion, the gas parcels will experience small

oscillations where the  $\mu$  gradient acts as a damping force. This is the so called **semiconvection**.

Viceversa, if a layer is Schwarzschild stable but Ledoux unstable (this can occur depending on  $\delta$  and  $\phi$ , that is depending on the EOS and the chemical composition):

$$\nabla_{\text{ad}} + \frac{\varphi}{\delta} \nabla_{\mu} \leq \nabla_{\text{rad}} \leq \nabla_{\text{ad}} \quad , \quad (9)$$

then the mean molecular weight gradient acts to *destabilize* the layer. In this case, in our simplistic picture, a blob of gas will slowly start moving because of  $\nabla_{\mu}$  but there will be no restoring forces, and we obtain the so called **thermohaline mixing** or **double diffusion instability**. The name double diffusion comes from the fact that for the gas parcel to move it has to diffuse thermal energy to its environment (which otherwise would stabilize it), as its different chemical composition also diffuses. This leads to the formation of long "fingers", as you can [prove in a kitchen experiment](#):

The thermohaline mixing is obviously not only a stellar phenomenon: it can occur for example in the sea close to the equator, where the surface is heated by the Sun and evaporates faster, leading to a layer with hotter and saltier water (higher  $\mu$ ) on top of colder and less salty water below.

An example where it occurs in stars are accretors in binaries which may receive helium enriched material from the outer layers of the core of the donor star, putting helium rich higher  $\mu$  gas on top of the lower  $\mu$  envelope.

### Which instability criterion should one use?

Naively, one may think that the Ledoux convection is more physically accurate, since it requires one less hypothesis (which we know to not always be correct inside a star). However, when calculating stellar models, what we are interested in is the *long-term* evolution of the star: as you can see from the thermohaline mixing figure above, this is not a 1D process (the "fingers" end with "mushrooms"), and we typically care about timescales very long compared to the timescales for these processes. These are in fact *thermal* processes and their timescales are proportional to the *local* thermal timescale, which as we have already seen is generally short compared to the evolutionary timescale.

Especially for convection in the stellar cores (where we will see energy is generated, therefore  $L$  can be very large, especially in massive stars and drive convection), convection may flatten the  $\mu$  gradient on a timescale short compared to the main sequence lifetime, therefore erasing the ingredient that





Figure 2: Double-diffusive fingers in hot salty water on top of cold fresh water. Credits: M. Cantiello.

differentiates the two (see for instance [Anders et al. 2022a](#)). This is an active topic of debate in the recent literature!

## Mixing length theory

Let's now consider what happens in an unstable layer: we need to model how the energy is transported in these layers, where radiative diffusion is insufficient and the gas will start moving. An ideal solution to this problem would follow the dynamics of buoyant parcels of gas over the (long) thermal timescale, which is in general *not* possible: hydrodynamic simulations can only compute the much shorter *dynamical* timescales!

Physically, in the unstable situation we have described above, we should expect macroscopic motion of gas (the "bubbles") to start, and these "bubbles" would move upward adiabatically, maintaining hydrostatic balance with the surroundings, until they release their excess heat, cooling down and finally falling back. This obviously is *not* a one dimensional problem, since we have some bubbles moving upwards and some moving downwards simultaneously so that the net mass flux is zero, but the net energy flux is non-zero. Moreover this typically leads to turbulence in the flow which is an inherently multi-dimensional problem.

[Erika Bohm-Vitense](#) developed in 1957 an effective mean-field theory to describe the space- and time-averaged steady state at which convective energy transport would saturate. This is the so-called *mixing length theory* (MLT) that is widely applied in stellar evolution still today, and is based on the simplified "bubble picture" from Prandtl we already used to derive a stability criterion.

Before deriving the energy flux, and ultimately temperature gradient in a convective region according to MLT, let's try to get an intuition for what this very successful, albeit very approximate picture tries to describe. We can consider a more familiar example of convection for that, such as a self-sustaining flame:

MLT is meant to describe the spatial and temporal average of the gas flow in the convective region driven by the (chemical) energy release from the flame. Intuitively, it's like taking a picture (assuming any snapshot in time is statistically equivalent to any other), and then averaging across the horizontal cross section of this flame to obtain an approximation to the time- and space- averaged vertical flow of energy and temperature. With all the limitations that this entails, MLT is a very successful theory that is sufficient for *most* stellar evolution applications since those typically are



Figure 3: A flame sustains itself by driving convection that brings in more oxygen to allow combustion to happen.

concerned with timescales that are very long w.r.t. the convective turnover timescale (i.e., in the fire analogy, very long compared to the "flickering" of the flames).

