5. Diffusive energy transport, Ionisation/excitation, Opacity

Textbook: §9.2, 8.1, and parts of 9.3, 10.4

Radiative and conductive energy transport

Radiative flux

$$F_{\rm rad} = -\frac{1}{3}\frac{c}{\kappa\rho}\frac{{\rm d}U_{\rm rad}}{{\rm d}r} = -\frac{4ac}{3}\frac{T^3}{\kappa\rho}\frac{{\rm d}T}{{\rm d}r}. \tag{5.1}$$

Eddington equation

$$\frac{{\rm d}T}{{\rm d}r} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{L_r}{4\pi r^2}, \tag{5.2}$$

where $L_r = 4\pi r^2 F_{\rm rad}$.

Rosseland mean

$$\frac{1}{\overline{\kappa}} = \frac{1}{\kappa_{\rm R}} \equiv \frac{\pi}{acT^3}\int_0^\infty \frac{1}{\kappa_\nu}\frac{{\rm d}B_\nu}{{\rm d}T}\,{\rm d}\nu\,, \qquad \text{where} \quad B_\nu \equiv \frac{c}{4\pi}U(\nu) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/kT}-1} \tag{5.3}$$

Since $\int({\rm d}B_\nu/{\rm d}T){\rm d}\nu = acT^3/\pi$, the Rosseland mean is the harmonic mean of κ_ν weighted by ${\rm d}B_\nu/{\rm d}T$.

Conduction

Usually, conduction is irrelevant. The exception is degenerate cores, where it dominates, making the cores isothermal. One can combine conduction with radiative transport by defining

$$F = F_{\rm rad} + F_{\rm cond} = -(k_{\rm rad} + k_{\rm cond})\nabla T. \tag{5.4}$$

If we define a conductive opacity $\kappa_{\rm cond}$ via

$$k_{\rm cond} \equiv \frac{4ac}{3}\frac{T^3}{\kappa_{\rm cond}}, \tag{5.5}$$

and by redefining $1/\overline{\kappa} = 1/\kappa_{\rm R} + 1/\kappa_{\rm cond}$, we can include conduction also in this way in the radiative transport equation.

Excitation and Ionisation

In general, the different states of ions and atoms will be populated according to the *Boltzmann equation,*

$$\frac{N_b}{N_a} = \frac{g_b}{g_a}{\rm e}^{-(\chi_b-\chi_a)/kT}. \tag{5.6}$$

Here, $g_{a,b}$ are the statistical weights (e.g., $g = 2n^2$ for level n in Hydrogen), and $\chi_{a,b}$ are the excitation potentials.

Comparing the ground state of one ionisation stage with the ground state of the next one, one has to take into account that the electron can have a range of kinetic energies and associated states. One finds

$$\frac{{\rm d}n_{i+1,0}(p)}{n_{i,0}} = \frac{g_{i+1,0}{\rm d}g_{\rm e}(p)}{g_{i,0}}{\rm e}^{-(\chi_i + p_{\rm e}/2m_{\rm e})/kT} \tag{5.7}$$

where $\mathrm{d}n_{i+1,0}(p)$ is the number density of atoms in the ground state of ionisation stage $i+1$ with an electron with momentum p, $n_{i,0}$ the number density of atoms in the ground state of ionisation stage i, and $g_\mathrm{e}(p)$ the statistical weight of the electron at momentum p. The latter is given by

$$\mathrm{d}g_\mathrm{e}(p) = \frac{2}{h^3}\frac{1}{n_\mathrm{e}} 4\pi p^2 \,\mathrm{d}p. \tag{5.8}$$

Integrating over all possible electron momenta and summing over all possible excitation states n (using the "partition function" $\mathcal{Z} = \sum_n g_n \exp(-\chi_n/kT)$), one finds the *Saha equation*,

$$\frac{n_{i+1}}{n_i} n_\mathrm{e} = \frac{\mathcal{Z}_{i+1}}{\mathcal{Z}_i} 2 \frac{(2\pi m_\mathrm{e} kT)^{3/2}}{h^3} \mathrm{e}^{-\chi_i/kT}. \tag{5.9}$$

Opacity

In general, the opacity is a complicated function of density, temperature and abundances. Three main processes dominate the continuum opacity at temperatures typically encountered in stars.

Electron scattering

$$\sigma_\mathrm{T} = \frac{8\pi}{3}\left(\frac{e^2}{m_\mathrm{e}c^2}\right)^2 = 6.65\ 10^{-25}\,\mathrm{cm}^2 \quad \Rightarrow \quad \kappa_\mathrm{es} = \sigma_\mathrm{T}\frac{1+X}{2m_\mathrm{H}} = 0.200(1+X)\,\mathrm{cm}^2\,\mathrm{g}^{-1}. \tag{5.10}$$

Free-free absorption

The free-free cross section for a certain ion i is given by

$$\sigma^\mathrm{ff}_{\nu,i} = \left(\frac{2m_\mathrm{e}}{\pi kT}\right)^{1/2} n_\mathrm{e} \frac{4\pi}{3\sqrt{3}} \frac{Z_i^2 e^6}{hcm_\mathrm{e}^2\nu^3} g^\mathrm{ff}_\nu = 3.69\,10^8\ \mathrm{cm}^2\ n_\mathrm{e}\frac{Z_i^2}{T^{1/2}\nu^3} g^\mathrm{ff}_\nu. \tag{5.11}$$

For a general mixture of ions, one has to add over all constituents and their corresponding Z_i^2:

$$\overline{n_\mathrm{ion}Z^2} = \sum \frac{\rho X_i}{m_\mathrm{H} A_i} Z_i^2 = \frac{\rho}{m_\mathrm{H}}\left[X + Y + \sum_{i\geq 3}\frac{X_i}{A_i}Z_i^2\right], \tag{5.12}$$

where hydrogen and helium are assumed to be completely ionised.

In the integration over frequency required to calculate the Rosseland mean, one finds that the dependence on ν leads to the introduction of a T^{-3} term. The result is the so-called *Kramers free-free opacity*,

$$\kappa_\mathrm{ff} = 3.8\ 10^{22}\,\mathrm{cm}^2\,\mathrm{g}^{-1}\ \rho T^{-7/2} g_\mathrm{ff}(1+X)\,(X+Y+B). \tag{5.13}$$

The Gaunt factor g_ff is a suitably averaged value of g^ff_ν.

Bound-free absorption

The semi-classical Kramers cross section for an ion with charge Z_i with an electron in state n is given by

$$\sigma^\mathrm{bf}_{\nu,i,n} = \frac{64\pi^4}{3\sqrt{3}}\frac{m_\mathrm{e}e^{10}}{ch^6}\frac{Z_i^4}{n^5\nu^3} g^\mathrm{bf}_{\nu,i,n} = 2.82\,10^{29}\,\mathrm{cm}^2\ \frac{Z_i^4}{n^5\nu^3} g^\mathrm{bf}_{\nu,i,n}. \tag{5.14}$$

Most of the ions will be in an ionisation state $i+1$ which cannot be ionised by a typical photon with $h\nu \simeq kT \ll \chi_{i+1}$; the relevant ions for the opacity are the somewhat rarer ions in ionisation state i. Combining the Boltzmann and Saha equations, and writing $n_{i,n}$ explicitly in terms of $n_{i+1,1}$,

$$n_{i,n} = n_{i+1,1} n_\mathrm{e} \frac{n^2}{2} \left(\frac{h^2}{2\pi m_\mathrm{e} kT} \right)^{3/2} \mathrm{e}^{\chi_{i,n}/kT}, \tag{5.15}$$

where the hydrogenic approximation ($g_n = 2n^2$) was made.

For the Rosseland mean, one needs to add all states of all ions. For stellar interiors, hydrogen and helium will be completely ionised, so the mean opacity will be proportional to the metallicity Z. One finds the *Kramers bound-free opacity*,

$$\kappa_\mathrm{bf} = 4.3\ 10^{25}\,\mathrm{cm}^2\,\mathrm{g}^{-1}\ \frac{g_\mathrm{bf}}{t} Z(1+X)\rho T^{-7/2}, \tag{5.16}$$

where g is a mean Gaunt factor and t the "guillotine" factor that accounts for the number of different ions being available.

Negative hydrogen ion

Hydrogen atom has a bound state for a second electron in the field of the proton, though it has a very low ionisation potential, $\chi_{\mathrm{H}^-} = 0.75\,\mathrm{eV}$. The number density of negative hydrogen ions will be proportional to the electron density, which, in all but the most metal-poor stars, will be set by ionisation of the metals (which have much lower ionisation potentials that hydrogen and helium). Thus, the H^- opacity will scale as $\kappa_{\mathrm{H}^-} \propto \rho X Z$ at low temperatures; H^- is of course easily ionized at higher temperatures, and at it very low temperatures even metals will not be ionized, so there will be no electrons to form H^- by combining with H.

For next time

– Read ahead about convection (10.4)

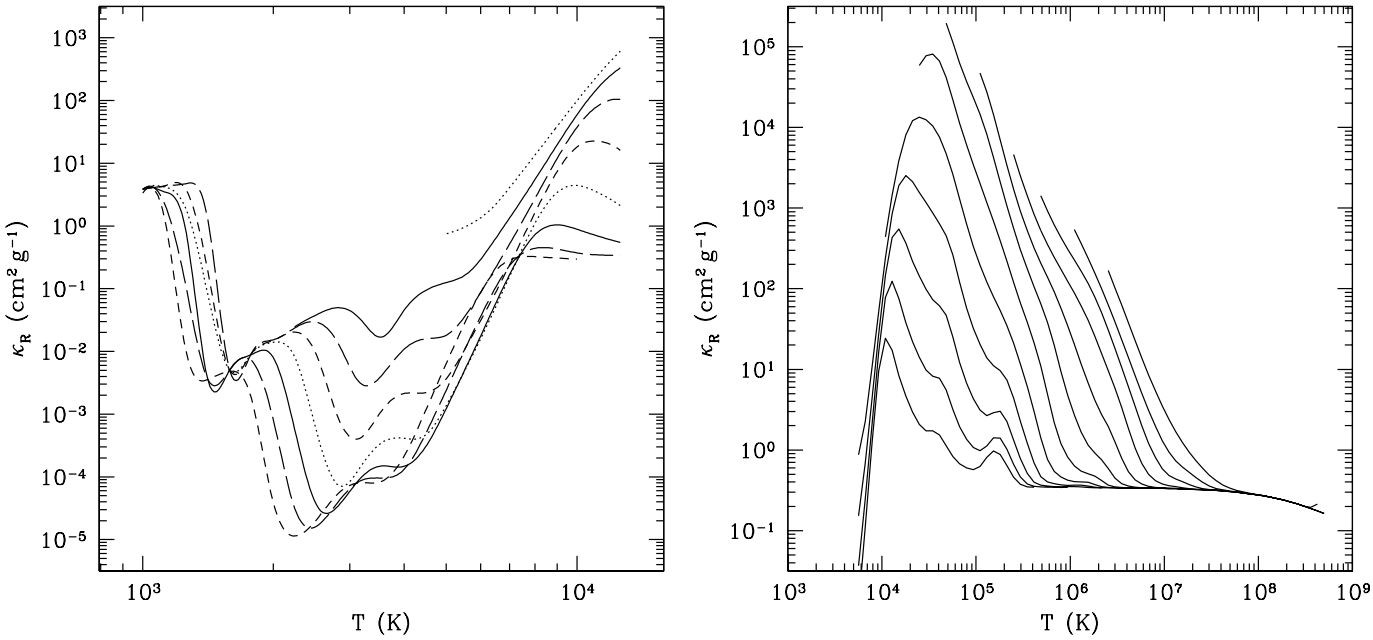


Fig. 5.1. Opacities as a function of temperature. *(left)* Low-temperature regime, from Alexander & Ferguson (1994, ApJ 437, 879). Opacities are shown for densities from 10^{-13} to $10^{-6}\,\mathrm{g\,cm^{-3}}$ in factors of ten, with lower densities corresponding to lower opacities. The sequence in line types is short-dashed, long-dashed, solid, dotted. The bump on the left is due to dust, that in the middle mostly to water, and that on the right to H^-. *(right)* High-temperature regime, for densities from 10^{-9} to $10^{2}\,\mathrm{g\,cm^{-3}}$, from the OPAL group (Iglesias & Rogers, 1996, ApJ 464, 943). The bump at the right is due to bound-free and free-free absorption, and the lower level at the left to electron scattering. Note the difference in scale between the two panels.

