Chapter 9

Early stages of evolution and the main sequence phase

In this and the following chapters, an account will be given of the evolution of stars as it follows from full-scale, detailed numerical calculations. Because the stellar evolution equations are highly non-linear, they have complicated solutions that cannot always be anticipated on the basis of fundamental principles. We must accept the fact that simple, intuitive explanations cannot always be given for the results that emerge from numerical computations. As a consequence, the account of stellar evolution that follows will be more descriptive and less analytical than previous chapters.

This chapter deals with early phases in the evolution of stars, as they evolve towards and during the main-sequence phase. We start with a very brief (and incomplete) overview of the formation of stars.

9.1   Star formation and pre-main sequence evolution

The process of star formation constitutes one of the main problems of modern astrophysics. Compared to our understanding of what happens *after* stars have formed out of the interstellar medium – that is, stellar evolution – star formation is a very ill-understood problem. No predictive theory of star formation exists, or in other words: given certain initial conditions, e.g. the density and temperature distributions inside an interstellar cloud, it is as yet not possible to predict with certainty, for example, the *star formation efficiency* (which fraction of the gas is turned into stars) and the resulting *initial mass function* (the spectrum and relative probability of stellar masses that are formed). We rely mostly on observations to answer these important questions.

This uncertainty might seem to pose a serious problem for studying stellar evolution: if we do not know how stars are formed, how can we hope to understand their evolution? The reason that stellar evolution is a much more quantitative and predictive branch of astrophysics than star formation was already alluded to in Chapter 7. Once a recently formed star settles into hydrostatic and thermal equilibrium on the main sequence, its structure is determined by the four structure equations and only depends on the initial composition. Therefore all the uncertain details of the formation process are wiped out by the time its nuclear evolution begins.

In the context of this course we can thus be very brief about star formation itself, as it has very little effect on the properties of stars themselves (at least as far as we are concerned with individual stars – it does of course have an important effect on stellar *populations*).

Observations indicate that stars are formed out of molecular clouds, typically giant molecular clouds with masses of order $10^5\,M_\odot$. These clouds have typical dimensions of $\sim 10\,\mathrm{parsec}$, temperatures of

$10 - 100\,\mathrm{K}$ and densities of $10 - 300$ molecules/cm^3 (where the lowest temperatures pertain to the densest parts of the cloud). A certain fraction, about 1 %, of the cloud material is in the form of dust which makes the clouds very opaque to visual wavelengths. The clouds are in pressure equilibrium (hydrostatic equilibrium) with the surrounding interstellar medium. Roughly, we can distinguish six stages in the star formation process.

Interstellar cloud collapse Star formation starts when a perturbation, e.g. due to a shock wave originated by a nearby supernova explosion or a collision with another cloud, disturbs the pressure equilibrium and causes (part of) the cloud to collapse under its self-gravity. The condition for pressure equilibrium to be stable against such perturbations is that the mass involved should be less than a critical mass, the *Jeans mass*, which is given by

$$ M_{\mathrm{J}} \approx 4 \times 10^4 \, M_\odot \left(\frac{T}{100\,\mathrm{K}} \right)^{3/2} \left(\frac{n}{\mathrm{cm}^{-3}} \right)^{-1/2} \tag{9.1} $$

where n is the molecular density by number (see e.g. MAEDER Sec. 18.2.1 for a derivation). For typical values of T and n in molecular clouds $M_{\mathrm{J}} \sim 10^3 - 10^4 \, M_\odot$. Cloud fragments with a mass exceeding the Jeans mass cannot maintain hydrostatic equilibrium and will undergo essentially free-fall collapse. Although the collapse is dynamical, the timescale $\tau_{\mathrm{dyn}} \propto \rho^{-1/2}$ (eq. 2.18) is of the order of millions of years because of the low densities involved. The cloud is transparent to far-infrared radiation and thus cools efficiently, so that the early stages of the collapse are *isothermal*.

Cloud fragmentation As the density of the collapsing cloud increases, its Jeans mass decreases by eq. (9.1). The stability criterion within the cloud may now also be violated, so that the cloud starts to fragment into smaller pieces, each of which continues to collapse. The fragmentation process probably continues until the mass of the smallest fragments (dictated by the decreasing Jeans mass) is less than $0.1 \, M_\odot$.

Formation of a protostellar core The increasing density of the collapsing cloud fragment eventually makes the gas *opaque* to infrared photons. As a result, radiation is trapped within the central part of the cloud, leading to heating and an increase in gas pressure. As a result the cloud core comes into hydrostatic equilibrium and the dynamical collapse is slowed to a quasi-static contraction. At this stage we may start to speak of a *protostar*.
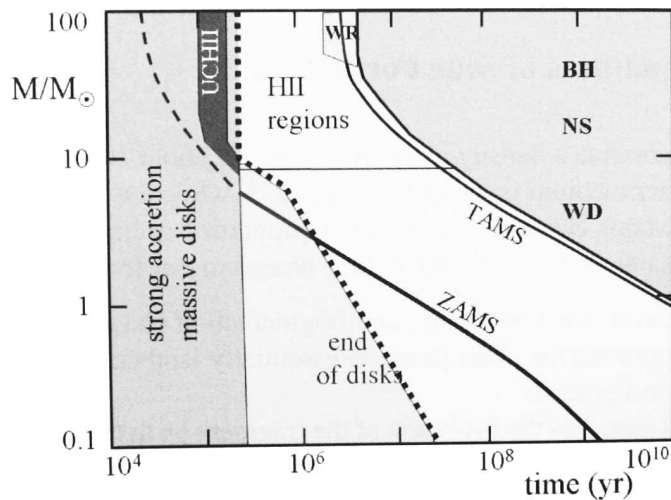

Figure 9.1. Timescales and properties of stars of mass M on the main sequence. Time along the abscissa is in logarithmic units to highlight the early phases, $t = 0$ corresponds to the formation of a hydrostatic core (stage 3 in the text). Initially the star is embedded in a massive accretion disk for $(1 - 2) \times 10^5$ years. In low-mass stars the disk disappears before the star settles on the zero-age main sequence (ZAMS). Massive stars reach the ZAMS while still undergoing strong accretion. These stars ionize their surroundings and excite an HII region around themselves. TAMS stands for terminal-age main sequence. Figure from MAEDER.
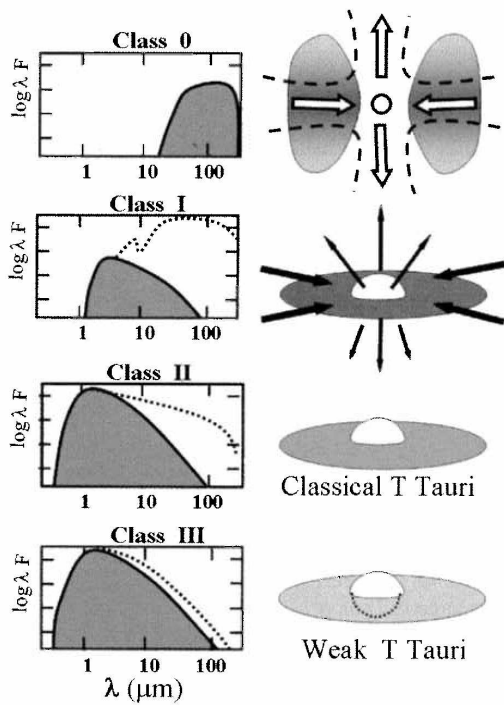
Figure 9.2. Schematic illustration of four stages in the evolution of protostars and their circumstellar disks. On the left, the stellar flux is depicted (shaded area) and the contribution from the disk (dotted line). On the right the corresponding geometry of the object is shown.

Class 0 objects are very young protostars ($\lesssim 10^4$ yrs) with almost spherical accretion at a high rate, emitting in the far-IR and sub-mm range. Class I protostars correspond to an advanced stage of accretion (age $\sim 10^5$ yrs), where the star is still embedded in a massive accretion disk, while jets or bipolar outflows are also observed. In class II the protostar has become visible as a classical T Tauri star on the pre-main sequence (age $\sim 10^6$ yrs), while the accretion disk is still optically thick giving rise to a large IR excess. Class III stars are already close to the main sequence (age $\sim 10^7$ yrs), with an optically thin accretion disk and weak emission lines. Figure from MAEDER.

Accretion The surrounding gas keeps falling onto the protostellar core, so that the next phase is dominated by accretion. Since the contracting clouds contain a substantial amount of angular momentum, the infalling gas forms an accretion disk around the protostar. These *accretion disks* are a ubiquitous feature of the star formation process and are observed around most very young stars, mostly at infrared and sub-millimeter wavelengths (see Fig. 9.2).

The accretion of gas generates gravitational energy, part of which goes into further heating of the core and part of which is radiated away, providing the luminosity of the protostar, so that

$$L \sim L_{\text{acc}} = \frac{GM\dot{M}}{2R} \tag{9.2}$$

where M and R are the mass and radius of the core and $\dot{M}$ is the mass accretion rate. The factor $\frac{1}{2}$ originates from the fact that half of the potential energy is dissipated in the accretion disk. Meanwhile he core heats up almost adiabatically since the accretion timescale $\tau_{\text{acc}} = M/\dot{M}$ is much smaller than the thermal timescale τ_{KH}.

Dissociation and ionization The gas initially consists of molecular hydrogen and behaves like an ideal gas, such that $\gamma_{\text{ad}} > \frac{4}{3}$ and the protostellar core is dynamically stable. When the core temperature reaches $\sim 2000\,\text{K}$ molecular hydrogen starts to dissociate, which is analogous to ionization and leads to a strong increase of the specific heat and a decrease of γ_{ad} below the critical value of $\frac{4}{3}$ (Sect. 3.5). Hydrostatic equilibrium is no longer possible and a renewed phase of *dynamical collapse* follows, during which the gravitational energy release is absorbed by the dissociating molecules without a significant rise in temperature. When H_2 is completely dissociated into atomic hydrogen HE is restored and the temperature rises again. Somewhat later, further dynamical collapse phases follow when first H and then He are ionized at $\sim 10^4$ K. When ionization of the protostar is complete it settles back into hydrostatic equilibrium at a much reduced radius (see below).

Pre-main sequence phase Finally, the accretion slows down and eventually stops and the protostar is revealed as a *pre-main sequence star*. Its luminosity is now provided by gravitational contraction and, according to the virial theorem, its internal temperature rises as $T \propto M^{2/3}\rho^{1/3}$ (Chapter 8). The surface cools and a temperature gradient builds up, transporting heat outwards. Further evolution takes place on the thermal timescale τ_{KH}.

A rough estimate of the radius R_p of a protostar after the dynamical collapse phase can be obtained by assuming that all the energy released during the collapse was absorbed in dissociation of molecular hydrogen (requiring $\chi_{H_2} = 4.48$ eV per H_2 molecule) and ionization of hydrogen ($\chi_H = 13.6$ eV) and helium ($\chi_{He} = 79$ eV). Because the final radius will be much smaller than the initial one, we can take the collapse to start from infinity. After the collapse the protostar is in hydrostatic equilibrium and must satisfy the virial theorem, $E_{tot} = \frac{1}{2}E_{gr}$. Taking E_{gr} as given by eq. (2.28), we can write

$$\frac{\alpha}{2}\frac{GM^2}{R_p} \approx \frac{M}{m_u}\left(\frac{X}{2}\chi_{H_2} + X\chi_H + \frac{Y}{4}\chi_{He}\right) \equiv \frac{M}{m_u}\chi. \tag{9.3}$$

Taking $X = 0.72$ and $Y = 1 - X$, we have $\chi = 16.9$ eV per baryon. For a polytrope of index n, $\alpha = 3/(5 - n)$ (eq. 4.19). We will shortly see that the protostar is completely convective and thus we can take $n = \frac{3}{2}$ and $\alpha = \frac{6}{7}$, such that

$$R_p \approx \frac{\alpha}{2}\frac{GMm_u}{\chi} \approx 50 R_\odot\left(\frac{M}{M_\odot}\right). \tag{9.4}$$

The average internal temperature can also be estimated from the virial theorem (eq. 2.29),

$$\bar{T} \approx \frac{\alpha}{3}\frac{\mu}{\mathcal{R}}\frac{GM}{R_p} = \frac{2}{3}\frac{\mu}{k}\chi \approx 8 \times 10^4 \text{ K}, \tag{9.5}$$

which is independent of the mass of the protostar. At these low temperatures the opacity is very high, rendering radiative transport inefficient and making the protostar convective throughout. The properties of such *fully convective stars* must be examined more closely.

9.1.1 Fully convective stars: the Hayashi line

We have seen in Sect. 7.2.3 that as the effective temperature of a star decreases the convective envelope gets deeper, occupying a larger and larger part of the mass. If T_{eff} is small enough stars can therefore become completely convective. In that case, as we derived in Sect. 5.5.2, energy transport is very efficient throughout the interior of the star, and a tiny superadiabaticity $\nabla - \nabla_{ad}$ is sufficient to transport a very large energy flux. The structure of such a star can be said to be *adiabatic*, meaning that the temperature stratification (the variation of temperature with depth) as measured by $\nabla = d\log T/d\log P$ is equal to ∇_{ad}. Since an almost arbitrarily high energy flux can be carried by such a temperature gradient, the *luminosity* of a fully convective star is practically *independent of its structure* – unlike for a star in radiative equilibrium, for which the luminosity is strongly linked to the temperature gradient.