- **Q:** based on this, can you guess where/when MLT will be an insufficient approximation?

### Convective energy flux

To calculate the energy flux carried by convection within the framework of MLT, let's consider the difference in temperature between a bubble that is displaced upwards by an amount  $\ell$  in an unstable layer w.r.t. the surrounding environment:

$$\Delta T = T_{\text{Bubble}} - T_{\text{env}} = \left( T_1 + \frac{dT}{dr} \Big|_{\text{Bubble}} \ell \right) - \left( T_1 + \frac{dT}{dr} \Big|_{\text{env}} \ell \right) = \left( \frac{dT}{dr} \Big|_{\text{Bubble}} - \frac{dT}{dr} \Big|_{\text{env}} \right) \ell . \quad (10)$$

To put this in the form of the gradients that we have defined above for the stability, we notice that  $dT/dr = T \times d \log(T) / d \log(P) \times d \log(P) / dr$ ,

and assume that  $T_{\text{Bubble}} \simeq T_{\text{env}} \equiv T$ , that is effectively consider only the zeroth order of the Taylor series of the temperature expansion, and rewrite for the temperature difference:

$$\Delta T = \frac{\ell}{H_p} T (\nabla_{\text{rad}} - \nabla_{\text{ad}}) \quad , \quad (11)$$

where we use the assumption that the environment is characterized by a radiative gradient and the bubble by an adiabatic gradient.

**N.B.:** The sign has changed because of the minus in the definition of the pressure scale height  $H_p$ . The stability criterion obtained above says that the unstable situation is when  $\nabla_{\text{rad}}$  is steeper than  $\nabla_{\text{ad}}$ , so the form above also guarantees that  $\Delta T$  is actually positive, as we expect for a bubble raising and carrying an excess thermal energy compared to the background.

The excess energy per unit volume carried by the raising bubble is then  $c_p \rho \Delta T$ , where  $c_p$  is the specific heat at constant pressure. Multiplying by the velocity of the bubble we get the **convective flux** (as you can verify by dimensional analysis!):

$$F_{\text{conv}} = c_p \rho \frac{\ell}{H_p} T (\nabla_{\text{rad}} - \nabla_{\text{ad}}) v_{\text{conv}} \quad . \quad (12)$$

- **Q:** why do we use the constant pressure  $c_p$  here? **Hint:** think of the assumptions we have discussed above.

Here there are two things we don't know yet: how far the bubble goes  $\ell$  and the convective velocity.

**N.B.:** In general, to maintain the net-zero mass flux, for each bubble of mass  $\delta m$  raising there is one of the same mass sinking. The raising one carries excess thermal energy w.r.t. the radiative background, and the sinking one carries a deficiency in thermal energy w.r.t. the background so the total convective flux should be twice what we have derived. On the other hand, by taking the difference in the gradients across the entire (as of yet unknown) travel path  $\ell$ , we are overestimating the gradient difference, and on average it should be roughly half of that, canceling out the mistake we make by neglecting the sinking bubbles.

## Convective velocity

To estimate the convective velocity  $v_{\text{conv}}$  we can consider the work done by the buoyancy forces (per unit volume) on the bubble.

- **Q:** before we even do this calculation, can you imagine an upper-limit for  $v_{\text{conv}}$  in the approximated picture we are developing? (**Hint:** we have assumed that any pressure imbalance between the bubble and the environment would be quickly be washed out)

The buoyancy force per unit volume acts in the direction opposite of gravity and has amplitude equal to weight of the displaced fluid minus the weight of the bubble itself,  $B = -\Delta\rho \times |g|$ , where  $\Delta\rho = \rho_{\text{Bubble}} - \rho_{\text{env}}$  is the difference in density between the rising fluid element and the environment. Doing a Taylor expansion and keeping only the first order in  $\ell$  we have:

$$\Delta\rho \simeq \left( \rho_1 + \frac{d\rho}{dr}|_{\text{ad}}\ell \right) - \left( \rho_1 + \frac{d\rho}{dr}|_{\text{rad}}\ell \right) , \quad (13)$$

This is the difference in density between the bubble and the environment at the end of the whole (yet unknown) travel path  $\ell$ . On average throughout the path, since the difference was  $\Delta\rho = 0$  at the beginning, we only have half of that, so let's just consider 1/2 of this to estimate the work done by buoyancy. We can further express the density gradients as a function of  $\nabla_i = \partial \log(T)/\partial \log(P)$ .

