Chapter 4

Energy transport in stellar interiors

The energy that a star radiates from its surface is generally replenished from sources or reservoirs located in its hot central region. We have seen that most stars are in a long-lived state of thermal equilibrium, in which these terms exactly balance. What would happen if the nuclear energy source in the centre is suddenly quenched? The answer is: very little, at least initially. Photons that carry the energy are continually scattered, absorbed and re-emitted in random directions. Because the photon mean free path is so small (typically $\sim$ 1 cm, see Sect. 3.1), the time it takes radiation to escape from the centre of the Sun by this random walk process is roughly 10^7 years, despite the fact that photons travel at the speed of light (in contrast, *neutrinos* produced in the centre of the Sun escape in a matter of seconds). Changes in the Sun's luminosity would only occur after millions of years, on the timescale for radiative energy transport, which you may recognise as the Kelvin-Helmholtz timescale for thermal readjustment.

The transport of energy in stars is the subject of this chapter. It will lead us to two additional differential equations for stellar structure. Radiation is the most important means of energy transport, and it is always present. However, it is not the only means. In stellar interiors, where matter and radiation are always in local thermodynamic equilibrium (Chapter 3), energy (heat) can be transported from hot to cool regions in two basic ways:

- Random thermal motions of the particles, a process that can be called *heat diffusion*. In this context we can consider both the gas and the radiation to consist of particles. In the case of photons, the process is known as *radiative diffusion*. In the case of gas particles (atoms, ions, electrons) it is usually called *heat conduction*.

- Collective (bulk) motions of the gas particles, which is known as *convection*. This is an important process in stellar interiors, not only because it can transport energy very efficiently, it also results in rapid mixing. Unfortunately, convection is one of the least understood ingredients of stellar physics.

4.1 Local energy conservation

In Chapter 2 we considered the global energy budget of a star, regulated by the virial theorem. We have still to take into account the conservation of energy on a local scale in the stellar interior. To do this we turn to the first law of thermodynamics (Sect. 3.4), which states that the internal energy of a system can be changed by two forms of energy transfer: heat and work. By δf we denote a change in a quantity f occurring in a small time interval δt. For a gas element of unit mass the first law can be

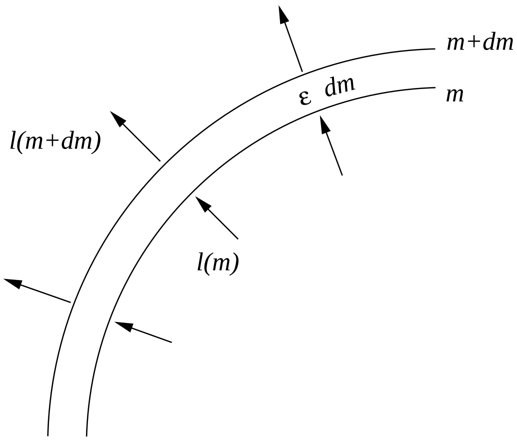


Figure 4.1. Energy generation and heat flow into and out of a spherical mass shell.

written as (see eq. 3.46)

$$\delta u = \delta q + \frac{P}{\rho^2}\,\delta\rho. \tag{4.1}$$

The first term is the heat added or extracted, and second term represents the work done on (or performed by) the element. We note that compression ($\delta\rho > 0$) involves an addition of energy, and expansion is achieved at the expense of the element's own energy.

Consider a spherical, Lagrangian shell inside the star of constant mass dm. Changes in the heat content of the shell ($\delta Q = \delta q\, dm$) can occur due to a number of sources and sinks:

- Heat is added by the release of nuclear energy, if available. The rate at which nuclear energy is produced per unit mass and per second is written as $\epsilon_{\rm nuc}$. The details of nuclear energy generation will be treated in Ch. 5.

- Heat can be removed by the release of energetic neutrinos, which escape from the stellar interior without interaction. Neutrinos are released as a by-product of some nuclear reactions, in which case they are often accounted for in $\epsilon_{\rm nuc}$. But neutrinos can also be released by weak interaction processes in very hot and dense plasmas. This type of neutrino production plays a role in late phases of stellar evolution, and the rate at which these neutrinos take away energy per unit mass is written as ϵ_ν.

- Finally, heat is absorbed or emitted according to the balance of heat fluxes flowing into and out of the shell. We define a new variable, the *local luminosity* l, as the rate at which energy in the form of heat flows outward through a sphere of radius r (see Fig. 4.1). In spherical symmetry, l is related to the radial energy flux F (in erg s^{-1} cm^{-2}) as

  $$l = 4\pi r^2 F. \tag{4.2}$$

  Therefore at the surface $l = L$ while at the centre $l = 0$. Normally heat flows outwards, in the direction of decreasing temperature. Therefore l is usually positive, but under some circumstances (e.g. cooling of central regions by neutrino emission) heat can flow inwards, meaning that l is negative. (We note that the energy flow in the form of neutrinos is treated separately and is *not* included in the definition of l and of the stellar luminosity L.)

We can therefore write:

$$\delta Q = \epsilon_{\rm nuc}\, dm\, \delta t - \epsilon_\nu\, dm\, \delta t + l(m)\, \delta t - l(m + dm)\, \delta t,$$

with $l(m + dm) = l(m) + (dl/dm) \cdot dm$, so that after dividing by dm,

$$\delta q = \left(\epsilon_{\rm nuc} - \epsilon_\nu - \frac{dl}{dm}\right)\delta t. \tag{4.3}$$

Combining eqs. (4.3) and (4.1) and taking the limit $\delta t \to 0$ yields:

$$\boxed{\frac{dl}{dm} = \epsilon_{\rm nuc} - \epsilon_\nu - \frac{\partial u}{\partial t} + \frac{P}{\rho^2}\frac{\partial \rho}{\partial t}} \tag{4.4}$$

This is the third equation of stellar evolution. The terms containing time derivatives are often combined into a function $\epsilon_{\rm gr}$:

$$\begin{aligned} \epsilon_{\rm gr} &= -\frac{\partial u}{\partial t} + \frac{P}{\rho^2}\frac{\partial \rho}{\partial t} \\ &= -T\frac{\partial s}{\partial t} \end{aligned} \tag{4.5}$$

where s is the specific entropy of the gas. One can then write

$$\frac{dl}{dm} = \epsilon_{\rm nuc} - \epsilon_\nu + \epsilon_{\rm gr} \tag{4.6}$$

If $\epsilon_{\rm gr} > 0$, energy is released by the mass shell, typically in the case of contraction. If $\epsilon_{\rm gr} > 0$, energy is absorbed by the shell, typically in the case of expansion.

In *thermal equilibrium* (see Sec. 2.3.2), the star is in a stationary state and the time derivatives vanish ($\epsilon_{\rm gr} = 0$). We then obtain a much simpler stellar structure equation,

$$\frac{dl}{dm} = \epsilon_{\rm nuc} - \epsilon_\nu. \tag{4.7}$$

If we integrate this equation over the mass we obtain

$$L = \int_0^M \epsilon_{\rm nuc}\, dm - \int_0^M \epsilon_\nu\, dm \equiv L_{\rm nuc} - L_\nu \tag{4.8}$$

which defines the nuclear luminosity $L_{\rm nuc}$ and the neutrino luminosity L_ν. Neglecting the neutrino losses for the moment, we see that thermal equilibrium implies that $L = L_{\rm nuc}$, that is, energy is radiated away at the surface at the same rate at which it is produced by nuclear reactions in the interior. This is indeed what we defined as thermal equilibrium in Sec. 2.3.2.

4.2 Energy transport by radiation and conduction

In Sect. 3.1 we briefly considered the photon mean free path $\ell_{\rm ph}$ and concluded that it is much smaller than important length scales within the star. In the deep interior of the Sun, for example, $\ell_{\rm ph} \sim 1\,{\rm cm} \ll R_\odot$. In other words: stellar matter is very *opaque* to radiation. As a result, radiation is trapped within the stellar interior, and photons very slowly diffuse outwards by a 'random walk' process. We also saw that the temperature difference over a distance $\ell_{\rm ph}$ is only $\Delta T \sim 10^{-4}\,{\rm K}$. This means that the radiation field is extremely close to black-body radiation with $U = u\rho = aT^4$ (Sec. 3.3.6). Black-body radiation is isotropic and as a result no net energy transport would take place. However, a tiny anisotropy is still present due to $\Delta T/T \sim 10^{-11}$. This small anisotropy is enough to carry the entire energy flux in the Sun (see Exercises).

4.2.1 Heat diffusion by random motions

The above estimates show that radiative energy transport in stellar interiors can be described as a diffusion process. This yields a great simplification of the physical description.

Consider a unit surface area and particles (e.g. photons) crossing the surface in either direction. Let z be a coordinate in the direction perpendicular to the surface. The number of particles crossing in the positive z direction (say upward) per unit area per second is

$$\frac{dN}{dt} = \tfrac{1}{6} n \bar{v},$$

where n is the particle density and $\bar{v}$ is their average velocity. The factor $\frac{1}{6}$ comes from the fact that $\frac{1}{2}$ of the particles cross the surface in one direction, and because their motions are isotropic the average velocity perpendicular to the surface is $\frac{1}{3}\bar{v}$. These particles have a slightly higher energy density U than the particles crossing the surface in the other (negative z) direction. If the mean free path of the particles is ℓ, then the excess energy carried up by the upward particle flow is $\delta U = -(dU/dz)\,\ell$ (since dU/dz is assumed to be negative). There is a similar, *negative* excess $-\delta U$ carried *down* by the particles moving in the other direction. Therefore the energy flux carried by this diffusion process is

$$F = \tfrac{1}{6}\bar{v}\,\delta U - (-\tfrac{1}{6}\bar{v}\,\delta U)$$

which can be generalised to[1]

$$\boldsymbol{F} = -D\,\boldsymbol{\nabla} U \tag{4.9}$$

where D is the diffusion coefficient

$$D = \tfrac{1}{3}\bar{v}\ell \tag{4.10}$$

Since $\boldsymbol{\nabla} U = (\partial U/\partial T)_V\,\boldsymbol{\nabla} T = C_V\,\boldsymbol{\nabla} T$ we can write

$$\boldsymbol{F} = -K\,\boldsymbol{\nabla} T \tag{4.11}$$

where $K = \frac{1}{3}\bar{v}\ell\, C_V$ is the conductivity. This description is valid for all particles in LTE, including photons.

4.2.2 Radiative diffusion of energy

For photons, we can take $\bar{v} = c$ and $U = aT^4$. Hence the specific heat (per unit volume) is $C_V = dU/dT = 4aT^3$. The photon mean free path can be obtained from the equation of radiative transfer, which states that the intensity I_ν of a radiation beam (in a medium without emission) is diminished over a length s by

$$\frac{dI_\nu}{ds} = -\kappa_\nu \rho\, I_\nu, \tag{4.12}$$

where κ_ν is the mass absorption coefficient or opacity coefficient (in cm^2 g^{-1}) at frequency ν. The mean free path is the distance over which the intensity decreases by a factor of e, which obviously depends on the frequency. If we make a proper average over frequencies (see Sec. 4.2.3), we can write

$$\ell_{\rm ph} = \frac{1}{\kappa\rho}. \tag{4.13}$$

[1] This can be compared to Fick's law for the diffusive flux $\boldsymbol{J}$ of particles (per unit area per second) between places with different particle densities, $\boldsymbol{J} = -D\,\boldsymbol{\nabla} n$, where the particle density n is replaced by the energy density U.

The quantity κ is simply called the *opacity*. We can then compute the radiative conductivity

$$K_{\rm rad} = \tfrac{4}{3}\frac{acT^3}{\kappa\rho}, \tag{4.14}$$

such that the radiative energy flux is

$$\boldsymbol{F}_{\rm rad} = -\tfrac{4}{3}\frac{acT^3}{\kappa\rho}\boldsymbol{\nabla}T. \tag{4.15}$$

In spherical symmetric star the flux is related to the local luminosity, $F_{\rm rad} = l/4\pi r^2$ (eq. 4.2). We can thus rearrange the equation to obtain the temperature gradient

$$\frac{dT}{dr} = -\frac{3\kappa\rho}{16\pi acT^3}\frac{l}{r^2} \tag{4.16}$$

or when combined with eq. (2.5) for dr/dm,

$$\boxed{\frac{dT}{dm} = -\frac{3}{64\pi^2 ac}\frac{\kappa l}{r^4 T^3}} \tag{4.17}$$

This is the temperature gradient required to carry the entire luminosity l by radiation. It gives the fourth stellar structure equation, for the case that energy is transported only by radiation. A star, or a region inside a star, in which this holds is said to be in *radiative equilibrium*, or simply *radiative*.

Eq. (4.17) is valid as long as $\ell_{\rm ph} \ll R$, i.e. as long as the LTE conditions hold. This breaks down when the stellar surface, the photosphere, is approached: this is where the photons escape, i.e. $\ell_{\rm ph} \gtrsim R$. Near the photosphere the diffusion approximation is no longer valid and we need to solve the full, and much more complicated, equations of radiative transfer. This is the subject of the study of *stellar atmospheres*. Fortunately, the LTE conditions and the diffusion approximation hold over almost the entire stellar interior.

In hydrostatic equilibrium, we can combine eqs. (4.17) and (2.12) as follows

$$\frac{dT}{dm} = \frac{dP}{dm}\cdot\frac{dT}{dP} = -\frac{Gm}{4\pi r^4}\,\frac{T}{P}\cdot\frac{d\log T}{d\log P}$$

so that we can define the dimensionless *radiative temperature gradient*

$$\boxed{\nabla_{\rm rad} = \left(\frac{d\log T}{d\log P}\right)_{\rm rad} = \frac{3}{16\pi acG}\frac{\kappa l P}{mT^4}} \tag{4.18}$$

This describes the logarithmic variation of T with depth (where depth is now expressed by the *pressure*) for a star in HE if energy is transported only by radiation.