It turns out that:

> Fully convective stars of a given mass occupy an almost vertical line in the H-R diagram (i.e. with $T_{eff} \approx$ constant). This line is known as the *Hayashi line*. The region to the right of the Hayashi line in the HRD (i.e. at lower effective temperatures) is a *forbidden region* for stars in hydrostatic equilibrium. On the other hand, stars to the left of the Hayashi line (at higher T_{eff}) cannot be fully convective but must have some portion of their interior in radiative equilibrium.

Since these results are important, not only for pre-main sequence stars but also for later phases of evolution, we will do a simplified derivation of the properties of the Hayashi line in order to make the above-mentioned results plausible.

Simple derivation of the Hayashi line

For any luminosity L, the interior structure is given by $\nabla = \nabla_{\text{ad}}$. For an ideal gas we have a constant $\nabla_{\text{ad}} = 0.4$, if we ignore the variation of ∇_{ad} in partial ionization zones. We also ignore the non-zero superadiabaticity of ∇ in the sub-photospheric layers (Sect. 5.5.2). The temperature stratification throughout the interior can then be described by a power law $T \propto P^{0.4}$. Using the ideal gas law, $P \propto \rho T$, we can eliminate T from both expressions and write

$$P = K\rho^{5/3},$$

which describes a polytrope of index $n = \frac{3}{2}$. Indeed, for an ideal gas the adiabatic exponent $\gamma_{\text{ad}} = \frac{5}{3}$. The constant K for a polytropic stellar model of index n is related to the mass M and radius R by eq. (4.15). For our fully convective star with $n = \frac{3}{2}$ we have $N_{3/2} = 0.42422$ (Table 4.1) and therefore

$$K = 0.42422 \, GM^{1/3}R. \tag{9.6}$$

Since the luminosity of a fully convective star is not determined by its interior structure, it must follow from the conditions (in particular the *opacity*) in the thin radiative layer from which photons escape, the photosphere. We approximate the photosphere by a spherical surface of negligible thickness, where we assume the photospheric boundary conditions (7.9) to hold. Writing the pressure, density and opacity in the photosphere (at $r = R$) as P_R, ρ_R and κ_R and the photospheric temperature as T_{eff}, we can write the boundary conditions as

$$\kappa_R P_R = \frac{2}{3}\frac{GM}{R^2}, \tag{9.7}$$

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4, \tag{9.8}$$

and we assume a power-law dependence of κ on ρ and T so that

$$\kappa_R = \kappa_0 \rho_R{}^a T_{\text{eff}}^b. \tag{9.9}$$

The equation of state in the photospheric layer is

$$P_R = \frac{\mathcal{R}}{\mu}\rho_R T_{\text{eff}}. \tag{9.10}$$

The interior, polytropic structure must match the conditions in the photosphere so that (using eq. 9.6)

$$P_R = 0.42422 \, GM^{1/3}R\rho_R{}^{5/3}. \tag{9.11}$$

For a given mass M, eqs. (9.7-9.11) constitute five equations for six unknowns, P_R, ρ_R, κ_R, T_{eff}, L and R. The solution thus always contains one free parameter, that is, the solution is a relation between two quantities, say L and T_{eff}. This relation describes the *Hayashi line* for a fully convective star of mass M.

Since we have assumed power-law expressions in all the above equations, the set of equations can be solved straightforwardly (involving some tedious algebra) to give a power-law relation between L and T_{eff} after eliminating all other unknowns. The solution can be written as

$$\log T_{\text{eff}} = A \log L + B \log M + C \tag{9.12}$$

where the constants A and B depend on the exponents a and b in the assumed expression for the opacity (9.9),

$$A = \frac{\frac{3}{2}a - \frac{1}{2}}{9a + 2b + 3} \qquad \text{and} \qquad B = \frac{a + 3}{9a + 2b + 3}. \tag{9.13}$$
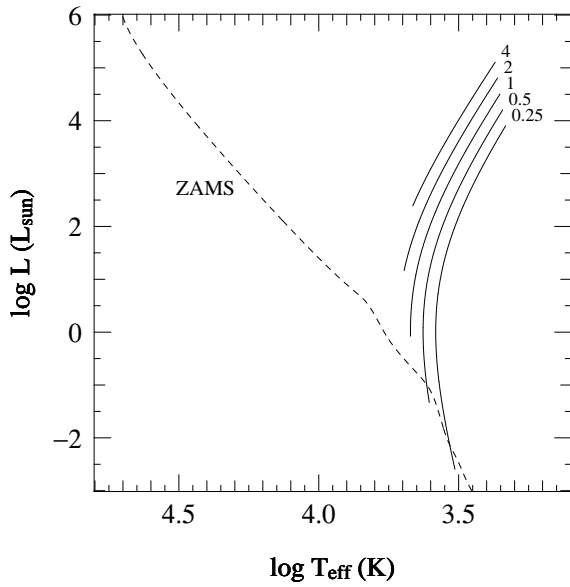
Figure 9.3. The position of the Hayashi lines in the H-R diagram for masses M = 0.25, 0.5, 1.0, 2.0 and 4.0 $M_\odot$ as indicated. The lines are analytic fits to detailed models computed for composition X = 0.7, Z = 0.02. The zero-age main sequence (ZAMS) for the same composition is shown as a dashed line, for comparison.

Note that the Hayashi lines do not have a constant slope, as expected from the simple analysis, but have a convex shape where the constant A (eq. 9.12) changes sign and becomes negative for high luminosities. The main reason is our neglect of ionization zones (where ∇_{ad} < 0.4) and the non-zero superadiabaticity in the outer layers, both of which have a larger effect in more extended stars.

Therefore the shape of the Hayashi line in the HRD is determined by how the opacity in the photosphere depends on ρ and T. Since fully convective stars have very cool photospheres, the opacity is mainly given by H^- absorption (Sect. 5.3) which increases strongly with temperature. According to eq. (5.34), $a \approx 0.5$ and $b \approx 9$ (i.e. $\kappa \propto T^9$!) in the the relevant range of density and temperature, which gives $A \approx 0.01$ and $B \approx 0.14$. Therefore (see Fig. 9.3)

- for a certain mass the Hayashi line is a very steep, almost vertical line in the HRD,

- the position of the Hayashi line depends on the mass, being located at higher T_{eff} for higher mass.

We can intuitively understand the steepness of the Hayashi line from the strong increase of H^- opacity with temperature. Suppose such a fully convective star would increase its radius slightly while attempting to keep L constant. Then the temperature in the photosphere would decrease and the photosphere would become much more transparent. Hence energy can escape much more easily from the interior, in other words: the luminosity will in fact *increase* strongly with a slight decrease in photospheric temperature.

The forbidden region in the H-R diagram

Consider models in the neighbourhood of the Hayashi line in the H-R diagram for a star of mass M. These models cannot have $\nabla = \nabla_{ad}$ throughout, because otherwise they would be *on* the Hayashi line. Defining $\bar{\nabla}$ as the average value of $d \log T / d \log P$ over the entire star, models on either side of the Hayashi line (at lower or higher T_{eff}) have either $\bar{\nabla} > \nabla_{ad}$ or $\bar{\nabla} < \nabla_{ad}$. It turns out (after more tedious analysis of the above equations and their dependence on polytropic index n) that models with $\bar{\nabla} < \nabla_{ad}$ lie at higher T_{eff} than the Hayashi line (to its left in the HRD) while models with $\bar{\nabla} > \nabla_{ad}$ lie at lower T_{eff} (to the right in the HRD).

Now consider the significance of $\bar{\nabla} \neq \nabla_{ad}$. If on average $\bar{\nabla} < \nabla_{ad}$ then some part of the star must have $\nabla < \nabla_{ad}$, that is, a portion of the star must be radiative. Since models in the vicinity of the Hayashi line still have cool outer layers with high opacity, the radiative part must lie in the deep interior. Therefore stars located (somewhat) to the *left* of the Hayashi line have radiative cores surrounded by convective envelopes (if they are far to the left, they can of course be completely radiative).
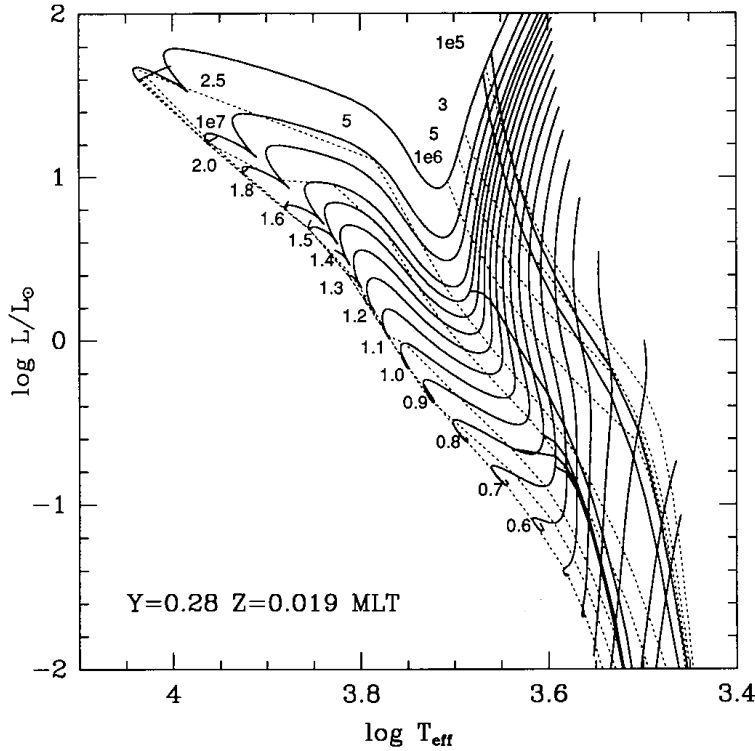
Figure 9.4. Pre-main-sequence evolution tracks for $0.3 - 2.5\,M_\odot$, according to the calculations of D'Antona & Mazzitelli (1994). The dotted lines are isochrones, connecting points on the tracks with the same age (between $t = 10^5$ yrs and 10^7 yrs, as indicated). Also indicated as solid lines that cross the tracks are the approximate locations of deuterium burning (between the upper two lines, near the $t \sim 10^5$ yr isochrone) and lithium burning (crossing the tracks at lower luminosity, at $t > 10^6$ yr).

On the other hand, if $\bar{\nabla} > \nabla_{ad}$ then a significant part of the star must have a *superadiabatic* temperature gradient (that is to say, apart from the outermost layers which are always superadiabatic). According to the analysis of Sect. 5.5.2, a significantly positive $\nabla - \nabla_{ad}$ will give rise to a very large convective energy flux, far exceeding normal stellar luminosities. Such a large energy flux very rapidly (on a dynamical timescale) transports heat outwards, thereby decreasing the temperature gradient in the superadiabatic region until $\nabla = \nabla_{ad}$ again. This restructuring of the star will quickly bring it back to the Hayashi line. Therefore the region to the right of the Hayashi line, with $T_{eff} < T_{eff,HL}$, is a *forbidden region* for any star in hydrostatic equilibrium.

9.1.2 Pre-main-sequence contraction

As a newly formed star emerges from the dynamical collapse phase it settles on the Hayashi line appropriate for its mass, with a radius roughly given by eq. (9.4). From this moment on we speak of the *pre-main sequence* phase of evolution. The pre-main sequence (PMS) star radiates at a luminosity determined by its radius on the Hayashi line. Since it is still too cool for nuclear burning, the energy source for its luminosity is gravitational contraction. As dictated by the virial theorem, this leads to an increase of its internal temperature. As long as the opacity remains high and the PMS star remains fully convective, it contracts along its Hayashi line and thus its luminosity decreases. Since fully convective stars are accurately described by $n = 1.5$ polytropes, this phase of contraction is indeed homologous to a very high degree! Thus the central temperature increases as $T_c \propto \rho_c^{1/3} \propto 1/R$.

As the internal temperature rises the opacity (and thus ∇_{rad}) decreases, until at some point $\nabla_{rad} < \nabla_{ad}$ in the central parts of the star and a radiative core develops. The PMS star then moves to the left in the H-R diagram, evolving away from the Hayashi line towards higher T_{eff} (see Fig. 9.4). As it keeps on contracting the extent of its convective envelope decreases and its radiative core grows in mass. (This phase of contraction is no longer homologous, because the density distribution must adapt itself to the radiative structure.) The luminosity no longer decreases but increases somewhat.

129

Once the star is mainly radiative further contraction is again close to homologous. The luminosity is now related to the temperature gradient and mostly determined by the mass of the protostar (see Sect. 7.4.2). This explains why PMS stars of larger mass turn away from the Hayashi line at a higher luminosity than low-mass stars, and why their luminosity remains roughly constant afterward.

Contraction continues, as dictated by the virial theorem, until the central temperature becomes high enough for nuclear fusion reactions. Once the energy generated by hydrogen fusion compensates for the energy loss at the surface, the star stops contracting and settles on the *zero-age main sequence* (ZAMS) if its mass is above the hydrogen burning limit of $0.08\,M_\odot$ (see Chapter 8). Since the nuclear energy source is much more concentrated towards the centre than the gravitational energy released by overall contraction, the transition from contraction to hydrogen burning again requires a (non-homologous) rearrangement of the internal structure.