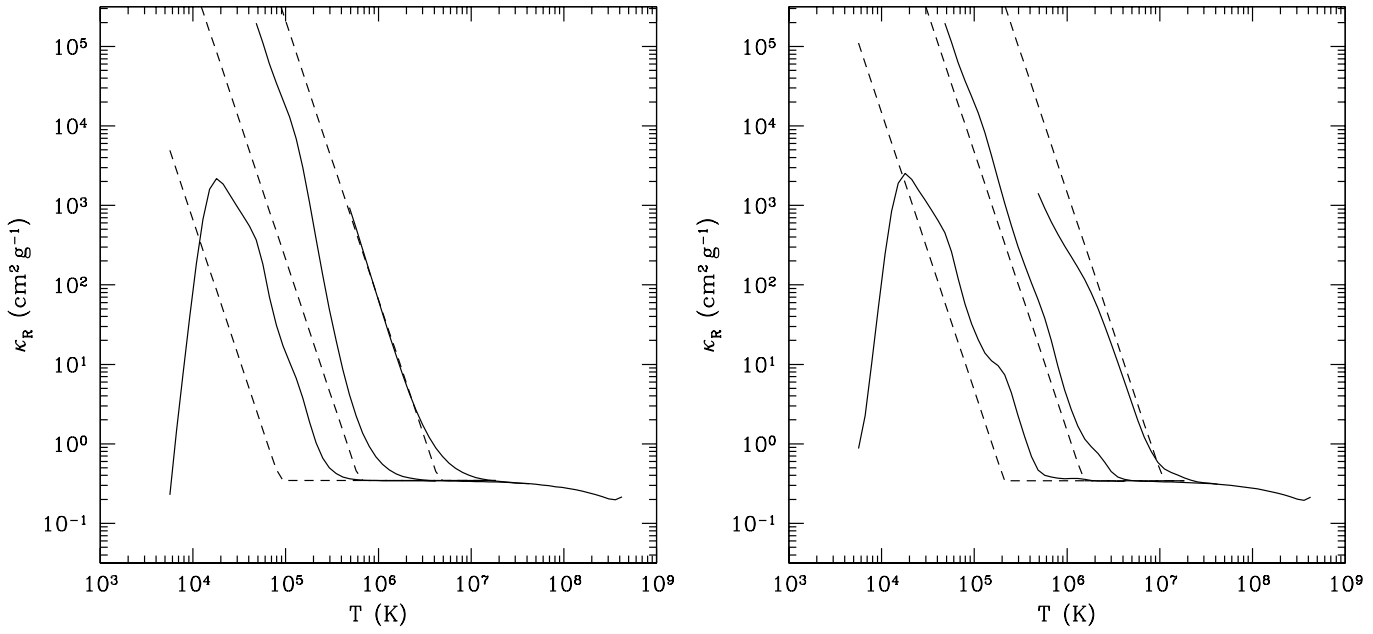


Fig. 5.2. Opacities as a function of temperature as estimated with the Kramers formulae (short-dashed lines) compared to those calculated by the OPAL group, for densities 10^{-6}, 10^{-3}, and $1\,\mathrm{g\,cm^{-3}}$. *(left)* $Z = 0$: OPAL vs. the Kramers free-free opacity; *(right)* $Z = 0.02$: OPAL vs. the Kramers bound-free opacity.

6. Convection, Mixing Length Theory

Textbook: §10.4

General stability criterion

$$-\frac{1}{\gamma}\frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}r} > -\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r}. \tag{6.1}$$

Schwarzschild instability criterion

$$\left.\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{ad}} < \left.\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{rad}} \quad \Leftrightarrow \quad \nabla_{\mathrm{ad}} < \nabla_{\mathrm{rad}}. \tag{6.2}$$

Ledoux instability criterion

$$\frac{\gamma-1}{\gamma} < \frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} - \frac{(\partial\ln\rho/\partial\ln\mu)}{(-\partial\ln\rho/\partial\ln T)}\frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln P}, \quad \Leftrightarrow \quad \nabla_{\mathrm{ad}} < \nabla_{\mathrm{rad}} - \frac{(\partial\ln\rho/\partial\ln\mu)}{(-\partial\ln\rho/\partial\ln T)}\nabla_\mu, \tag{6.3}$$

where we have defined $\nabla_\mu = \mathrm{d}\ln\mu/\mathrm{d}\ln P$ to be the changes in μ due to changes in composition X_i only, and where for a fully-ionised ideal gas, the term with the partial derivatives equals unity.

Efficiency of convection

A general expression for the convective flux is

$$F_{\mathrm{conv}} = \rho\overline{v}_{\mathrm{conv}}\Delta q = \rho\overline{v}_{\mathrm{conv}}c_P\Delta T = \rho\overline{v}_{\mathrm{conv}}c_P T\frac{\ell_{\mathrm{mix}}}{2H_P}\left(\nabla-\nabla_{\mathrm{ad}}\right), \tag{6.4}$$

where ℓ_{mix} is the *mixing length*, usually parametrized as a fraction of the scale height, i.e., $\ell_{\mathrm{mix}} \equiv \alpha_{\mathrm{mix}}H_P$, with α_{mix} the *mixing length parameter*.

To estimate v_{conv}, we use a method different from that used in the textbook: balance buoyancy ($Vg\Delta\rho = \rho Vg\Delta T/T$) and friction ($-A\rho v^2$); evaluate velocity at $l_{\mathrm{mix}}/2$; define $V/A = \beta\ell_{\mathrm{mix}}$, where β is a shape factor; and find

$$v_{\mathrm{conv}}^2 = \frac{\beta g}{H_P}\frac{\ell_{\mathrm{mix}}^2}{2}\left(\nabla-\nabla_{\mathrm{ad}}\right). \tag{6.5}$$

This leads to a convective flux given by

$$F_{\mathrm{conv}} = \rho c_P T\alpha_{\mathrm{mix}}^2\sqrt{\frac{\beta g H_P}{8}}\left(\nabla-\nabla_{\mathrm{ad}}\right)^{3/2}. \tag{6.6}$$

7. Completely convective stars and the Hayashi line

Textbook: –

Generalities

For completely convective stars, the temperature gradient needs to be only very slightly superadiabatic for substantial luminosities to be transported. The implication is that whatever luminosity the star manages to radiate away, will be brought to the surface without any problem by a corresponding energy flux in the convective regions. Thus, the actual luminosity of the star is determined in the only radiative region in the star, the photosphere.

A completely convective star

To find a solution for the whole star, we need to match a photosphere to the interior solution, where the latter is given by a polytrope $P = K\rho^{5/3}$. Matching the two solutions will set K, and for fixed K one knows how the radius depends on mass. For the run of pressure in the atmosphere, we have

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM}{R^2}\rho \quad \text{or} \quad \frac{\mathrm{d}P}{\mathrm{d}h} = -g\rho,$$

where h is the height above some reference level. For the photosphere, $\tau = \kappa\rho h = \frac{2}{3}$, or

$$h = \frac{2}{3\kappa\rho} \quad \Rightarrow \quad P_{\rm phot} = \frac{2g}{3\kappa}. \tag{7.1}$$

Now, assume that the opacity is given by a law of the form

$$\kappa = \kappa_0 P^a T^b, \tag{7.2}$$

where in general a will be a positive number of order unity, while for cool temperatures b will be a relatively large positive number. Given this general opacity law, one has

$$P_{\rm phot}^{1+a} = \frac{2}{3\kappa_0 T_{\rm eff}^b} g \quad \Rightarrow \quad P_{\rm phot} = \left(\frac{2}{3\kappa_0 G}\frac{M}{R^2 T_{\rm eff}^b}\right)^{1/(1+a)}. \tag{7.3}$$

For the interior, we write the polytropic relation in terms of pressure and temperature, and combine it with the mass-radius relation for polytropes of $n = 1.5$ (Table 4.1),

$$\left.\begin{array}{l} P = K\rho^{5/3} \\ P = \dfrac{\rho}{\mu m_{\rm H}} kT \\ K = C_{1.5} G M^{1/3} R \quad \text{with} \quad C_{1.5} = 0.42422 \end{array}\right\} \quad \Rightarrow \quad P = K\left(\frac{P\mu m_{\rm H}}{kT}\right)^{5/3} \quad \Rightarrow \quad P_{\rm int} = \frac{M^{-1/2}}{(RC_{1.5}G)^{-3/2}}\left(\frac{kT}{\mu m_{\rm H}}\right)^{5/2}. \tag{7.4}$$

Equating $P_{\rm int}$ with $P_{\rm phot}$, raising to the $2(1+a)$ power, and sorting, one finds

$$\left(\frac{2}{3\kappa_0}\right)^2 G^{1+3a} M^{3+a} R^{-1+3a} = C_{1.5}^{-3-3a}\left(\frac{k}{\mu m_{\rm H}}\right)^{5+5a} T^{5+5a+2b}. \tag{7.5}$$

Solving for $T_{\rm eff}$,

$$T_{\rm eff} = \mathcal{C}_R M^{\frac{3+a}{5+5a+2b}} R^{\frac{-1+3a}{5+5a+2b}} \quad \text{with} \quad \mathcal{C}_R = \left[\left(\frac{2}{3\kappa_0}\right)^2 G^{1+3a} C_{1.5}^{3+3a}\left(\frac{k}{\mu m_{\rm H}}\right)^{-5-5a}\right]^{\frac{1}{5+5a+2b}}. \tag{7.6}$$

For a of order unity and large positive b one thus sees that $T_{\rm eff}$ depends only very weakly on the mass and radius. With $L = 4\pi R^2 \sigma T_{\rm eff}^4$, we can determine the dependencies on M and L, and thus where the star would be in the HRD. One finds

$$\left(\frac{2}{3\kappa_0}\right)^2 G^{1+3a} M^{3+a} \left(\frac{L}{4\pi\sigma}\right)^{(3a-1)/2} = C_{1.5}^{-3-3a} \left(\frac{k}{\mu m_{\rm H}}\right)^{5+5a} T^{3+11a+2b}, \tag{7.7}$$

$$T_{\rm eff} = \mathcal{C}_L M^{\frac{6+2a}{6+22a+4b}} L^{\frac{3a-1}{6+22a+4b}} \quad \text{with} \quad \mathcal{C}_L = \left[\left(\frac{2}{3\kappa_0}\right)^2 G^{1+3a} C_{1.5}^{3+3a} \left(\frac{k}{\mu m_{\rm H}}\right)^{-5-5a}\right]^{\frac{2}{6+22a+4b}}. \tag{7.8}$$

Again, for a of order unity and large b, $T_{\rm eff}$ depends extremely weakly on the luminosity, and thus one expects nearly vertical lines in the HRD. Given the slight positive dependence on M, one expects the lines to move slightly towards higher temperatures for larger masses.

Complications

The scaling that one finds from the above relations is reasonable. If one were to calculate numerical values, however, the answers would be very puzzling. The reason is that the assumption of a polytrope breaks down near the surface. Going towards the surface, it first fails in the ionisation zone, where recombination is an additional source of heat. Due to the recombination, the temperature of an adiabatically expanding blob does not decrease as it would otherwise, and therefore, above the ionisation zone the temperatures will be higher than would be the case if recombination were ignored. The effect can be seen Fig. 7.1.

Just below the photosphere, the convective energy transport becomes much less efficient, i.e., the superadiabatic gradient becomes substantial, while in the assumption of a $n = 1.5$ polytrope it is assumed to be negligible. With less efficient energy transport, the temperature will decrease more rapidly than adiabatic. Thus, the substantially superadiabatic region near the photosphere counteracts the effects of the ionisation zone. Net, the ionisation zone is more important.