4.2.3 The Rosseland mean opacity

The radiative diffusion equations derived above are independent of frequency ν, since the flux F is integrated over all frequencies. However, in general the opacity coefficient κ_ν depends on frequency, such that the κ appearing in eq. (4.16) or (4.17) must represent a proper average over frequency. This average must be taken in a particular way.

If $\boldsymbol{F}_\nu\,d\nu$ represents the radiative flux in the frequency interval $[\nu, \nu + d\nu]$, then eq. (4.9) must be replaced by

$$\boldsymbol{F}_\nu = -D_\nu\,\boldsymbol{\nabla}U_\nu = -D_\nu\,\frac{\partial U_\nu}{\partial T}\,\boldsymbol{\nabla}T \tag{4.19}$$

where

$$D_\nu = \tfrac{1}{3} c \ell_\nu = \frac{c}{3\kappa_\nu \rho}. \tag{4.20}$$

The energy density U_ν in the same frequency interval follows from eq. (3.41), $U_\nu = h\nu\, n(\nu)$,

$$U_\nu = \frac{8\pi h}{c^3} \frac{\nu^3}{e^{h\nu/kT} - 1} \tag{4.21}$$

which is proportional to the Planck function for the intensity of black-body radiation. The total flux is obtained by integrating eq. (4.19) over all frequencies,

$$\boldsymbol{F} = -\left[\frac{c}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, d\nu\right] \boldsymbol{\nabla} T. \tag{4.22}$$

This is eq. (4.11) but with conductivity

$$K_{\rm rad} = \frac{c}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, d\nu. \tag{4.23}$$

Comparing with eq. (4.14) shows that the proper average of opacity as it appears in eq. (4.16) or (4.17) is

$$\frac{1}{\kappa} = \frac{1}{4aT^3} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, d\nu. \tag{4.24}$$

This is the so-called *Rosseland mean opacity*. The factor $4aT^3$ appearing in eq. (4.24) is equal to $\int_0^\infty (\partial U_\nu/\partial T)\, d\nu$, so that the Rosseland mean can be seen as the harmonic mean of κ_ν with weighting function $\partial U_\nu/\partial T$. In other words, $1/\kappa$ represents an average *transparency* of the stellar gas rather than an average opacity.

4.2.4 Conductive transport of energy

Collisions between the gas particles (ions and electrons) can also transport heat. Under normal (ideal gas) conditions, however, the collisional conductivity is very much smaller than the radiative conductivity. The collisional cross sections are typically $10^{-18} - 10^{-20}\,\mathrm{cm}^2$ at the temperatures in stellar interiors, giving a mean free path collisions that is several orders of magnitude smaller than $\ell_{\rm ph}$. Furthermore the average particle velocity $\overline{v} = \sqrt{3kT/m} \ll c$. So normally we can neglect heat conduction compared to radiative diffusion of energy.

However, the situation can be quite different when the electrons become degenerate. In that case both their velocities increase (their momenta approach the Fermi momentum, see Sec. 3.3.5) and, more importantly, their mean free paths increase (most of the quantum cells of phase space are occupied, so an electron has to travel further to find an empty cell and transfer its momentum). At very high densities, when $\ell_e \gg \ell_{\rm ph}$, electron conduction becomes a much more efficient way of transporting energy than radiative diffusion. This is important for stars in late stages of evolution with dense degenerate cores and for white dwarfs, in which efficient electron conduction results in almost isothermal cores.

The energy flux due to heat conduction can be written as

$$\boldsymbol{F}_{\rm cd} = -K_{\rm cd}\, \boldsymbol{\nabla} T \tag{4.25}$$

such that the sum of radiative and conductive fluxes is

$$\boldsymbol{F} = \boldsymbol{F}_{\rm rad} + \boldsymbol{F}_{\rm cd} = -(K_{\rm rad} + K_{\rm cd})\, \boldsymbol{\nabla} T. \tag{4.26}$$

Hence if we write the conductivity in the same form as eq. (4.14), we can define a *conductive opacity* $\kappa_{\rm cd}$ by analogy with the radiative opacity, from

$$K_{\rm cd} = \frac{4acT^3}{3\kappa_{\rm cd}\,\rho}. \tag{4.27}$$

Hence we can write the combined flux in the same form as the radiative flux, eq. (4.15),

$$\boldsymbol{F} = -\frac{4acT^3}{3\rho}\left(\frac{1}{\kappa_{\rm rad}} + \frac{1}{\kappa_{\rm cd}}\right)\boldsymbol{\nabla}T, \tag{4.28}$$

if we replace $1/\kappa$ by $1/\kappa_{\rm rad} + 1/\kappa_{\rm cd}$. This result simply means that the transport mechanism with the largest flux will dominate, that is, the mechanism for which the stellar matter has the highest transparency.

4.3 Opacity

The opacity coefficient κ appearing in eq. (4.17) determines how large the temperature gradient must be in order to carry a given luminosity l by radiation. Therefore κ is an important quantity that has a large effect on the structure of a star.

4.3.1 Sources of opacity

In the following subsections we briefly describe the different physical processes that contribute to the opacity in stellar interiors, and give some simple approximations.

Electron scattering

An electromagnetic wave that passes an electron causes it to oscillate and radiate in other directions, like a classical dipole. This scattering of the incoming wave is equivalent to the effect of absorption, and can be described by the Thomson cross-section of an electron

$$\sigma_{\rm e} = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2 = 6.652\times10^{-25}\,{\rm cm}^2 \tag{4.29}$$

The associated opacity coefficient is the combined cross-section of all electrons in a unit mass of gas, obtained by dividing $\sigma_{\rm Th}$ by $\rho/n_e = \mu_{\rm e} m_{\rm u}$,

$$\kappa_{\rm es} = \frac{\sigma_{\rm e}}{\mu_{\rm e} m_{\rm u}} = 0.20\,(1+X)\ \ {\rm cm}^2/{\rm g} \tag{4.30}$$

Since the electron scattering opacity is independent of frequency, this expression is also the Rosseland mean. In the last equality we have assumed that the gas is completely ionized and μ_e is given by eq. (3.20). Electron scattering is an important opacity source in an ionized gas that is not too dense. When the degree of ionization drops (typically when $T \lesssim 10^4$ K in hydrogen-rich gas) the electron density becomes so small that the electron scattering opacity is strongly reduced below eq. (4.30).

When the photon energy becomes a significant fraction of the rest mass of the electron, $h\nu \gtrsim 0.1 m_e c^2$, the exchange of momentum between photon and electron must be taken into account (Compton scattering). This occurs at high temperature, since the Planck function has a maximum at $h\nu = 4.965\,kT$ (Wien's law), i.e. when $kT \gtrsim 0.02 m_e c^2$ or $T \gtrsim 10^8$ K. At such temperatures the electron scattering opacity is smaller than given by eq. (4.30).

Free-free absorption

A free electron cannot absorb a photon because this would violate momentum and energy conservation. If a charged ion is in its vicinity, however, absorption is possible because of the electromagnetic coupling between the ion and electron. This is the inverse process of bremsstrahlung, where an electron passing by and interacting with an ion emits a photon.

The full derivation of the absorption coefficient for this *free-free absorption* process is a quantum-mechanical problem. However, an approximate calculation has been done classically by Kramers. He derived that the absorption efficiency of such a temporary electron-ion system is proportional to ${Z_i}^2\nu^{-3}$, where Z_i is the ion charge. To obtain the cross-section of a certain ion i, this has to be multiplied by the electron density n_e and by the time during which the electron and ion will be close enough. This can be estimated from the mean velocity of the electrons, $\upsilon \sim T^{1/2}$, so that $\Delta t \sim 1/\upsilon \sim T^{-1/2}$, i.e. $\sigma_{\mathrm{ff},i} \sim n_e T^{-1/2} {Z_i}^2 \nu^{-3}$. Finally the opacity coefficient follows by multiplying by n_i/ρ, where n_i is the ion number density, and summing over all ions in the mixture:

$$\kappa_{\nu,\mathrm{ff}} \propto \frac{n_e}{\rho} \Sigma_i n_i\, {Z_i}^2\, T^{-1/2}\, \nu^{-3}.$$

In a completely ionized gas, $n_e/\rho = 1/(\mu_e m_u) = (1{+}X)/2m_u$, while the sum $\Sigma_i n_i {Z_i}^2 = (\rho/m_u)\,\Sigma(X_i {Z_i}^2/A_i) = (\rho/m_u)\,(X + Y + B)$, where $B = \Sigma_{i>2}(X_i {Z_i}^2/A_i)$ is the contribution of elements heavier than helium. As long as their abundance is small, one can take $X + Y + B \approx 1$ to a reasonable approximation.

When we take the Rosseland mean, the factor ν^{-3} becomes a factor T^{-3} (this can be verified by performing the integration of eq. 4.24 with $\kappa_\nu \propto \nu^\alpha$). We thus obtain

$$\kappa_{\mathrm{ff}} \propto \rho\, T^{-7/2}. \tag{4.31}$$

An opacity law of this form is called *Kramers opacity*. Putting in the numerical factors and the compositional dependence for an ionized gas, the following approximate expression is obtained,

$$\kappa_{\mathrm{ff}} \approx 3.8 \times 10^{22} (1 + X)\, \rho\, T^{-7/2} \quad \mathrm{cm^2/g}. \tag{4.32}$$

This formula has to be used with caution: it can give some insight in simplifying approaches but should not be used in serious applications. One omission is a correction factor for quantum-mechanical effects, the so-called Gaunt factor g_{ff}.

Bound-free and bound-bound absorption

Bound-free absorption is the absorption of a photon by a bound electron whereby the photon energy exceeds the ionization energy of the ion or atom. The computation of the opacity due to this process requires carefully taking into account the atomic physics of all the ions and atoms present in the mixture, and is thus very complicated. Classical considerations, similar to those for free-free absorption, show that the frequency dependence is again $\sim \nu^{-3}$, as long as $h\nu > \chi_{\mathrm{ion}}$. Therefore, in rough approximation the total bound-free opacity is also of the Kramers form. A very approximate formula is

$$\kappa_{\mathrm{bf}} \approx 4.3 \times 10^{25} (1 + X) Z\, \rho\, T^{-7/2} \quad \mathrm{cm^2/g}. \tag{4.33}$$

One should not apply this formula for $T \lesssim 10^4$ K because most of the photons are not energetic enough to ionize the electrons. The bound-free opacity is seen to depend directly on the metallicity Z. One thus has, again very approximately, $\kappa_{\mathrm{bf}} \approx 10^3 Z \times \kappa_{\mathrm{ff}}$. We may thus expect bound-free absorption to dominate over free-free absorption for $Z \gtrsim 10^{-3}$.

Bound-bound absorption is related to photon-induced transitions between bound states in atoms or ions. Although this is limited to certain transition frequencies, the process can be efficient because the absorption lines are strongly broadened by collisions. Again, the computation is complex because one has to include a detailed treatment of line profiles under a wide variety of conditions. Bound-bound absorption is mainly important for $T \lesssim 10^6$ K, at higher temperatures its contribution to the total opacity is small.

The negative hydrogen ion

An important source of opacity in relatively cool stars (e.g. in the solar atmosphere) is formed by bound-free and bound-bound absorption of the negative hydrogen ion, H^-. Neutral hydrogen is easily polarized by a nearby charge and can then form a bound state with another electron, with an ionization potential of 0.75 eV. The resulting H^- is very fragile and is easily ionized at temperatures of a few thousand K. However, making the ion requires the presence of both neutral hydrogen and free electrons. The free electron come mainly from singly ionized metals such as Na, K, Ca or Al. The resulting opacity is therefore sensitive to metallicity and to temperature. A very approximate formula in the range $T \sim (3-6) \times 10^3$ K, $\rho \sim (10^{-10} - 10^{-5})$ g/cm^3 and $0.001 < Z < 0.02$ is

$$\kappa_{\mathrm{H}-} \approx 2.5 \times 10^{-31} \left(\frac{Z}{0.02}\right) \rho^{1/2}\, T^9 \ \ \mathrm{cm^2/g} \tag{4.34}$$

At very low metal abundance and/or $T \lesssim 3000$ K the H^- opacity becomes ineffective. At $T \gtrsim 10^4$ K most of the H^- has disappeared and the Kramers opacity and electron scattering take over.

Molecules and dust

In very cool stars with $T_{\mathrm{eff}} \lesssim 4000$ K opacity sources arising from molecules and (at even lower temperatures) dust grains become dominant. Here one has to deal with the complex molecular chemistry and dust formation processes, which still contains a lot of uncertainty.

Conductive opacities

As we saw in Sec. 4.2.4, energy transport by means of heat conduction can also be described by means of a conductive opacity coefficient κ_{cd}. Under ideal gas conditions, conduction is very inefficient compared to radiative transport of energy ($\kappa_{\mathrm{cd}} \gg \kappa_{\mathrm{rad}}$). Therefore we only need to consider the case of a degenerate electron gas. In this case the following approximation holds

$$\kappa_{\mathrm{cd}} = 4.4 \times 10^{-3} \frac{\Sigma Z_i^{5/3} X_i/A_i}{(1+X)^2} \frac{(T/10^7\ \mathrm{K})^2}{(\rho/10^5\ \mathrm{g/cm^3})^2} \ \ \mathrm{cm^2/g}. \tag{4.35}$$

At high densities and low temperatures, the conductive opacity becomes very small because of the large electron mean free path in a highly degenerate gas. This is why degenerate stellar regions are highly conductive and rapidly become isothermal.

