

## Basic Assumptions About Convection

The calculation of convection requires that we adopt some assumptions about the process. In general, these assumptions are false (sometimes by several orders of magnitude), but we are stuck with them.

- There exist discrete “blobs” of material inside a star that have sizes of the order of  $l_m$ , where  $l_m$  is the **mixing length**. These blobs can move distances of the order of  $l_m$  before losing their identity. (This variable  $l_m$  is somewhat of a fudge-factor, and many of the uncertainties will get hidden in it.)
- The mixing length is small, compared to other scale lengths in the star. In particular, if  $\lambda_P$  is the pressure scale height, then  $l_m \ll \lambda_P$ .
- The blob maintains pressure balance with its surroundings. Thus the time for a pressure wave to cross the blob is much shorter than the timescale for blob motion,  $l_m/v_s \ll l_m/v$ .
- Acoustic waves, shocks, magnetic fields, rotation, are all negligible.
- The temperatures and densities inside a blob are only slightly different than those surrounding it.

These assumptions are sometimes called “Boussinesq”; they are valid for almost incompressible fluids with small density and temperature variations.

## Convective Energy Transport

In a radiative region of a star, the temperature gradient is given by  $\nabla_{\text{rad}}$ ; in the limiting case of a fully convective region (*i.e.*, where all the energy is transported by convection), the gradient will be  $\nabla_{\text{ad}}$ . In the outer envelopes of stars, however, the temperature gradient can be  $\nabla_{\text{rad}} > \nabla > \nabla_{\text{ad}}$ .

The computation of  $\nabla$  in the presence of convection requires some work. Recall that the flux transported by radiative (and conductive) diffusion is

$$F_{\text{rad}} = -\frac{4acT^3}{3\kappa\rho} \frac{dT}{dr} \quad (3.1.4)$$

When combined with the hydrostatic equilibrium equation (2.2.2) and the definition of  $\nabla$  (2.4.1), this becomes

$$F_{\text{rad}} = \frac{4acGM^4}{3\kappa Pr^2} \nabla \quad (8.1.1)$$

This gives a relation between  $F_{\text{rad}}$  and  $\nabla$ . Unfortunately, neither is known, since  $F_{\text{rad}}$  carries only part of the star's flux. The other part is carried through blobs of matter, which move with velocity  $v$  and have density  $\rho$ . If these blobs are in pressure equilibrium with their surroundings, then they will be capable of depositing

$$F_{\text{cnv}} \sim \rho v c_P \Delta T \quad (8.1.2)$$

where  $\Delta T$  is the temperature difference between the blob and the surrounding material. Since  $\nabla_{\text{rad}}$  is defined as the value needed to transport all of the star's flux, we have

$$F_{\text{tot}} = F_{\text{rad}} + F_{\text{cnv}} = \frac{\mathcal{L}}{4\pi r^2} = \frac{4acGM^4}{3\kappa Pr^2} \nabla_{\text{rad}} \quad (8.1.3)$$

The task is therefore to estimate  $\rho$ ,  $v$ , and  $\Delta T$  for the blobs.

First, let's find an expression for  $\Delta T$  in terms of the temperature gradient  $\nabla$ . To do this, consider the behavior of a single blob of material. Assume that a typical blob will move a radial distance  $l_m$  before losing its identity, and that, at a given location  $r$  inside the star, a typical blob will have already moved half this distance. At this time, the temperature difference between this typical blob and its surroundings will be

$$T_i - T_s = \left( \frac{dT_i}{dr} - \frac{dT_s}{dr} \right) \frac{l_m}{2}$$

where the subscripts  $i$  and  $s$  refer to the material inside and outside the blob. (We will now drop the subscript  $s$ , since those are just the global variables of the star.) If we divide each side by  $T$  and multiply by the pressure scale height,  $\lambda_P = -dr/d \ln P$ , then

$$\left( \frac{T_i - T}{T} \right) \lambda_P = \frac{d}{dr} \left( \frac{dT_i}{T} - \frac{dT}{T} \right) \left( -\frac{dr}{d \ln P} \right) \frac{l_m}{2}$$

or

$$\frac{T_i - T}{T} = \frac{\Delta T}{T} = (\nabla - \nabla_i) \frac{l_m}{2\lambda_P} \quad (8.1.4)$$

Note that at this point, the variables  $\nabla$  and  $\nabla_i$  are still unknown.