By energy conservation, the work done by buoyancy on the bubble is equal to the kinetic energy (per unit volume) acquired by the bubble, which is what we will use to make  $v_{\text{conv}}$  appear:

$$E_{\text{kin,Bubble}} = B \cdot \ell \Rightarrow \frac{1}{2}\rho v_{\text{conv}}^2 = \frac{\rho}{4H_p}(\nabla_{\text{rad}} - \nabla_{\text{ad}})\ell^2 g . \quad (14)$$

**N.B.:** the buoyancy force and the displacement vector are antiparallel, which, together with the definition of  $H_p$ , adjusts the minus signs.

Conveniently in Eq. 14 both  $v_{\text{conv}}$  and  $\ell$  are squared: the same exact reasoning applies to the bubbles sinking and those rising! Eq. 14 is a relation between the two unknowns we have in the convective energy flux,  $v_{\text{conv}}$  and  $\ell$ , which allows us to eliminate one for the other:

$$v_{\text{conv}} = \sqrt{\frac{\ell^2 g}{2H_p}(\nabla_{\text{rad}} - \nabla_{\text{ad}})} , \quad (15)$$

## The mixing length and $\alpha_{\text{MLT}}$

At this point enters the heuristic hypothesis proposed by [Bohm-Vitense 1958](#) (**N.B.:** the original paper is in German): let's assume that the length scale  $\ell$

traveled *on average* by a convectively moving bubble before losing its identity by releasing its excess heat to the surroundings (or absorbing the amount of heat it was lacking, in the case of a sinking bubble), is proportional to the local pressure scale height. This heuristic hypothesis is sensible, since the pressure scale height tells us something about the thermal stratification of the gas, and we are discussing an instability that needs to transport energy when diffusion is insufficient, and it is still very widely used today. It gives us the central hypothesis of MLT:

$$\ell = \alpha_{MLT} H_p \quad (16)$$

The average length traveled by a bubble  $\ell$  is the so-called mixing length that gives the name to this "theory", and the proportionality constant  $\alpha_{MLT}$  is one of the most infamous free parameters in stellar evolution that is calibrated on stellar observations. If the heuristic hypothesis underpinning this approach holds, it should be a quantity of order 1.

Putting all things together, we can now express the convective energy flux as a function of known quantities and this free parameter  $\alpha_{MLT}$ :

$$F_{\text{conv}} = \rho c_P T \alpha_{MLT}^2 \sqrt{\frac{1}{2} g H_P (\nabla_{\text{rad}} - \nabla_{\text{ad}})^{3/2}} \quad (17)$$

**N.B.:** The convective flux predicted by MLT in Eq. 17 is  $\propto$  to a power of the *superadiabaticity*  $(\nabla_{\text{env}} - \nabla_{\text{ad}}) \equiv (\nabla_{\text{rad}} - \nabla_{\text{ad}})$ , because of the assumption of an initially radiative background environment.

## Efficiency of convection

Convection is an *instability*, meaning once it kicks in, it grows exponentially fast towards a saturated state. We have neglected the growth phase (see also [below](#)), and found an approximate description for the steady state depending on a free parameter  $\alpha_{MLT}$ . We can now ask, in such steady state, how big is the superadiabaticity needed such that the convective flux carries *all* the energy? We can estimate this equating the convective flux  $F_{\text{conv}}$  to the entire flux that needs to be carried throughout a layer at radius  $r$ :

$$F_{\text{conv}} \equiv \frac{L(r)}{4\pi r^2} \quad (18)$$

To obtain an order of magnitude estimate, we can substitute in  $F\{\text{conv}\}$  the average density of the star,  $T$  from the virial theorem estimate, assume a monoatomic gas for  $c_P$ , and using  $L(r)/4\pi r^2 \sim L/R^2$  we obtain:

$$\nabla_{\text{rad}} - \nabla_{\text{ad}} \simeq \left(\frac{LR}{M}\right)^{2/3} \frac{R}{GM} \simeq 10^{-8} \quad , \quad (19)$$

where in the last one we use the numerical values for the Sun. This of course is an estimate valid in the interior of the Sun (because we have used implicitly assumptions of LTE, and used averaged values).