4.3.2 An overall view of stellar opacities

In general, $\kappa = \kappa(\rho, T, X_i)$ is a complicated function of density, temperature and composition. While certain approximations can be made, as in the examples shown above, these are usually too simplified and inaccurate to apply in detailed stellar models.

In practical stellar structure calculations one usually interpolates in pre-computed opacity tables, e.g. as calculated in the 1990s by the OPAL project. An example is shown in Fig. 4.2 for a quasi-solar mixture of elements. One may recognize the various regions in the density-temperature plane where

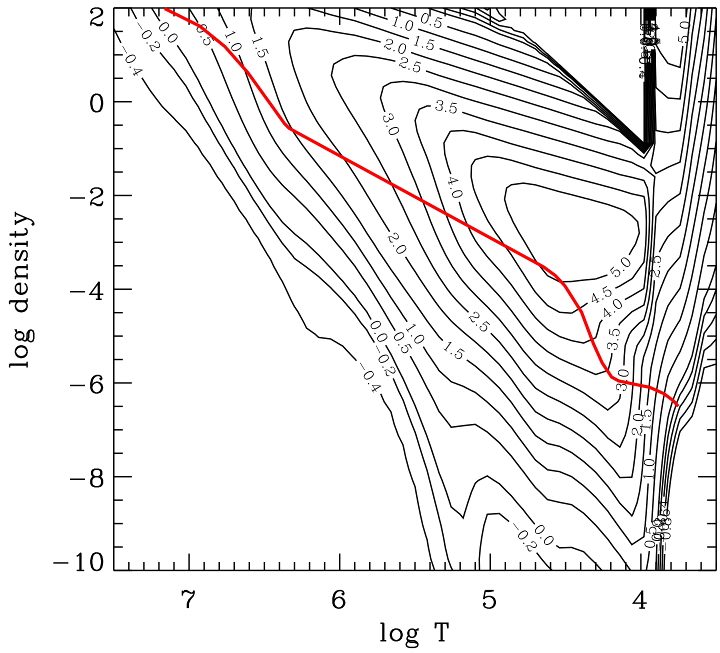


Figure 4.2. Contours of the Rosseland mean opacity as function of T and ρ, for a mixture of elements representative of solar abundances ($X = 0.7, Z = 0.02$), calculated by OPAL. The contour values are $\log\kappa$. The thick (red) line is a structure model for the Sun. At low density and high temperature (lower left part), κ has a constant value given by electron scattering. At higher ρ and lower T, opacity increases due to free-free and bound-free absorptions. The many ridges and wiggles show that the Kramers power law is in fact a rather poor approximation of the actual opacity. For $T < 10^4$ K opacity decreases drastically due to recombination of hydrogen, the main opacity source here is the H^- ion. At lower temperatures still (not shown), κ rises again due to molecules and dust formation. Finally, at very high density the opacity is dominated by the conductivity of degenerate electrons and decreases drastically with increasing ρ.

one of the processes discussed above dominates. It should be clear that there is much more structure in the function $\kappa(\rho, T)$ than in the simple power-law approximations, such as the Kramers law.

For comparison, an interior structure model for the Sun is also shown. The opacity in the solar interior is dominated by free-free and bound-free absorption, and is very high (up to 10^5 cm/g) in the envelope, at temperatures between 10^4 and 10^5 K. In the surface layers the opacity rapidly decreases due to the H^- opacity. More massive stars are located at lower densities than the Sun, and generally have much lower opacities in the envelope. In the most massive stars the opacity is dominated by electron scattering, at low ρ and high T.

Note that the chemical composition, inparticular the metallicity Z, can have a large effect on κ. This provides the most important influence of composition on stellar structure.

4.4 The Eddington luminosity

We have seen that radiative transport of energy inside a star requires a temperature gradient dT/dr, the magnitude of which is given by eq. (4.16). Since $P_{\rm rad} = \frac{1}{3}aT^4$, this means there is also a gradient in the radiation pressure:

$$\frac{dP_{\rm rad}}{dr} = -\frac{4}{3}aT^3\frac{dT}{dr} = -\frac{\kappa\rho}{4\pi c}\,\frac{l}{r^2}. \tag{4.36}$$

This radiation pressure gradient represents an outward force due to the net flux of photons outwards. Of course, for a star in hydrostatic equilibrium this outward radiation force must be smaller than the inward force of gravity, as given by the pressure gradient necessary for HE, eq. (2.11). In other words,

$$\left|\frac{dP_{\rm rad}}{dr}\right| < \left|\left(\frac{dP}{dr}\right)_{\rm HE}\right| \quad \Rightarrow \quad \frac{\kappa\rho}{4\pi c}\,\frac{l}{r^2} < \frac{Gm\rho}{r^2}.$$

This gives an upper limit to the local luminosity, which is known as the (local) *Eddington luminosity*,

$$l < \frac{4\pi c G m}{\kappa} = l_{\rm Edd}. \tag{4.37}$$

This is the maximum luminosity that can be carried by radiation, inside a star in hydrostatic equilibrium.

The inequality expressed by eq. (4.37) can be violated in the case of a very large heat flux (large l), which may result from intense nuclear burning, or in the case of a very high opacity κ. As we saw in Sec. 4.3, high opacities are encountered at relatively low temperatures, near the ionization temperature of hydrogen and helium (and for example in the outer layers of the Sun). In such cases hydrostatic equilibrium (eq. 2.12) and radiative equilibrium (eq. 4.17) cannot hold simultaneously. Therefore, if the star is to remain in HE, energy must be transported by a different means than radiative diffusion. This means of transport is *convection*, the collective motion of gas bubbles that carry heat and can distribute it efficiently. We shall consider convection in detail in Sec. 4.5. It will turn out that eq. (4.37) is a necessary, but not a sufficient condition for a region of a star to be stable against convection.

The surface layer of a star is always radiative, since it is from here that energy escapes the star in the form of photons. Applying eq. (4.37) at the surface of the star ($m = M$) we get

$$L < L_{\rm Edd} = \frac{4\pi c G M}{\kappa}, \tag{4.38}$$

where κ is the opacity in the photosphere. Violation of this condition now means violation of hydrostatic equilibrium: matter is accelerated away from the star by the photon pressure, giving rise to violent mass loss. The Eddington luminosity expressed by eq. (4.38) is a critical stellar luminosity that cannot be exceeded by a star in hydrostatic equilibrium. If we assume κ to be approximately constant (in very luminous main-sequence stars it is dominated by electron scattering, so this assumption is not bad) than $L_{\rm Edd}$ is only dependent on M. It can be expressed as follows

$$L_{\rm Edd} = 3.2 \times 10^4 \left(\frac{M}{M_\odot}\right)\left(\frac{0.4\,{\rm cm^2/g}}{\kappa}\right) L_\odot. \tag{4.39}$$

Since $L_{\rm Edd}$ is proportional to M, while stars (at least on the main sequence) follow a mass-luminosity relation $L \propto M^x$ with $x > 1$ (Sec. 1.2.2), this implies that for stars of increasing mass L will at some point exceed $L_{\rm Edd}$. Hence, we can expect a *maximum mass* to exist for main-sequence stars.

4.5 Convection

For radiative diffusion to carry to transport energy outwards, a certain temperature gradient is needed, given by eq. (4.16) or eq. (4.17). The larger the luminosity that has to be carried, the larger the temperature gradient required. There is, however, an upper limit to to the temperature gradient inside a star – if this limit is exceeded an instability in the gas sets in. This instability leads to cyclic macroscopic motions of the gas, known as *convection*. Convection can be regarded as a type of dynamical instability, although (as we shall see later in this section) it does not have disruptive consequences. In particular, it does not lead to an overall violation of hydrostatic equilibrium. Convection affects the structure of a star only as an efficient means of heat transport and as an efficient mixing mechanism.

In Sec. 4.4 we already derived an upper limit to the luminosity that can be transported by radiation. We will now derive a more stringent criterion for convection to occur, based on considerations of dynamical stability.

4.5.1 Criteria for stability against convection

So far we have assumed strict spherical symmetry in our description of stellar interiors, i.e. assuming all variables are constant on concentric spheres. In reality there will be small fluctuations, arising for example from the thermal motions of the gas particles. If these small perturbations do not grow they can safely be ignored. However, if the perturbations do grow they can give rise to macroscopic motions, such as convection. We therefore need to consider the *dynamical stability* of a layer inside a star.

Consider a mass element that, due to a small perturbation, is displaced upwards by a small distance as depicted in Fig. 4.3. At its original position (at radius r) the density and pressure are ρ_1 and P_1, and at its new position ($r + \Delta r$) the ambient density and pressure are ρ_2 and P_2. Since pressure decreases outwards, $P_2 < P_1$ and the gas element will expand to restore pressure equilibrium with its surroundings. Hence the pressure of the gas element at position 2 is $P_e = P_2$, but its new density after expansion ρ_e is not necessarily equal to ρ_2. If $\rho_e > \rho_2$, the gas element will experience a net buoyancy force downwards (by Archimedes' law), which pushes it back towards its original position. Then the small perturbation is quenched, and the situation is stable. On the other hand, if $\rho_e < \rho_2$ then there is a net buoyancy force upwards and we have an *unstable* situation that leads to convection.

The expansion of the gas element as it rises over Δr occurs on the local dynamical timescale (i.e. with the speed of sound), which is typically much shorter than the local timescale for heat exchange, at least in the deep interior of the star. The displacement and expansion of the gas element will therefore be very close to adiabatic. We have seen in Sec. 3.4 that the adiabatic exponent $\gamma_{\rm ad}$ defined by eq. (3.54) describes the logarithmic response of the pressure to an adiabatic change in the density. Writing as $\delta\rho_e$ and δP_e the changes in the density and pressure of the element when it is displaced

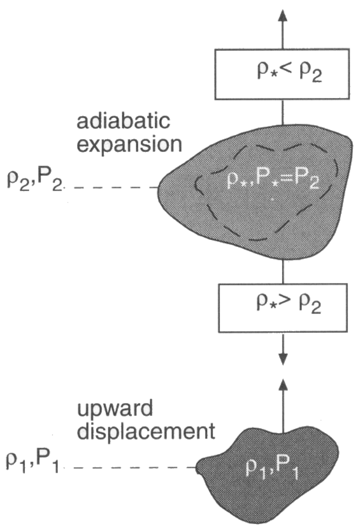


Figure 4.3. Schematic illustration of the Schwarzschild criterion for stability against convection. A gas element is perturbed and displaced upwards from position 1 to position 2, where it expands adiabatically to maintain pressure equilibrium with its surroundings. If its density is larger than the ambient density, it will sink back to its original position. If its density is smaller, however, buoyancy forces will accelerate it upwards: convection occurs. Figure reproduced from PRIALNIK.

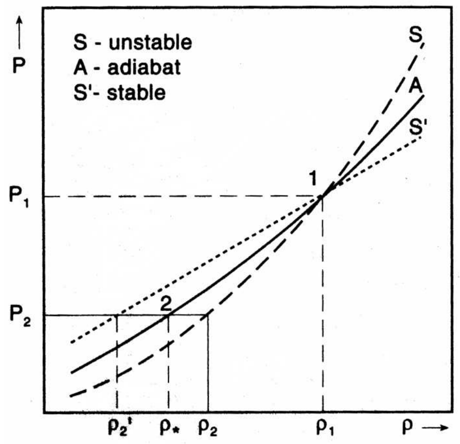


Figure 4.4. Schematic pressure-density diagram corresponding to the situation depicted in Fig. 4.3. A layer is stable against convection if the density varies more steeply with pressure than for an adiabatic change. Figure reproduced from PRIALNIK.

over a small distance Δr, we therefore have

$$\frac{\delta P_e}{P_e} = \gamma_{\rm ad}\,\frac{\delta \rho_e}{\rho_e}. \tag{4.40}$$

Here δP_e is determined by the pressure gradient dP/dr inside the star because $P_e = P_2$, i.e. $\delta P_e = P_2 - P_1 = (dP/dr)\,\Delta r$. Therefore the change in density $\delta\rho_e$ follows from eq. (4.40)

$$\delta\rho_e = \frac{\rho_e}{P_e}\,\frac{1}{\gamma_{\rm ad}}\,\frac{dP}{dr}\Delta r. \tag{4.41}$$

We can write $\rho_e = \rho_1 + \delta\rho_e$ and $\rho_2 = \rho_1 + (d\rho/dr)\,\Delta r$, where $d\rho/dr$ is the density gradient inside the star. We can then express the criterion for stability against convection, $\rho_e > \rho_2$, as

$$\delta\rho_e > \frac{d\rho}{dr}\Delta r, \tag{4.42}$$

which combined with eq. (4.41) yields an upper limit to the density gradient for which a layer inside the star is stable against convection,

$$\frac{1}{\rho}\frac{d\rho}{dr} < \frac{1}{P}\frac{dP}{dr}\,\frac{1}{\gamma_{\rm ad}}, \tag{4.43}$$

where we have replaced P_e and ρ_e by P and ρ, since the perturbations are assumed to be very small. Remember, however, that both $d\rho/dr$ and dP/dr are negative. If we divide eq. (4.43) by dP/dr we obtain the general criterion for stability against convection, as depicted in Fig. 4.4,

$$\frac{d\log\rho}{d\log P} > \frac{1}{\gamma_{\rm ad}}. \tag{4.44}$$

If condition (4.44) is violated then convective motions will develop. Gas bubbles that due to a small perturbation are slightly hotter than their surroundings will move up, transporting their heat content upwards until they are dissolved. Other bubbles, which due to a small perturbation are slightly cooler than their environment will move down and have a smaller heat content than their surroundings. When these elements finally dissolve, they absorb heat from their surroundings. Therefore, both the upward and downward moving convective bubbles effectively transport heat in the upward direction. Hence there is a *net upward heat flux*, even though there is no net mass flux (upward and downward moving bubbles carry equal amounts of mass). This is the principle behind convective heat transport.