Now let's estimate the velocity of the blob. According to the definition of  $\alpha$ ,  $\delta$ , and  $\varphi$ ,

$$\frac{d\rho}{\rho} = \alpha \frac{dP}{P} - \delta \frac{dT}{T} + \varphi \frac{d\mu}{\mu} \quad (3.2.2)$$

However, if the blob remains in pressure equilibrium with its surroundings,  $dP = 0$ , and if the chemical composition does not change,  $d\mu = 0$ . Hence

$$\frac{d\rho}{\rho} = -\delta \frac{dT}{T}$$

which implies

$$\frac{\rho_i - \rho}{\rho} = -\delta (\nabla - \nabla_i) \frac{l_m}{2\lambda_P} \quad (8.1.5)$$

This difference in density leads to a net buoyancy force of  $f_b = -g(\rho_i - \rho)$ ; the work performed by this force while moving the blob is then

$$f_b \cdot \frac{l_m}{2} = -g(\rho_i - \rho) \frac{l_m}{2} = g\rho\delta(\nabla - \nabla_i) \frac{l_m^2}{4\lambda_P}$$

Now suppose that  $\sim 1/2$  this work goes into the kinetic energy of the blob while the other half operates on the surroundings. The mean velocity of the material passing the blob is then given by

$$\frac{1}{2} \rho v^2 = \frac{1}{2} g\rho\delta(\nabla - \nabla_i) \frac{l_m^2}{4\lambda_P}$$

or

$$v = \frac{l_m}{2} \left\{ \frac{GM\delta(\nabla - \nabla_i)}{r^2\lambda_P} \right\}^{1/2} \quad (8.1.6)$$

With this expression, we can write the convective flux as a function of the temperature gradients that are internal and external to the blob:

$$F_{\text{cnv}} = \rho v c_P \Delta T = \rho \left\{ \frac{GM\delta(\nabla - \nabla_i) l_m^2}{4\lambda_P r^2} \right\}^{1/2} c_P \left\{ \frac{(\nabla - \nabla_i) l_m T}{2\lambda_P} \right\}$$

or

$$F_{\text{cnv}} = \rho c_P T \left( \frac{GM\delta}{r^2} \right)^{1/2} \frac{l_m^2}{4} \left( \frac{\nabla - \nabla_i}{\lambda_P} \right)^{3/2} \quad (8.1.7)$$

Finally, let's estimate  $\nabla_i$  for a blob with volume  $V$ , surface area  $S$ , and total heat content (over its surroundings)  $Q$ . Clearly, if the blob remains in pressure equilibrium, then

$$Q = \rho V c_P \Delta T \quad (8.1.8)$$

As the blob moves, some of this heat will be radiated away. From (3.1.4), the rate at which this will happen will be

$$\frac{dQ}{dt} = -\frac{4 a c T^3}{3 \kappa \rho} \frac{dT}{dx} S \quad (8.1.9)$$

where  $x$  is the direction normal to the surface of the blob. To evaluate this, let's make the simplest approximation that the temperature gradient normal to the blob depends on the size of the blob, *i.e.*,

$$\frac{dT}{dx} \sim \frac{T_i - T}{l_m}$$

Through (8.1.4), this implies

$$\frac{dT}{dx} = \frac{T_i - T}{l_m} = T (\nabla - \nabla_i) \frac{l_m}{2\lambda_P} \frac{1}{l_m} = \frac{T (\nabla - \nabla_i)}{2\lambda_P}$$

and

$$\frac{dQ}{dt} = -\frac{2 a c T^4}{3 \kappa \rho} \frac{(\nabla - \nabla_i)}{\lambda_P} S \quad (8.1.10)$$