Before thermal equilibrium on the ZAMS is reached, however, several nuclear reactions have already set in. In particular, a small quantity of *deuterium* is present in the interstellar gas out of which stars form, with a mass fraction $\sim 10^{-5}$. Deuterium is a very fragile nucleus that reacts easily with normal hydrogen ($^2\text{H} + {}^1\text{H} \rightarrow {}^3\text{He} + \gamma$, the second reaction in the pp chain). This reaction destroys all deuterium present in the star when $T \approx 1.0 \times 10^6\,\text{K}$, while the protostar is still on the Hayashi line. The energy produced ($5.5\,\text{MeV}$ per reaction) is large enough to halt the contraction of the PMS star for a few times $10^5\,\text{yr}$. (A similar but much smaller effect happens somewhat later at higher T when the initially present lithium, with mass fraction $\lesssim 10^{-8}$, is depleted). Furthermore, the $^{12}\text{C}(p,\gamma)^{13}\text{N}$ reaction is already activated at a temperature below that of the full CNO-cycle, due to the relatively large initial ^{12}C abundance compared to the equilibrium CNO abundances. Thus almost all ^{12}C is converted into ^{14}N before the ZAMS is reached. The energy produced in this way also halts the contraction temporarily and gives rise to the wiggles in the evolution tracks just above the ZAMS location in Fig, 9.4. Note that this occurs even in low-mass stars, $\lesssim 1\,M_\odot$, even though the pp chain takes over the energy production on the main sequence in these stars once CN equilibrium is achieved (see Sect. 9.2).

Finally, the time taken for a protostar to reach the ZAMS depends on its mass. This time is basically the Kelvin-Helmholtz contraction timescale (eq. 2.36). Since contraction is slowest when both R and L are small, the pre-main sequence lifetime is dominated by the final stages of contraction, when the star is already close to the ZAMS. We can therefore estimate the PMS lifetime by putting ZAMS values into eq. (2.36) which yields $\tau_{\text{PMS}} \approx 10^7 (M/M_\odot)^{-2.5}\,\text{yr}$. Thus massive protostars reach the ZAMS much earlier than lower-mass stars (and the term 'zero-age' main sequence is somewhat misleading in this context, although it hardly makes a difference to the total lifetime of a star). Indeed in young star clusters (e.g. the Pleiades) only the massive stars have reached the main sequence while low-mass stars still lie above and to the right of it.

9.2 The zero-age main sequence

Stars on the zero-age main sequence are (nearly) homogeneous in composition and are in complete (hydrostatic and thermal) equilibrium. Detailed models of ZAMS stars can be computed by solving the four differential equations for stellar structure numerically. It is instructive to compare the properties of such models to the simple main-sequence homology relations derived in Sect. 7.4.

From the homology relations we expect a homogeneous, radiative star in hydrostatic and thermal equilibrium with constant opacity and an ideal-gas equation of state to follow a mass-luminosity and mass-radius relation (7.32 and 7.36),

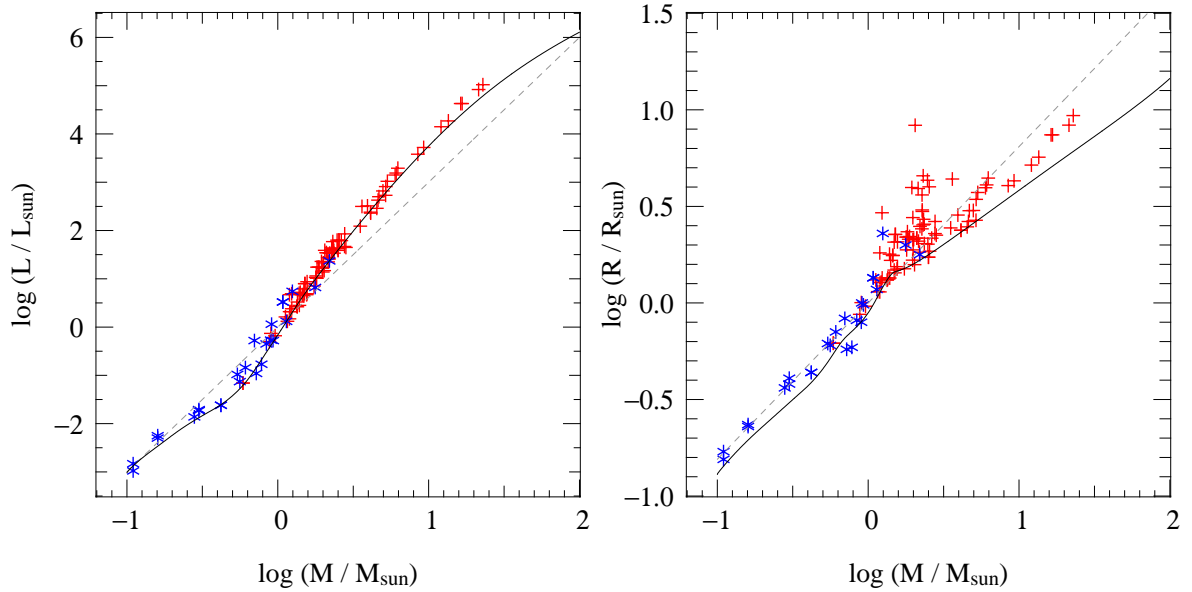$$L \propto \mu^4 M^3, \qquad R \propto \mu^{\frac{\nu-4}{\nu+3}} M^{\frac{\nu-1}{\nu+3}}.$$

Figure 9.5. ZAMS mass-luminosity (left) and mass-radius (right) relations from detailed structure models with $X = 0.7, Z = 0.02$ (solid lines) and from homology relations scaled to solar values (dashed lines). For the radius homology relation, a value $\nu = 18$ appropriate for the CNO cycle was assumed (giving $R \propto M^{0.81}$); this does not apply to $M < 1 \, M_\odot$ so the lower part should be disregarded. Symbols indicate components of double-lined eclipsing binaries with accurately measured M, R and L, most of which are MS stars.

These relations are shown as dashed lines in Fig. 9.5, where they are compared to observed stars with accurately measured M, L and R (see Chapter 1) and to detailed ZAMS models. The mass-radius homology relation depends on the temperature sensitivity (ν) of the energy generation rate, and is thus expected to be different for stars in which the pp chain dominates ($\nu \approx 4$, $R \propto M^{0.43}$) and stars dominated by the CNO cycle ($\nu \approx 18$, $R \propto \mu^{0.67} M^{0.81}$, as was assumed in Fig. 9.5).

Homology predicts the qualitative behaviour rather well, that is, a steep L-M relation and a much shallower R-M relation. However, it is not quantitatively accurate and it cannot account for the changes in slope ($d \log L / d \log M$ and $d \log R / d \log M$) of the relations. This was not to be expected, given the simplifying assumptions made in deriving the homology relations. The slope of the L-M relation is shallower than the homology value of 3 for masses below $1 \, M_\odot$, because such stars have large convective envelopes (as illustrated in Sect. 5.5; see also Sect. 9.2.2 below). The slope is significantly steeper than 3 for masses between 1 and $10 \, M_\odot$: in these stars the main opacity source is free-free and bound-free absorption, which increases outward rather than being constant through the star. In very massive stars, radiation pressure is important which results in flattening the L-M relation. The reasons for the changes in $d \log R / d \log M$ are similar. Note that for low masses we should have used the homology relation for the pp chain (for reasons explained in Sect. 9.2.1 below), which has a *smaller* slope – the opposite of what is seen in the detailed ZAMS models. The occurrence of convective regions (see Sect. 9.2.2) is the main reason for this non-homologous behaviour.

The detailed ZAMS models do reproduce the observed stellar luminosities quite well. The models trace the lower boundary of observed luminosities, consistent with the expected increase of L with time during the main sequence phase (see Sect. 9.3). The same can be said for the radii (right panel of Fig. 9.5), although the scatter in observed radii appears much larger. Partly this is due to the much finer scale of the ordinate in this diagram compared to the luminosity plot. The fact that most of the observed stellar radii are larger than the detailed ZAMS models is explained by expansion during (and after) the main sequence (see Sect. 9.3).
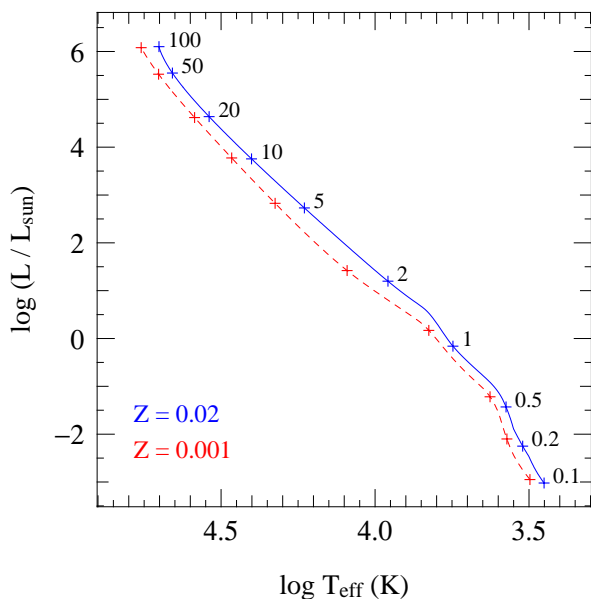
131

Figure 9.6. The location of the zero-age main sequence in the Hertzsprung-Russell diagram for homogeneous, detailed stellar models with $X = 0.7, Z = 0.02$ (blue solid line) and with $X = 0.757, Z = 0.001$ (red dashed line). Plus symbols indicate models for specific masses (in units of $M_\odot$). ZAMS models for metal-poor stars are hotter and have smaller radii. Relatively low-mass stars at low metallicity are also more luminous than their metal-rich counterparts.

The location of the detailed ZAMS models in the H-R diagram is shown in Fig. 9.6. The solid (blue) line depicts models for quasi-solar composition, which were also used in Fig. 9.5. The increase of effective temperature with stellar mass (and luminosity) reflects the steep mass-luminosity relation and the much shallower mass-radius relation – more luminous stars with similar radii must be hotter, by eq. (1.1). The slope of the ZAMS in the HRD is not constant, reflecting non-homologous changes in structure as the stellar mass increases.

The effect of *composition* on the location of the ZAMS is illustrated by the dashed (red) line, which is computed for a metal-poor mixture characteristic of Population II stars. Metal-poor main sequence stars are hotter and have smaller radii. Furthermore, relatively low-mass stars are also more luminous than their metal-rich counterparts. One reason for these differences is a lower bound-free opacity at lower Z (eq. 5.33), which affects relatively low-mass stars (up to about $5\,M_\odot$). On the other hand, higher-mass stars are dominated by electron-scattering opacity, which is independent of metallicity. These stars are smaller and hotter for a different reason (see Sect. 9.2.1).

9.2.1 Central conditions

We can estimate how the central temperature and central density scale with mass and composition for a ZAMS star from the homology relations for homogeneous, radiative stars in thermal equilibrium (Sec. 7.4.2, see eqs. 7.37 and 7.38 and Table 7.1). From these relations we may expect the central temperature to increase with mass, the mass dependence being larger for the pp chain ($T_c \propto M^{0.57}$) than for the CNO cycle ($T_c \propto M^{0.21}$). Since the CNO cycle dominates at high T, we can expect low-mass stars to power themselves by the pp chain and high-mass stars by the CNO cycle. This is confirmed by detailed ZAMS models, as shown in Fig. 9.7. For solar composition, the transition occurs at $T \approx 1.7 \times 10^7$ K, corresponding to $M \approx 1.3\,M_\odot$. Similarly, from the homology relations, the central density is expected to decrease strongly with mass in stars dominated by the CNO cycle ($\rho_c \propto M^{-1.4}$), but much less so in pp-dominated low-mass stars ($\rho_c \propto M^{-0.3}$). Also this is borne out by the detailed models in Fig. 9.7; in fact the central density increases slightly with mass between 0.4 and $1.5\,M_\odot$. The abrupt change in slope at $0.4\,M_\odot$ is related to the fact that stars with $M \lesssim 0.4\,M_\odot$ are completely convective. For these lowest-mass stars one of the main assumptions made in the homology relations (radiative equilibrium) breaks down.
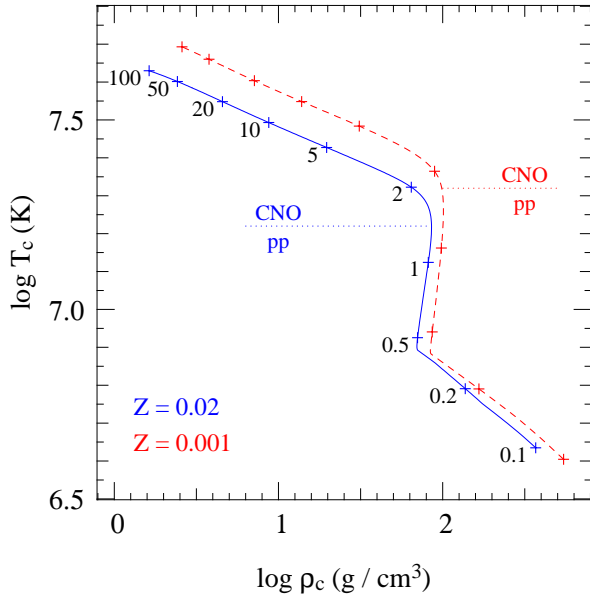
Figure 9.7. Central temperature versus central density for detailed ZAMS models with $X = 0.7, Z = 0.02$ (blue solid line) and with $X = 0.757, Z = 0.001$ (red dashed line). Plus symbols indicate models for specific masses (in units of $M_\odot$). The dotted lines indicate the approximate temperature border between energy production dominated by the CNO cycle and the pp chain. This gives rise to a change in slope of the T_c, ρ_c relation.