The Schwarzschild and Ledoux criteria

The stability criterion (4.44) is not of much practical use, because it involves computation of a density gradient which is not part of the stellar structure equations. We would rather have a criterion for the temperature gradient, because this also appears in the equation for radiative transport. We can rewrite eq. (4.44) in terms of temperature by using the equation of state. We write the equation of state in its general, differential form (eq. 3.47) but now also take into account a possible variation in composition. If we characterize the composition by the mean molecular weight μ then $P = P(\rho, T, \mu)$ and we can write

$$\frac{dP}{P} = \chi_T \frac{dT}{T} + \chi_\rho \frac{d\rho}{\rho} + \chi_\mu \frac{d\mu}{\mu}, \tag{4.45}$$

with χ_T and χ_ρ defined by eqs. (3.48) and (3.49), and χ_μ is defined as

$$\chi_\mu = \left(\frac{\partial \log P}{\partial \log \mu}\right)_{\rho,T}. \tag{4.46}$$

For an ideal gas $\chi_\mu = -1$. With the help of eq. (4.45) we can write the variation of density with pressure through the star as

$$\frac{d\log\rho}{d\log P} = \frac{1}{\chi_\rho}\left(1 - \chi_T \frac{d\log T}{d\log P} - \chi_\mu \frac{d\log\mu}{d\log P}\right) = \frac{1}{\chi_\rho}\left(1 - \chi_T \nabla - \chi_\mu \nabla_\mu\right) \tag{4.47}$$

where $\nabla = d\log T/d\log P$ and $\nabla_\mu = d\log\mu/d\log P$ represent the actual gradients of temperature and of mean molecular weight through the star, regarding P as the variable that measures depth. For the displaced gas element the composition does not change, and from eq. (3.61) we can write

$$\frac{1}{\gamma_{\rm ad}} = \frac{1}{\chi_\rho}(1 - \chi_T \nabla_{\rm ad}),$$

so that the stability criterion (4.44) becomes

$$\nabla < \nabla_{\rm ad} - \frac{\chi_\mu}{\chi_T}\nabla_\mu. \tag{4.48}$$

If all the energy is transported by radiation then $\nabla = \nabla_{\rm rad}$ as given by eq. (4.18). Hence we can replace ∇ by $\nabla_{\rm rad}$ in eq. (4.48) and thus arrive at the *Ledoux criterion* which states that a layer is stable against convection if

$$\boxed{\nabla_{\rm rad} < \nabla_{\rm ad} - \frac{\chi_\mu}{\chi_T}\nabla_\mu} \qquad \text{(Ledoux)} \tag{4.49}$$

In chemically homogeneous layers $\nabla_\mu = 0$ and eq. (4.49) reduces to the simple *Schwarzschild criterion* for stability against convection

$$\boxed{\nabla_{\rm rad} < \nabla_{\rm ad}} \qquad \text{(Schwarzschild)} \tag{4.50}$$

(N.B. Note the difference in meaning of the various ∇ symbols appearing in the above criteria: $\nabla_{\rm rad}$ and ∇_μ represent a *spatial* gradient of temperature and mean molecular weight, respectively. On the other hand, $\nabla_{\rm ad}$ represents the temperature variation in a gas element undergoing a pressure variation.)

For an ideal gas ($\chi_T = 1, \chi_\mu = -1$) the Ledoux criterion reduces to

$$\nabla_{\rm rad} < \nabla_{\rm ad} + \nabla_\mu. \tag{4.51}$$

Molecular weight normally increases when going deeper into the star, because nuclear reactions produce more heavier elements in deeper layers. Therefore normally $\nabla_\mu > 0$, so that according to the Ledoux criterion a composition gradient has a stabilizing effect. This is plausible because an upwards displaced element will then have a higher μ than its surroundings, so that even when it is hotter than its new environment (which would make it unstable according to the Schwarzschild criterion) it has a higher density and the buoyancy force will push it back down.

We can relate the convection criterion to the Eddington limit derived in Sec. 4.4. By writing $\nabla_{\rm rad}$ in terms of l, $l_{\rm Edd}$ (defined in eq. 4.37) and $P_{\rm rad} = (1-\beta)P$ we can rewrite the Schwarzschild criterion for stability as

$$l < 4(1-\beta)\nabla_{\rm ad}\, l_{\rm Edd} \tag{4.52}$$

(see one of the exercises). For $\beta > 0$ and $\nabla_{\rm ad} > 0.25$ we see that convection already sets in before the Eddington limit is reached.

Occurrence of convection

According to the Schwarzschild criterion, we can expect convection to occur if

$$\nabla_{\rm rad} = \frac{3}{16\pi acG}\frac{P}{T^4}\frac{\kappa l}{m} > \nabla_{\rm ad}. \tag{4.53}$$

This requires one of following:

- A large value of κ, that is, convection occurs in opaque regions of a star. Examples are the outer envelope of the Sun (see Fig. 4.2) and of other cool stars, because opacity increases with decreasing temperature. Since low-mass stars are cooler than high-mass stars, we may expect low-mass stars to have convective envelopes.

- A large value of l/m, i.e. regions with a large energy flux. We note that towards the centre of a star $l/m \approx \epsilon_{\rm nuc}$ by eq. (4.4), so that stars with nuclear energy production that is strongly peaked towards the centre can be expected to have convective cores. We shall see that this is the case for relatively massive stars.

- A small value of $\nabla_{\rm ad}$, which as we have seen in Sec. 3.5 occurs in partial ionization zones at relatively low temperatures. Therefore, even if the opacity is not very large, the surface layers of a star may be unstable to convection. It turns out that stars of all masses have shallow surface convection zones at temperatures where hydrogen and helium are partially ionized.

These effects are illustrated in Fig. 4.5.

4.5.2 Convective energy transport

We still have to address the question how much energy can be transported by convection and, related to this, what is the actual temperature gradient ∇ inside a convective region. To answer these questions properly requires a detailed theory of convection, which to date remains a very difficult problem in astrophysics that is still unsolved. Even though convection can be simulated numerically, this requires solving the equations of hydrodynamics in three dimensions over a huge range of length scales and time scales, and of pressures, densities and temperatures. Such simulations are therefore very time-consuming and still limited in scope, and cannot be applied in stellar evolution calculations. We have to resort to a very simple one-dimensional 'theory' that is based on rough estimates, and is known as the *mixing length theory* (MLT).

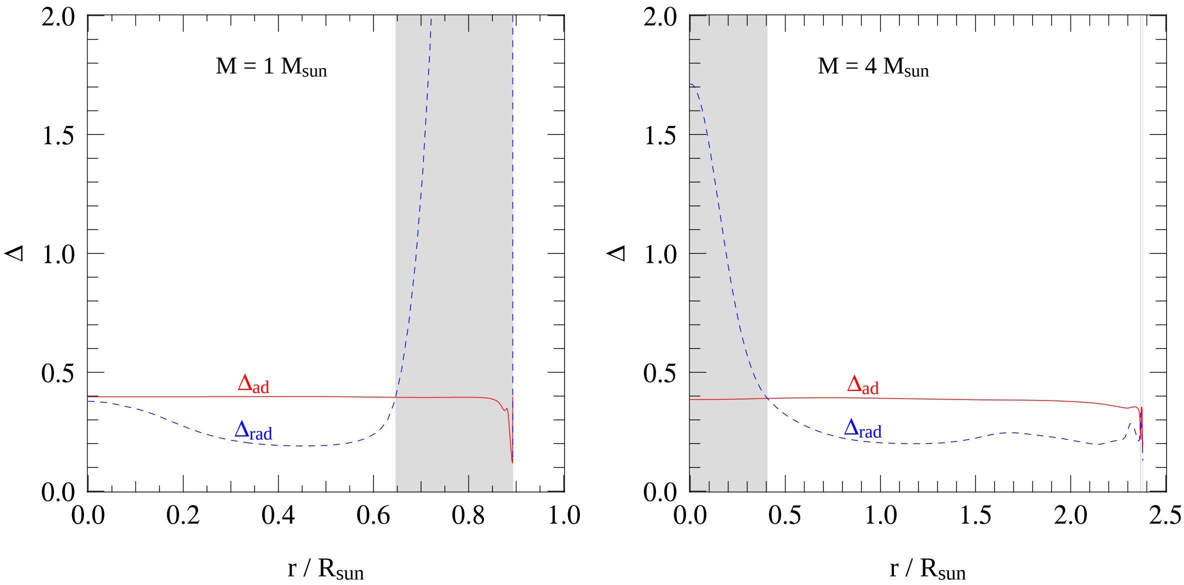


Figure 4.5. The variation of $\nabla_{\rm ad}$ (red, solid line) and $\nabla_{\rm rad}$ (blue, dashed line) with radius in two detailed stellar models of $1\,M_\odot$ and $4\,M_\odot$ at the start of the main sequence. The solar-mass model has a very large opacity in its outer layers, resulting in a large value of $\nabla_{\rm rad}$ which gives rise to a convective envelope where $\nabla_{\rm rad} > \nabla_{\rm ad}$ (indicated by gray shading). On the other hand, the $4\,M_\odot$ model has a hotter outer envelope with lower opacity so that $\nabla_{\rm rad}$ stays small. The large energy generation rate in the centre now results in a large $\nabla_{\rm rad}$ and a convective core extending over the inner $0.4\,R_\odot$. In both models $\nabla_{\rm ad} \approx 0.4$ since the conditions are close to an ideal gas. In the surface ionization zones, however, $\nabla_{\rm ad} < 0.4$ and a thin convective layer appears in the $4\,M_\odot$ model.

In the MLT one approximates convective motions by blobs of gas that travel up or down over a radial distance $\ell_{\rm m}$ (the mixing length), after which they dissolve in their surroundings and lose their identity. In this process it releases its excess heat to its surroundings (or, in the case of downward moving blob, absorbs its heat deficit from its surroundings). The mixing length $\ell_{\rm m}$ is an unknown free parameter in this very schematic model. One presumes that $\ell_{\rm m}$ is of the order of the local pressure scale height, which is the radial distance over which the pressure changes by an e-folding factor,

$$H_P = \left|\frac{dr}{d\ln P}\right| = \frac{P}{\rho g}. \tag{4.54}$$

The last equality holds for a star in hydrostatic equilibrium. The assumption that $\ell_{\rm m} \sim H_P$ is not unreasonable considering that a rising gas blob will expand. Supposing that in a convective region in a star, about half of a spherical surface area is covered by rising blobs and the other half by sinking blobs, the expanding rising blobs will start covering most of the surface area after rising over one or two pressure scale heights.

The convective energy flux

Within the framework of MLT we can calculate the convective energy flux, and the corresponding temperature gradient required to carry this flux, as follows. After rising over a radial distance $\ell_{\rm m}$ the temperature difference between the gas element and its surroundings is

$$\Delta T = T_e - T_{\rm surroundings} = \left[\left(\frac{dT}{dr}\right)_e - \frac{dT}{dr}\right]\ell_{\rm m} = \Delta\left(\frac{dT}{dr}\right)\ell_{\rm m}.$$

Here dT/dr is the ambient temperature gradient, $(dT/dr)_e$ is the variation of temperature with radius that the element experiences as it rises and expands adiabatically, and $\Delta(dT/dr)$ is the difference between these two. We can write ΔT in terms of ∇ and $\nabla_{\rm ad}$ by noting that

$$\frac{dT}{dr} = T\,\frac{d\ln T}{dr} = T\,\frac{d\ln T}{d\ln P}\,\frac{d\ln P}{dr} = -\frac{T}{H_P}\,\nabla \qquad \text{and} \qquad \left(\frac{dT}{dr}\right)_e = -\frac{T}{H_P}\,\nabla_{\rm ad},$$

noting that the $-$ sign appears because $dP/dr < 0$ in eq. (4.54). Hence

$$\Delta T = T\,\frac{\ell_{\rm m}}{H_P}\,(\nabla - \nabla_{\rm ad}). \tag{4.55}$$

The excess of internal energy of the gas element compared to its surrounding is $\Delta u = c_P \Delta T$ per unit mass. If the convective blobs move with an average velocity $\upsilon_{\rm c}$, then the energy flux carried by the convective gas elements is

$$F_{\rm conv} = \upsilon_{\rm c}\,\rho\Delta u = \upsilon_{\rm c}\,\rho c_P \Delta T \tag{4.56}$$

We therefore need an estimate of the average convective velocity. If the difference in density between a gas element and its surroundings is $\Delta\rho$, then the buoyancy force will give an acceleration

$$a = -g\,\frac{\Delta\rho}{\rho} \approx g\,\frac{\Delta T}{T},$$

where the last equality is exact for an ideal gas for which $P \propto \rho T$ and $\Delta P = 0$. The blob is accelerated over a distance $\ell_{\rm m}$, i.e. for a time t given by $\ell_{\rm m} = \frac{1}{2}at^2$. Therefore its average velocity is $\upsilon_{\rm c} \approx \ell_{\rm m}/t = \sqrt{\frac{1}{2}\ell_{\rm m} a}$, that is

$$\upsilon_{\rm c} \approx \sqrt{\tfrac{1}{2}\ell_{\rm m} g \frac{\Delta T}{T}} \approx \sqrt{\frac{\ell_{\rm m}{}^2 g}{2H_P}\,(\nabla - \nabla_{\rm ad})}. \tag{4.57}$$

Combining this with eq. (4.56) gives

$$F_{\rm conv} = \rho c_P T \left(\frac{\ell_{\rm m}}{H_P}\right)^2 \sqrt{\tfrac{1}{2} g H_P}\,(\nabla - \nabla_{\rm ad})^{3/2}. \tag{4.58}$$

The above two equations relate the convective velocity and the convective energy flux to the so-called *superadiabaticity* $\nabla - \nabla_{\rm ad}$, the degree to which the actual temperature gradient ∇ exceeds the adiabatic value.