Now, as the blob moves, its internal temperature will change, both from the radiative loss of heat, and from adiabatic expansion (or contraction). The change in temperature is

$$\begin{aligned} \left(\frac{dT}{dr}\right)_i &= \left(\frac{dT}{dr}\right)_{\text{ad}} + \left(\frac{dT}{dr}\right) \\ &= \left(\frac{dT}{dr}\right)_{\text{ad}} - \frac{1}{\rho V c_P} \left(\frac{dQ}{dr}\right) \\ &= \left(\frac{dT}{dr}\right)_{\text{ad}} - \frac{1}{\rho V c_P} \frac{dt}{dr} \left(\frac{dQ}{dt}\right) \\ &= \left(\frac{dT}{dr}\right)_{\text{ad}} - \frac{1}{\rho V c_P v} \left(\frac{2 a c T^4}{3 \kappa \rho} \frac{(\nabla - \nabla_i)}{\lambda_P} S\right) \end{aligned}$$

$$\left(\frac{dT}{dr}\right)_i = \left(\frac{dT}{dr}\right)_{\text{ad}} - \frac{2 a c T^4 (\nabla - \nabla_i)}{3 \kappa \rho^2 c_P v \lambda_P} \left(\frac{S}{V}\right) \quad (8.1.11)$$

Now let's choose some shape for the blob. The ratio of surface area to volume for a sphere of diameter  $l_m$  is  $S/V = 6/l_m$ , but other shapes will have different values. For the moment, let's just assign  $S/V = \zeta/l_m$ . If we now substitute this in, and multiply by  $\lambda_P/T = -(dr/d \ln P)/T$ , we get

$$\left(\frac{dT}{dr}\right)_i \left(-\frac{dr}{d \ln P} \frac{1}{T}\right) = \left(\frac{dT}{dr}\right)_{\text{ad}} \left(-\frac{dr}{d \ln P} \frac{1}{T}\right) - \frac{2 a c T^4 (\nabla - \nabla_i) \zeta}{3 \kappa \rho^2 c_P v \lambda_P l_m} \left(\frac{\lambda_P}{T}\right)$$

$$\left(\frac{d \ln T}{d \ln P}\right)_i = \left(\frac{d \ln T}{d \ln P}\right)_{\text{ad}} + \frac{2 \zeta a c T^3 (\nabla - \nabla_i)}{3 \kappa \rho^2 c_P v l_m}$$

or, if we let  $\zeta = 9$  for compatibility with the literature

$$\frac{\nabla_i - \nabla_{\text{ad}}}{\nabla - \nabla_i} = \frac{6 a c T^3}{\kappa \rho^2 c_P v l_m} \quad (8.1.12)$$

This last equation provides us the information we need to solve the problem. Aside from the local variables ( $P$ ,  $T$ ,  $\rho$ , *etc.*), we have five new variables associated with the convection problem:  $F_{\text{rad}}$ ,  $F_{\text{cnv}}$ ,  $v$ ,  $\nabla_i$ , and  $\nabla$ . However, we also have five equations: (8.1.1), (8.1.3), (8.1.6), (8.1.7), and (8.1.12). Thus, all the variables can be computed (although it takes a bit of math). The mixing length  $l_m$  is a free parameter, and must be chosen reasonably to match observations as best as possible.

## Solving the Convection Equations

The five convection equations are

$$F_{\text{rad}} = \frac{4acGM T^4}{3\kappa P r^2} \nabla \quad (8.1.1)$$

$$F_{\text{rad}} + F_{\text{cnv}} = \frac{4acGM T^4}{3\kappa P r^2} \nabla_{\text{rad}} \quad (8.1.3)$$

$$v = l_m \left\{ \frac{GM\delta (\nabla - \nabla_i)}{4r^2 \lambda_P} \right\}^{1/2} \quad (8.1.6)$$

$$F_{\text{cnv}} = \rho c_P T \left( \frac{GM\delta}{r^2} \right)^{1/2} \frac{l_m^2}{4} \left( \frac{\nabla - \nabla_i}{\lambda_P} \right)^{3/2} \quad (8.1.7)$$

$$\frac{\nabla_i - \nabla_{\text{ad}}}{\nabla - \nabla_i} = \frac{6acT^3}{\kappa \rho^2 c_P v l_m} \quad (8.1.12)$$

To solve these equations, we first consolidate the constants and location condition variables into two dimensionless quantities.

$$U = \frac{3acT^3}{c_P \rho^2 \kappa l_m^2} \left( \frac{4\lambda_P r^2}{GM\delta} \right)^{1/2} \quad (8.1.13)$$

and

$$W = \nabla_{\text{rad}} - \nabla_{\text{ad}} \quad (8.1.14)$$

All the quantities in these variables are known from the stellar structure, and can be taken as constants.