The energy generation rate of the CNO cycle depends on the total CNO abundance. At lower metallicity, the transition between pp chain and CNO cycle therefore occurs at a higher temperature. As a consequence, the mass at which the transition occurs is also larger. Furthermore, high-mass stars powered by the CNO cycle need a higher central temperature to provide the same total nuclear power. Indeed, comparing metal-rich and metal-poor stars in Figs. 9.6 and 9.7, the luminosity of two stars with the same mass is similar, but their central temperature is higher. As a consequence of the virial theorem (eq. 2.29 or 7.28), their radius must be correspondingly smaller.

9.2.2 Convective regions

An overview of the occurrence of convective regions on the ZAMS as a function of stellar mass is shown in Fig. 9.8. For any given mass M, a vertical line in this diagram shows which conditions are encountered as a function of depth, characterized by the fractional mass coordinate m/M. Gray shading indicates whether a particular mass shell is convective (gray) or radiative (white). We can thus distinguish three types of ZAMS star:

- completely convective, for $M < 0.35\,M_\odot$,

- radiative core + convective envelope, for $0.35\,M_\odot < M < 1.2\,M_\odot$,

- convective core + radiative envelope, for $M > 1.2\,M_\odot$.

This behaviour can be understood from the Schwarzschild criterion for convection, which tells us that convection occurs when $\nabla_{\rm rad} > \nabla_{\rm ad}$ (eq. 5.50). As discussed in Sec. 5.5.1, a large value of $\nabla_{\rm rad}$ is found when the opacity κ is large, or when the energy flux to be transported (in particular the value of l/m) is large, or both. Starting with the latter condition, this is the case when a lot of energy is produced in a core of relatively small mass, i.e. when the energy generation rate $\epsilon_{\rm nuc}$ is strongly peaked towards the centre. This is certainly the case when the CNO-cycle dominates the energy production, since it is very temperature sensitive ($\nu \approx 18$) which means that $\epsilon_{\rm nuc}$ rapidly drops as the temperature decreases from the centre outwards. It results in a steep increase of $\nabla_{\rm rad}$ towards the centre and thus to a convective core. This is illustrated for a $4\,M_\odot$ ZAMS star in Fig. 5.4. The size of the convective core increases with stellar mass (Fig. 9.8), and it can encompass up to 80% of the mass
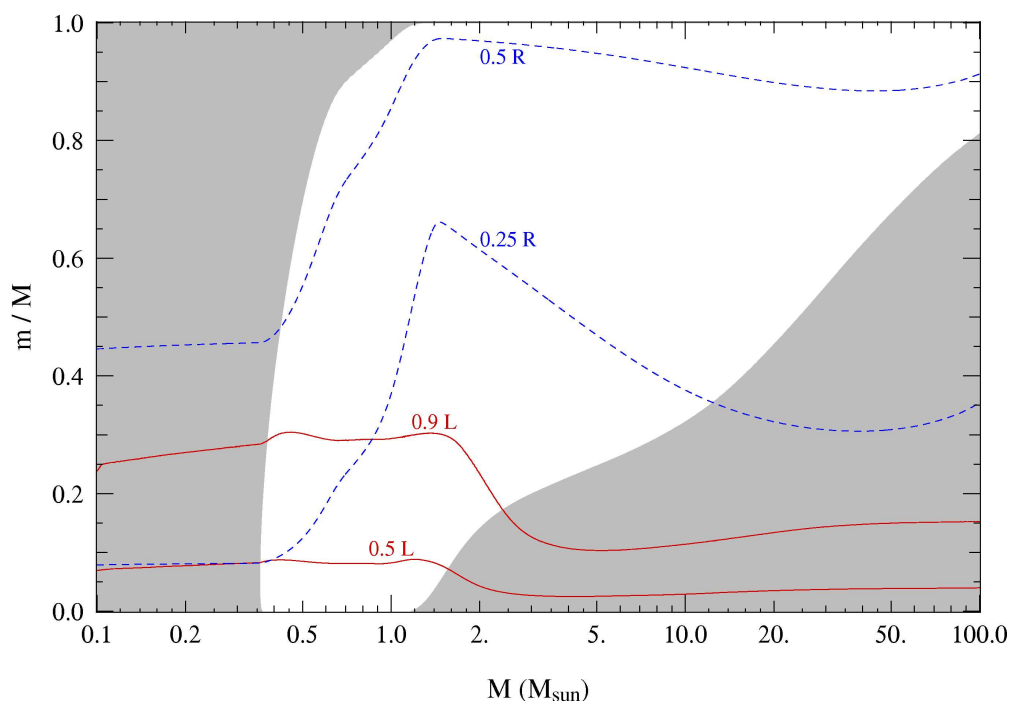
133

Figure 9.8. Occurrence of convective regions (gray shading) on the ZAMS in terms of fractional mass coordinate m/M as a function of stellar mass, for detailed stellar models with a composition $X = 0.70$, $Z = 0.02$. The solid (red) lines show the mass shells inside which 50% and 90% of the total luminosity are produced. The dashed (blue) lines show the mass coordinate where the radius r is 25% and 50% of the stellar radius R. (After KIPPENHAHN & WEIGERT.)

of the star when M approaches $100\,M_\odot$. This is mainly related with the fact that at high mass, ∇_{ad} is depressed below the ideal-gas value of 0.4 because of the growing importance of radiation pressure. At $100\,M_\odot$ radiation pressure dominates and $\nabla_{ad} \approx 0.25$.

In low-mass stars the pp-chain dominates, which has a much smaller temperature sensitivity. Energy production is then distributed over a larger area, which keeps the energy flux and thus ∇_{rad} low in the centre and the core remains radiative (see the $1\,M_\odot$ model in Fig. 5.4). The transition towards a more concentrated energy production at $M > 1.2\,M_\odot$ is demonstrated in Fig. 9.8 by the solid lines showing the location of the mass shell inside which most of the luminosity is generated.

Convective envelopes can be expected to occur in stars with low effective temperature, as discussed in Sec. 7.2.3. This is intimately related with the rise in opacity with decreasing temperature in the envelope. In the outer envelope of a $1\,M_\odot$ star for example, κ can reach values of $10^5\ \mathrm{cm}^2/\mathrm{g}$ which results in enormous values of ∇_{rad} (see Fig. 5.4). Thus the Schwarzschild criterion predicts a convective outer envelope. This sets in for masses less than $\approx 1.5\,M_\odot$, although the amount of mass contained in the convective envelope is very small for masses between 1.2 and $1.5\,M_\odot$. Consistent with the discussion in Sec. 7.2.3, the depth of the convective envelope increases with decreasing T_{eff} and thus with decreasing M, until for $M < 0.35\,M_\odot$ the entire star is convective. Thus these very low-mass stars lie on their respective Hayashi lines.

9.3 Evolution during central hydrogen burning

Fig. 9.9 shows the location of the ZAMS in the H-R diagram and various evolution tracks for different masses at Population I composition, covering the central hydrogen burning phase. Stars evolve away

from the ZAMS towards higher luminosities and larger radii. Low-mass stars ($M \lesssim 1\,M_\odot$) evolve towards higher T_{eff}, and their radius increase is modest. Higher-mass stars, on the other hand, evolve towards lower T_{eff} and strongly increase in radius (by a factor 2 to 3). Evolved main-sequence stars are therefore expected to lie above and to the right of the ZAMS. This is indeed confirmed by comparing the evolution tracks to observed stars with accurately determined parameters.

As long as stars are powered by central hydrogen burning they remain in hydrostatic and thermal equilibrium. Since their structure is completely determined by the four (time-independent) structure equations, the evolution seen in the HRD is due to the changing composition inside the star (i.e. due to chemical evolution of the interior). How can we understand these changes?

Nuclear reactions on the MS have two important effects on the structure:

- Hydrogen is converted into helium, therefore the mean molecular weight μ increases in the core of the star (by more than a factor two from the initial H-He mixture to a pure He core by the end of central hydrogen burning). The increase in luminosity can therefore be understood from the homology relation $L \propto \mu^4 M^3$. It turns out that the μ^4 dependence of this relation describes the luminosity increase during the MS quite well, if μ is taken as the mass-averaged value over the whole star.

- The nuclear energy generation rate ϵ_{nuc} is very sensitive to the temperature. Therefore nuclear reactions act like a *thermostat* on the central regions, keeping the central temperature almost constant. Since approximately $\epsilon_{\text{pp}} \propto T^4$ and $\epsilon_{\text{CNO}} \propto T^{18}$, the CNO cycle is a better thermostat than the pp chain. Since the luminosity increases and at the same time the hydrogen abundance decreases during central H-burning, the central temperature must increase somewhat to keep up the energy production, but the required increase in T_c is very small.

Since μ increases while $T_c \approx$ constant, the ideal-gas law implies that $P_c/\rho_c \propto T_c/\mu$ must decrease. This means that either the central density must increase, or the central pressure must decrease. The latter possibility means that the layers surrounding the core must expand, as explained below. In
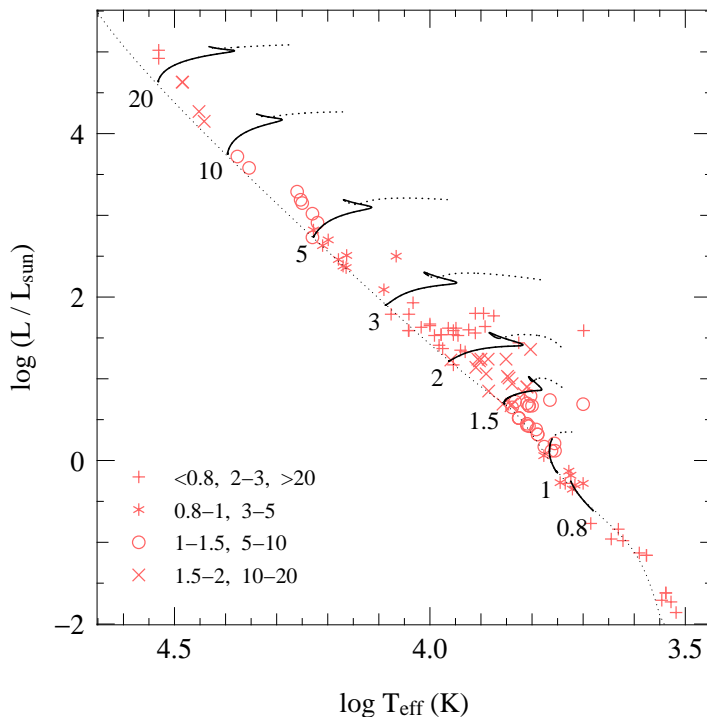

Figure 9.9. Evolution tracks in the H-R diagram during central hydrogen burning for stars of various masses, as labelled (in $M_\odot$), and for a composition $X = 0.7, Z = 0.02$. The dotted portion of each track shows the continuation of the evolution after central hydrogen exhaustion; the evolution of the $0.8\,M_\odot$ star is terminated at an age of 14 Gyr. The thin dotted line in the ZAMS. Symbols show the location of binary components with accurately measured mass, luminosity and radius (as in Fig. 9.5). Each symbol corresponds to a range of measured masses, as indicated in the lower left corner (mass values in $M_\odot$).

either case, the density contrast between the core and the envelope increases, so that evolution during central H-burning causes *non-homologous* changes to the structure.

9.3.1 Evolution of stars powered by the CNO cycle

We can understand why rather massive stars ($M \gtrsim 1.3 \, M_\odot$) expand during the MS by considering the pressure that the outer layers exert on the core:

$$P_{\text{env}} = \int_{m_c}^{M} \frac{Gm}{4\pi r^4} \, dm \tag{9.14}$$

Expansion of the envelope (increase in r of all mass shells) means a decrease in the envelope pressure on the core. This decrease in pressure is needed because of the sensitive thermostatic action of the CNO cycle, $\epsilon_{\text{CNO}} \propto \rho T^{18}$, which allows only very small increases in T_c and ρ_c. Since μ_c increases as H being is burned into He, the ideal-gas law dictates that P_c must decrease. This is only possible if P_{env} decreases, i.e. the outer layers must expand to keep the star in HE ($\rho_{\text{env}} \downarrow$ and $R \uparrow$). This self-regulating envelope expansion mechanism is the only way for the star to adapt itself to the composition changes in the core while maintaining both HE and TE.

Another important consequence of the temperature sensitivity the CNO cycle is the large concentration of ϵ_{nuc} towards the centre. This gives rise to a large central $\nabla_{\text{rad}} \propto l/m$ and hence to *convective cores*, which are mixed homogeneously ($X(m)$ = constant within the convective core mass M_{cc}). This increases the amount of fuel available and therefore the lifetime of central hydrogen burning (see Fig. 9.10). In general M_{cc} decreases during the evolution, which is a consequence of the fact that $\nabla_{\text{rad}} \propto \kappa$ and since $\kappa \propto 1 + X$ for the main opacity sources (see Sect. 5.3) the opacity in the core decreases as the He abundance goes up.

Towards the end of the main sequence phase, as X_c becomes very small, the thermostatic action of the CNO reactions diminishes and T_c has to increase substantially to keep up the energy production. When hydrogen is finally exhausted, this occurs within the whole convective core of mass M_{cc} and ϵ_{nuc} decreases. The star now loses more energy at its surface than is produced in the centre, it gets out of thermal equilibrium and it will undergo an overall contraction. This occurs at the red point of the evolution tracks in Fig. 9.9, after which T_{eff} increases. At the blue point of the hook feature in the HRD, the core has contracted and heated up sufficiently that at the edge of the former convective core the temperature is high enough for the CNO cycle to ignite again in a shell around the helium core. This is the start of the *hydrogen-shell burning* phase which will be discussed in Chapter 10.