Estimate of the convective temperature gradient

Which value of $\nabla - \nabla_{\rm ad}$ is required to carry the whole energy flux of a star by convection, i.e. $F_{\rm conv} = l/4\pi r^2$. To make a rough estimate, we take typical values for the interior making use of the virial theorem and assuming an ideal gas:

$$\rho \approx \bar{\rho} = \frac{3M}{4\pi R^3} \qquad T \approx \bar{T} \sim \frac{\mu}{\mathcal{R}}\frac{GM}{R} \qquad c_P = \frac{5}{2}\frac{\mathcal{R}}{\mu} \qquad \sqrt{gH_P} = \sqrt{\frac{P}{\rho}} = \sqrt{\frac{\mathcal{R}}{\mu}T} \sim \sqrt{\frac{GM}{R}}$$

noting that the last expression is also approximately equal to the average speed of sound $\upsilon_{\rm s}$ in the interior. We then obtain, neglecting factors of order unity,

$$F_{\rm conv} \sim \frac{M}{R^3}\left(\frac{GM}{R}\right)^{3/2} (\nabla - \nabla_{\rm ad})^{3/2}. \tag{4.59}$$

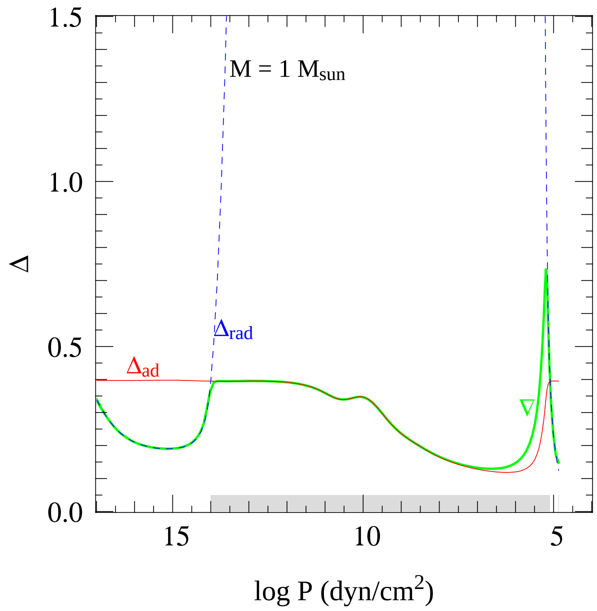


Figure 4.6. The variation of ∇_{ad} (red, solid line) and ∇_{rad} (blue, dashed line) in the same detailed model of 1 $M_\odot$ as shown in Fig. 4.5, but now plotted against $\log P$ rather than radius to focus on the outermost layers (where the pressure gradient is very large). The thick green line shows the actual temperature gradient ∇. The partial ionization zones are clearly visible as depressions in ∇_{ad} (compare to Fig. 3.5b). The convection zone stretches from $\log P \approx 14$ to 5 (indicated by a gray bar along the bottom). In the deep interior (for $\log P > 8$) convection is very efficient and $\nabla = \nabla_{ad}$. Higher up, at lower pressures, convection becomes less and less efficient at transporting energy and requires a larger T-gradient, $\nabla > \nabla_{ad}$. In the very outer part of the convection zone convection is very inefficient and $\nabla \approx \nabla_{rad}$.

If we substitute $F_{conv} = l/4\pi r^2 \sim L/R^2$ then we can rewrite the above to

$$\nabla - \nabla_{ad} \sim \left(\frac{LR}{M}\right)^{2/3} \frac{R}{GM} \tag{4.60}$$

Putting in typical numbers, i.e. solar luminosity, mass and radius, we obtain the following rough estimate for the superadiabaticity in the deep interior of a star like the Sun

$$\nabla - \nabla_{ad} \sim 10^{-8}$$

Convection is so efficient at transporting energy that only a tiny superadiabaticity is required. This means that $F_{conv} \gg F_{rad}$ in convective regions. A more accurate estimate yields $\nabla - \nabla_{ad} \sim 10^{-5} \ldots 10^{-7}$, which is still a very small number. We can conclude that in the deep stellar interior the actual temperature stratification is nearly adiabatic, and independent of the details of the theory. Therefore a detailed theory of convection is not needed for energy transport by convection and we can simply take

$$\frac{dT}{dm} = -\frac{Gm}{4\pi r^4}\frac{T}{P}\nabla \qquad \text{with} \qquad \nabla = \nabla_{ad}. \tag{4.61}$$

However in the outermost layers the situation is different, because $\rho \ll \bar{\rho}$ and $T \ll \bar{T}$. Therefore F_{conv} is much smaller and the superadiabaticity becomes substantial ($\nabla > \nabla_{ad}$). The actual temperature gradient then depends on the details of the convection theory. Within the context of MLT, the T-gradient depends on the assumed value of $\alpha_m = \ell_m/H_P$. In practice one often calibrates detailed models computed with different values of α_m to the radius of the Sun and of other stars with well-measured radii. The result of this procedure is that the best match is obtained for $\alpha_m \approx 1.5 \ldots 2$.

As the surface is approached, convection becomes very inefficient at transporting energy. Then $F_{\rm conv} \ll F_{\rm rad}$ so that radiation effectively transports all the energy, and $\nabla \approx \nabla_{\rm rad}$ despite convection taking place. These effects are shown in Fig. 4.6 for a detailed solar model.

4.5.3 Convective mixing

Besides being an efficient means of transporting energy, convection is also a very efficient *mixing mechanism*. We can see this by considering the average velocity of convective cells, eq. (4.57), and taking $\ell_{\rm m} \approx H_P$ and $\sqrt{gH_P} \approx v_{\rm s}$, so that

$$v_{\rm c} \approx v_{\rm s}\,\sqrt{\nabla - \nabla_{\rm ad}}. \tag{4.62}$$

Because $\nabla - \nabla_{\rm ad}$ is only of the order 10^{-6} in the deep interior, typical convective velocities are strongly subsonic, by a factor $\sim 10^{-3}$, except in the very outer layers where $\nabla - \nabla_{\rm ad}$ is substantial. This is the main reason why convection has no disruptive effects, and overall hydrostatic equilibrium can be maintained in the presence of convection.

By substituting into eq. (4.62) rough estimates for the interior of a star, i.e. $v_{\rm s} \sim \sqrt{GM/R}$ and eq. (4.60) for $\nabla - \nabla_{\rm ad}$, we obtain $v_{\rm c} \sim (LR/M)^{1/3} \approx 5 \times 10^3$ cm/s for a star like the Sun. These velocities are large enough to mix a convective region on a small timescale. We can estimate the timescale on which a region of radial size $d = qR$ is mixed as $\tau_{\rm mix} \approx d/v_{\rm c} \sim q(R^2M/L)^{1/3}$, which is $\sim q \times 10^7$ sec for solar values. Depending on the fractional extent q of a convective region, the convective mixing timescale is of the order of weeks to months. Hence $\tau_{\rm mix} \ll \tau_{\rm KH} \ll \tau_{\rm nuc}$, so that over a thermal timescale, and certainly over a nuclear timescale, a convective region inside a star will be mixed homogeneously. (Note that convective mixing remains very efficient in the outer layers of a star, even though convection becomes inefficient at transporting energy.)

This has important consequences for stellar evolution, which we will encounter in future chapters. Briefly, the large efficiency of convective mixing means that:

- A star in which nuclear burning occurs in a *convective core* will homogenize the region inside the core by transporting burning ashes (e.g. helium) outwards and fuel (e.g. hydrogen inwards). Such a star therefore has a larger fuel supply and can extend its lifetime compared to the hypothetical case that convection would not occur.

- A star with a deep *convective envelope*, such that it extends into regions where nuclear burning has taken place, will mix the burning products outwards towards the surface. This process (often called 'dredge-up'), which happens when stars become red giants, can therefore modify the surface composition, and in such a star measurements of the surface abundances provide a window into nuclear processes that have taken place deep inside the star.

Composition changes inside a star will be discussed in the next chapter.

4.5.4 Convective overshooting

To determine the extent of a region that is mixed by convection, we need to look more closely at what happens at the boundary of a convective zone. According to the Schwarzschild criterion derived in Sec. 4.5.1, in a chemically homogeneous layer this boundary is located at the surface where $\nabla_{\rm rad} = \nabla_{\rm ad}$. At this point the acceleration due to the buoyancy force, $a \approx g(\nabla - \nabla_{\rm ad})$, vanishes. Just outside this boundary, the acceleration changes sign and a convective bubble will be strongly braked – even more so when the non-mixed material outside the convective zone has a lower μ and hence a lower density. However, the convective eddies have (on average) a non-zero velocity when they cross the

Schwarzschild boundary, and will *overshoot* by some distance due to their inertia. A simple estimate of this overshooting distance shows that it should be much smaller than a pressure scale height, so that the Schwarzschild criterion should determine the convective boundary quite accurately. However the convective elements also carry some heat and mix with their surroundings, so that both $|\nabla - \nabla_{\rm ad}|$ and the μ-gradient decrease. Thus also the effective buoyancy force that brakes the elements decreases, and a positive feedback loop can develop that causes overshooting elements to penetrate further and further. This is a highly non-linear effect, and as a result the actual overshooting distance is very uncertain and could be substantial.

Convective overshooting introduces a large uncertainty in the extent of mixed regions, with important consequences for stellar evolution. A convectively mixed core that is substantially larger will generate a larger fuel supply for nuclear burning, and thus affects both the hydrogen-burning lifetime and the further evolution of a star. In stellar evolution calculations one usually parametrizes the effect of overshooting by assuming that the distance $d_{\rm ov}$ by which convective elements penetrate beyond the Schwarzschild boundary is a fixed fraction of the local pressure scale height, $d_{\rm ov} = \alpha_{\rm ov} H_P$. Here $\alpha_{\rm ov}$ is a free parameter, that can be calibrated against observations (see later chapters).

Suggestions for further reading

The contents of this chapter are also covered by Chapters 3, 5 and 8 of Maeder, by Chapters 4, 5, 7 and 17 of Kippenhahn and by Chapters 4 and 5 of Hansen.

Exercises

4.1 Radiation transport

The most important way to transport energy form the interior of the star to the surface is by radiation, i.e. photons traveling from the center to the surface.

(a) How long does it typically take for a photon to travel from the center of the Sun to the surface? [Hint: estimate the mean free path of a photon in the central regions of the Sun.] How does this relate to the thermal timescale of the Sun?

(b) Estimate a typical value for the temperature gradient dT/dr. Use it to show that the difference in temperature ΔT between to surfaces in the solar interior one photon mean free path $\ell_{\rm ph}$ apart is

$$\Delta T = \ell_{\rm ph} \frac{dT}{dr} \approx 2 \times 10^{-4}\,\mathrm{K}.$$

In other words the anisotropy of radiation in the stellar interior is very small. This is why radiation in the solar interior is close to that of a black body.

(c) Verify that a gas element in the solar interior, which radiates as a black body, emits $\approx 6 \times 10^{23}\,\mathrm{erg\,cm^{-2}\,s^{-1}}$.

If the radiation field would be exactly isotropic, then the same amount of energy would radiated into this gas element by the surroundings and so there would be no net flux.

(d) Show that the minute deviation from isotropy between two surfaces in the solar interior one photon mean free path apart at $r \sim R_\odot/10$ and $T \sim 10^7$ K, is sufficient for the transfer of energy that results in the luminosity of the Sun.

(e) Why does the diffusion approximation for radiation transport break down when the surface (photosphere) of a star is approached?

4.2 Mass-luminosity relation

Without solving the stellar structure equations, we can already derive useful scaling relations. In this question you will use the equation for radiative energy transport with the equation for hydrostatic equilibrium to derive a scaling relation between the mass and the luminosity of a star.

(a) Derive how the central temperature, T_c, scales with the mass, M, radius, R, and luminosity, L, for a star in which the energy transport is by radiation. To do this, use the stellar structure equation (4.16) for the temperature gradient in radiative equilibrium. Assume that $r \sim R$ and that the temperature is proportional to T_c, $l \sim L$ and estimating $dT/dr \sim -T_c/R$.

(b) Derive how T_c scales with M and R, using the hydrostatic equilibrium equation, and assuming that the ideal gas law holds.

(c) Combine the results obtained in (a) and (b), to derive how L scales with M and R for a star whose energy transport is radiative.

You have arrived at a mass-luminosity relation without assuming anything about how the energy is *produced*, only about how it is *transported* (by radiation). It shows that the luminosity of a star is *not* determined by the rate of energy production in the centre, but by how fast it can be transported to the surface!

(d) Compare your answer to the relation between M and L which you derived from observations (Exercise 1.3). Why does the derived power-law relation starts to deviate from observations for low mass stars? Why does it deviate for high mass stars?

4.3 Conceptual questions: convection

(a) Why does convection lead to a net heat flux upwards, even though there is no net mass flux (upwards and downwards bubbles carry equal amounts of mass)?

(b) Explain the Schwarzschild criterion

$$\left(\frac{d\ln T}{d\ln P}\right)_{\rm rad} > \left(\frac{d\ln T}{d\ln P}\right)_{\rm ad}$$

in simple physical terms (using Archimedes law) by drawing a schematic picture . Consider both cases $\nabla_{\rm rad} > \nabla_{\rm ad}$ and $\nabla_{\rm rad} < \nabla_{\rm ad}$. Which case leads to convection?