We can then proceed to eliminate the variables. If we substitute the expression for velocity (8.1.6) in (8.1.12), then

$$\frac{\nabla_i - \nabla_{\text{ad}}}{\nabla - \nabla_i} = \left( \frac{6acT^3}{\kappa\rho^2c_P l_m} \right) \left( \frac{4\lambda_P r^2}{GM\delta l_m^2} \right)^{1/2} = \frac{2U}{(\nabla - \nabla_i)^{1/2}}$$

or

$$\nabla_i - \nabla_{\text{ad}} = 2U (\nabla - \nabla_i)^{1/2} \quad (8.1.15)$$

Next, we can use (8.1.1) and (8.1.3) to eliminate  $F_{\text{rad}}$

$$F_{\text{cnv}} = \frac{4acGM T^4}{3\kappa P r^2} (\nabla_{\text{rad}} - \nabla) \left( \frac{dP}{dr} \right) \left( \frac{r^2}{GM\rho} \right) \left( \frac{dr}{d \ln P} \right) \left( \frac{1}{\lambda_P} \right)$$

leaving

$$F_{\text{cnv}} = \frac{4acT^4}{3\kappa\rho\lambda_P} (\nabla_{\text{rad}} - \nabla) \quad (8.1.16)$$

and use this with (8.1.7) to eliminate  $F_{\text{cnv}}$

$$\frac{4acT^4}{3\kappa\rho\lambda_P} (\nabla_{\text{rad}} - \nabla) = \rho c_P T \left( \frac{GM\delta}{r^2} \right)^{1/2} \frac{l_m^2}{4} \left( \frac{\nabla - \nabla_i}{\lambda_P} \right)^{3/2}$$

for

$$\left( \frac{\nabla - \nabla_i}{\lambda_P} \right)^{3/2} = \frac{8act^3}{3c_P\rho^2\kappa l_m^2 \lambda_P^{3/2}} \left( \frac{4\lambda_P r^2}{GM\delta} \right)^{1/2} (\nabla_{\text{rad}} - \nabla)$$

or

$$(\nabla - \nabla_i)^{3/2} = \frac{8}{9} U (\nabla_{\text{rad}} - \nabla) \quad (8.1.17)$$



Finally, we can combine (8.1.15) and (8.1.17) by noting that if we re-write (8.1.15) as

$$(\nabla - \nabla_{\text{ad}}) - (\nabla - \nabla_i) - 2U (\nabla - \nabla_i)^{1/2} = 0$$

then the equation is actually a quadratic in  $(\nabla - \nabla_i)^{1/2}$ , with the (positive) solution

$$(\nabla - \nabla_i)^{1/2} = -U + (U^2 + \nabla - \nabla_{\text{ad}})^{1/2} \quad (8.1.18)$$

if we assign

$$\xi = U^2 + \nabla - \nabla_{\text{ad}}$$

then (8.1.18) is simply

$$(\nabla - \nabla_i)^{1/2} = \xi - U \quad (8.1.19)$$

Now if we substitute this into (8.1.17), then

$$(\xi - U)^3 = \frac{8}{9}U (\nabla_{\text{rad}} - \nabla)$$

or, since  $\nabla = \xi^2 - \nabla_{\text{ad}} - U^2$  and  $W = \nabla_{\text{rad}} - \nabla_{\text{ad}}$ ,

$$(\xi - U)^3 - \frac{8}{9}U (\xi^2 - U^2 - W) = 0 \quad (8.1.20)$$

This is straightforward cubic equation in  $\xi$ , that should have only one real root. Since  $U$  and  $\nabla_{\text{ad}}$  are both known quantities the solution of this equation gives  $\nabla$ , which in turn, can be used to calculation  $F_{\text{rad}}$  and  $F_{\text{cnv}}$ .