9.3.2 Evolution of stars powered by the pp chain

In stars with $M \lesssim 1.3 \, M_\odot$ the central temperature is too low for the CNO cycle and the main energy-producing reactions are those of the pp chain. The lower temperature sensitivity $\epsilon_{\text{pp}} \propto \rho T^4$ means that T_c and ρ_c increase more than was the case for the CNO cycle. Therefore the outer layers need to expand less in order to maintain hydrostatic equilibrium in the core. As a result, the radius increase in low-mass stars is modest and they evolve almost parallel to the ZAMS in the H-R diagram (see Fig. 9.9).

Furthermore, the lower T-sensitivity of the pp chains means that low-mass stars have radiative cores. The rate of change of the hydrogen abundance in each shell is then proportional to the overall reaction rate of the pp chain (by eq. 6.41), and is therefore highest in the centre. Therefore a hydrogen abundance gradient builds up gradually, with $X(m)$ increasing outwards (see Fig. 9.10). As a result, hydrogen is depleted gradually in the core and there is a smooth transition to hydrogen-shell burning. The evolution tracks for low-mass stars therefore do not show a hook feature.
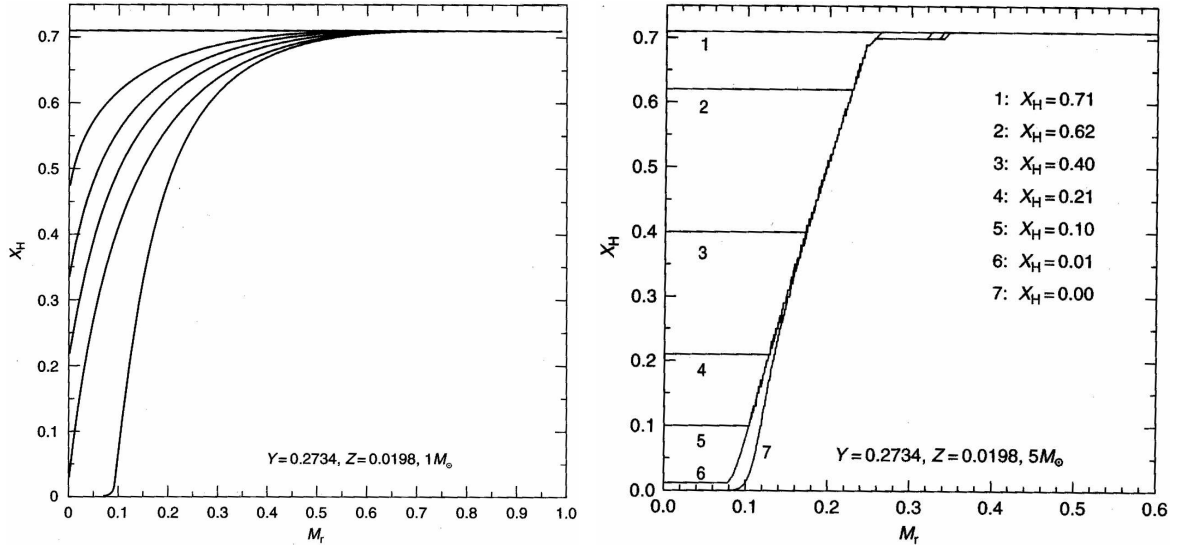
136

Figure 9.10. Hydrogen abundance profiles at different stages of evolution for a 1 $M_\odot$ star (left panel) and a 5 $M_\odot$ star (right panel) at quasi-solar composition. Figures reproduced from SALARIS & CASSISI.

Note that stars in the approximate mass range $1.1 - 1.3\ M_\odot$ (at solar metallicity) undergo a transition from the pp chain to the CNO cycle as their central temperature increases. Therefore these stars at first have radiative cores and later develop a growing convective core. At the end of the MS phase such stars also show a hook feature in the HRD.

9.3.3 The main sequence lifetime

The timescale τ_{MS} that a star spends on the main sequence is essentially the nuclear timescale for hydrogen burning, given by eq. (2.37). Another way of deriving essentially the same result is by realizing that, in the case of hydrogen burning, the rate of change of the hydrogen abundance X is related to the energy generation rate ϵ_{nuc} by eq. (6.43),

$$\frac{dX}{dt} = -\frac{\epsilon_{nuc}}{q_H}. \tag{9.15}$$

Here $q_H = Q_H/4m_u$ is the effective energy release per unit mass of the reaction chain ($4\,^1H \rightarrow {}^4He + 2\,e^+ + 2\,\nu$), corrected for the neutrino losses. Hence q_H is somewhat different for the pp chain and the CNO cycle. Note that q_H/c^2 corresponds to the factor ϕ used in eq. (2.37). If we integrate eq. (9.15) over all mass shells we obtain, for a star in thermal equilibrium,

$$\frac{dM_H}{dt} = -\frac{L}{q_H}, \tag{9.16}$$

where M_H is the total mass of hydrogen in the star. Note that while eq. (9.15) only strictly applies to regions where there is no mixing, eq (9.16) is also valid if the star has a convective core, because convective mixing only redistributes the hydrogen supply. If we now integrate over the main sequence lifetime we obtain for the total mass of hydrogen consumed

$$\Delta M_H = \frac{1}{q_H} \int_0^{\tau_{MS}} L\,dt = \frac{\langle L \rangle\, \tau_{MS}}{q_H}, \tag{9.17}$$

where $\langle L \rangle$ is the time average of the luminosity over the main-sequence lifetime. We can write $\Delta M_{\rm H} = f_{\rm nuc} M$ by analogy with eq. (2.37), and write $f_{\rm nuc}$ as the product of the initial hydrogen mass fraction X_0 and an effective core mass fraction q_c inside which all hydrogen is consumed, so that

$$\tau_{\rm MS} = X_0\, q_{\rm H}\, \frac{q_c M}{\langle L \rangle}. \tag{9.18}$$

We have seen that the luminosity of main-sequence stars increases strongly with mass. Since the variation of L during the MS phase is modest, we can assume the same relation between $\langle L \rangle$ and M as for the ZAMS. The other factors appearing in eq. (9.18) do not or only weakly depend on the mass of the star (see below) and can in a first approximation be taken as constant. For a mass-luminosity relation $\langle L \rangle \propto M^\eta$ – where η depends on the mass range under consideration with $\eta \approx 3.8$ on average – we thus obtain $\tau_{\rm MS} \propto M^{1-\eta}$. Hence $\tau_{\rm MS}$ decreases strongly towards larger masses.

This general trend has important consequences for the observed H-R diagrams of star clusters. All stars in a cluster can be assumed to have formed at approximately the same time and therefore now have the same age $\tau_{\rm cl}$. Cluster stars with a mass above a certain limit $M_{\rm to}$ have main-sequence lifetimes $\tau_{\rm MS} < \tau_{\rm cl}$ and have therefore already left the main sequence, while those with $M < M_{\rm to}$ are still on the main sequence. The main sequence of a cluster has an upper end (the 'turn-off point') at a luminosity and effective temperature corresponding to $M_{\rm to}$, the so-called *turn-off mass*, determined by the condition $\tau_{\rm MS}(M_{\rm to}) = \tau_{\rm cl}$. The turn-off mass and luminosity decrease with cluster age (e.g. see Fig. 1.2). This the basis for the *age determination* of star clusters.

The actual main-sequence lifetime depend on a number of other factors. The effective energy release per gram $q_{\rm H}$ depends on which reactions are involved in energy production and therefore has a slight mass dependence. More importantly, the exact value of q_c is determined by the hydrogen profile left at the end of the main sequence. This is somewhat mass-dependent, especially for massive stars in which the relative size of the convective core tends to increase with mass (Fig. 9.8). A larger convective core mass means a larger fuel reservoir and a longer lifetime. Our poor understanding of convection and mixing in stars unfortunately introduces considerable uncertainty in the size of this reservoir and therefore both in the main-sequence lifetime of a star of a particular mass and in its further evolution.

9.3.4 Complications: convective overshooting and semi-convection

As discussed in Sect. 5.5.4, the size of a convective region inside a star is expected to be larger than predicted by the Schwarzschild (or Ledoux) criterion because of convective *overshooting*. However, the extent $d_{\rm ov}$ of the overshooting region is not known reliably from theory. In stellar evolution calculations this is usually parameterized in terms of the local pressure scale height, $d_{\rm ov} = \alpha_{\rm ov} H_P$. In addition, other physical effects such as stellar rotation may contribute to mixing material beyond the formal convective core boundary. Detailed stellar evolution models in which the effects of convective overshooting are taken into account generally provide a better match to observations. For this reason, overshooting (or perhaps a variety of enhanced mixing processes) is thought to have a significant effect in stars with sizable convective cores on the main sequence.

Overshooting has several important consequences for the evolution of a star:

1. a longer main-sequence lifetime, because of the larger hydrogen reservoir available;

2. a larger increase in luminosity and radius during the main sequence, because of the larger region inside which μ increases which enhances the effects on L and R discussed earlier in this section;

3. the hydrogen-exhausted core mass is larger at the end of the main sequence, which in turn leads to (a) larger luminosities during all evolution phases after the main sequence (e.g. see Fig. 10.2 in the next chapter) and, as a result, (b) *shorter* lifetimes of these post-main sequence phases.
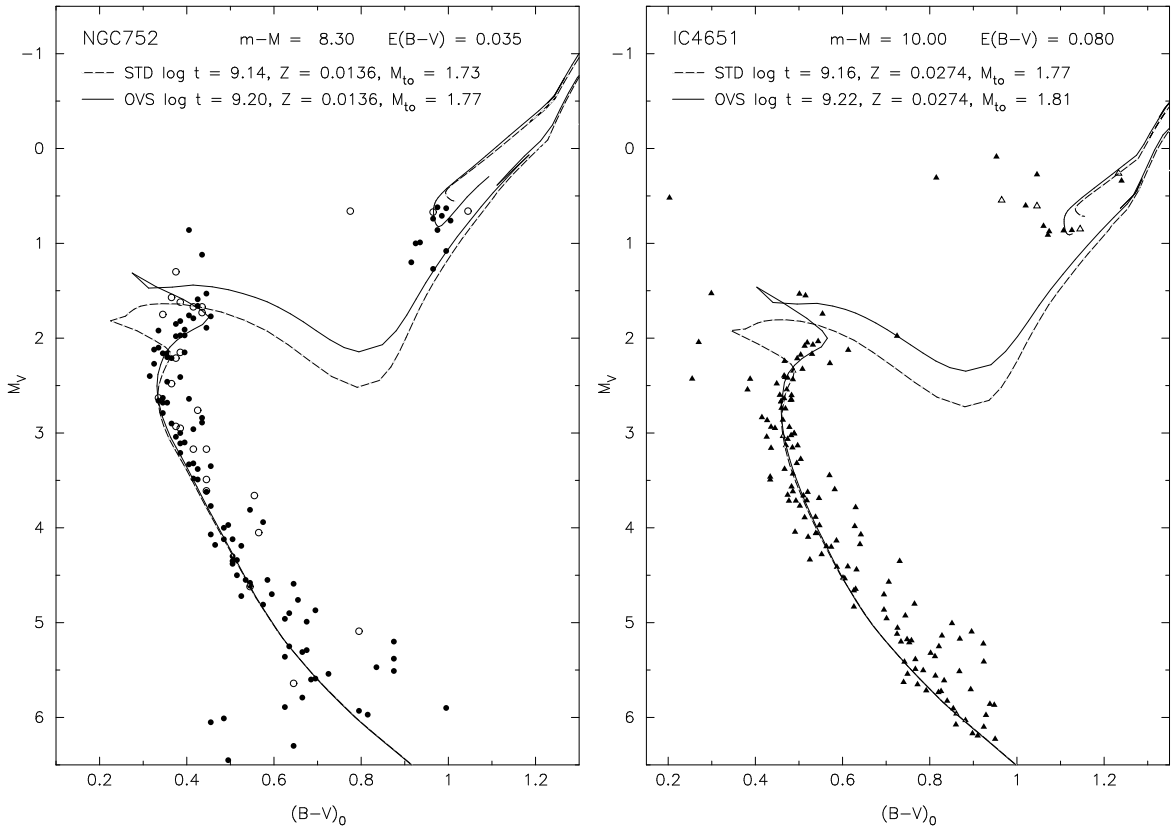
Figure 9.11. Two examples of *isochrone fitting* to the colour-magnitude diagrams of open clusters, NGC 752 and IC 4651. The distribution of stars in the turn-off region is matched to isochrones for standard stellar evolution models (STD) and for models with convective overshooting (OVS). The overshooting models are better able to reproduce the upper extension of the main sequence band in both cases.

Some of these effects, particularly (2) and (3a), provide the basis of observational tests of overshooting. Stellar evolution models computed with different values of α_{ov} are compared to the observed width of the main sequence band in star clusters (see for example Fig. 9.11), and to the luminosities of evolved stars in binary systems. If the location in the HRD of the main sequence turn-off in a cluster is well determined, or if the luminosity difference between binary components can be accurately measured, a quantitative test is possible which allows a calibration of the parameter α_{ov}. Such tests indicate that $\alpha_{ov} \approx 0.25$ is appropriate in the mass range $1.5 - 8\,M_\odot$. For larger masses, however, α_{ov} is poorly constrained.