(c) What is meant by the *superadiabaticity* of a convective region? How is it related to the convective energy flux (qualitatively)? Why is it very small in the interior of a star, but can be large near the surface?

4.4 Applying Schwarzschild's criterion

(a) Low-mass stars, like the Sun, have convective envelopes. The fraction of the mass that is convective increases with decreasing mass. A 0.1 $M_\odot$ star is completely convective. Can you qualitatively explain why?

(b) In contrast higher-mass stars have radiative envelopes and convective cores, for reasons we will discuss in the coming lectures. Determine if the energy transport is convective or radiative at two different locations ($r = 0.242R_\odot$ and $r = 0.670R_\odot$) in a $5M_\odot$ main sequence star. Use the data of a 5 $M_\odot$ model in the table below. You may neglect the radiation pressure and assume that the average particle has a mass $\bar{m} = 0.7m_{\rm u}$.

$r/R_\odot$	$m/M_\odot$	$L_r/L_\odot$	T [K]	ρ [g cm^{-3}]	κ [g^{-1} cm^2]
0.242	0.199	3.40×10^2	2.52×10^7	18.77	0.435
0.670	2.487	5.28×10^2	1.45×10^7	6.91	0.585

4.5 Comparing radiative and convective cores

Consider a H-burning star of mass $M = 3M_\odot$, with a luminosity L of $80L_\odot$, and an initial composition $X = 0.7$ and $Z = 0.02$. The nuclear energy is generated only in the central 10% of the mass, and the energy generation rate per unit mass, ϵ_{nuc}, depends on the mass coordinate as

$$\epsilon_{nuc} = \epsilon_c \left(1 - \frac{m}{0.1M}\right)$$

(a) Calculate and draw the luminosity profile, l, as a function of the mass, m. Express ϵ_c in terms of the known quantities for the star.

(b) Assume that all the energy is transported by radiation. Calculate the H-abundance as a function of mass and time, $X = X(m,t)$. What is the central value for X after 100 Myr? Draw X as a function of m. (Hint: the energy generation per unit mass is $Q = 6.3 \times 10^{18}$ erg g^{-1}).

(c) In reality, $\epsilon_{\rm nuc}$ is so high that the inner 20% of the mass is unstable to convection. Now, answer the same question as in (b) and draw the new X profile as a function of m. By how much is the central H-burning lifetime extended as a result of convection?

4.6 The Eddington luminosity

The Eddington luminosity is the maximum luminosity a star (with radiative energy transport) can have, where radiation force equals gravity.

(a) Show that

$$l_{\rm max} = \frac{4\pi cGm}{\kappa}.$$

(b) Consider a star with a uniform opacity κ and of uniform parameter $1 - \beta = P_{\rm rad}/P$. Show that $L/L_{\rm Edd} = 1 - \beta$ for such a star.

(c) Show that the Schwarzschild criterion for stability against convection $\nabla_{\rm rad} < \nabla_{\rm ad}$ can be rewritten as:

$$\frac{l}{l_{\rm Edd}} < 4\frac{P_{\rm rad}}{P}\nabla_{\rm ad}$$

(d) Consider again the star of question (b). By assuming that it has a convective core, and no nuclear energy generation outside the core, show that the mass fraction of this core is given by

$$\frac{M_{\rm core}}{M} = \frac{1}{4\nabla_{ad}}.$$

Chapter 5

Nuclear reactions in stars

For a star in thermal equilibrium, an internal energy source is required to balance the energy loss from the surface in the form of radiation. This energy source is provided by *nuclear reactions* that take place in the deep interior, where the temperature and density are sufficiently high. Apart from energy generation, another important effect of nuclear reactions is that they change the composition by transmutations of one element into another.

5.1 Nuclear reactions and composition changes

Consider a nuclear reaction whereby two nuclei (denoted I and J) combine to produce two other nuclei (K and L):

$$I + J \rightarrow K + L \tag{5.1}$$

Many, though not all, reactions are of this type, and the general principles discussed here also apply to reactions involving different numbers of nuclei. Each nucleus is characterized by two integers, the baryon number or mass number A_i and the charge Z_i. Nuclear charges and baryon numbers must be conserved during a reaction, i.e. for the example above:

$$Z_i + Z_j = Z_k + Z_l \qquad \text{and} \qquad A_i + A_j = A_k + A_l \tag{5.2}$$

Some reactions involve electrons or positrons, or neutrinos or antineutrinos. In that case the lepton number must also be conserved during the reaction. Therefore any three of the reactants uniquely determine the fourth. The reaction rate – the number of reactions taking place per cm^3 and per second – can therefore be identified by three indices. We denote the reaction rate of reaction (5.1) as $r_{ij,k}$.

The rate at which the number density n_i of nuclei I changes with time can generally be the result of different nuclear reactions, some of which consume I as above, and others that proceed in the reverse direction and thereby produce I. If we denote the rate of reactions of the latter type as $r_{kl,i}$, we can write

$$\frac{dn_i}{dt} = -\sum_{j,k} r_{ij,k} + \sum_{k,l} r_{kl,i} \tag{5.3}$$

The number density of nucleus I is related to the mass fraction X_i by $n_i = X_i\,\rho/(A_i m_u)$, so that we can write the rate of change of the mass fraction due to nuclear reactions as

$$\boxed{\frac{dX_i}{dt} = A_i \frac{m_u}{\rho}\Big(-\sum_{j,k} r_{ij,k} + \sum_{k,l} r_{kl,i}\Big)} \tag{5.4}$$

For each nuclear species i one can write such an equation, describing the composition change at a particular mass shell inside the star (with density ρ and temperature T) resulting from nuclear reactions. In the presence of internal mixing (in particular of *convection*, Sec. 4.5.3) the redistribution of composition between different mass shells should also be taken into account.

5.2 Nuclear energy generation

The energy released (or absorbed) per reaction of type $I + J \to K + L$ (eq. 5.1) is

$$Q_{ij,k} = (m_i + m_j - m_k - m_l)c^2, \tag{5.5}$$

where m_i are the actual masses of the nuclei involved. Note that $Q_{ij,k} \neq 0$ since $m_i \neq A_i m_u$ (except, per definition, for ^{12}C, e.g. see Table 5.1). When $Q_{ij,k} > 0$ energy is released and one speaks of an *endothermic* reaction. $Q_{ij,k}$, often called the Q-value of a reaction and usually expressed in MeV, corresponds to the difference in binding energy of the nuclei involved:

$$E_B(A_Z) := [(A - Z)m_n + Zm_p - m(A_Z)]\, c^2, \tag{5.6}$$

where $m(A_Z)$ is the mass of a nucleus with mass A and proton number Z, and m_n abd m_p are the masses of a free neutron and proton respectively. The binding energy is related to the 'mass defect' of a nucleus,

$$\Delta m = (A - Z)m_n + Zm_p - m(A_Z).$$

The binding energy per nucleon, displayed in Fig. 5.1 against mass number A, is an informative quantity. Energy can be gained from the fusion of light nuclei into heavier ones if E_B/A increases.

The energy generation rate (in erg g^{-1} s^{-1}) from the reaction $I + J \to K + L$ is

$$\epsilon_{ij,k} = \frac{Q_{ij,k} r_{ij,k}}{\rho}. \tag{5.7}$$

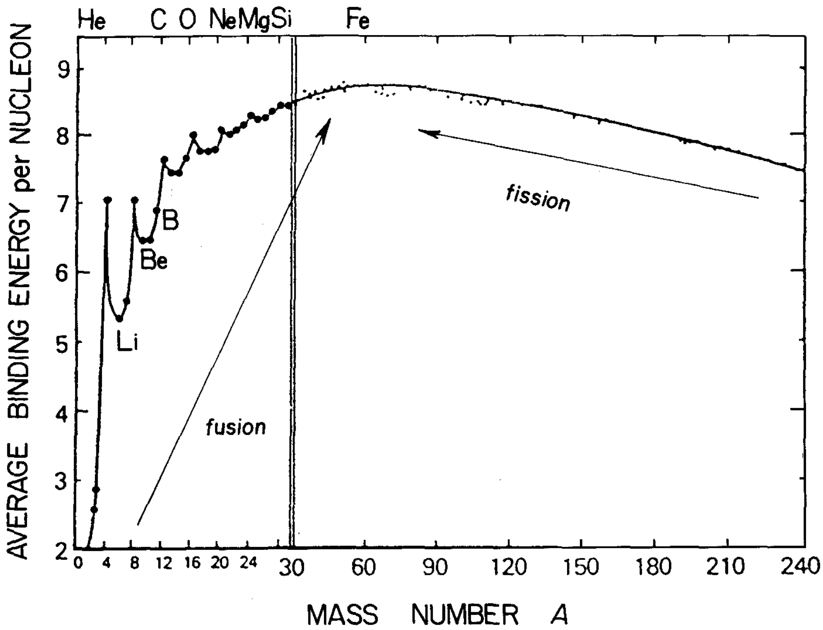


Figure 5.1. Binding energy per nucleon E_B/A

Table 5.1. Masses of several important nuclei

element	Z	A	M/m_u	element	Z	A	M/m_u	element	Z	A	M/m_u
n	0	1	1.008665	C	6	12	12.000000	Ne	10	20	19.992441
H	1	1	1.007825		6	13	13.003354	Mg	12	24	23.985043
	1	2	2.014101	N	7	13	13.005738	Si	14	28	27.976930
He	2	3	3.016029		7	14	14.003074	Fe	26	56	55.934940
	2	4	4.002603		7	15	15.000108	Ni	28	56	55.942139
Li	3	6	6.015124	O	8	15	15.003070				
	3	7	7.016003		8	16	15.994915				
Be	4	7	7.016928		8	17	16.999133				
	4	8	8.005308		8	18	17.999160				

Therfore the total nuclear energy gerenation rate is

$$\epsilon_{\rm nuc} = \sum_{i,j,k} \epsilon_{ij,k} = \sum_{i,j,k} \frac{Q_{ij,k}}{\rho}\, r_{ij,k} \tag{5.8}$$

Note the similarity between the expressions for the nuclear energy generation rate and the equation for composition changes, eq. (5.4). Both are proportional to $r_{ij,k}$, so that for a simple case where only *one* reaction occurs (or one reaction dominates a reaction chain) one has $\epsilon_{\rm nuc} = (Q/A_i m_{\rm u})\, dX_i/dt$.

5.3 Thermonuclear reaction rates

Consider nuclei of type i and j with number densities n_i and n_j. Suppose their relative velocity is υ, and their reaction cross section at this velocity is $\sigma(\upsilon)$. The rate (per second) at which a particular nucleus i captures nuclei of type j is then $n_j \upsilon \sigma(\upsilon)$. The nuclear reaction rate (number of reactions s^{-1} cm^{-3}) at relative velocity υ is therefore

$$r_{ij}(\upsilon) = \frac{1}{1+\delta_{ij}}\, n_i\, n_j\, \upsilon\, \sigma(\upsilon). \tag{5.9}$$

The factor $1/(1+\delta_{ij}) = \frac{1}{2}$ for reactions between nuclei of the same type, which prevents such reactions to be counted twice.

In a stellar gas there is a distribution of velocities which (in the case of an ideal gas in LTE) is given by the Maxwell-Boltzmann distribution, eq. (3.13). If each particle velocity follows a M-B distribution, then also their *relative* velocity follows a M-B distribution,

$$\phi(\upsilon) = 4\pi\upsilon^2 \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{m\upsilon^2}{2kT}\right), \tag{5.10}$$

where m is the reduced mass in the centre-of-mass frame

$$m = \frac{m_1 m_2}{m_1 + m_2}. \tag{5.11}$$

The reaction rate is then

$$r_{ij} = \frac{1}{1+\delta_{ij}}\, n_i n_j \langle\sigma\upsilon\rangle = \frac{1}{1+\delta_{ij}}\, n_i n_j \int_0^\infty \phi(\upsilon)\upsilon\, \sigma(\upsilon)\, d\upsilon \tag{5.12}$$

We see that the dependence on velocity in eq. (5.9) now becomes a dependence on the *temperature* in the expression for the overall reaction rate. This temperature dependence is expressed in the product $\langle\sigma\upsilon\rangle$,

$$\langle\sigma\upsilon\rangle = \left(\frac{8}{\pi m}\right)^{1/2} (kT)^{-3/2} \int_0^\infty \sigma(E)\, E\, \exp\left(-\frac{E}{kT}\right) dE, \tag{5.13}$$

where we have replaced the relative velocity υ by the kinetic energy in the centre-of-mass frame, $E = \frac{1}{2}m\upsilon^2$.

5.3.1 Coulomb barrier

Nuclear reactions normally take place between charged nuclei which experience a repulsive Coulomb potential

$$E_C = \frac{Z_1 Z_2 e^2}{r} \tag{5.14}$$

if the nuclear charges are Z_1 and Z_2. To experience the attractive nuclear force the particles have to approach each other within a typical distance $r_n \sim 10^{-13}$ cm. They must therefore overcome a typical Coulomb barrier $E_C(r_n) \approx Z_1 Z_2$ MeV, see Fig. 5.2.

We can compare this to typical particle energies at 10^7 K), $\langle E_{\rm kin}\rangle = \frac{3}{2}kT \approx 1$ keV. This falls short of the Coulomb barrier by a factor of more than 1000! With purely classical considerations the probability of nuclear reactions happening at such temperatures (typical of the centre of the Sun and other main-sequence stars) is vanishingly small. We need to turn to quantum mechanics to see how nuclear reactions are possible at these temperatures.