7.1. Contraction along the Hayashi track

The star needs to contract in order to provide the energy it radiates away. Since it is completely convective, the entropy remains constant through the star, but decreases (increasing the entropy of the universe in order not to violate the second law). Since $\mathrm{d}q = T\mathrm{d}s$, the energy generated per gram is proportional to the local temperature. Therefore, the increase in luminosity in a shell $\mathrm{d}M_r$ is $\mathrm{d}L_r \propto T\mathrm{d}M_r$. With this, and with $P \propto T^{5/2}$, we can estimate whether the radiative gradient decreases towards the surface or towards the centre of the star. We assume again an opacity law of the form $\kappa \propto P^a T^b$, with $a = 1$, $b = -4.5$ for a Kramers-type law. We find

$$\begin{aligned}\frac{\mathrm{d}\ln\nabla_{\rm rad}}{\mathrm{d}\ln r} &= \frac{\mathrm{d}\ln L_r}{\mathrm{d}\ln r} - \frac{\mathrm{d}\ln M_r}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln\kappa}{\mathrm{d}\ln r} + \frac{\mathrm{d}\ln P}{\mathrm{d}\ln r} - 4\frac{\mathrm{d}\ln T}{\mathrm{d}\ln r} \\ &= \frac{\mathrm{d}\ln L_r}{\mathrm{d}\ln r} - \frac{\mathrm{d}\ln M_r}{\mathrm{d}\ln r} + [b - 4 + 2.5(a+1)]\,\frac{\mathrm{d}\ln T}{\mathrm{d}\ln r}.\end{aligned} \tag{7.9}$$

Generally, one has $M_r = \int \rho r^2 \mathrm{d}r$ and $L_r \propto \int T\rho r^2 \mathrm{d}r$. With the polytropic relations, therefore, $M_r \propto \int \theta^n \xi^2 \mathrm{d}\xi$ and $L_r \propto \int \theta^{n+1}\xi^2\mathrm{d}\xi$. Thus, one can use the solution $\theta(\xi)$ for a polytropic star to calculate $\mathrm{d}\ln(M_r, L_r, T)/\mathrm{d}\ln r$. The result for $n = 1.5$ is shown in Fig. 7.2. Also drawn is $\mathrm{d}\ln\nabla_{\rm rad}/\mathrm{d}\ln r$, assuming $a = 1$ and $b = -4.5$. One sees that it is always larger than zero, i.e., the radiative gradient decreases inwards. This is true for any reasonable opacity law. In consequence, the interior is always the first part of the star to become radiative.

We can also estimate how the radiative gradient scales scales with the stellar parameters in the core. There, the temperature hardly varies, and one has $L_r \propto (L/M)T_{\rm c}M_r$. Furthermore, for

any two stars with the same structure, $T_c \propto M/R$ and $P_c \propto M^2/R^4$, with the same constants of proportionality. Taking again $\kappa \propto P^a T^b$, one finds for the radiative gradient in the core,

$$\nabla_{\rm rad,c} \propto \frac{L_r \kappa P}{M_r T^4} \propto \frac{L}{M} T^{1+b-4} P^{a+1} \propto L M^{-2+b+2a} R^{-1-b-4a} \propto L M^{-4.5} R^{-0.5} \tag{7.10}$$

where in the last proportionality we used $a = 1$, $b = -4.5$ (Kramers). From Eqs. 7.6, 7.8, one sees that for given mass, $L \propto R^\alpha$, with $\alpha = (6 + 22a + 4b)/(5 + 5a + 2b)$, where a and b are now the coefficients in the atmospheric opacity law. Generally, $a \simeq 1$ and b large, hence, $\alpha \simeq 2$. Thus, the radiative gradient decreases as one descends the Hayashi track. At constant luminosity, one has $R \propto M^\beta$, with $\beta = (6 + 2a)/(7 - a + 2b) \lesssim 1$. Hence, the radiative gradient is smaller for larger masses, and more massive stars will become radiative in their core sooner.

For next time

- Think about why a star cannot be to the right of Hayashi limit.
- Read ahead on stellar energy sources (§10.3)

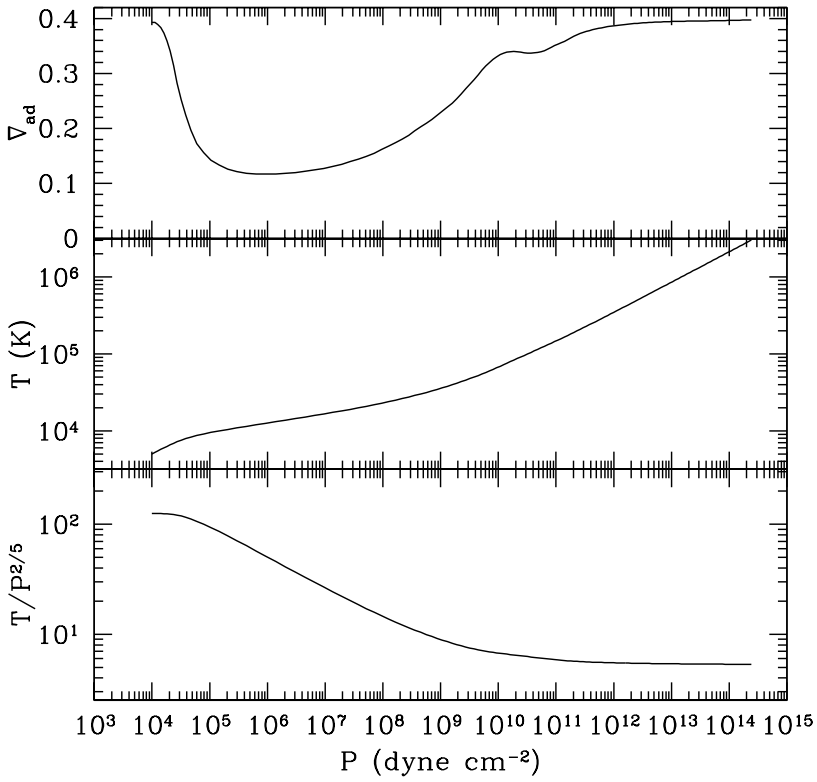


Fig. 7.1. Adiabatic gradient *(top)*, temperature *(middle)* and $T/P^{2/5}$ *(bottom)* as a function of pressure, calculated using the OPAL equation of state for a solar mixture. The effect of the hydrogen and helium ionisation zones is clearly seen in the depressions in $\nabla_{\rm ad}$ and the changes in slope in the other panels. As a result, a completely convective star will have a higher surface temperature than would be expected if the ionisation zone were ignored. The effect is partly undone by the superadiabatic gradient becoming substantial just below the photosphere.

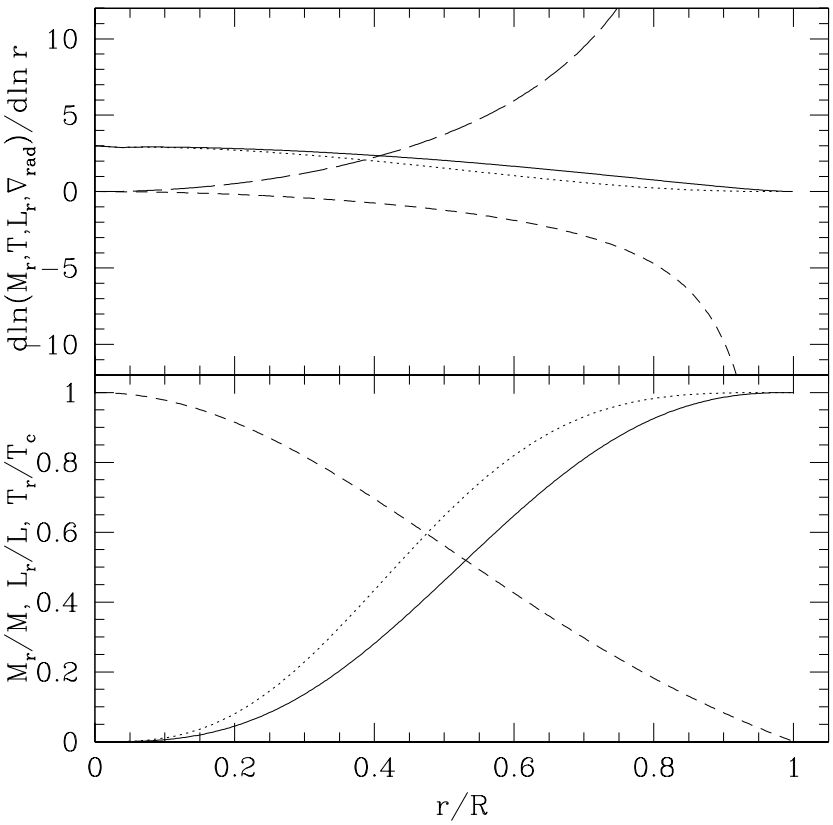


Fig. 7.2. *(Bottom)* Run of mass (solid line), luminosity (dotted line), and temperature (dashed line) as a function of radius for a contracting polytrope with $n = 1.5$ (i.e., the local energy generation per unit mass is proportional to temperature). *(Top)* Logarithmic derivatives of mass (solid line), luminosity (dotted line), temperature (short-dashed line), and radiative gradient (long-dashed line) as a function of radius. A Kramers-type opacity law was assumed.

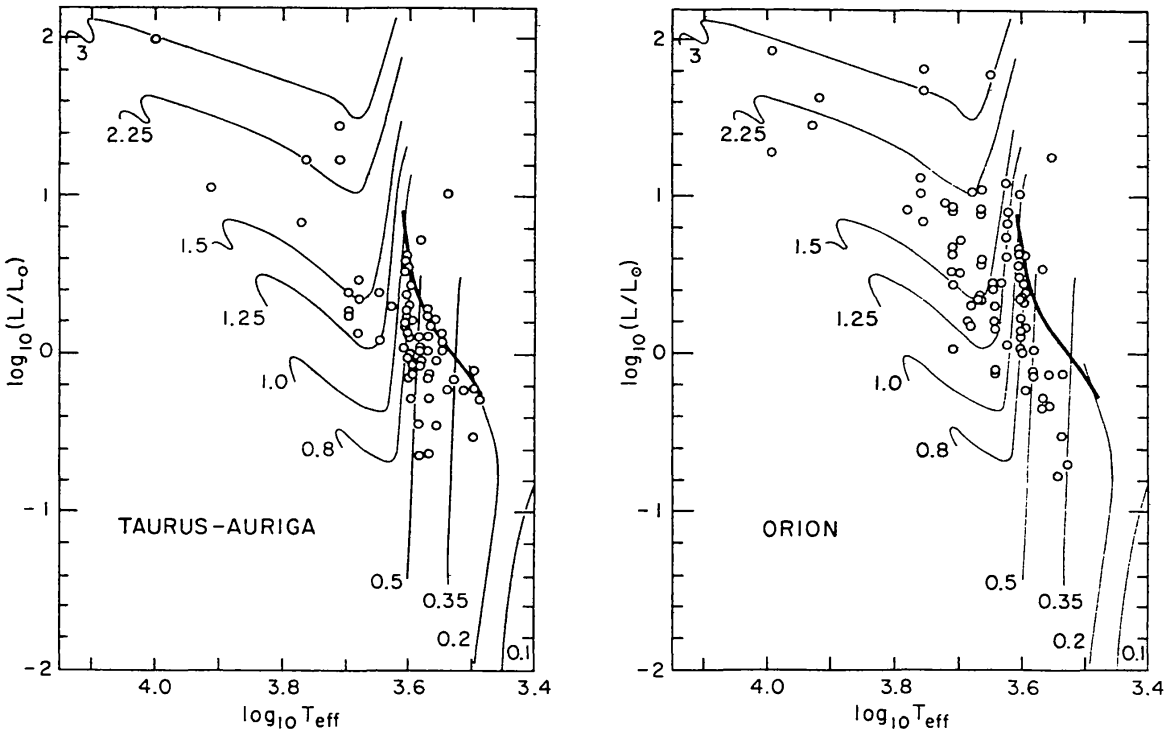


Fig. 7.3. Theoretical tracks for the pre-main sequence contraction phase for several different masses (as indicated). Overdrawn are observed temperatures and luminosities for pre-main sequence stars in two star-forming regions with rather different properties. In both, stars first appear along a very similar "birth line" (indicated with the thick line).

8. Energy balance: Contraction/expansion and nuclear processes

Textbook: §10.3

Energy Balance

$$\frac{\mathrm{d}L_r}{\mathrm{d}r} = 4\pi r^2 \rho\, \epsilon \quad \Leftrightarrow \quad \frac{\mathrm{d}L_r}{\mathrm{d}M_r} = \epsilon, \tag{8.1}$$

where ϵ is the energy generated per unit mass. In general,

$$\epsilon = \epsilon_{\rm grav} + \epsilon_{\rm nuc} - \epsilon_\nu, \tag{8.2}$$

where $\epsilon_{\rm grav}$ is the energy liberated or lost by contraction or expansion, $\epsilon_{\rm nuc}$ is the energy produced (or lost) in nuclear processes, and ϵ_ν is that part of the latter that escapes the star immediately in the form of neutrinos.

Contraction or expansion

The energy gained or lost in mass movements inside the star can be derived from the first law of thermodynamics, and written in various equivalent forms as

$$\epsilon_{\rm grav} = -\frac{\mathrm{d}Q}{\mathrm{d}t} = -T\frac{\mathrm{d}S}{\mathrm{d}t} = -\frac{\mathrm{d}u}{\mathrm{d}t} - P\frac{\mathrm{d}\mathcal{V}}{\mathrm{d}t}, \tag{8.3}$$

where $\mathcal{V} \equiv 1/\rho$ and u is the energy density per unit mass.