Another phenomenon that introduces an uncertainty in stellar evolution models is related to the difference between the Ledoux and Schwarzschild criterion for convection (see Sect. 5.5.1). Outside the convective core a composition gradient (∇_μ) develops, which can make this region dynamically stable according to the Ledoux criterion while it would have been convective if the Schwarzschild criterion were applied. In such a region an over-stable oscillation pattern can develop on the thermal timescale, which slowly mixes the region and thereby smooths out the composition gradient. This process is called *semi-convection*. Its efficiency and the precise outcome are uncertain. Semi-convective situations are encountered during various phases of evolution, most importantly during central hydrogen burning in stars with $M > 10\,M_\odot$ and during helium burning in low- and intermediate-mass stars.

139

Suggestions for further reading

The process of star formation and pre-main sequence evolution is treated in much more detail in Chapters 18–20 of MAEDER, while the properties and evolution on the main sequence are treated in Chapter 25. See also KIPPENHAHN & WEIGERT Chapters 22 and 26–30.

Exercises

9.1 Kippenhahn diagram of the ZAMS

Figure 9.8 indicates which regions in zero-age main sequence stars are convective as a function of the mass of the star.

(a) Why are the lowest-mass stars fully convective? Why does the mass of the convective envelope decrease with M and disappear for $M \gtrsim 1.3\,M_\odot$?

(b) What changes occur in the central energy production around $M = 1.3\,M_\odot$, and why? How is this related to the convection criterion? So why do stars with $M \approx 1.3\,M_\odot$ have convective cores while lower-mass stars do not?

(c) Why is it plausible that the mass of the convective core increases with M?

9.2 Conceptual questions

(a) What is the Hayashi line? Why is it a line, in other words: why is there a whole range of possible luminosities for a star of a certain mass on the HL?

(b) Why do no stars exist with a temperature cooler than that of the HL? What happens if a star would cross over to the cool side of the HL?

(c) Why is there a mass-luminosity relation for ZAMS stars? (In other words, why is there a unique luminosity for a star of a certain mass?)

(d) What determines the shape of the ZAMS is the HR diagram?

9.3 Central temperature versus mass

Use the homology relations for the luminosity and temperature of a star to derive how the central temperature in a star scales with mass, and find the dependence of T_c on M for the pp-chain and for the CNO-cycle. To make the result quantitative, use the fact that in the Sun with $T_c \approx 1.3 \times 10^7$ K the pp-chain dominates, and that the CNO-cycle dominates for masses $M \gtrsim 1.3\,M_\odot$. (Why does the pp-chain dominate at low mass and the CNO-cycle at high mass?)

9.4 Mass-luminosity relation

Find the relation between L and M and the slope of the main sequence, assuming an opacity law $\kappa = \kappa_0 \rho T^{-7/2}$ (the Kramers opacity law) and that the energy generation rate per unit mass $\epsilon_{nuc} \propto \rho T^\nu$, where $\nu = 4$.

Chapter 10

Post-main sequence evolution through helium burning

After the main-sequence phase, stars are left with a hydrogen-exhausted core surrounded by a still hydrogen-rich envelope. To describe the evolution after the main sequence, it is useful to make a division based on the mass:

low-mass stars are those that develop a degenerate helium core after the main sequence, leading to a relatively long-lived *red giant branch* phase. The ignition of He is unstable and occurs in a so-called *helium flash*. This occurs for masses between $0.8\,M_\odot$ and $\approx 2\,M_\odot$ (this upper limit is sometimes denoted as M_{HeF}).

intermediate-mass stars develop a helium core that remains non-degenerate, and they ignite helium in a stable manner. After the central He burning phase they form a carbon-oxygen core that becomes degenerate. Intermediate-mass stars have masses between M_{HeF} and $M_{\mathrm{up}} \approx 8\,M_\odot$. Both low-mass and intermediate-mass stars shed their envelopes by a strong stellar wind at the end of their evolution and their remnants are CO white dwarfs.

massive stars have masses larger than $M_{\mathrm{up}} \approx 8\,M_\odot$ and ignite carbon in a non-degenerate core. Except for a small mass range ($\approx 8-11\,M_\odot$) these stars also ignite heavier elements in the core until an Fe core is formed which collapses.

In this chapter the evolution between the end of the main sequence and the development of a carbon-oxygen core is discussed. We concentrate on low-mass and intermediate-mass stars, but the principles are equally valid for massive stars. The evolution of massive stars in the H-R diagram is, however, also strongly affected by mass loss and we defer a more detailed discussion of massive stars until Chapter 12.

10.1 The Schönberg-Chandrasekhar limit

During central hydrogen burning on the main sequence, we have seen that stars are in thermal equilibrium ($\tau_{\mathrm{nuc}} \gg \tau_{\mathrm{KH}}$) with the surface luminosity balanced by the nuclear power generated in the centre. After the main sequence a hydrogen-exhausted core is formed inside which nuclear energy production has ceased. This inert helium core is surrounded by a hydrogen-burning shell and a H-rich envelope. For such an inert core to be in thermal equilibrium requires a zero net energy flow, $l(m) = \int_m \epsilon_{\mathrm{nuc}}\,dm = 0$ and hence $dT/dr \propto l = 0$. This implies that the core must be *isothermal* to remain in TE. Such a stable situation is possible only under certain circumstances.

A star composed of ideal gas at constant temperature corresponds to a polytrope with $\gamma = 1$, i.e. with $n \to \infty$. Such a polytrope would have infinite radius (Chapter 4) or, if its radius were finite, would have infinitely high central density, both of which are unphysical. In other words, *completely isothermal stars made of ideal gas cannot exist.* The reason is that the pressure gradient needed to support such a star against its own gravity is produced only by the density gradient, $dP/dr = (\mathcal{R}T/\mu)\,d\rho/dr$, with no help from a temperature gradient. Thus hydrostatic equilibrium in an isothermal star would require a very large density gradient.

It turns out, however, that if only the core of the star is isothermal, and the mass M_c of this isothermal core is only a small fraction of the total mass of the star, then a stable configuration is possible. If the core mass exceeds this limit, then the pressure within the isothermal core cannot sustain the weight of the overlying envelope. This was first discovered by Schönberg and Chandrasekhar in 1942, who computed the maximum core mass fraction $q_c = M_c/M$ to be

$$\frac{M_c}{M} < q_{SC} = 0.37 \left(\frac{\mu_{env}}{\mu_c} \right)^2 \approx 0.10 \tag{10.1}$$

where μ_c and μ_{env} are the mean molecular weight in the core and in the envelope respectively. This limit is known as the *Schönberg-Chandrasekhar limit*. The typical value $q_{SC} \approx 0.10$ is appropriate for a helium core with $\mu_c = 1.3$ and a H-rich envelope. (A simple, qualitative derivation of eq. 10.1 can be found in MAEDER Section 25.5.1.)
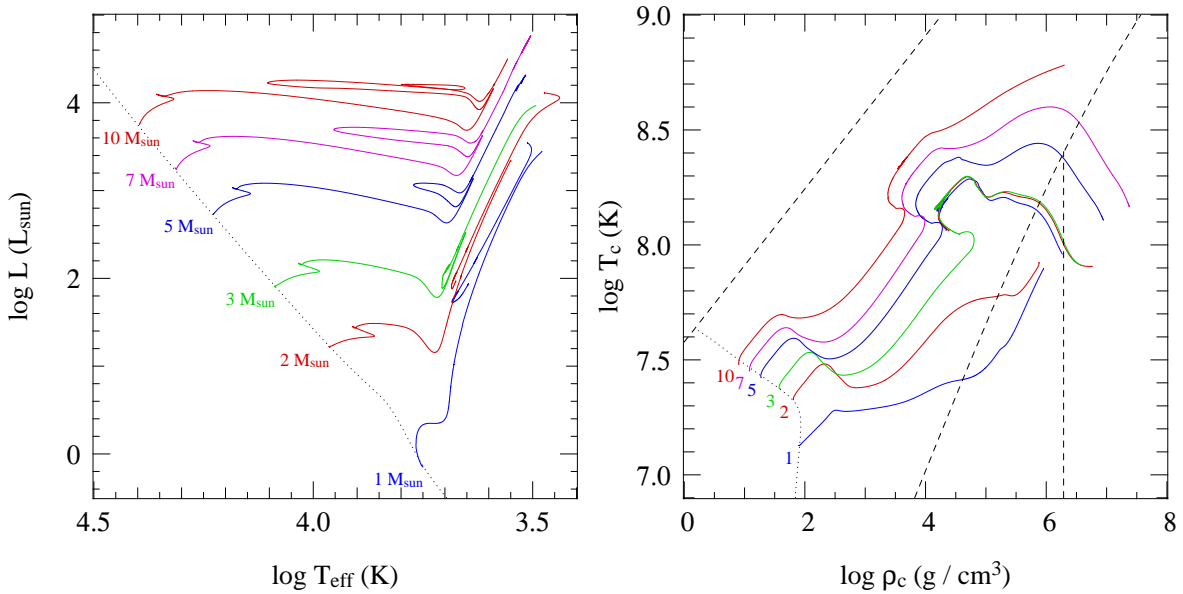

Figure 10.1. Evolution tracks for stars of quasi-solar composition ($X = 0.7$, $Z = 0.02$) and masses of 1, 2, 3, 5, 7 and 10 $M_\odot$ in the H-R diagram (left panel) and in the central temperature versus density plane (right panel). Dotted lines in both diagrams show the ZAMS, while the dashed lines in the right-hand diagram show the borderlines between equation-of-state regions (as in Fig. 3.4). The 1 $M_\odot$ model is characteristic of low-mass stars: the central core becomes degenerate soon after leaving the main sequence and helium is ignited in an unstable flash at the top of the red giant branch. When the degeneracy is eventually lifted, He burning becomes stable and the star moves to the *zero-age horizontal branch* in the HRD, at $\log L \approx 1.8$. The 2 $M_\odot$ model is a borderline case that just undergoes a He flash. The He flash itself is not computed in these models, hence a gap appears in the tracks. The 5 $M_\odot$ model is representative of intermediate-mass stars, undergoing quiet He ignition and He burning in a loop in the HRD. The appearance of the 7 and 10 $M_\odot$ models in the HRD is qualitatively similar. However, at the end of its evolution the 10 $M_\odot$ star undergoes carbon burning in the centre, while the cores of lower-mass stars become strongly degenerate. (Compare to Fig. 8.4.)

Stars that leave the main sequence with a helium core mass below the Schönberg-Chandrasekhar limit can therefore remain in complete equilibrium (HE and TE) during hydrogen-shell burning. This is the case for stars with masses up to about $8\,M_\odot$, if convective overshooting is neglected. Overshooting increases the core mass at the end of central H-burning, and therefore the upper mass limit for stars remaining in TE after the main sequence decreases to about $2\,M_\odot$ in calculations that include moderate overshooting.

When the mass of the H-exhausted core exceeds the Schönberg-Chandrasekhar limit – either immediately after the main sequence in relatively massive stars, or in lower-mass stars after a period of H-shell burning during which the helium core mass increases steadily – thermal equilibrium is no longer possible. The helium core then contracts and builds up a temperature gradient. This temperature gradient adds to the pressure gradient that is needed to balance gravity and keep the star in HE. However, the temperature gradient also causes an outward heat flow from the core, such that it keeps contracting and heating up in the process (by virtue of the virial theorem). This contraction occurs on the thermal (Kelvin-Helmholtz) timescale in a quasi-static way, always maintaining a state very close to HE.

Low-mass stars ($M \lesssim 2\,M_\odot$) have another way of maintaining both HE and TE during hydrogen-shell burning. In such stars the helium core is relatively dense and cool and electron degeneracy can become important in the core after the main sequence. Degeneracy pressure is independent of temperature and can support the weight of the envelope even in a relatively massive core, as long as the degenerate core mass does not exceed the Chandrasekhar mass.[1] In that case the Schönberg-Chandrasekhar limit no longer applies. Inside such degenerate helium cores efficient energy transport by *electron conduction* (Sec. 5.2.4) can keep the core almost isothermal.

Effects of core contraction: the 'mirror principle'

The following principle appears to be generally valid, and provides a way of interpreting the results of detailed numerical calculations:

> Whenever a star has an *active shell-burning source*, the burning shell acts as a *mirror* between the core and the envelope:
> | | | |
> |---|---|---|
> | core contraction | $\Rightarrow$ | envelope expansion |
> | core expansion | $\Rightarrow$ | envelope contraction |

This 'mirror principle' can be understood by the following argument. To maintain thermal equilibrium, the burning shell must remain at approximately constant temperature due to the thermostatic action of nuclear burning. Contraction of the burning shell would entail heating, so the burning shell must also remain at roughly constant radius. As the core contracts, ρ_{shell} must therefore decrease and hence also the pressure in the burning shell must decrease. Therefore the pressure P_{env} of the overlying envelope must decrease, so the layers above the shell must expand (an example of this behaviour can be seen in Fig. 10.4, to be discussed in the next section).