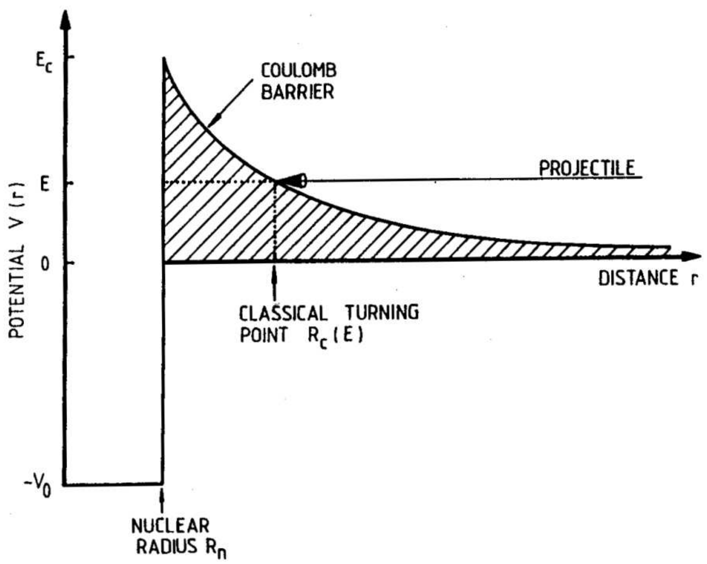


Figure 5.2. Depiction of the combined nuclear and Coulomb potential.

5.3.2 Tunnel effect

- even if $E \ll E_C(r_n) \Rightarrow$ finite probability that projectile penetrates repulsive Coulomb barrier
- tunneling probability: $P = \exp(-b \cdot E^{-1/2})$

  approximately
  $$P \approx \exp(-r_c/\lambda)$$
  $$r_c = \frac{Z_1 Z_2 e^2}{E}$$
  $$\lambda = \frac{\hbar}{p} = \frac{\hbar}{\sqrt{2mE}} \quad \text{(de Broglie wavelength)}$$
- quantitatively: $P = \exp(-2\pi\eta)$
  $$2\pi\eta = 2\pi \frac{Z_1 Z_2 e^2 (\frac{1}{2}m)^{1/2}}{\hbar E^{1/2}} = 31.29 \cdot Z_1 Z_2 \left(\frac{A}{E}\right)^{\frac{1}{2}} \qquad (E \text{ in keV})$$
  $$A = \frac{A_1 A_2}{A_1 + A_2} \text{ (reduced mass in } m_\mathrm{u})$$

5.3.3 Cross sections

- classically: $\sigma = \pi(R_1 + R_2)^2$
- QM: $\sigma = \pi\lambda^2$ with $\lambda = \dfrac{\hbar}{p} \propto \dfrac{1}{\sqrt{E}}$
- tunneling probability: $P = \exp(-b/\sqrt{E})$

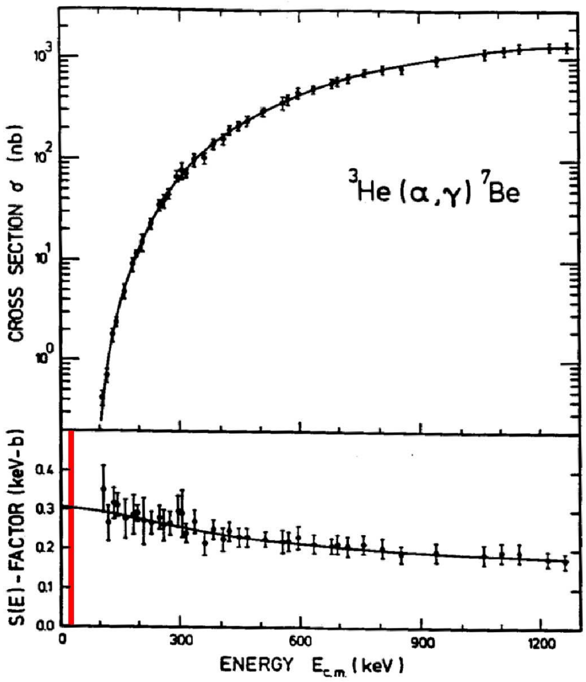


Figure 5.3. Example of the dependence of the reaction cross section with energy E for the ^{3}He+^{4}He $\rightarrow$ ^{7}Be+γ reaction. Although σ varies very strongly with energy, and becomes immeasurable at very low E, the $S(E)$-factor is only a very weak function of E and (in the case of this reaction) can be safely extrapolated to the low energies that are relevant for nuclear reactions in stars (red vertical bar).

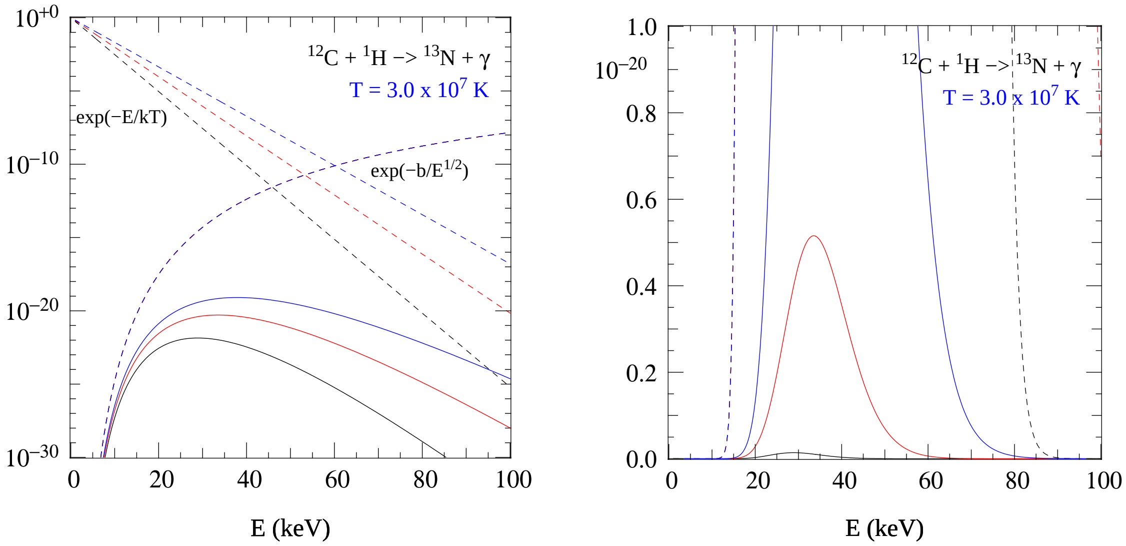


Figure 5.4. Example of the Gamow peak for the $^{12}C + p \rightarrow {}^{13}N + \gamma$ reaction. The left panel shows as dasked lines the tunnel probablity factor $(\exp(-b/E^{1/2}))$ and the tail of the Maxwell distribution $(\exp(-E/kT))$ for three values of temperature: $T = 2.0 \times 10^7$ K (lower, black), 2.5×10^7 K (middle, red) and 3.0×10^7 K (upper, blue). The solid lines are the resulting Gamow peaks. To appreciate the sharpness of the peak, and the enormous sensitivity to temperature, the same curves are plotted on a linear scale in the right panel. The Gamow peak for $T = 2.0 \times 10^7$ K is barely visible at ~ 30 keV, while that for 3.0×10^7 K is off the scale.

Define $S(E)$ as follows:

$$\boxed{\sigma(E) := S(E)\frac{\exp(-b/\sqrt{E})}{E}} \tag{5.15}$$

$S(E)$ = "astrophysical S-factor"

- slowly varying function of E (unlike σ !)
- intrinsic nuclear properties (e.g. resonances)
- extrapolated from measurements at high E

5.3.4 The Gamow peak

Combining eqs. (5.13) and (5.15) $\Rightarrow$

$$\langle\sigma v\rangle = (8/\pi m)^{\frac{1}{2}}(kT)^{-\frac{3}{2}}\int_0^\infty S(E)\exp\left(-\frac{E}{kT}-\frac{b}{\sqrt{E}}\right)dE \tag{5.16}$$

$$\approx (8/\pi m)^{\frac{1}{2}}(kT)^{-\frac{3}{2}}\,S(E_0)\int_0^\infty \exp\left(-\frac{E}{kT}-\frac{b}{\sqrt{E}}\right)dE \tag{5.17}$$

where the integrand, with two competing exponentials, describes the so-called *Gamow peak* (see figure below). See K&W chapter 18 for a discussion of the properties.

For different reactions, $b \propto Z_1 Z_2\sqrt{\frac{A_1 A_2}{A_1+A_2}}$, so that a reaction between heavier nuclei (larger A and Z) will have a much lower rate at constant temperature. In other words, it will need a higher T to occur at an appreciable rate.

5.3.5 Temperature dependence of reaction rates

Approximate Gamow peak by Gaussian:

$$\text{height} = \exp(-\frac{3E_0}{kT}) \propto \exp(-\frac{b^{\frac{2}{3}}}{T^{\frac{1}{3}}})$$

$$\text{width} = 4\sqrt{\frac{E_0 kT}{3}} \propto b^{\frac{1}{3}} T^{\frac{5}{6}}$$

So that from eq. (5.17):

$$\langle\sigma\upsilon\rangle \propto b^{\frac{1}{3}} T^{-\frac{2}{3}} \exp(-\frac{b^{\frac{2}{3}}}{T^{\frac{1}{3}}}) \tag{5.18}$$

This means that

- $\langle\sigma\upsilon\rangle$ drops strongly with increasing Coulomb barrier
- $\langle\sigma\upsilon\rangle$ increases *very strongly with temperature*

Approximate $\langle\sigma\upsilon\rangle \approx \langle\sigma\upsilon\rangle_0 T^\nu$, where $\nu := \frac{d\log\langle\sigma\upsilon\rangle}{d\log T} \gg 1$

example ($T = 1.5 \cdot 10^7 K$):

reaction	$\langle\sigma\upsilon\rangle \propto$	E_C
p + p	$T^{3.9}$	0.55 MeV
p + ^{14}N	T^{20}	2.27 MeV
α + ^{12}C	T^{42}	3.43 MeV
^{16}O + ^{16}O	T^{182}	14.07 MeV

Sensitivity of reaction rates to T and to $Z_1 Z_2$ implies:

- nuclear burning of H, He, C, etc are well separated in temperature
- only few reactions occur at same time

Energy generation rate

combining expressions for $\epsilon_{ij} = Q_{ij}\frac{r_{ij}}{\rho}$ and $r_{ij} = n_i n_j \langle\sigma\upsilon\rangle \Rightarrow$
energy generation rate (for two-particle reactions):

$$\epsilon_{ij} = \frac{Q_{ij}}{m_u^2}\frac{X_i X_j}{A_i A_j}\rho\langle\sigma\upsilon\rangle \approx \epsilon_0 X_i X_j \rho T^\nu \tag{5.19}$$

where ϵ_0 depends strongly on the Coulomb barrier ($Z_i Z_j$)

5.4 The main nuclear burning cycles

5.4.1 Hydrogen burning

Net effect of hydrogen burning:

$$4\,{}^{1}\mathrm{H} \rightarrow {}^{4}\mathrm{He} + 2\,\mathrm{e}^{+} + 2\,\nu \quad (Q = 26.73\,\mathrm{MeV})$$

energy release per gram: $Q/m({}^{4}\mathrm{He}) = 6.4 \times 10^{18}$ erg/g

- two neutrinos released per ^{4}He (weak interactions, $\mathrm{p} \rightarrow \mathrm{n}$) $\Rightarrow$ decrease of the effective Q value
- four-particle reaction extremely unlikely $\Rightarrow$ chain of reactions necessary
- first reaction is the pp reaction:

  $$\mathrm{p} + \mathrm{p} \rightarrow \mathrm{D} + \mathrm{e}^{+} + \nu$$

  involves simultaneous β-decay of p $\Rightarrow$ extremely small cross section, $10^{-20}\times$ that of strong interaction. The reaction rate is unmeasurably small $\Rightarrow$ only known from theory

5.4.2 The p-p chains

After some D is produced, several reaction paths are possible:

$${}^{1}\mathrm{H} + {}^{1}\mathrm{H} \rightarrow {}^{2}\mathrm{H} + \mathrm{e}^{+} + \nu$$
$${}^{2}\mathrm{H} + {}^{1}\mathrm{H} \rightarrow {}^{3}\mathrm{He} + \gamma$$

$${}^{3}\mathrm{He} + {}^{3}\mathrm{He} \rightarrow {}^{4}\mathrm{He} + 2\,{}^{1}\mathrm{H}$$

ppI

$${}^{3}\mathrm{He} + {}^{4}\mathrm{He} \rightarrow {}^{7}\mathrm{Be} + \gamma$$

$${}^{7}\mathrm{Be} + \mathrm{e}^{-} \rightarrow {}^{7}\mathrm{Li} + \nu$$
$${}^{7}\mathrm{Li} + {}^{1}\mathrm{H} \rightarrow {}^{4}\mathrm{He} + {}^{4}\mathrm{He}$$

ppII

$${}^{7}\mathrm{Be} + {}^{1}\mathrm{H} \rightarrow {}^{8}\mathrm{B} + \gamma$$
$${}^{8}\mathrm{B} \rightarrow {}^{8}\mathrm{Be}^{*} + \mathrm{e}^{+} + \nu$$
$${}^{8}\mathrm{Be}^{*} \rightarrow {}^{4}\mathrm{He} + {}^{4}\mathrm{He}$$