## When is Convection Important?

The quantity  $U$  is a measure of how efficient convection is at transporting energy. The convective cells have two characteristic timescales: one giving the length of time required to move the distance  $l_m$  and deliver its energy, and the other describing how long it takes for the cell to lose its heat via radiation. The former timescale can be estimated from the free-fall equation

$$\tau_{\text{ff}} \sim \left( \frac{2l_m}{g} \right)^{1/2} = \left( \frac{2l_m r^2}{GM} \right)^{1/2}$$

The latter can be derived using our estimates of the total heat content (8.1.8) and rate of heat loss (8.1.9) of each blob.

$$\tau_{\text{adj}} \sim Q \Big/ \frac{dQ}{dt} \sim \rho V c_P \Delta T \Big/ \left( \frac{4acT^3}{3\kappa\rho} \frac{\Delta T}{l_m} S \right)$$

Thus the ratio of these numbers

$$\begin{aligned} \frac{\tau_{\text{ff}}}{\tau_{\text{adj}}} &\sim \left( \frac{4acT^3}{3c_P\rho^2\kappa l_m} \right) \left( \frac{V}{S} \right) \left( \frac{2l_m r^2}{GM} \right)^{1/2} \\ &\sim \left( \frac{2\sqrt{2}\zeta acT^3}{3\rho^2 c_P \kappa l_m^2} \right) \left( \frac{4l_m r^2}{GM} \right)^{1/2} \\ &\sim \left( \frac{\delta\lambda_P}{8l_m} \right)^{1/2} U \end{aligned} \tag{8.1.21}$$

Thus, to order unity,  $U$  gives the efficiency of energy transport. If  $U \rightarrow 0$ , then the timescale for heat transport is much less than that for heat loss, and  $\nabla$  should equal  $\nabla_{\text{ad}}$ . This can be

seen via (8.1.13) and (8.1.15), since for small  $U$ ,  $\nabla_i \approx \nabla_{\text{ad}}$ , and thus  $\nabla \approx \nabla_{\text{ad}}$ . This occurs in dense regions (*i.e.*, in convective cores). In the outer envelopes of stars, however, the convective blobs can radiate a significant amount of their heat while they are in transit. In this case,  $U \rightarrow \infty$ , so through (8.1.15),  $\nabla \rightarrow \nabla_i$ , and therefore through (8.1.17),  $\nabla \rightarrow \nabla_{\text{rad}}$ . The full mixing length theory is usually only needed in the outer parts of cool stars, where *superadiabatic* conditions can exist *i.e.*, where both radiative and convective flux transportation are important. Most of the time, one or the other dominates, so you can use  $\nabla_{\text{rad}}$  or  $\nabla_{\text{ad}}$ , whichever is appropriate.

The results you get for convective energy transport depend on the mixing length. In the literature, mixing length is not quoted; instead the value usually given is

$$\alpha = \frac{l_m}{\lambda_P} \quad (8.1.22)$$

Traditionally,  $\alpha \sim 1$ ; currently, the “best fit” seems to be  $\alpha \sim 1.7$ . (Note that this violates one of the tenants of mixing length theory.)

Convection theory is only an approximation. Convection similar to that occurring in stars, cannot be studied in the laboratory. If we define the ratio of radiative and thermal diffusion as

$$\nu_T = \frac{4 a c T^3}{3 \kappa \rho} \bigg/ \rho c_P = \frac{4 a c T^3}{3 \kappa \rho^2 c_P}$$

and let  $\nu$  represent the viscosity of the gas, then the Rayleigh number

$$\mathbf{Ra} = \frac{\delta g l_m^4}{\nu_T \nu} (\nabla - \nabla_{\text{ad}})$$

is a dimensionless quantity that represents how well convection is proceeding. For laboratory fluids,  $\mathbf{Ra} \sim 10^{11}$ , but in stars,  $\mathbf{Ra} \sim 10^{21}$ . Thus, our experiments are 10 orders of magnitude away from reality. (In fact, the closest driven convection on earth is convection heating in the air above deserts.)