Nuclear processes

The main source of energy in stars is nuclear fusion, which we will now treat in more detail than in CO, § 10.3 (Kippenhahn & Weigert, chapter 18, was used extensively below).

Basic considerations

The energy gained or lost in nuclear processes is related to the mass defect Δm:

$$E = \Delta m\, c^2 = \left(\sum_i m_{{\rm init},i} - \sum_j m_{{\rm final},j} \right) c^2. \tag{8.4}$$

The mass defect reflects the different binding energies per nucleon in different nuclei,

$$\frac{E_{\rm bind}}{A} = \frac{1}{A}\left(Z m_{\rm p} + (A-Z) m_{\rm n} - m_{\rm nucleus}\right) c^2 \ . \tag{8.5}$$

The binding energy per nucleon increases steeply from hydrogen, then flattens out and starts to decrease, having reached a maximum at ^{56}Fe; see Fig. 8.1. Defining hydrogen to have zero binding energy, helium has 7.07 MeV per nucleon, carbon 7.68 MeV, and iron 8.73 MeV.

For fusion, nuclei must be brought close enough together that the short-range strong nuclear force can dominate over the weaker, but long-range repulsive Coulomb force. The range of the strong nuclear force is set by the Compton wavelength of its carrier, the pi meson, $\hbar/m_\pi c = 1.41\,$fm. The repulsive Coulomb potential at a distance of $\sim 1\,$fm ($10^{-13}\,$cm) is $E_{\rm Coul} = Z_1 Z_2 e^2/r \simeq 1.44\,{\rm MeV}\,\left(\frac{1\,{\rm fm}}{r}\right) Z_1 Z_2$, where Z_1 and Z_2 are the atomic numbers of the colliding nuclei. This should be compared with typical kinetic energy of a particle, of order $kT = 0.86\, T_7\,$keV, where T_7 is the

temperature in units of 10^7 K. Thus, classically, in the centre of the Sun (where $T_7 \approx 1.5$), particles trying to interact should be turned around by the Coulomb force at $\sim 10^3$ fm; as a result, no reactions would be expected.

From quantum mechanics, however, a particle has a certain finite probability of "tunneling" through the Coulomb barrier (see CO, p. 147–148, which is perhaps more insightful than the motivation on p. 335-338). The reaction cross section per nucleus is usually written as,

$$\sigma(E) = \frac{S(E)}{E} \mathrm{e}^{-b/\sqrt{E}} \quad \text{with} \quad b = \frac{1}{h} 2\pi^2 \sqrt{2m'} Z_1 Z_2 e^2 \quad \text{and} \quad m' = \frac{m_1 m_2}{m_1 + m_2}. \tag{8.6}$$

Here, the term $1/E$ reflect the effective area for the interaction (for which one can take $\pi\lambda^2 \propto 1/p^2 \propto 1/E$), and the exponential term the penetration probability; effects from the nuclear force are absorbed into a function $S(E)$ which is, under most conditions, a relatively slowly varying function of the interaction energy E (but see "resonances" below).

The fusion product is at first a compound nucleus in an excited state with positive total energy. Often, this compound nucleus will decay into the same particles that formed it – i.e., the incoming particle is just scattered by the collision. The cases in which the decay products are different define the net reaction rate, the details of which are hidden in $S(E)$. The rates $S(E)$ can be calculated (with great difficulty!), or one can extrapolate from measurements (which are typically done at far larger energies than those relevant to stellar conditions).

In general, the compound nucleus has several discrete bound states at negative energies in the nuclear potential well, the stable ground state of the nucleus and some excited states that can decay into lower-energy states by emission of photons (γ-rays). These states are similar to the bound states of electrons in an atom, but comprising nucleons instead of electrons. However, the compound nucleus may also have quasi-stable excited states of positive energy (below the top of the Coulomb barrier), which can decay by emission of particles (by quantum tunnelling *outwards* through the Coulomb barrier) as well as by emission of a photon. Incoming particles with "resonant" energy corresponding to such a quasi-stable state can form a compound nucleus much more easily, leading to a greatly enhanced reaction rate.

Given the cross section $\sigma(E)$, the reaction rate between particles of types a and b (at a given energy E) is given by

$$r_{a,b}(E)\,\mathrm{d}E = n_a n_b v \sigma(E) f(E)\,\mathrm{d}E\,, \tag{8.7}$$

where n_a and n_b are the number densities of a and b, v is the relative velocity between a and b (corresponding to energy E), $f(E)$ is the energy probability distribution, and $\sigma(E)$ is the cross section defined above. The factor v accounts for the fact that for larger velocities v, more particles pass each other per unit time. Note that if particles a and b are identical, we need to multiply the above by $\frac{1}{2}$ in order to avoid counting double. Including that in the integrated reaction rate, we find a rate

$$r_{a,b} = \frac{1}{1+\delta_{a,b}} n_a n_b \langle \sigma v \rangle\,, \qquad \text{where} \quad \langle \sigma v \rangle \equiv \int_0^\infty v(E)\sigma(E)f(E)\,\mathrm{d}E \tag{8.8}$$

is the average reaction rate per pair of particles, i.e., $\langle \sigma v \rangle$ is an effective cross-section.

If the velocity probability distributions are Maxwellian for both particles (i.e., particles have momenta as in Eq. 3.8, divided by n), the distribution of the *relative velocity* of the particles is also Maxwellian, but with $m = m' = m_a m_b/(m_a + m_b)$ [verify this]. We can rewrite the Maxwell distribution in Eq. 3.8 as a function of energy using $p = \sqrt{2mE}$ and $\mathrm{d}p = \frac{1}{2}\sqrt{2m/E}\,\mathrm{d}E$,

$$f(E)\,\mathrm{d}E = \frac{2\pi\sqrt{E}}{(\pi kT)^{3/2}} \mathrm{e}^{-E/kT}\,\mathrm{d}E\,. \tag{8.9}$$

Hence, for the effective cross section (using $v(E) = p/m = \sqrt{2E/m}$),

$$\langle \sigma v \rangle = \left(\frac{8}{m'\pi}\right)^{1/2} \left(\frac{1}{kT}\right)^{3/2} \int_0^\infty S(E) \mathrm{e}^{-E/kT} \mathrm{e}^{-b/\sqrt{E}} \, \mathrm{d}E \; . \tag{8.10}$$

The integrand will be small everywhere but near where the two exponentials cross, which is called the "Gamow peak"; see Fig. 8.2. Assuming $S(E)$ is a slowly varying function, the maximum of the integrand will be where the term $h(E) \equiv -E/kT - b/\sqrt{E}$ in the exponential reaches a maximum; this position E_0 is thus obtained via:

$$\begin{aligned} \frac{\mathrm{d}h(E)}{\mathrm{d}E} &= \frac{\mathrm{d}}{\mathrm{d}E}(-E/kT - b/\sqrt{E}) = 0 \quad \Rightarrow \\ E_0 &= \left(\frac{bkT}{2}\right)^{2/3} = 5.665\,\mathrm{keV}\,(Z_1 Z_2)^{2/3} \left(\frac{m'}{m_\mathrm{u}}\right)^{1/3} T_7^{2/3} \; , \end{aligned} \tag{8.11}$$

where m_u is the atomic unit mass. Using a Taylor expansion of $h(E)$ around its maximum,

$$h(E) = h_0 + h_0'(E - E_0) + \frac{1}{2} h_0''(E - E_0)^2 + \ldots \simeq -\tau - \frac{1}{4}\tau \left(\frac{E}{E_0} - 1\right)^2 + \ldots \; , \tag{8.12}$$

where we have used the fact that the first derivative h_0' must be zero (since we are expanding around the maximum), and where we have defined

$$\tau = \frac{3E_0}{kT} = 19.721\,(Z_1 Z_2)^{2/3} \left(\frac{m'}{m_\mathrm{u}}\right)^{1/3} T_7^{-1/3} \; . \tag{8.13}$$

Using this in the integral, the exponential is approximately a Gaussian, as one can see by substituting $\xi = (E/E_0 - 1)\sqrt{\tau}/2$,

$$\int_0^\infty \mathrm{e}^{h(E)} \, \mathrm{d}E = \int_0^\infty \mathrm{e}^{-\tau - \frac{1}{4}\tau (E/E_0 - 1)^2} \, \mathrm{d}E = \frac{2}{3} kT \tau^{1/2} \mathrm{e}^{-\tau} \int_{-\sqrt{\tau}/2}^\infty \mathrm{e}^{-\xi^2} \, \mathrm{d}\xi \; . \tag{8.14}$$

Since τ is relatively large and the main contribution to the integral comes from the range close to E_0 (i.e., $\xi = 0$), the error introduced by extending the integration to $-\infty$ is small, i.e., the integral is approximately $\sqrt{\pi}$. For the Gaussian, the fractional full width at half maximum $\Delta E/E_0$ is

$$\frac{\Delta E}{E_0} = 4 \left(\frac{\ln 2}{\tau}\right)^{1/2} = 0.750\,(Z_1 Z_2)^{-1/3} \left(\frac{m'}{m_\mathrm{u}}\right)^{-1/6} T_7^{1/6} \; . \tag{8.15}$$

Doing the integration using the Gaussian and inserting the result in Eq. 8.10 (after taking out the slowly varying $S(E)$), one obtains

$$\langle \sigma v \rangle = \frac{4}{3} \left(\frac{2}{m'}\right)^{1/2} \left(\frac{1}{kT}\right)^{1/2} S_0 \tau^{1/2} \mathrm{e}^{-\tau} \; , \tag{8.16}$$

where $S_0 = S(E_0)$. Since $T \propto \tau^{-3}$ (Eq. 8.13), one thus has that $\langle \sigma v \rangle \propto \tau^2 \mathrm{e}^{-\tau}$. It is the exponential, however, that really determines the reaction speeds. The dependences on Z_1, Z_2, and m' ensure that more massive, more highly charged ions hardly react at all as long as the fusion processes of the lighter elements still are taking place.

It is often useful to know the temperature dependence of the reaction rate, given by

$$\nu \equiv \frac{\partial \ln \langle \sigma \nu \rangle}{\partial \ln T} = \frac{1}{3}(\tau - 2) = 6.574\,(Z_1 Z_2)^{2/3} \left(\frac{m'}{m_\mathrm{u}}\right)^{1/3} T_7^{-1/3} - \frac{2}{3} \tag{8.17}$$

(note that, for a given reaction, ν usually becomes smaller with increasing temperature). For the fusion of two protons in the centre of the Sun, $Z_1 = Z_2 = 1$, $m' = \frac{1}{2}$, $T_7 \simeq 1.5$, hence $\nu \simeq 4$, which is a relatively mild temperature dependence. For other fusion processes, we will find exponents of $\nu \sim 20$ and above, making these processes among the most strongly varying functions in physics.

Corrections to the above rate formulae

A few corrections are usually made in more detailed derivations. The first is a small correction factor $g_{a,b}$ to account for any temperature dependence of S_0 and for the inaccuracy of approximating the Gamow peak by a Gaussian. The second is more physical, and is a correction $f_{a,b}$ for the effect of electron screening — due to the presence of electrons, the effective potential that two ions see is slightly reduced ("screened"); as a result, the reaction will be faster. This correction is more important at higher densities, and at *very* high densities burning starts to depend more sensitively on the density than on the temperature. (For this case, one speaks of *pycnonuclear reactions*.) Also, separate terms may be added to account for resonances.