The fact that the superadiabaticity is so small implies that *when convection is efficient, the temperature gradient in the star can be approximated with adiabatic*. This comes from a relatively rough estimate, and validates a posteriori many of the questionable assumptions we have made in this lecture: since in the end the gradient is very nearly adiabatic when convection is efficient, the details do not matter that much.

**N.B.:** In the outer layers of the star, where  $\rho \ll \langle \rho \rangle$  and  $T \ll \langle T \rangle$ , this estimate breaks down, convection is not necessarily efficient and the gradient is not necessarily adiabatic. This is important for many stellar applications, for example eruptive mass loss of massive stars, and dynamical stability of mass transfer in binaries.

### On the convective velocity and chemical mixing

Eq. 15 derived above tells us that  $v\{\text{conv}\} \propto (\nabla_{\text{rad}} - \nabla_{\text{ad}})^{1/2} \times (gH_P)^{1/2} \simeq (\nabla_{\text{rad}} - \nabla_{\text{ad}})^{1/2} c_{\text{sound}} \simeq (\nabla_{\text{rad}} - \nabla_{\text{ad}})^{1/2} v_{\text{Thermal}}$ . We have just seen that for efficient convection, the superadiabaticity is small, implying that the convective velocities are much smaller than the *local* sound speed (which guarantees that hydrostatic equilibrium is verified along the path  $\ell$ ), and, equivalently, the local *thermal* velocity.

Nevertheless, even a velocity of  $v\{\text{conv}\} \simeq 10^{-4} c_{\text{sound}}$  is sufficient to mix material very efficiently over the evolutionary timescales (**N.B.:** the sound crossing time of a star is  $\sim \tau_{\text{ff}} \ll$  evolutionary timescales by a factor smaller than  $10^{-4}$ . For the Sun it's hours/billions of years - for now we are using geological evidence on Earth to estimate the age of the Sun). Therefore, *convection not only carries energy flux, but can also mix the chemical composition!*

Similar argument apply to semiconvection and thermohaline mixing (and even the kitchen experiment can clearly show that thermohaline mixing can result in mixing of the composition).

This is not always important, as we will see: in the Sun's envelope for example, convection mixes homogeneous material. However, in the core of a massive star, it mixes the material in the burning region (where hydrogen is turned into helium) into material that is non-burning and thus initially more hydrogen rich. As we will see, this determines qualitative morphological differences in the end of the main sequence evolution of massive stars w.r.t. low mass stars.

One can derive from MLT a diffusion coefficient for the mixing of chemicals by convection  $D_{\text{conv}} = 1/3 \times v_{\text{conv}} \ell = 1/3 \times v_{\text{conv}} \alpha_{\text{MLT}} H_p$  (and similarly for thermohaline and semiconvective mixing), allowing for a diffusive approximation of convective mixing. In reality convective mixing is an advective process: the macroscopic motion of the fluid carries around chemicals, and then they diffuse from the "bubble" into the environment after having being advected. A diffusion approximation is still possible however because of the very high efficiency of this mixing (and the fact that when we apply MLT we do not attempt to describe faithfully the details of the local dynamics of each gas parcel).

## Limitations of MLT

### Convective boundary mixing (a.k.a. "overshooting")

The stability criteria derived from buoyancy argument only determine the location where one can expect that radiative diffusion is insufficient to carry energy, and therefore small perturbations will result in macroscopic motion of matter at  $v \simeq v_{\text{conv}} \ll v_{\text{thermal}}$ . However, what happens when a convective element of gas reaches the instability boundary with a non-zero velocity? There, the buoyant force (and hence the acceleration) goes to zero, but the element has already a non-zero velocity! Thus, we should expect it to "overshoot" this boundary, decelerate outside of it, extending further the convective mixing.

This simplistic picture of overshooting really requires a multi-dimensional treatment. As you can see in the numerical simulations by for example [Anders et al. 2022](#) (with the usual caveat that simulations  $\neq$  physical world, but at least these do not assume spherical symmetry), at the outer boundary the velocity perturbations of the gas will turn over and necessarily acquire non-radial components. More in general, the gradient within the convective boundary mixing region may remain adiabatic (convective penetration) and deviate from it only progressively, and the picture of overshooting alone (which does not specify the temperature gradient in this region) is again





Figure 4: Overshooting top in a thunderstorm on Earth seen from the ISS. Convection carries energy, generate turbulence (think of when an airplane crosses clouds!), and facilitates condensation of water into drops forming the cloud and raining down below this thunderstorm. The "anvil" spreads at the stability boundary, but in the center you can see the overshooting top.

an oversimplification. Significant work is presently dedicated to better understanding the convective boundary and mixing processes active in these regions.