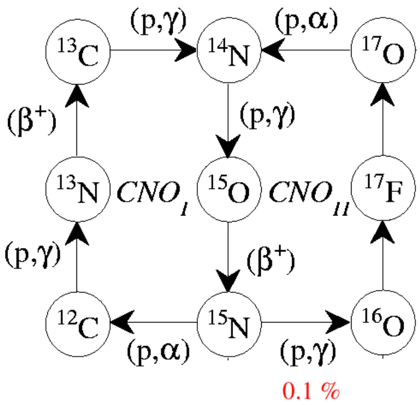
ppIII

For $T > 8 \times 10^{6}$ K $\Rightarrow$ reactions in equilibrium $\Rightarrow$ overall reaction rate = r_{pp} (slowest reaction)

Energy production by pp chain

Neutrino release:
ppI: 0.53 MeV per ^{4}He
ppII: 0.81 MeV per ^{4}He
ppIII: 6.71 MeV per ^{4}He
$\Rightarrow$ effective Q value depends on temperature

5.4.3 The CNO cycle



- net effect: $4\,{}^{1}\mathrm{H} \rightarrow {}^{4}\mathrm{He} + 2\,\mathrm{e}^{+} + 2\,\nu$
- at high enough T: reactions in equilibrium
- total number of CNO-nuclei is conserved
- define lifetime of nuclear species i against reacting with particle j:

$$\tau_j(i) \;:=\; \left|\frac{n_i}{[dn_i/dt]_j}\right| \;=\; \frac{1}{n_j\langle\sigma\upsilon\rangle_{ij}} \tag{5.20}$$

at $T \simeq 2\cdot 10^7$ K: $\tau_\mathrm{p}(^{15}\mathrm{N}) \ll \tau_\mathrm{p}(^{13}\mathrm{C}) < \tau_\mathrm{p}(^{12}\mathrm{C}) \ll \tau_\mathrm{p}(^{14}\mathrm{N}) \ll \tau_\mathrm{star}$
i.e. 35 yr 1600 yr 6600 yr 9×10^5 yr

- at high enough T, ^{12}C, ^{13}C, ^{14}N and ^{15}N will seek equilibrium abundances, given by $dn(^{12}\mathrm{C})/dt = dn(^{13}\mathrm{C})/dt = ...$, etc $\Rightarrow$

$$\left(\frac{n(^{12}\mathrm{C})}{n(^{13}\mathrm{C})}\right)_\mathrm{eq} = \frac{\langle\sigma v\rangle_{13}}{\langle\sigma v\rangle_{12}} = \frac{\tau_\mathrm{p}(^{12}\mathrm{C})}{\tau_\mathrm{p}(^{13}\mathrm{C})}, \qquad \text{etc} \tag{5.21}$$

- overall rate of CNO-cycle deteremined by the $^{14}\mathrm{N}(\mathrm{p},\gamma)^{15}\mathrm{O}$ reaction
- apart from ^{4}He, ^{14}N is the major product of the CNO-cycle

CNO cycle vs pp chain

See K&W figure 18.8.

Temperature sensitivity approximately:

$$\epsilon_\mathrm{pp} \propto X^2\rho\, T^4 \tag{5.22}$$

$$\epsilon_\mathrm{CNO} \propto X\, X_{14}\,\rho\, T^{18} \tag{5.23}$$

5.4.4 Helium burning

No stable nucleus with $A = 8 \Rightarrow$ helium burning occurs in two steps:

1. $\alpha + \alpha \leftrightarrow {}^8\mathrm{Be}$ builds up small equilibrium concentration of ^{8}Be
2. ${}^8\mathrm{Be} + \alpha \rightarrow {}^{12}\mathrm{C}^* \rightarrow {}^{12}\mathrm{C} + \gamma$ becomes possible at $T \approx 10^8$ K due to a resonance in ^{12}C (predicted by Fred Hoyle in 1954 on the basis of the existence of carbon in the Universe)

Net effect is called *triple-α* reaction:

$$3\alpha \rightarrow {}^{12}\mathrm{C} + \gamma \quad (Q = 7.275\,\mathrm{MeV})$$

energy release per gram: $Q/m({}^{12}\mathrm{C}) = 5.9 \times 10^{17}$ erg/g ($\approx$ 1/10 of H-burning)

When sufficient ^{12}C has been created, also:

$${}^{12}\mathrm{C} + \alpha \rightarrow {}^{16}\mathrm{O} + \gamma \quad (Q = 7.162\,\mathrm{MeV})$$

reaction rate of ${}^{12}\mathrm{C}(\alpha,\gamma){}^{16}\mathrm{O}$ is uncertain $\Rightarrow$ affects final C/O ratio...

Approximately, at $T \approx 10^8$ K:

$$\epsilon_{3\alpha} \propto Y^3 \rho^2 T^{40} \text{ (!)}$$

$$\epsilon_{\alpha\mathrm{C}} \propto Y X_{12} \rho T^{20}$$

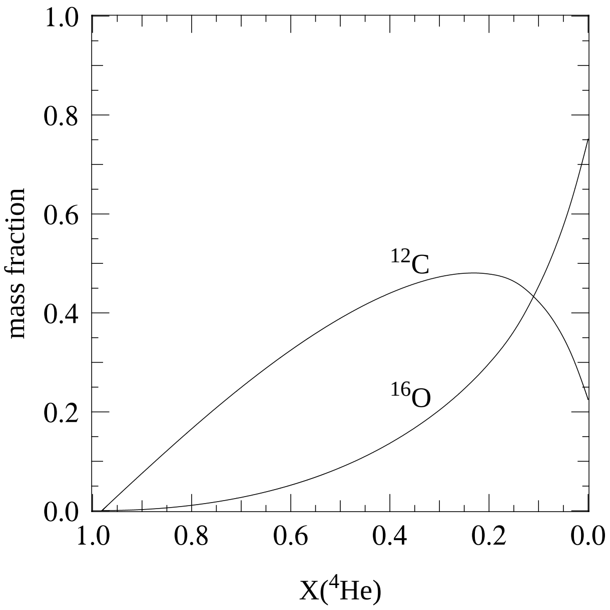


Figure 5.5. Dependence of the mass fractions of ^{12}C and ^{16}O on ^{4}He during He-burning, for typical conditions in intermediate-mass stars.

5.4.5 Carbon burning and beyond

The following nuclear burning processes occur at successively higher temperatures in the cores of massive stars:

Carbon burning occurs at $T_9(\equiv T/10^9\,\mathrm{K}) \gtrsim 0.5$ mainly by the following reaction channels:

$$\begin{aligned} {}^{12}\mathrm{C} + {}^{12}\mathrm{C} \to {}^{24}\mathrm{Mg}^* &\to {}^{20}\mathrm{Ne} + \alpha \quad + 4.6\,\mathrm{MeV} \quad (\sim 50\%) \\ &\to {}^{23}\mathrm{Na} + \mathrm{p} \quad + 2.2\,\mathrm{MeV} \quad (\sim 50\%) \end{aligned}$$

Many side reactions occur with the released p and α particles, e.g. ${}^{23}\mathrm{Na}(\mathrm{p},\alpha){}^{20}\mathrm{Ne}$, ${}^{20}\mathrm{Ne}(\alpha,\gamma){}^{24}\mathrm{Mg}$, and many others. The composition after carbon burning is mostly ${}^{16}\mathrm{O}$, ${}^{20}\mathrm{Ne}$ and ${}^{24}\mathrm{Mg}$ (together 95% by mass fraction). These most abundant nuclei have equal numbers of protons and neutron, but some of the side reactions produce neutron-rich isotopes like ${}^{21,22}\mathrm{Ne}$, ${}^{23}\mathrm{Na}$ and ${}^{25,26}\mathrm{Mg}$, so that after C burning the overall composition has a 'neutron excess' ($n/p > 1$, or $\mu_e > 2$).

Neon burning is the next cycle, occurring at lower temperature ($T_9 \approx 1.5$) than oxygen burning because ${}^{16}\mathrm{O}$ is a very stable 'double-magic' nucleus ($N = Z = 8$). It occurs by a combination of photodisintegration and α-capture reactions:

$$\begin{aligned} &{}^{20}\mathrm{Ne} + \gamma \leftrightarrow {}^{16}\mathrm{O} + \alpha \quad - 4.7\,\mathrm{MeV} \\ &{}^{20}\mathrm{Ne} + \alpha \to {}^{24}\mathrm{Mg} + \gamma \quad + 9.3\,\mathrm{MeV} \end{aligned}$$

The first reaction is endothermic, but effectively the two reactions combine to $2\,{}^{20}\mathrm{Ne} \to {}^{16}\mathrm{O} + {}^{24}\mathrm{Mg}$ with a net energy generation > 0. The composition after neon burning is mostly ${}^{16}\mathrm{O}$ and ${}^{24}\mathrm{Mg}$ (together 95% by mass).

Oxygen burning occurs at $T_9 \approx 2.0$ mainly by the reaction channels:

$$\begin{aligned} {}^{16}\mathrm{O} + {}^{16}\mathrm{O} \to {}^{32}\mathrm{S}^* &\to {}^{28}\mathrm{Si} + \alpha \quad + 9.6\,\mathrm{MeV} \quad (\sim 60\%) \\ & {}^{31}\mathrm{P} + \mathrm{p} \quad + 7.7\,\mathrm{MeV} \quad (\sim 40\%) \end{aligned}$$

together with many side reactions with the released p and α particles, as was the case for C-burning. The composition after oxygen burning is mostly ${}^{28}\mathrm{Si}$ and ${}^{32}\mathrm{S}$ (together 90% by mass)

Silicon burning does not occur by ${}^{28}\mathrm{Si} + {}^{28}\mathrm{Si}$, but instead by a series of photodisintegration (γ,α) and α-capture (α,γ) reactions when $T_9 \gtrsim 3$. Part of the silicon 'melts' into lighter nuclei, while another part captures the released ${}^{4}\mathrm{He}$ to make heavier nuclei:

$$\begin{aligned} &{}^{28}\mathrm{Si}\,(\gamma,\alpha)\,{}^{24}\mathrm{Mg}\,(\gamma,\alpha)\,{}^{20}\mathrm{Ne}\,(\gamma,\alpha)\,{}^{16}\mathrm{O}\,(\gamma,\alpha)\,{}^{12}\mathrm{C}\,(\gamma,\alpha)\,2\alpha \\ &{}^{28}\mathrm{Si}\,(\alpha,\gamma)\,{}^{32}\mathrm{S}\,(\alpha,\gamma)\,{}^{36}\mathrm{Ar}\,(\alpha,\gamma)\,{}^{40}\mathrm{Ca}\,(\alpha,\gamma)\,{}^{44}\mathrm{Ti}\,(\alpha,\gamma)\,\ldots\,{}^{56}\mathrm{Ni} \end{aligned}$$

Most of these reactions are in equilibrium with each other, and their abundances can be described by nuclear equivalents of the Saha equation for ionization equilibrium. For $T > 4\cdot 10^9\,\mathrm{K}$ a state close to *nuclear statistical equilibrium (NSE)* can be reached, where the most abundant nuclei are those with the lowest binding energy, constrained by the total number of neutrons and protons present. The final composition is then mostly ${}^{56}\mathrm{Fe}$ because $p/n < 1$ (due to β-decays and e^--captures).

Suggestions for further reading

The contents of this chapter are also covered by Chapter 9 of MAEDER and by Chapter 18 of KIPPENHAHN.

Exercises

5.1 Conceptual questions: Gamow peak

N.B. Discuss your answers to this question with your fellow students or with the assistant.

In the lecture (see eq. 5.17) you saw that the reaction rate is proportional to

$$\langle\sigma\upsilon\rangle = \left(\frac{8}{m\pi}\right)^{1/2} \frac{S(E_0)}{(kT)^{3/2}} \int_0^\infty e^{-E/kT} e^{-b/E^{1/2}} dE,$$

where the factor $b = \pi(2m)^{1/2} Z_1 Z_2 e^2/\hbar$, and $m = m_1 m_2/(m_1 + m_2)$ is the reduced mass.

(a) Explain in general terms the meaning of the terms $e^{-E/kT}$ and $e^{-b/E^{1/2}}$.

(b) Sketch both terms as function of E. Also sketch the product of both terms.

(c) The reaction rate is proportional to the area under the product of the two terms. Draw a similar sketch as in question (b) but now for a higher temperature. Explain why and how the reaction rate depends on the temperature.

(d) Explain why hydrogen burning can take place at lower temperatures than helium burning.

(e) Elements more massive than iron, can be produced by neutron captures. Neutron captures can take place at low temperatures (even at terrestial temperatures). Can you explain why?

5.2 Hydrogen burning

(a) Calculate the energy released per reaction in MeV (the Q-value) for the three reactions in the pp1 chain. (Hint: first calculate the equivalent of $m_u c^2$ in MeV.)

(b) What is the total effective Q-value for the conversion of four ^{1}H nuclei into ^{4}He by the pp1 chain? Note that in the first reaction ($^1\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{H} + \mathrm{e}^+ + \nu$) a neutrino is released with (on average) an energy of 0.263 MeV.

(c) Calculate the energy released by the pp1 chain in erg/g.

(d) Will the answer you get in (c) be different for the pp2 chain, the pp3 chain or the CNO cycle? If so, why? If not, why not?

5.3 Relative abundances for CN equilibrium

Estimate the CN-equilibrium relative abundances if their lifetimes against proton capture at $T = 2{\times}10^7$ K are:

$\tau(^{15}N) \approx 30$ yr.
$\tau(^{13}C) \approx 1600$ yr.
$\tau(^{12}C) \approx 6600$ yr.
$\tau(^{14}N) \approx 6 \times 10^5$ yr.

5.4 Helium burning

(a) Calculate the energy released per gram for He burning by the 3α reaction and the $^{12}\mathrm{C} + \alpha$ reaction, if the final result is a mixture of 50% carbon and 50% oxygen (by mass fraction).

(b) Compare the answer to that for H-burning. How is this related to the duration of the He-burning phase, compared to the main-sequence phase?