Timescales

For a reaction of particles a and b, the number densities decrease with time. We define a timescale for each type of particle

$$\tau_a \equiv -\frac{n_a}{\mathrm{d}n_a/\mathrm{d}t} = \frac{n_a}{(1+\delta_{a,b})r_{a,b}} = \frac{1}{n_b \langle\sigma v\rangle_{a,b}} \,. \tag{8.18}$$

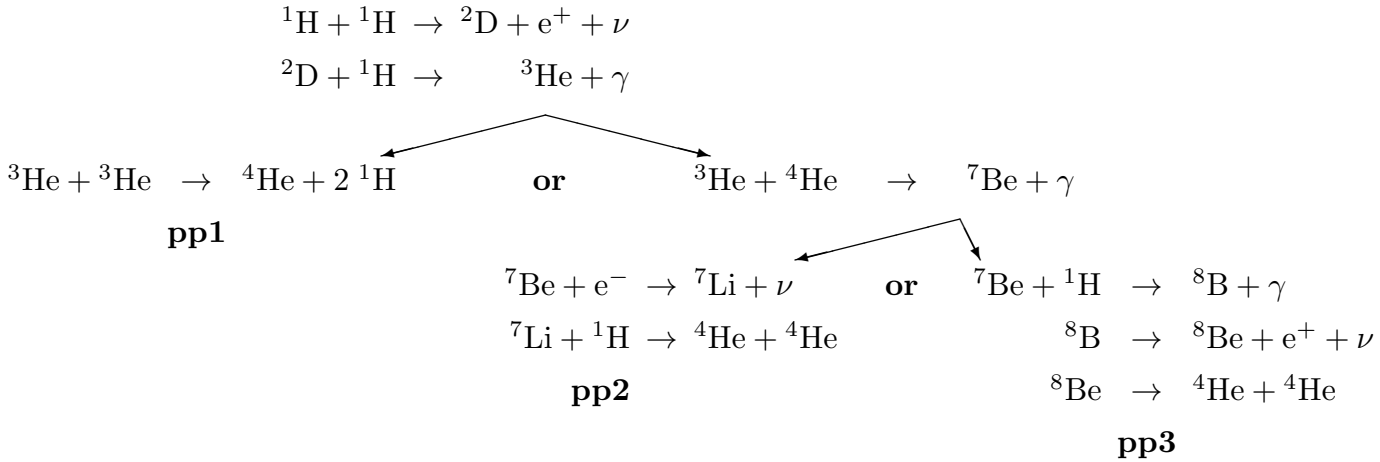
With this definition, $n_a \propto \mathrm{e}^{-t/\tau_a}$. Note that when two particles of the same type react (i.e., when $b = a$), the rate as defined in Eq. 8.7 above is a factor two smaller, but two particles of type a are destroyed per reaction, so the final expression for the timescale does not contain the factor $1 + \delta_{a,b}$. [Show that when there are multiple reactions, the timescale is given by $\tau_a^{-1} = \sum_b (1/\tau_{a,b})$.]

Hydrogen burning

In principle, many nuclear reactions can occur at the same time. As we saw above, however, the weighting of the exponential with $(Z_1 Z_2)^{2/3}$ strongly inhibits processes involving more massive, more highly charged particles. In combination with the initial abundances of stars, with the largest fraction of the mass being hydrogen, generally only a small number of fusion processes turn out to be relevant in a given evolutionary stage.

P-P chain

In less massive stars ($M \lesssim 1.2\ M_\odot$), the fusion of hydrogen to helium on the main sequence is mostly by the *proton-proton chain* (p-p chain). The possible variants of the p-p chain are:



In these chains, the positrons made will meet an electron and annihilate, adding 1.022 MeV of photon energy. Note that while the total energy released (per ^{4}He produced) for the three chains

is equal, the fraction of that energy put in *neutrinos* is not the same. The net energy put into the local medium per ^{4}He nucleus produced is 26.20 MeV for pp1, 25.67 for pp2, and 19.20 for pp3.

The relative frequency of the branches depends on the temperature, density, and chemical composition. Since the reduced mass is slightly larger for the ^{3}He+^{4}He reaction than it is for the ^{3}He+^{3}He reaction, it will have a slightly larger temperature sensitivity. With increasing temperature, pp2 and pp3 will therefore start to dominate over pp1 if ^{4}He is present in appreciable amounts. Similarly, with increasing temperature, the importance of proton capture on ^{7}Be will start to dominate over the electron capture.

For low temperatures, say $T_7 \lesssim 0.8$, one has to calculate all the reactions independently and keep track of relative abundances. For higher temperatures, the intermediate reactions will be in equilibrium, and the energy generation can be taken to be proportional to the first step, which is the slowest. This is because it involves the weak nuclear force in the decay of a proton to a neutron during the short time the two protons are together. Indeed, in by far most cases, the compound two-proton nucleus that is formed at first, will simply break apart into two protons again. As a result, the effective cross-section is very small, $\sim 10^{-47}$ cm^2. For the energy, one finds

$$\epsilon_{\rm pp} = 2.54\,10^6 \ {\rm erg\ s^{-1}\ g^{-1}}\ \psi f_{1,1}\, g_{1,1}\, X_1^2\, \rho\, T_6^{-2/3}\, {\rm e}^{-33.81/T_6^{1/3}}, \tag{8.19}$$

with an uncertainty of about 5%. Here, $g_{1,1} \simeq 1 + 0.00382 T_6$, $f_{1,1} \simeq 1$ for electron screening, and ψ corrects for the relative contributions of the different chains. At $T_7 \lesssim 1$, $\psi \simeq 1$, but at $T_7 = 2$, it varies between 1.4 for $Y = 0.3$ to nearly 2 for $Y = 0.9$. At still higher temperatures, when pp3 starts to dominate, it goes to 1.5 almost independent of Y. The temperature dependence of the reaction, as calculated from Eq. 8.17, is relatively mild: $\nu \simeq 4$ (i.e., $\epsilon_{\rm pp} \propto T^4$, much less steep than we will find below for other reactions).

CNO cycle

At sufficiently high temperatures, hydrogen can be burned to helium via the CNO cycle, in which carbon, nitrogen, and oxygen act more or less as catalysts (these have to be present, of course). The reactions are split in a main cycle (CN cycle) and a secondary cycle (ON cycle), as follows:

M1. $^{12}{\rm C} + {}^1{\rm H} \rightarrow {}^{13}{\rm N} + \gamma$

M2. $^{13}{\rm N} \rightarrow {}^{13}{\rm C} + {\rm e}^+ + \nu$

M3. $^{13}{\rm C} + {}^1{\rm H} \rightarrow {}^{14}{\rm N} + \gamma$

M4. $^{14}{\rm N} + {}^1{\rm H} \rightarrow {}^{15}{\rm O} + \gamma$

M5. $^{15}{\rm O} \rightarrow {}^{15}{\rm N} + {\rm e}^+ + \nu$

M6. $^{15}{\rm N} + {}^1{\rm H} \rightarrow \begin{cases} {}^{12}{\rm C} + {}^4{\rm He} & \text{and back to line M1 (main CN cycle).} \\ {}^{16}{\rm O} + \gamma & \text{(secondary ON cycle):} \end{cases}$

S1. $^{16}{\rm O} + {}^1{\rm H} \rightarrow {}^{17}{\rm F} + \gamma$

S2. $^{17}{\rm F} \rightarrow {}^{17}{\rm O} + {\rm e}^+ + \nu$

S3. $^{17}{\rm O} + {}^1{\rm H} \rightarrow \begin{cases} {}^{14}{\rm N} + {}^4{\rm He} & \text{and back to line M4.} \\ {}^{18}{\rm F} + \gamma \end{cases}$

S4. $^{18}{\rm F} \rightarrow {}^{18}{\rm O} + {\rm e}^+ + \nu$

S5. $^{18}{\rm O} + {}^1{\rm H} \rightarrow {}^{15}{\rm N} + {}^4{\rm He}$ and back to line M6.

The branch to the ON cycle (at line M6) is roughly 10^{-3} to 10^{-4} times less likely than the main branch back to the beginning of the CN cycle. The ON cycle is important, however, since it results

in oxygen being converted to nitrogen (which takes part in the CN cycle) — the branching *inside* the ON cycle (at line S3) does not strongly favor one branch over the other, but both branches lead to the CN cycle. The beta-decay times are of order $10^2 \ldots 10^3$ seconds, *much* shorter than typical nuclear reaction timescales.

Again, for high enough temperatures the reaction cycle will reach equilibrium, and the reaction rate will be set by the slowest link in the CN cycle, which is the proton-capture on ^{14}N. Because of this bottleneck in the CN cycle, and due to the small branching ratio into the ON cycle, most of the CNO originally present will be turned into ^{14}N. The energy gain of the whole cycle, after taking out neutrino losses, is 24.97 MeV, and one finds

$$\epsilon_{\rm CNO} = 7.48\,10^{27}\,{\rm erg\,s^{-1}\,g^{-1}}\;g_{14,1}\,f_{14,1}\,X_{\rm CNO}\,X_1\,\rho\,T_6^{-2/3}\,{\rm e}^{-152.31/T_6^{1/3}-(T_6/800.)^2} \tag{8.20}$$

(with an uncertainty of $\pm 10\%$), where $g_{14,1} \simeq 1 - 0.002T_6$, $f_{14,1} \sim 1$ for electron screening, and $X_{\rm CNO} = X_{\rm C} + X_{\rm N} + X_{\rm O}$. At somewhat lower temperatures, the CN cycle can reach equilibrium, but the burning of ^{16}O proceeds slowly; Eq. 8.20 is still quite a good approximation, but with $X_{\rm CNO} = X_{\rm C} + X_{\rm N} + |\Delta X_{\rm O\to N}(t)|$, where $|\Delta X_{\rm O\to N}(t)|$ is the amount of ^{16}O that has been burned to nitrogen as of time t (note that the intermediate ^{17}O stage may also slow down the conversion of ^{16}O to nitrogen, since the reaction rates of ^{16}O and ^{17}O may be comparable).

Inside stars that burn predominantly via the CNO cycle, the nitrogen abundance will be far larger than it normally is, while carbon and oxygen will be correspondingly underabundant. Indeed, such abundance patterns are observed in massive stars which have lost a lot of mass, so that processed material reaches the surface. Examples of these are the ON stars and Wolf-Rayet stars of type WN. (In carbon-rich Wolf-Rayet stars, one even sees the products of helium fusion.) Also, in lower-mass red giants, some CNO-processed material is mixed to the surface.

For the CNO cycle, the temperature sensitivity is high, $\nu = 23 \ldots 13$ for $T_6 = 10 \ldots 50$. As a result, the p-p chain dominates at low temperatures, and the CNO cycle at high temperatures, as is illustrated in Fig. 8.3. Furthermore, because of the steep temperature dependence, the energy production will be highly concentrated towards the centre. Therefore, L_r/r^2 will be large, and thus $\nabla_{\rm rad}$ will be large as well. This is why massive stars have convective cores.

Helium burning

When all the hydrogen has been fused into helium, it is difficult to continue, because until one reaches carbon, the elements following helium have lower binding energy per nucleon (see Fig. 8.1). As a result, the fusion of two helium nuclei leads to a ^{8}Be nucleus whose ground state is nearly 100 keV lower in energy; therefore, it decays back into two alpha particles in a few 10^{-16} s. Nevertheless, this is still about 10^5 times longer than the encounter time — in fact, a ^{8}Be abundance of about 10^{-9} builds up in stellar matter. Occasionally, it will happen that another alpha particle comes by so that a carbon nucleus can be formed. This whole process is called the *triple-alpha reaction* because it almost is a three-body interaction. Writing out the reactions,

$$^4{\rm He} + {}^4{\rm He} \;\rightleftharpoons\; {}^8{\rm Be}$$

$$^8{\rm Be} + {}^4{\rm He} \;\rightarrow\; {}^{12}{\rm C} + \gamma$$

The total energy released per carbon nucleus formed is 7.274 MeV. For these reactions, it is much less straightforward to derive an energy generation rate, because "*resonances*" (as described above) are important for both the above steps. Roughly, the energy generation rate is

$$\epsilon_{3\alpha} = 4.99\,10^{11}\,{\rm erg\,s^{-1}\,g^{-1}}\;f_{3\alpha}\,Y^3\,\rho^2\,T_8^{-3}\left(1 + 0.00354T_8^{-0.65}\right){\rm e}^{-43.92/T_8} \tag{8.21}$$

(with an uncertainty of $\pm 14\%$), where $f_{3\alpha} = f_{4,4}f_{8,4}$ is the combined electron screening factor. For this reaction, the temperature sensitivity is very high, $\nu = 40 \ldots 19$ for $T_8 = 1 \ldots 2$.