## Time-dependence of convection

By construction MLT (attempts to, arguably successfully) describe the *steady state reached at the saturation point of the convective instability*. This is usually sufficient for stars since the evolution takes much longer than the convective turnover timescale  $\tau_{\text{conv}}$ .

However, there are short phases and/or specific problems in stellar physics when one is concerned with timescales short or comparable to  $\tau_{\text{conv}}$ . For example:

- during explosions (e.g., helium flash, pulsational-pair instability).
- when looking at the interplay between convection and stellar oscillations
- ...

In this case, we need to model how convection turns on/off (having a model for  $d v_{\text{conv}} / dt$ , see e.g., [Arnett 1969](#), [Wood 1974](#), [Gough 1977](#), and more recently [Jermyn 2023](#)).

## Efficiency of semiconvection/thermohaline mixing

The processes we have discussed are all inherently multi-dimensional. Therefore, in their 1D formulation necessary to have stellar structure and evolution models, we introduce free parameters, such as  $\alpha_{\text{MLT}}$  discussed above. In a sense, these free parameters *encapsulate* our ignorance about details we cannot represent in spherical symmetry.

This is also true for thermohaline mixing and semiconvection, each coming with their poorly known efficiency parameter. Furthermore, we need to better understand how these mixing processes in the star interact with each other, with rotation, magnetic fields, etc. Entire conferences are being dedicated to these topics, e.g. [KITP "Probes of transport in stars"](#)!

## Turbulence

Allegedly, Erika Bohm Vitense said that had she been aware of the work of Kolmogorov on turbulence, she would never had proposed MLT as a theory

for convection! This is because in a convective layer one should expect a subsonic but highly turbulent flow. We can in fact estimate the Reynolds number:

$$\text{Re} = \frac{\ell v_{\text{conv}}}{\nu} \quad , \quad (20)$$

where  $\nu$  here is the kinematic viscosity of the stellar plasma, a quantity that is very complicated to compute from first principle, but is usually a very small quantity in practice. Plugging in numbers for the Sun's envelope, for example, one gets  $\text{Re} \simeq 10^{12}$  (see for example [Jermyn et al. 2022](#)). Such large value is usually associated with a high degree of turbulence - an inherently 3D phenomenon, associated with intermittence (i.e., time dependence!).

**N.B.:** multi-D simulations allow for 3D convective flows, which is one step better than MLT, but on the other hand, the effective Reynolds number they reach are much smaller than the ones estimated in the stars: take them as indicative of the limitations of MLT, but they have their own limitations and shortcomings!

## The biggest strength of MLT

MLT has been used for >50 years in 1D stellar evolution calculations, and despite decades of person-time in trying to improve many of its aspects, it remains, for better or worse, a cornerstone of stellar modeling. This is because it is an "effective mean field theory" that successfully describes the time and space averaged state of saturated convective instability in a stellar gas using one free parameter only ( $\alpha_{\text{MLT}}$ ) and allows us to make **evolutionary** calculations on timescales much longer than the convective turnover timescale.

## Summary of the key assumptions

- The gradient of the background environment is **radiative** (meaning energy is transported by radiation diffusion)
- We model the thermal flux tubes carrying energy up and down with no mass flux as bubbles moving adiabatically. These are an idealization of realistic convective flows which are turbulent and thus space- and time-dependent on very short timescales (think of a flickering flame).

- The "bubbles" maintain hydrostatic equilibrium w.r.t. the environment at any point in their travel
- We assume that the average distance travelled by a bubble is proportional to the local pressure scale height  $\ell = \alpha_{\text{MLT}} H_p$  (the "mixing length").
- $\alpha_{\text{MLT}}$  is *the* free parameter of this approach that effectively exists because convection is inherently a 3D phenomenon that we are trying to approximate in spherical symmetry.

## Homework

- Using the definition of pressure scale height and the hydrostatic equilibrium equation, show that  $P/\rho = 1/(\rho g H_p) \Rightarrow c_{\text{sound}}^2 \simeq g H_p$