## Semi-Convection

A phenomenon that takes place in the cores of massive stars is semi-convection. Consider a star with a convective core within which nuclear reactions are proceeding, and a radiative envelope. In the convective core,  $\nabla_{\text{rad}} > \nabla_{\text{ad}}$ ; outside the core,  $\nabla_{\text{rad}} < \nabla_{\text{ad}}$ , and, at the border between the two regions,  $\nabla_{\text{rad}} = \nabla_{\text{ad}}$ . As always,  $\nabla_{\text{rad}}$  is given by

$$\nabla_{\text{rad}} = \frac{3\kappa\mathcal{L}P}{16\pi acGM T^4} \quad (3.1.6)$$

Now assume that the major source of opacity is electron scattering, and that the matter is fully ionized. (Since we are talking about the core of a massive star, these are excellent assumptions.) As nuclear reactions proceed in the core, the hydrogen content of the region will decrease. Since the electron scattering opacity for fully ionized matter is given by

$$\kappa_e = 0.2(1 + X) \quad (6.1.7)$$

this results in a decrease in the stellar opacity.

Now let's examine the behavior of  $\nabla_{\text{rad}}$  at a distance  $\epsilon$  from the convective core. By definition, for convection to occur

$$\nabla_{\text{rad},i} > \nabla_{\text{ad}} \quad \text{inside the core}$$

and

$$\nabla_{\text{rad},o} < \nabla_{\text{ad}} \quad \text{outside the core}$$

Therefore, since  $\nabla_{\text{ad}}$  shouldn't change much over a distance  $\epsilon$ ,

$$\nabla_{\text{rad},i} > \nabla_{\text{rad},o} \quad \text{by definition of convection}$$

However,  $\nabla_{\text{rad}}$  is directly proportional to the hydrogen content, and  $X_i < X_o$ . Thus,

$$\nabla_{\text{rad},i} < \nabla_{\text{rad},o} \quad \text{from the opacity}$$

Thus there is a fundamental contradiction.

Another way to look at the problem is to consider the temperature and density difference between a convective blob and its surroundings. Recall that for a region to be dynamically stable,

$$\left(\frac{d\rho}{dr}\right)_i > \left(\frac{d\rho}{dr}\right)_s \quad (3.2.1)$$

or

$$\frac{\Delta\rho}{\Delta r} > 0$$

where  $\Delta$  represents the difference between the region internal to the blob and the blob's surroundings. For positive displacements of  $\Delta r$ , the blob will be denser than its surroundings, and will sink back to its original position.

Now consider the temperature difference between the blob and the surrounding.

$$\Delta T = \left[ \left(\frac{dT}{dr}\right)_i - \left(\frac{dT}{dr}\right)_s \right] \Delta r$$

If the blob is in pressure equilibrium (given) and the stellar region is chemically homogeneous, then from (3.2.2)

$$\frac{\Delta T}{T} = -\frac{\Delta\rho}{\delta\rho}$$

The negative sign indicates that the inside of the blob will be cooler than its surroundings. It will therefore gain some energy by radiation, which will result in the blob having a slightly lower density, and a slightly weaker restoring force due to buoyancy. Any motion of the blob will eventually be damped out.

Next, consider the situation where the surrounding medium is chemically inhomogeneous. We again assume that the region is dynamically stable, with  $\Delta\rho/\Delta r > 0$ . However, in this case

$$\frac{\Delta T}{T} = -\frac{\Delta\rho}{\delta\rho} + \frac{\varphi}{\delta} \frac{\Delta\mu}{\mu} \quad (8.2.1)$$

Consequently, if

$$\Delta\mu = \left(\frac{d\mu}{dr}\right)_i - \left(\frac{d\mu}{dr}\right)_s > \frac{\Delta\rho}{\rho} \frac{\mu}{\varphi}$$

then blobs that have a positive  $\Delta r$  displacement will have  $\Delta T > 0$ . These blobs will be hotter than their surroundings and will therefore lose heat as they move. When this happens, their density will increase, the restoring force will increase, and their oscillation will slowly increase in amplitude. The blob will be vibrationally unstable; the result will be a slow mixing, which will eventually destroy the chemical gradient. This semi-convection will occur when

$$\nabla_{\text{ad}} < \nabla < \nabla_{\text{ad}} + \frac{\varphi}{\delta} \left(\frac{d \ln \mu}{d \ln P}\right) \quad (8.2.2)$$

Note that from (8.2.1) and (8.2.2), it will only occur when the mean molecular weight is decreasing rapidly with radius, or equivalently,

$$\left(\frac{d \ln \mu}{d \ln P}\right) > 0$$

This occurs in high mass stars with convective cores. In these objects, the convective region gets smaller with time, leaving behind a chemical gradient.