Other fusion processes can occur simultaneously (energy gain in MeV is shown to the right):

$$
\begin{aligned}
&{}^{12}\mathrm{C} + {}^{4}\mathrm{He} \rightarrow {}^{16}\mathrm{O} + \gamma && && 7.162\\
&{}^{16}\mathrm{O} + {}^{4}\mathrm{He} \rightarrow {}^{20}\mathrm{Ne} + \gamma && && 4.730\\
&{}^{14}\mathrm{N} + {}^{4}\mathrm{He} \rightarrow {}^{18}\mathrm{F} + \gamma\,, && {}^{18}\mathrm{F} \rightarrow {}^{18}\mathrm{O} + \mathrm{e}^{+} + \nu && 5.635 \text{ (total, excluding neutrino energy)}\\
&{}^{18}\mathrm{O} + {}^{4}\mathrm{He} \rightarrow {}^{22}\mathrm{Ne} + \gamma && && 9.667
\end{aligned}
$$

The second of these is slow, and for the last two ^{14}N is not very abundant (and thus its product ^{18}O is not very abundant either). The first reaction is therefore the most important one. It is rather complicated (and has an uncertainty of $\pm 40\%$); approximately,

$$
\begin{aligned}
\epsilon_{12,\alpha} \simeq\; & 9.58\,10^{26}\,\mathrm{erg\,s^{-1}\,g^{-1}}\; f_{12,4}\, X_{12}\, Y \rho\, T_8^{-2} \Big[\left(1 + 0.254 T_8 + 0.00104 T_8^2 - 0.000226 T_8^3\right) \mathrm{e}^{-(T_8/46.)^2} \\
& + \left(0.985 + 0.9091 T_8 - 0.1349 T_8^2 + 0.00729 T_8^3\right) \mathrm{e}^{-(T_8/13.)^2} \Big]\, \mathrm{e}^{-71.361/T_8^{1/3}}\,. \qquad (8.22)
\end{aligned}
$$

Carbon burning and onward

After helium has been exhausted, the next processes to start are those of carbon burning, at temperatures of order $T_9 = 0.5 \ldots 1$. The situation is very complicated, since the excited ^{24}Mg nucleus that is produced is unstable and can decay in a number of different ways:

$$
\begin{aligned}
{}^{12}\mathrm{C} + {}^{12}\mathrm{C} &\rightarrow {}^{24}\mathrm{Mg} + \gamma && 13.931\\
&\rightarrow {}^{23}\mathrm{Mg} + \mathrm{n} && -2.605\\
&\rightarrow {}^{23}\mathrm{Na} + \mathrm{p} && 2.238\\
&\rightarrow {}^{20}\mathrm{Ne} + \alpha && 4.616\\
&\rightarrow {}^{16}\mathrm{O} + 2\alpha && -0.114
\end{aligned}
$$

The last column lists the energy gain in MeV. Here, the most probable reactions are those leaving ^{23}Na and ^{20}Ne. The next complication that arises, is that the proton and alpha particle produced in these two reactions immediately fuse with other particles (since for them, the temperatures are extremely high). As a result of these complications, the energy rate is rather uncertain. For some approximate values, see Kippenhahn & Weigert, p. 167.

For temperatures above 10^9 K, the photon energies become so large that they can lead to the break-up of not-so-tightly bound nuclei. Reaction rates analogous to the Saha equation for ionization can be written to determine equilibrium conditions. Generally, however, equilibrium will not be reached as time is most definitely running out if a star reaches these stages. A reaction which is important subsequent to Carbon burning is ${}^{20}\mathrm{Ne} + \gamma \rightarrow {}^{16}\mathrm{O} + \alpha$ (the *reverse* of the helium burning reaction). The alpha particles resulting from this photo-disintegration are captured faster by Neon (via ${}^{20}\mathrm{Ne} + \alpha \rightarrow {}^{24}\mathrm{Mg} + \gamma$) than by the Oxygen nuclei, and hence the net reaction is $2\,{}^{20}\mathrm{Ne} + \gamma \rightarrow {}^{16}\mathrm{O} + {}^{24}\mathrm{Mg} + \gamma$, with an energy gain of 4.583 MeV. This is called Neon burning.

The next phase is oxygen burning, for which temperatures in excess of 10^9 K are required. As for carbon burning, the reaction can proceed via a number of channels:

$$
\begin{aligned}
{}^{16}\mathrm{O} + {}^{16}\mathrm{O} &\rightarrow {}^{32}\mathrm{S} + \gamma && 16.541\\
&\rightarrow {}^{31}\mathrm{S} + \mathrm{n} && 1.453\\
&\rightarrow {}^{31}\mathrm{P} + \mathrm{p} && 7.677\\
&\rightarrow {}^{28}\mathrm{Si} + \alpha && 9.593\\
&\rightarrow {}^{24}\mathrm{Mg} + 2\alpha && -0.393
\end{aligned}
$$

For these reactions, the most frequent product is ^{31}P; next most frequent is ^{28}Si. Again, the small particles immediately lead to a multitude of other reactions. Among the end products will be a large amount of ^{28}Si.

At the end of Oxygen burning, photo-disintegration becomes more and more important. In particular, photo-disintegration of ^{28}Si leads to the ejection of protons, neutrons and alpha particles, which fuse with other ^{28}Si particles to form bigger nuclei that in turn are subjected to photo-disintegration. Still, gradually larger nuclei are built up, up to ^{56}Fe. Since iron is so strongly bound, it may survive as the dominant species. The whole process is called silicon burning.

For next time

- Think about what happens when the core has turned into Iron.
- Read ahead about stellar models: textbook §10.5 and Appendix H.

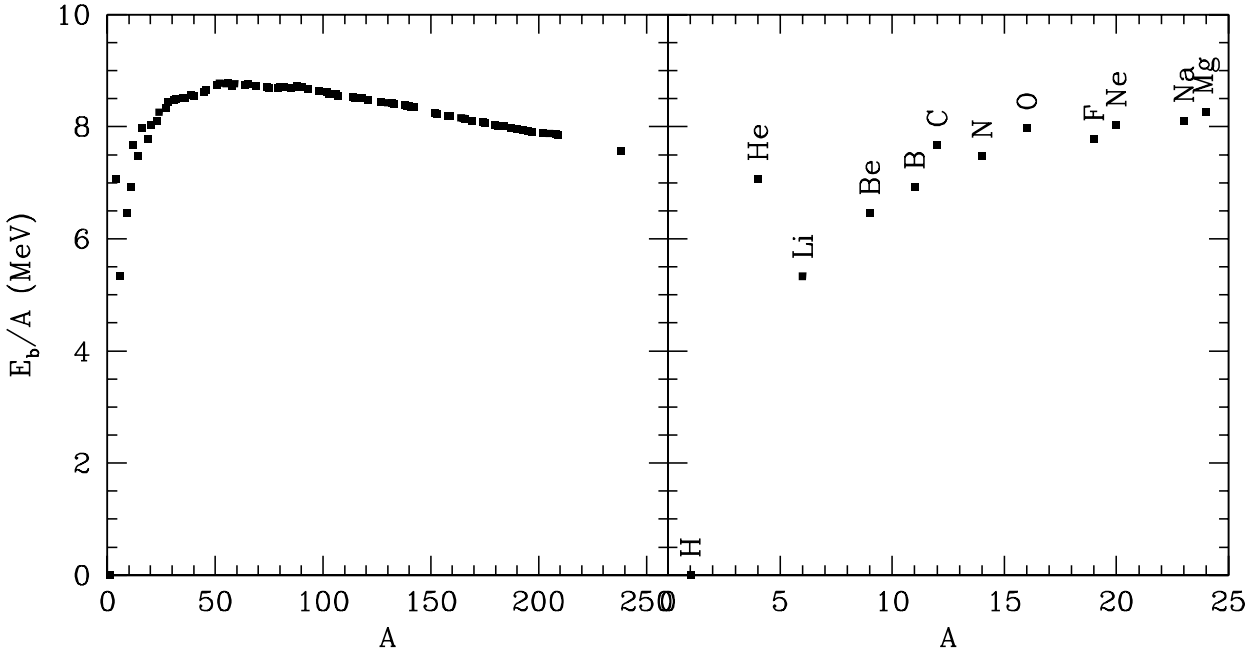


Fig. 8.1. Binding energy per nucleon for the different elements. In the right-hand panel an enlargement of the plot is shown and the elements are labeled. From Verbunt (2000, first-year lecture notes, Utrecht University).

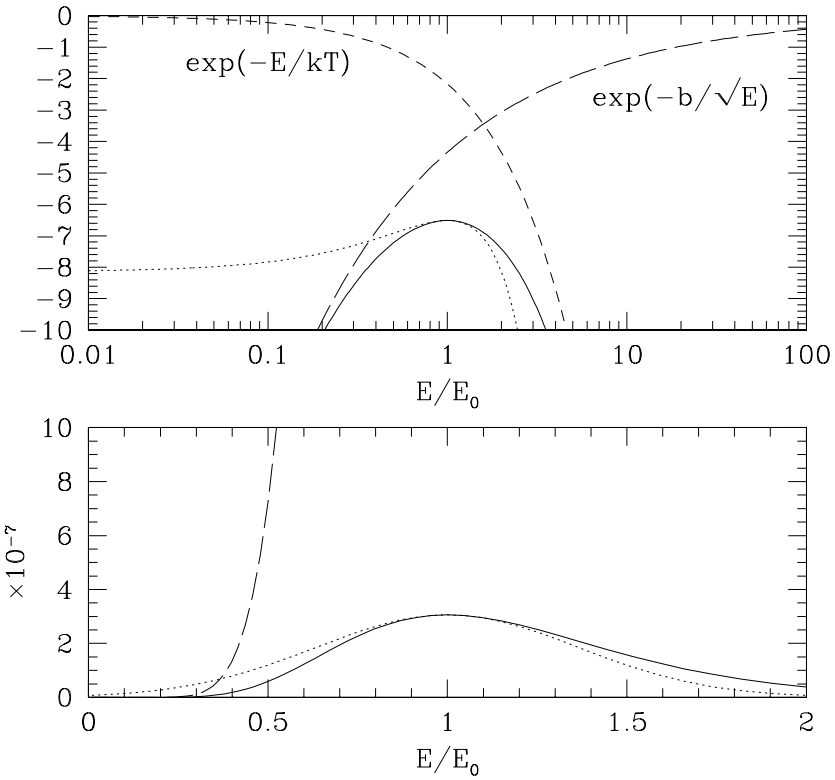


Fig. 8.2. Gamow peak resulting from the competing exponential terms: (1) from the Maxwellian (short-dashed line: $\propto \exp(-E/kT)$, with $kT = 0.2\,E_0$ here), and (2) from the penetration probability (long-dashed line: $\propto \exp(-b/\sqrt{E})$, with $b = 10\sqrt{E_0}$ here). The solid line indicates the product, and the dotted line the approximating Gaussian discussed in the text. *Upper panel:* logarithmic scale; *lower panel:* linear scale.

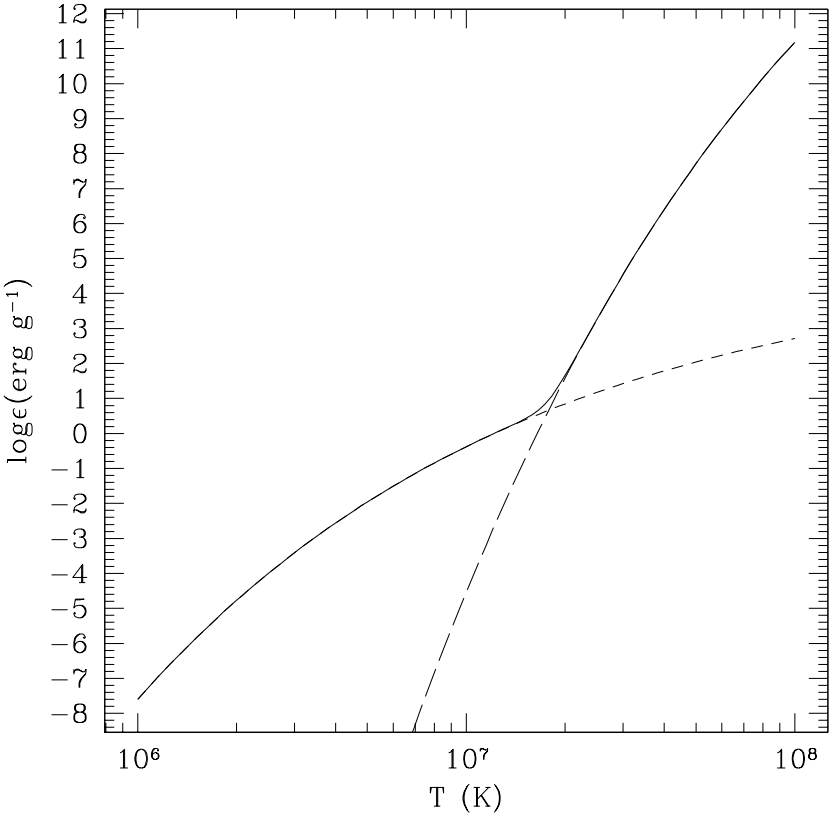


Fig. 8.3. Energy generation rates for matter with $\rho = 10\,\mathrm{g\,cm^{-3}}$, $X_1 = 0.7$, $X_{\mathrm{CNO}} = 0.01$, and a range of temperatures. The contributions from the p-p chain (short-dashed) and CNO cycle (long-dashed) are also indicated separately.