10.2 The hydrogen-shell burning phase

In this section we discuss in some detail the evolution of stars during hydrogen-shell burning, until the onset of helium burning. Based on the above section, qualitative differences are to be expected between low-mass stars ($M \lesssim 2\,M_\odot$) on the one hand and intermediate- and high-mass stars ($M \gtrsim$

---

[1]Note the very different physical meanings of the *Chandrasekhar mass* and the *Schönberg-Chandrasekhar limit*!
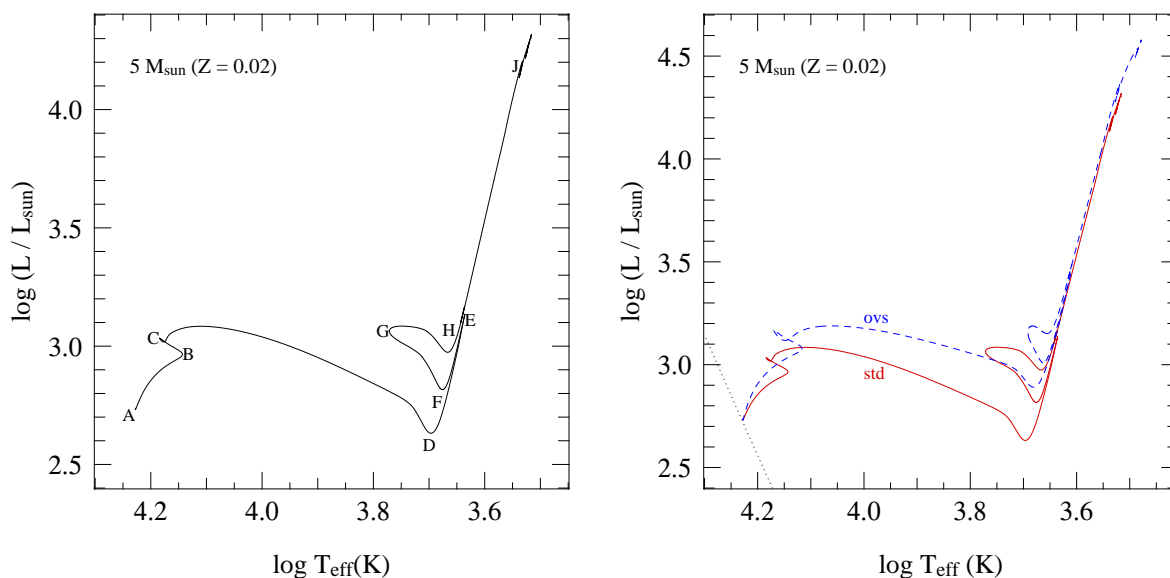
Figure 10.2. Evolution track in the Hertzsprung-Russell diagram of a 5 $M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. See text for details. The evolution track in the left panel was computed without convective overshooting. The right panel shows a comparison between this track and the evolution of the same star computed with moderate overshooting ($\alpha_{ov} = l_{ov}/H_P \approx 0.25$; dashed line), illustrating some of the effects discussed in Sec. 9.3.4.

2 $M_\odot$) on the other hand. Therefore we discuss these two cases separately, starting with the evolution of higher-mass stars because it is relatively simple compared to low-mass stars. We use two detailed stellar evolution sequences, for stars of 5 $M_\odot$ and 1 $M_\odot$ respectively, as examples for the general evolutionary behaviour of stars in these two mass ranges.

10.2.1 Hydrogen-shell burning in intermediate-mass and massive stars

Fig. 10.2 shows the evolution track of a 5 $M_\odot$ star of quasi-solar composition ($X = 0.7, Z = 0.02$) in the H-R diagram, and Fig. 10.3 shows some of the interior details of the evolution of this star as a function of time from the end of central hydrogen burning. Point B in both figures corresponds to the start of the overall contraction phase near the end of the main sequence (when the central H mass fraction $X_c \approx 0.03$) and point C corresponds to the exhaustion of hydrogen in the centre and the disappearance of the convective core. The hatched regions in the 'Kippenhahn diagram' (lower panel of Fig. 10.3) show the rapid transition at point C from hydrogen burning in the centre to hydrogen burning in a shell.

The H-exhausted core initially has a mass of about 0.4 $M_\odot$ which is below the Schönberg-Chandra-sekhar limit, so the star initially remains in TE and the first portion of the hydrogen-shell burning phase (C–D) is relatively slow, lasting about 2×10^6 yr. The temperature and density gradients between core and envelope are still shallow, so that the burning shell initially occupies a rather large region in mass. This phase is therefore referred to as *thick shell burning*. The helium core gradually grows in mass until it exceeds the S-C limit and the contraction of the core speeds up. The envelope expands at the same time, exemplifying the 'mirror principle' discussed above. This becomes more clear in Fig. 10.4 which shows the radial variations of several mass shells inside the star. After point C the layers below the burning shell contract while the layers above expand, at an accelerating rate towards the end of phase C–D. As a result the temperature and density gradients between core and
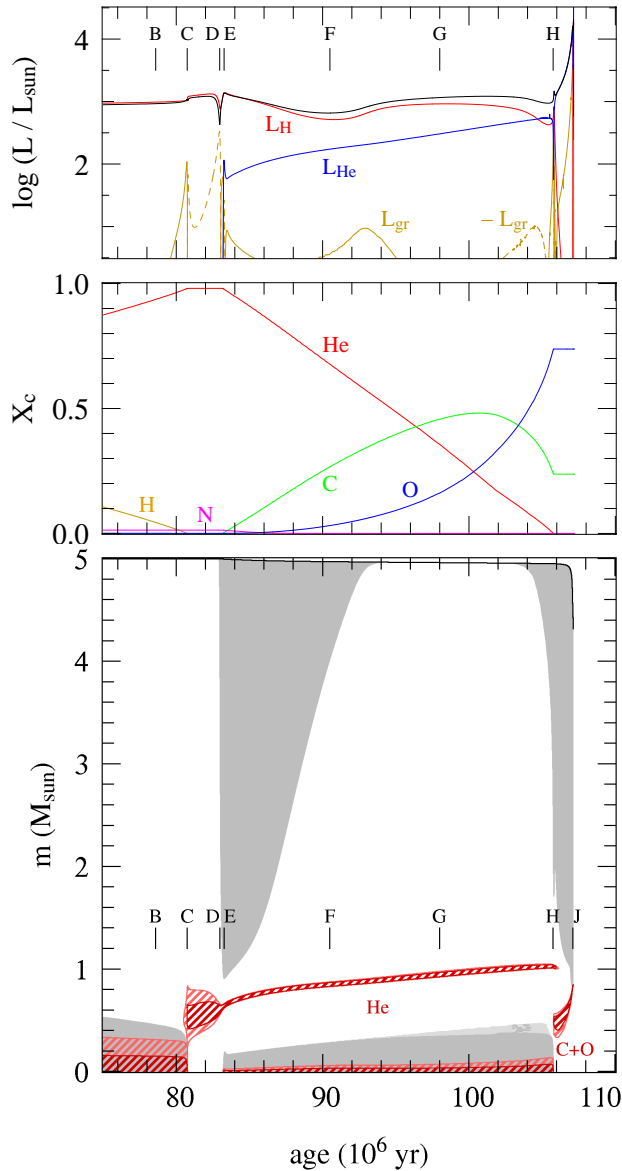
144

Figure 10.3. Internal evolution of a 5 $M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. The panels show various internal quantities as a function of time, from top to bottom:

(a) Contributions to the luminosity from hydrogen burning (red line), helium burning (blue) and gravitational energy release (orange; dashed parts show net *absorption* of gravitational energy). The black line is the surface luminosity.

(b) Central mass fractions of various elements (^{1}H, ^{4}He, ^{12}C, ^{14}N and ^{16}O) as indicated.

(c) Internal structure as a function of mass coordinate m, known as a 'Kippenhahn diagram'. A vertical line through the graph corresponds to a model at a particular time. Gray areas are convective, lighter-gray areas are semi-convective. The red hatched regions show areas of nuclear energy generation, where $\epsilon_{nuc} > 10\,L/M$ (dark red) and $\epsilon_{nuc} > 2\,L/M$ (light red). The letters B...J indicate the corresponding points in the evolution track in the H-R diagram, plotted in Fig. 10.2. See text for details.

envelope increase, and the burning shell occupies less and less mass (Fig. 10.3c). The latter portion of hydrogen-shell burning is therefore referred to as *thin shell burning*. Most of the time between C and D is spent in the thick shell burning phase at relatively small radii and $\log T_{eff} > 4.05$. The phase of expansion from $\log T_{eff} \approx 4.05$ to point D at $\log T_{eff} \approx 3.7$ occurs on the Kelvin-Helmholtz timescale and takes only a few times 10^5 yrs. A substantial fraction of the energy generated by shell burning is absorbed by the expanding envelope (dashed yellow line in Fig. 10.3a), resulting in a decrease of the surface luminosity between C and D.

The rapid evolution on a thermal timescale across the H-R diagram from the end of the main sequence to $T_{eff} \approx 5000$ K is characteristic of all intermediate-mass stars. The probability of detecting stars during this short-lived phase is very small, resulting in a gap in the distribution of stars in the H-R diagram known as the *Hertzsprung gap*.

As point D is approached the envelope temperature decreases and the opacity in the envelope rises, impeding radiative energy transport. The envelope grows increasingly unstable to convection, starting from the surface, until at D a large fraction of the envelope mass has become convective. During
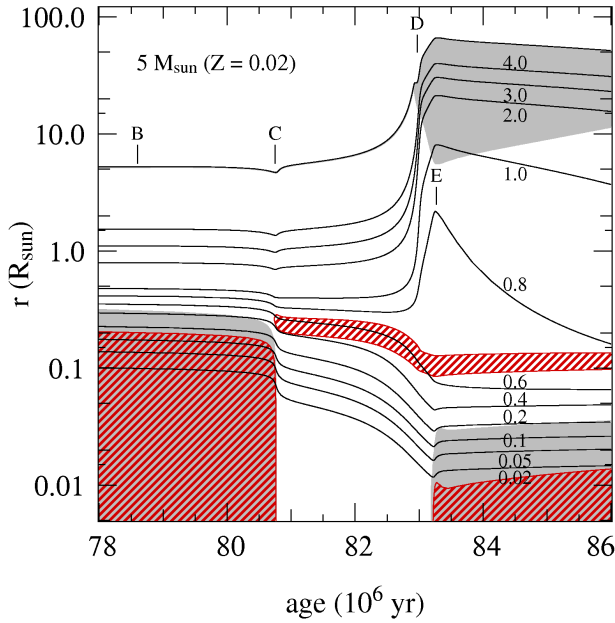
145

Figure 10.4. Radial variation of various mass shells (solid lines) in the 5 $M_\odot$ ($Z = 0.02$) star of Fig. 10.3, during the early post-main sequence evolution. Each line is labelled with its mass co-ordinate m in units of $M_\odot$; the top-most curve indicates the total radius R. Gray areas indicate convection and red cross-hatched areas have intense nuclear burning ($\epsilon_{\text{nuc}} > 10\,L/M$). Letters B...E correspond to those in Fig. 10.3.

phase D–E the star is a red giant with a deep convective envelope. The star is then located close to the Hayashi line in the H-R diagram, and while it continues to expand in response to core contraction, the luminosity increases as the effective temperature remains at the approximately constant value corresponding to the Hayashi line. The expansion of the star between D and E still occurs on the thermal timescale, so the H-shell burning phase of intermediate-mass stars on the red-giant branch is very short-lived.

At its deepest extent at point E, the base of the convective envelope is located at mass coordinate $m = 0.9\,M_\odot$ which is below the maximum extent of the former convective core during central H-burning (about 1.25 $M_\odot$ at the start of the main sequence). Hence material that was formerly inside the convective core, and has therefore been processed by hydrogen burning and the CNO-cycle, is mixed throughout the envelope and appears at the surface. This process is called *dredge-up* and occurs about halfway between D and E in Fig. 10.2. Dredge-up on the red giant branch also occurs in low-mass stars and we defer its discussion to Sec. 10.2.3.

The helium cores of intermediate-mass stars remain non-degenerate during the entire H-shell burning phase C–E, as can be seen in Fig. 10.1. These stars develop helium cores with masses larger than 0.3 $M_\odot$, the minimum mass for helium fusion discussed in Ch. 8. In the 5 $M_\odot$ star at point E the helium core mass is 0.6 $M_\odot$ when a central temperature of 10^8 K is reached and helium is ignited in the core. The ignition of helium halts further core contraction and envelope expansion and therefore corresponds to a local maximum in luminosity and radius. Evolution through helium burning will be discussed in Sec. 10.3.1.