9. Stellar Models

Textbook: §10.5, App. H

The problem

To calculate a star's structure, we need to solve the equations of hydrostatic equilibrium, mass continuity, energy balance, and energy transport. It makes most sense to write these in terms of fractional mass M_r rather than fractional radius r (since composition profiles are determined by the position in terms of M_r, which, unlike r, does not change when the star expands or contracts). The mass continuity equation (Eq. 1.3) can be used to put the equations into the following form:

$$\text{[mass continuity (Eq. 1.3)]:} \quad \frac{\mathrm{d}r}{\mathrm{d}M_r} = \frac{1}{4\pi r^2 \rho} \,, \tag{9.1}$$

$$\text{[hydrostatic equilibrium (Eq. 1.2)]:} \quad \frac{\mathrm{d}P}{\mathrm{d}M_r} = -\frac{GM_r}{4\pi r^4} \,, \tag{9.2}$$

$$\text{[energy balance (Eq. 8.1)]:} \quad \frac{\mathrm{d}L_r}{\mathrm{d}M_r} = \epsilon_{\rm nuc} - \epsilon_\nu + \epsilon_{\rm grav} \,, \tag{9.3}$$

$$\text{[generalized Eddington equation]:} \quad \frac{\mathrm{d}T}{\mathrm{d}M_r} = -\frac{GM_rT}{4\pi r^4 P}\nabla_* \,. \tag{9.4}$$

In Eq. 9.4, depending on whether the layer is radiative or convective, one has

$$\nabla_* = \begin{cases} \nabla_{\rm rad} = \dfrac{3}{16\pi acG}\dfrac{\kappa L_r P}{M_r T^4} & \text{(radiative layers)} \,, \\ \nabla_{\rm ad} + \nabla_{\rm sa} & \text{(convective layers)} \,. \end{cases} \tag{9.5}$$

Here, $\nabla_{\rm sa}$ is the super-adiabatic part of the gradient (i.e., $\nabla_{\rm sa} \equiv \nabla_{\rm conv} - \nabla_{\rm ad}$); $\nabla_{\rm sa}$ can be neglected in the interior (where $\nabla_{\rm conv} \simeq \nabla_{\rm ad}$) but not near the surface (where $\nabla_{\rm conv} > \nabla_{\rm ad}$). The condition for convection can either be the Ledoux or the Schwarzschild criterion.

Evolution consists of thermal adjustments (via $\epsilon_{\rm grav}$) and changes in the abundances, due to the fusion reactions that proceed with rates $r_{a,b}$ (Eq. 8.8 — note that $\langle\sigma v\rangle$ is a function of T):

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = \frac{m_i}{\rho}\left(\sum_{j,k} r_{j,k(\to i)} - \sum_{k'}(1+\delta_{i,k'})\, r_{i,k'}\right) , \qquad i = 1, \ldots, I \,, \tag{9.6}$$

where i labels all isotopes being considered, $r_{j,k(\to i)}$ are reactions that produce isotope i (from j and k), and $r_{i,k'}$ are reactions that destroy i (and also k'). One of the relations can be replaced by the normalization condition, $\sum_i X_i = 1$ (or this condition can be used to check that you have coded the nuclear reactions correctly!). Furthermore, the abundances should be mixed in convective (and semi-convective) zones, taking account of possible overshooting.

In the above equations, we assume that the equation of state, the opacity, and the nuclear reactions are known functions of composition, temperature, and either density or pressure – these are equivalent, as the usual expression of the equation of state $P = P(\rho, T, X_i)$ can be inverted and expressed as $\rho = \rho(P, T, X_i)$ instead. In other words, as functions of (ρ, T, X_i) or (P, T, X_i), we have:

Equation of state:	$\{\ P(\rho,T,X_i)$ or $\rho(P,T,X_i)\ \},\ \nabla_{\rm ad},\ s,\ C_V,\ C_P,\ \left(\frac{\partial \ln P}{\partial \ln T}\right)_\rho ,\ \left(\frac{\partial \ln P}{\partial \ln \rho}\right)_T$
Opacity (incl. conduction):	κ
Nuclear reaction rates:	$r_{j,k},\ \epsilon_{\rm nuc},\ \epsilon_\nu$

[Note that equation-of-state quantities s, C_V, C_P, $\left(\frac{\partial \ln P}{\partial \ln T}\right)_\rho$, and $\left(\frac{\partial \ln P}{\partial \ln \rho}\right)_T$ enter into $\epsilon_{\rm grav}$ and the formulae that can be used to obtain $\nabla_{\rm conv}$ in regions where $\nabla_{\rm sa}$ is not negligible.] With the above given, there are as many differential equations as unknowns.

While the equations can be expressed equally well in terms of (ρ, T, X_i), for simplicity, we will assume hereafter that the above are expressed as functions of (P, T, X_i). The unknowns are then $(P, r, L_r, T, X_1, \ldots, X_I)$, whose dependence as a function of M_r and t is to be determined. For this purpose, we need boundary conditions at $M_r = 0$ and $M_r = M$ and initial values for the composition X_i and gravitational energy (e.g., an entropy profile).

Boundary conditions

The inner boundary condition is simple: $r = 0$, $L_r = 0$ for $M_r = 0$. Unfortunately, we cannot put any a priori constraints on $P_{\rm c}$ and $T_{\rm c}$, so that integrating from the centre outwards we have families of two-parameter solutions $r(P_{\rm c}, T_{\rm c})$ and $L_r(P_{\rm c}, T_{\rm c})$. For small M_r, we can write these functions as expansions in M_r,

$$r(P_{\rm c}, T_{\rm c}) \;=\; \left(\frac{3}{4\pi\rho_{\rm c}}\right)^{1/3} M_r^{1/3} \;, \tag{9.7}$$

$$L_r(P_{\rm c}, T_{\rm c}) \;=\; (\epsilon_{\rm nuc,c} - \epsilon_{\nu,{\rm c}} + \epsilon_{\rm grav,c})\, M_r \;, \tag{9.8}$$

where $\rho_{\rm c}$ and the various $\epsilon_{\rm c}$ are known functions of $(P_{\rm c}, T_{\rm c})$. These expansions are often more useful than the $M_r = 0$ conditions, since Eqs. 9.1, 9.2, and 9.4 become indeterminate at $M_r = 0$.

At the surface, we will have conditions for P and T, but R and L are unknown a priori, leading to a situation similar to that in the centre: for given M, R, and L, one can calculate $\log g$ and $T_{\rm eff}$, which determine the run of pressure and temperature in the atmosphere. Thus, integrating from the surface downwards we have families of two-parameter solutions $P(R, L)$ and $T(R, L)$. Unfortunately, the surface condition is not simple. One could use $P = 0$, $T = 0$ for $M_r = M$, but for convective envelopes this leads to gross errors. Somewhat more elegant is to use the photosphere, where $T_{\rm eff} = (L/4\pi R^2\sigma)^{1/4}$ and $P_{\rm phot} = 2g/3\kappa$. The condition for the pressure is derived from requiring $\tau = \frac{2}{3}$ at the photosphere, as was done in the discussion of the Hayashi line (Eq. 7.1); for κ, a suitably chosen average of the opacity above the photosphere has to be used in order to get an accurate value for $P_{\rm phot}$ (see Fig. 9.1).

The main problem with these simple boundary conditions is that near the surface the assumptions underlying the energy transport equation break down: the photon mean-free path becomes substantial. In these regions, much more detailed radiative transfer calculations are required. One can use a simple "grey atmosphere" approximation (in which one assumes that the opacity κ_ν is equal to the Rosseland value, independent of wavelength) to perform an approximate integral over the atmosphere. An alternate solution to this problem is to leave it to those interested in detailed stellar atmospheres, and use a grid of their results. For given (R, L), one calculates $T_{\rm eff}$ and $\log g$, and uses this to to interpolate in the (R, L, M) grid of model atmosphere results to find P_*, T_* at the bottom of the atmosphere.

Computational methods

There are several ways one could attempt to calculate stellar models and evolution numerically. First consider the case where $X_i(M_r)$ and $\epsilon_{\rm grav}(M_r)$ are known, i.e., where we have to solve just the structure of the star.

In principle, one could simply start integrating from both sides for trial values of $(P_{\rm c}, T_{\rm c})$ and (R, L), and try to match the two solutions at some intermediate fitting point, by varying the trial values. This is called the *shooting method*. In general, given a good scheme, the solution converges quickly (the program `statstar` in CO, App. H, is a simple example; see Numerical Recipes, § 17.2 for more details). It is not very efficient, however, if one wants to calculate the evolution,

in which the star evolves through a series of spatial models which are very similar. For this case, it is better to use a method which uses the spatial model from a previous step as an initial guess and makes small adjustments in order to find the new equilibrium. Most commonly used for this purpose is the *Henyey method*, which is especially well-suited for solving differential equations with boundary conditions on both sides.

The method works as follows. Take a grid of points $M_r^{(j)}$, with $j = 1, \ldots, N$. Then, discretise the differential equations, bring both sides to the left-hand side, and call these $A_i^{(j)}$. Then, a solution will be given by

$$A_i^{(j)} = \frac{y_i^{(j+1)} - y_i^{(j)}}{M_r^{(j+1)} - M_r^{(j)}} - f_i(M_r^{(j+\frac{1}{2})}, y_1^{(j+\frac{1}{2})}, y_2^{(j+\frac{1}{2})}, y_3^{(j+\frac{1}{2})}, y_4^{(j+\frac{1}{2})}) = 0 \,, \qquad i = 1, \ldots, 4 \,, \quad j = 1, \ldots, N-1 \tag{9.9}$$

where $y_1, \ldots, y_4$ are the four variables of interest (e.g., $y_1 = r$, $y_2 = P$, $y_3 = L_r$, $y_4 = T$), the index i numbers the four equations, and $f_1, \ldots, f_4$ are the right–hand side functions in the differential equations. The superscript $j + \frac{1}{2}$ is meant to indicate that a suitable average of the values at grid points j and $j + 1$ is taken (e.g., just a straight mean).

At the inner and outer boundaries, we have

$$\begin{aligned} B_1^{(\mathrm{in})} &= r^{(1)} - r(P_\mathrm{c}, T_\mathrm{c}) = y_1^{(1)} - f_1^{(\mathrm{in})}(y_2^{(1)}, y_4^{(1)}) = 0 \,, \\ B_3^{(\mathrm{in})} &= L_r^{(1)} - L_r(P_\mathrm{c}, T_\mathrm{c}) = y_3^{(1)} - f_3^{(\mathrm{in})}(y_2^{(1)}, y_4^{(1)}) = 0 \,, \\ B_2^{(\mathrm{out})} &= P^{(N)} - P(R, L) = y_2^{(N)} - f_2^{(\mathrm{out})}(y_1^{(N)}, y_3^{(N)}) = 0 \,, \\ B_4^{(\mathrm{out})} &= T^{(N)} - T(R, L) = y_4^{(N)} - f_4^{(\mathrm{out})}(y_2^{(N)}, y_4^{(N)}) = 0 \,, \end{aligned} \tag{9.10}$$

where we assumed one could determine $(P_\mathrm{c}, T_\mathrm{c})$ from the values at the first grid point and (R, L) from those at the last. Note that for the simple case for which $M_r^{(1)} = 0$, the functions $r(P, T)$ and $L_r(P, T)$ are identical to zero. If one choses to work in logarithmic units for $\{\rho, P, r, T\}$, however, the first point cannot be at $M_r = 0$, and therefore the inner boundary conditions are written in their more general form above. Thus, with the above definitions of A, B, a solution for the problem requires $A_i^{(j)} = 0$, $B_i = 0$.

Considering the whole grid, we have $4N$ unknowns $y_i^{(j)}$ and $4(N-1) + 2 + 2 = 4N$ equations. Now suppose that we have a first approximation $y_i^{(j)}(1)$ to the solution. For this initial guess, the constraints will not be met, i.e., $A_i^{(j)}(1) \neq 0$, $B_i(1) \neq 0$, and we need to find corrections $\delta y_i^{(j)}$ such that a second approximation $y_i^{(j)}(2) = y_i^{(j)}(1) + \delta y_i^{(j)}$ does give a solution, i.e., we are looking for changes $\delta y_i^{(j)}$ that imply changes $\delta A_i^{(j)}$, δB_i, such that $A_i^{(j)}(1) + \delta A_i^{(j)} = 0$, $B_i(1) + \delta B_i = 0$, or

$$\begin{aligned} \delta B_i^{(\mathrm{in})} &= -B_i^{(\mathrm{in})}(1) \,, \qquad i = 1, 3 \\ \delta A_i^{(j)} &= -A_i^{(j)}(1) \,, \qquad i = 1, \ldots, 4 \,, \quad j = 1, \ldots, N-1 \\ \delta B_i^{(\mathrm{out})} &= -B_i^{(\mathrm{out})}(1) \,, \qquad i = 2, 4 \,. \end{aligned} \tag{9.11}$$

For small enough corrections, we can expand the A and B linearly in $\delta y_i^{(j)}$, and write

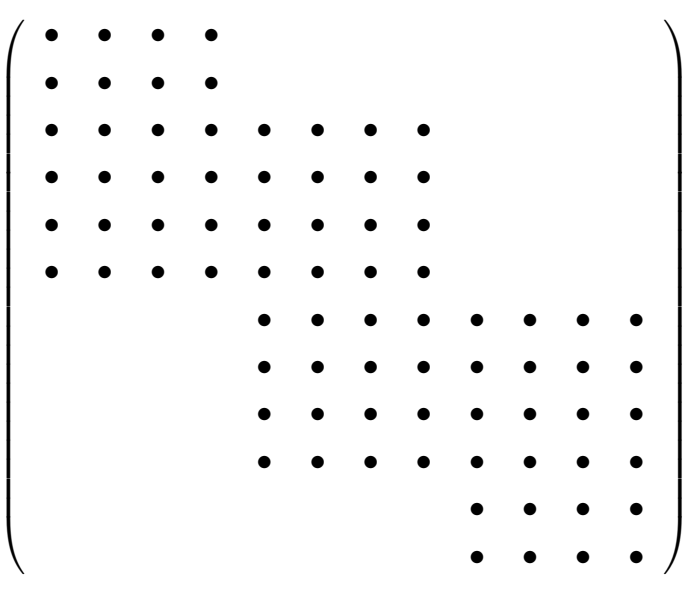
$$\begin{aligned}
\sum_{k=1}^{4} \frac{\partial B_i^{(\mathrm{in})}}{\partial y_k^{(1)}} \delta y_k^{(1)} &= -B_i^{(\mathrm{in})} \,, \qquad i = 1, 3 \\
\sum_{k=1}^{4} \frac{\partial A_i^{(j)}}{\partial y_k^{(j)}} \delta y_k^{(j)} + \sum_{k=1}^{4} \frac{\partial A_i^{(j)}}{\partial y_k^{(j+1)}} \delta y_k^{(j+1)} &= -A_i^{(j)} \,, \qquad i = 1, \ldots, 4 \,, \quad j = 1, \ldots, N-1 \\
\sum_{k=1}^{4} \frac{\partial B_i^{(\mathrm{out})}}{\partial y_k^{(N)}} \delta y_k^{(N)} &= -B_i^{(\mathrm{out})} \,, \qquad i = 2, 4
\end{aligned} \tag{9.12}$$

[we have dropped the (1) numbering the 1st approximation]. This system has $2 + 4(N-1) + 2 = 4N$ equations which need to be solved for the $4N$ unknown corrections $\delta y_i^{(j)}$. In matrix form,

$$\mathcal{H} \begin{pmatrix} \delta y_1^{(1)} \\ \vdots \\ \delta y_i^{(j)} \\ \vdots \\ \delta y_4^{(N)} \end{pmatrix} = - \begin{pmatrix} B_1^{(\mathrm{in})} \\ \vdots \\ A_i^{(j)} \\ \vdots \\ B_4^{(\mathrm{out})} \end{pmatrix} , \tag{9.13}$$

where $\mathcal{H}$ is called the *Henyey matrix*. Generally, this matrix equation can be solved ($\det \mathcal{H} \neq 0$), but since we used a first-order expansion, the next approximation $y + \delta y$ will still not fulfill the conditions accurately. Thus, one iterates, until a certain pre-set convergence criterion is met.

Note that Henyey matrix has a relatively simple form, as can be seen by writing out which elements are actually used for the case $N = 3$,



$$= - \begin{pmatrix} B_1^{(\mathrm{in})} \\ B_3^{(\mathrm{in})} \\ A_1^{(1)} \\ A_2^{(1)} \\ A_3^{(1)} \\ A_4^{(1)} \\ A_1^{(2)} \\ A_2^{(2)} \\ A_3^{(2)} \\ A_4^{(2)} \\ B_2^{(\mathrm{out})} \\ B_4^{(\mathrm{out})} \end{pmatrix} .$$

Here, the bullets indicate the elements that are used; all others are zero. Because of the simple structure, the solution can be found in a relative straightforward manner. See Numerical Recipes, § 17.3, for details, and for a method that is fast and minimizes storage.

Evolution

So far, we have ignored the chemical evolution and assumed that $\epsilon_{\rm grav}$ was a known function. The latter function can be estimated easily once we have made an initial model and want to compute a model one time step later, by approximating

$$\epsilon_{\rm grav}^{(j+\frac{1}{2})} = -T^{(j+\frac{1}{2})} \frac{\rm d}{{\rm d}t} s^{(j+\frac{1}{2})} = -\frac{T^{(j+\frac{1}{2})}}{\Delta t} \left(s^{(j+\frac{1}{2})} - s_{\rm prev}^{(j+\frac{1}{2})} \right), \tag{9.14}$$

Here, we expressed $\epsilon_{\rm grav}$ in terms of the entropy change, but the other expressions in Eq. 8.3 can be used in the same way. The point to note is that $s_{\rm prev}$, which is the entropy that the element had in the previous model, is known. Since the current entropy is a known function $s(P, T, X_i)$ (from the equation of state), also ${\rm d}s/{\rm d}t$ is a known function of (P, T, X_i). Thus, $\epsilon_{\rm grav}$ is a known function of (P, T, X_i) and can be used without problems in deriving the stellar structure. [A complication arises in convective regions, especially ones that are advancing into regions of different chemical composition. Mixing at constant pressure has no energy cost, but since it is an irreversible process, it results in an increase of entropy (which of course does *not* contribute towards $\epsilon_{\rm grav}$). On the other hand, when one is mixing the products of nuclear burning (i.e., *heavy* nuclei) outwards against gravity while mixing unburned stellar material (i.e., *light* nuclei) downwards, there is an energy cost involved in doing this (which is incurred throughout the region where material mixed upwards has a higher mean molecular weight than material mixed downwards). These effects may need to be accounted for correctly during stages when the star is evolving on a short timescale, which involves some modification of Eq. 8.3 for $\epsilon_{\rm grav}$.]

A scheme like the above for including a variable that changes in time is usually called an *implicit scheme*, since the time derivative is calculated implicitly, using parameters from the new model one is trying to determine. Schemes which rely only on previous model(s) are called *explicit*; these are often easier to code but in order to keep good accuracy small timesteps need to be taken.

For the abundances, an explicit scheme is simpler. In such a scheme, one determine the time derivatives $({\rm d}X_i/{\rm d}t)_{\rm prev}$ from the previous models according to Eq. 9.6 and then for the next model uses

$$X_i = X_{i,{\rm prev}} + \Delta t \left(\frac{{\rm d}X_i}{{\rm d}t} \right)_{\rm prev} . \tag{9.15}$$

Note that it is also possible to calculate the chemical evolution using an implicit scheme. For a more detailed but quite readable discussion, see Eggleton (1971, MNRAS 151, 351). In the same reference, another choice of independent grid variable is discussed, which allows one to regrid the model automatically so that fine grid spacing is used where required.

For next time

– Read ahead about the main sequence: textbook §10.6, 13.1.

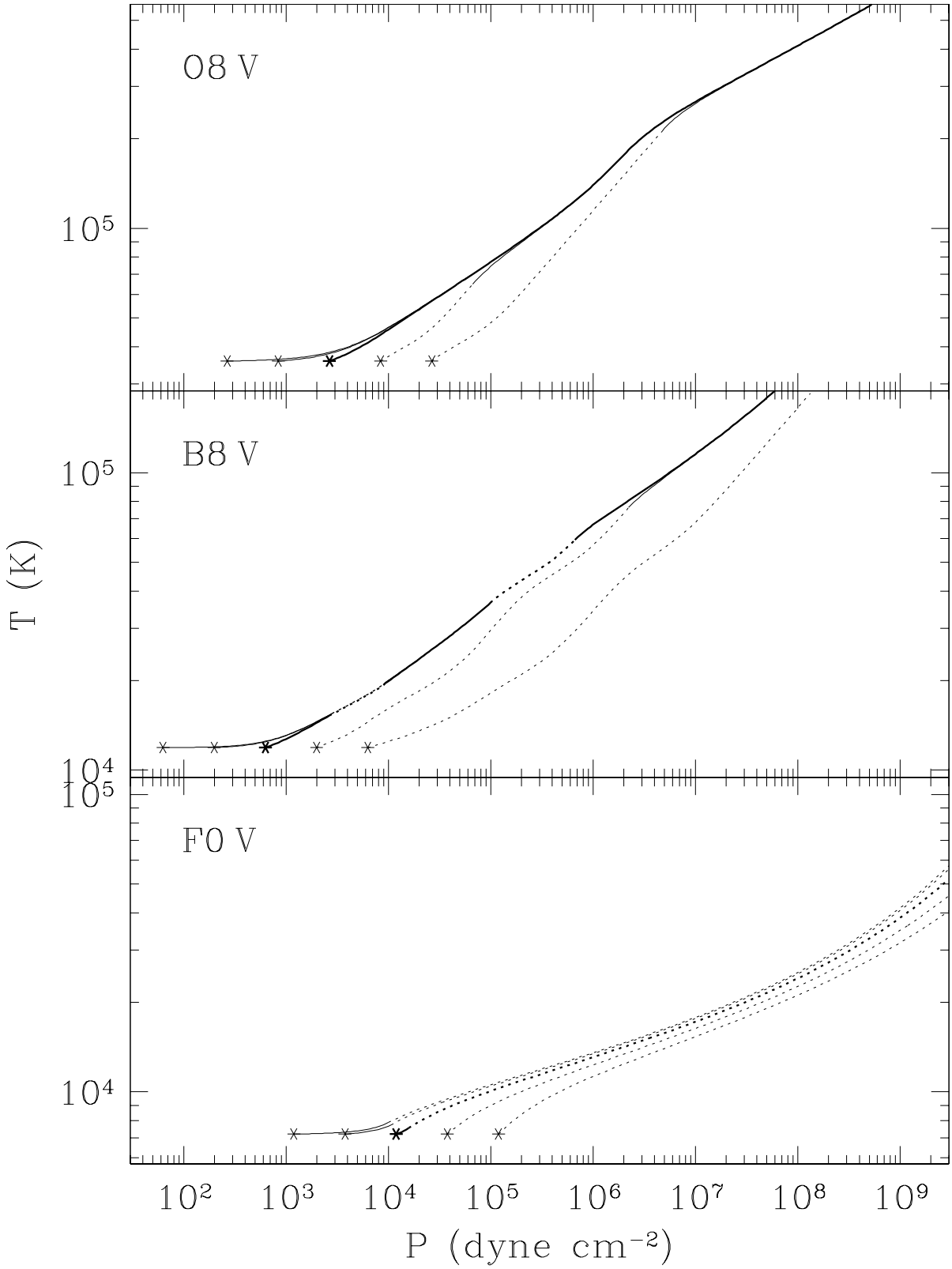


Fig. 9.1. Effect on the stellar envelope of choosing an incorrect value of $P_{\rm phot}$ in main sequence stars of solar metallicity, for a massive star with a fully radiative envelope (O8 V: $T_{\rm eff} \approx 37{,}000$ K, $M \approx 15\ M_\odot$), an intermediate mass star with very small convective zones in ionization regions (B8 V: $T_{\rm eff} \approx 12{,}000$ K, $M \approx 2.5\ M_\odot$), and a relatively low-mass star with a convective envelope of non-negligible extent (F0 V: $T_{\rm eff} \approx 7{,}200$ K, $M \approx 1.2\ M_\odot$). Star symbols ("$*$") indicate choices for $P_{\rm phot}$ at the relevant $T_{\rm eff}$, and lines indicate run of T with P inside the photosphere (solid lines indicate radiative regions, dashed lines indicate convective regions). **Heavy** symbols and lines indicate the **correct** models.