10.2.2 Hydrogen-shell burning in low-mass stars

Compared to intermediate-mass stars, low-mass stars (with $M \lesssim 2\,M_\odot$) have small or no convective cores during central hydrogen burning, and when they leave the main sequence their cores are relatively dense and already close to becoming degenerate (see Fig. 10.1). In stars with $M \lesssim 1.1\,M_\odot$ the transition from central to shell hydrogen burning is gradual and initially $M_c/M < 0.1$ so the star can remain in thermal equilibrium with an isothermal helium core. By the time the helium core has grown to $\approx 0.1M$, its density is large enough that electron degeneracy dominates the pressure and the
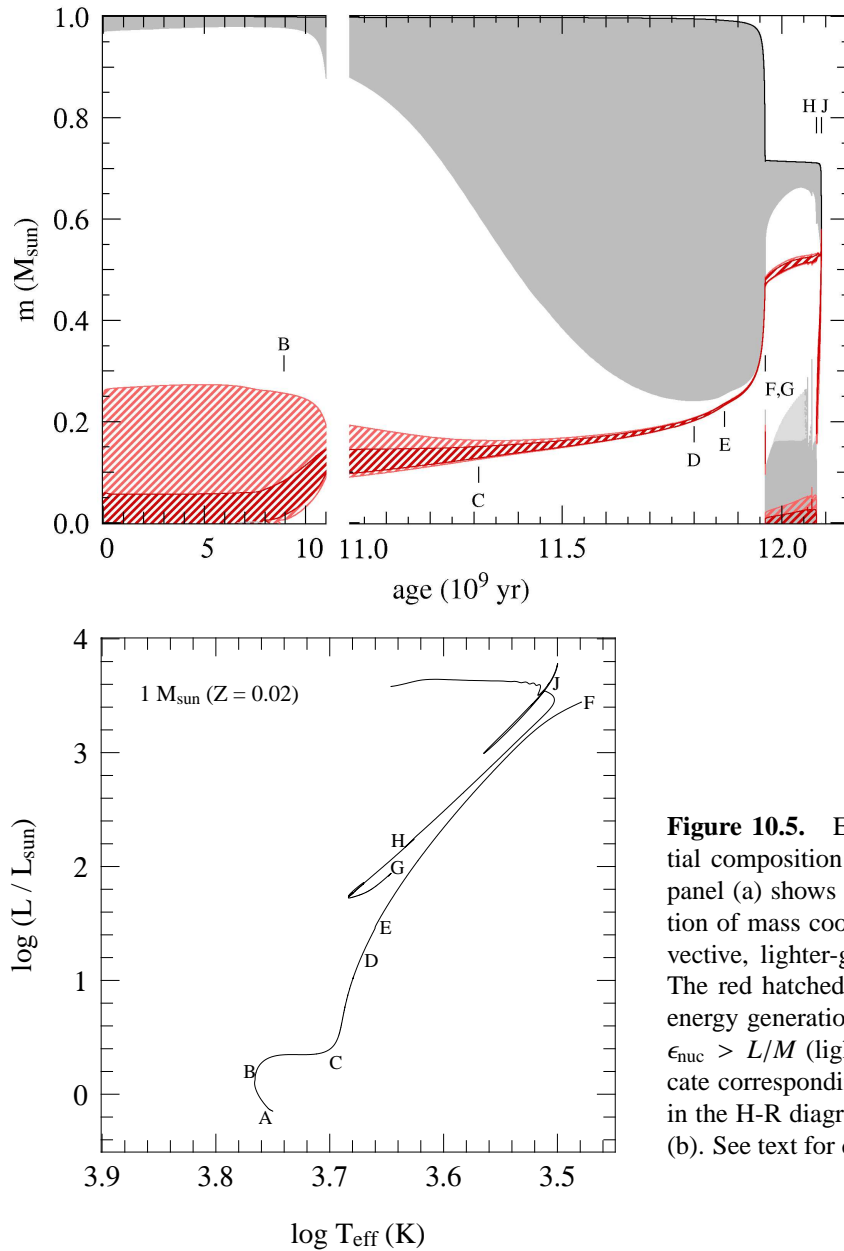
Figure 10.5. Evolution of a 1 $M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. The top panel (a) shows the internal structure as a function of mass coordinate m. Gray areas are convective, lighter-gray areas are semi-convective. The red hatched regions show areas of nuclear energy generation: $\epsilon_{nuc} > 5\,L/M$ (dark red) and $\epsilon_{nuc} > L/M$ (light red). The letters A...J indicate corresponding points in the evolution track in the H-R diagram, plotted in the bottom panel (b). See text for details.

Schönberg-Chandrasekhar limit has become irrelevant. Therefore low-mass stars can remain in HE and TE throughout hydrogen-shell burning and there is no Hertzsprung gap in the H-R diagram.

This can be seen in Fig. 10.5 which shows the internal evolution of a 1 $M_\odot$ star with quasi-solar composition in a Kippenhahn diagram and the corresponding evolution track in the H-R diagram. Hydrogen is practically exhausted in the centre at point B ($X_c = 10^{-3}$) after 9 Gyr, after which nuclear energy generation gradually moves out to a thick shell surrounding the isothermal helium core. Between B and C the core slowly grows in mass and contracts, while the envelope expands in response and the burning shell gradually becomes thinner in mass. By point C the helium core has become degenerate. At the same time the envelope has cooled and become largely convective, and the star finds itself at the base of the *red giant branch* (RGB), close to the Hayashi line. The star remains in thermal equilibrium throughout this evolution and phase B–C lasts about 2 Gyr for this 1 $M_\odot$ star. This long-lived phase corresponds to the well-populated *subgiant branch* in the H-R diagrams of old

star clusters.

Stars with masses in the mass range $1.1 - 1.5\,M_\odot$ show a very similar behaviour after the main sequence, the only difference being the small convective core they develop during core H-burning. This leads to a 'hook' in the evolution track at central H exhaustion (see Sec. 9.3). The subsequent evolution during H-shell burning is similar, the core remaining in TE until it becomes degenerate on the RGB and a correspondingly slow evolution across the subgiant branch. Stars with $1.5 \lesssim M/M_\odot \lesssim 2$ do exhibit a small Hertzsprung gap as they reach the Schönberg-Chandrasekhar limit before their cores become degenerate. After a period of slow, thick shell burning on the subgiant branch they undergo a phase of rapid, thermal-timescale expansion until they reach the giant branch. In this case the gap in T_{eff} to be bridged is narrow because the main sequence is already relatively close in effective temperature to the Hayashi line.

Regardless of these differences between stars of different mass during the early shell-H burning phase, all stars with $M \lesssim 2\,M_\odot$ have in common that their helium cores become degenerate before the central temperature is high enough for helium ignition, and they settle into TE on the red giant branch.

10.2.3 The red giant branch in low-mass stars

The evolution of low-mass stars along the red giant branch is very similar and almost independent of the mass of the star. The reason for this similarity is that by the time the helium core has become degenerate, a very strong density contrast has developed between the core and the envelope. The envelope is so extended that it exerts very little weight on the compact core, while there is a very large pressure gradient between core and envelope. The pressure at the bottom of the envelope (see eq. 9.14) is very small compared to the pressure at the edge of the core and in the hydrogen-burning shell separating core and envelope. Therefore the stellar structure depends almost entirely on the properties of the helium core. Since the core is degenerate, its structure is independent of its thermal properties (temperature) and only depends on its mass. Therefore the structure of a low-mass red giant is essentially a function of its *core mass*.

As a result there is a very tight relation between the helium core mass and the luminosity of a red giant, which is entirely due to the hydrogen shell-burning source. This *core-mass luminosity* relation is very steep for small core masses, $M_{\mathrm{c}} \lesssim 0.5\,M_\odot$ and can be approximately described by a power law

$$L \approx 2.3 \times 10^5 L_\odot \left(\frac{M_{\mathrm{c}}}{M_\odot}\right)^6 \tag{10.2}$$

Note that the luminosity of a low-mass red giant is independent of its total mass. Therefore the evolution of all stars with $M \lesssim 2\,M_\odot$ converges after the core becomes degenerate, which occurs when $M_{\mathrm{c}} \approx 0.1M$, i.e. later for larger M. From this point on also the central density and temperature start following almost the same evolution track (e.g. see Fig. 10.1b).

In the H-R diagram the star is located along the Hayashi line appropriate for its mass M. Higher-mass red giants therefore have slightly higher T_{eff} at the same luminosity.[2] Note that the location of the Hayashi line also depends on the *metallicity* of the star, since the effective temperature of a completely convective star is determined by the H^- opacity in the photosphere (Sec. 9.1.1). Because the H^- opacity increases with metallicity (Sec. 5.3), more metal-rich red giants of the same mass and luminosity are located at lower T_{eff}. This provides a means of deriving the metallicity of a globular cluster from the location of its RGB stars in the H-R diagram.

---

[2]This means there is also a *core-mass radius* relation, but it is less tight than the M_{c}-L relation and depends slightly on the total mass.
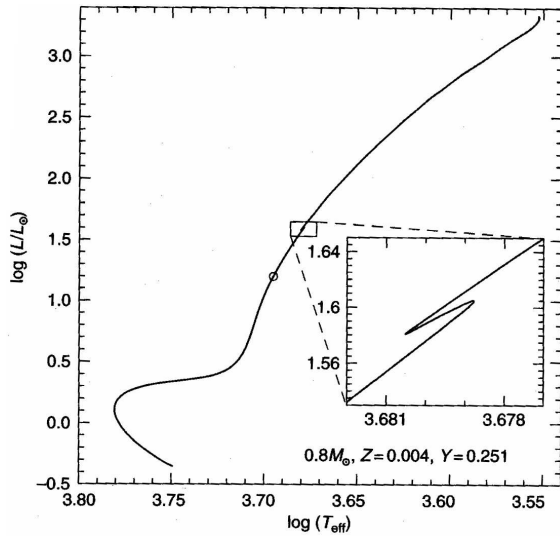
Figure 10.6. Evolution track of a $0.8\,M_\odot$ star of rather low metallicity, $Z = 0.004$. The inset shows the temporary decrease of luminosity when the H-burning shell crosses the hydrogen discontinuity left by the first dredge-up (corresponding to point E in Fig. 10.5). The open circle indicates where first dredge-up occurs. Figure from SALARIS & CASSISI.

As the H-burning shell adds mass to the degenerate helium core, the core slowly contracts and the radius and luminosity increase. The higher luminosity means the H-shell must burn at a higher rate, leading to faster core-mass growth. The evolution along the RGB thus speeds up as the luminosity increases (see Fig. 10.5). The density contrast between core and envelope increases and the mass within the burning shell decreases, to $\approx 0.001\,M_\odot$ near the tip of the RGB. Since less mass is contained in the burning shell while the luminosity increases, the energy generation rate per unit mass $\epsilon_{\rm nuc}$ increases strongly, which means the temperature within the burning shell also increases. With it, the temperature in the degenerate helium core increases. When the tip of the RGB is reached (at point F in Fig. 10.5) at $L \approx 2000\,L_\odot$ and a core mass of $\approx 0.45\,M_\odot$, the temperature in the degenerate core has reached a value close to 10^8 K and helium is ignited. This is an unstable process due to the strong degeneracy, and leads to a thermonuclear runaway known as the *helium flash* (see Sec. 10.3.2).

First dredge-up and the luminosity bump

When the convective envelope reaches its deepest extent at point D in Fig. 10.5, it has penetrated into layers that were processed by H-burning during the main sequence, and have been partly processed by the CN-cycle. Up to point D the surface He abundance increases and the H abundance decreases, but more noticeably the C/N ratio decreases by a large factor. This is called the *first dredge-up* phase (later dredge-ups occur after He burning).

Some time later, at point E in Fig. 10.5 the H-burning shell has eaten its way out to the discontinuity left by the convective envelope at its deepest extent. The shell suddenly finds itself in an environment with a higher H abundance (and a lower mean molecular weight). As a consequence it starts burning at a slightly lower rate, leading to a slight decrease in luminosity (see Fig. 10.6). The resulting loop (the star crosses this luminosity range three times) results in a larger number of stars in this luminosity range in a stellar population. This 'bump' in the luminosity function has been observed in many old star clusters.

Mass loss on the red giant branch

Another process that becomes important in low-mass red giants is *mass loss*. As the stellar luminosity and radius increase as a star evolves along the giant branch, the envelope becomes loosely bound and it is relatively easy for the large photon flux to remove mass from the stellar surface. The process

driving mass loss in red giants is not well understood. When calculating the effect of mass loss in evolution models an empirical formula due to Reimers is often used:

$$\dot{M} = -4 \times 10^{-13}\, \eta\, \frac{L}{L_\odot}\, \frac{R}{R_\odot}\, \frac{M_\odot}{M}\, M_\odot/\text{yr} \tag{10.3}$$

where η is a parameter of order unity. Note that the Reimers formula implies that a fixed fraction of the stellar luminosity is used to lift the wind material out of the gravitational potential well. However, the relation is based on observations of only a handful of stars with well-determined stellar parameters.

A value of $\eta \sim 0.25 - 0.5$ is often used because it gives the right amount of mass loss on the RGB to explain the morphology in the H-R diagram of stars in the subsequent helium-burning phase, on the *horizontal branch*. The $1\, M_\odot$ star of our example loses about $0.3\, M_\odot$ of its envelope mass by the time it reaches the tip of the giant branch.

10.3 The helium burning phase

As the temperature in the helium core approaches 10^8 K, the 3α reaction starts to produce energy at a significant rate. This is the onset of the *helium burning* phase of evolution. Unlike for hydrogen burning, the reactions involved in helium burning (see Sect. 6.4.2) are the same for all stellar masses. However, the conditions in the core at the ignition of helium are very different in low-mass stars (which have degenerate cores) from stars of higher mass (with non-degenerate cores). Therefore these cases will be discussed separately.

10.3.1 Helium burning in intermediate-mass stars

We again take the $5\, M_\odot$ star depicted in Figs. 10.2–10.3 as a typical example of an intermediate-mass star. The ignition of helium takes place at point E in these figures. Since the core is non-degenerate at this point ($\rho_c \approx 10^4$ g/cm^3, Fig. 10.1), nuclear burning is thermally stable and helium ignition proceeds quietly. Owing to the high temperature sensitivity of the He-burning reactions, energy production is highly concentrated towards the centre which gives rise to a convective core. The mass of the convective core is $0.2\, M_\odot$ initially and grows with time (unlike was the case for hydrogen burning).

Initially, the dominant reaction is the 3α reaction which converts ^{4}He into ^{12}C inside the convective core. As the ^{12}C abundance builds up, the ^{12}C$+\alpha$ reaction gradually takes over, so that ^{16}O is also produced at a rate that increases with time (see Fig. 10.3b and compare to Fig. 6.6). When the central He abundance $X_{\text{He}} < 0.2$ the mass fraction of ^{12}C starts decreasing as a result of the diminishing 3α rate (which is proportional to X_{He}^3). The final ^{12}C/^{16}O ratio is about 0.3, decreasing somewhat with stellar mass. This is related to the fact that in more massive stars the central temperature during He burning is larger. Note that the final ^{12}C/^{16}O ratio depends on the uncertain rate of the ^{12}C(α, γ) reaction, and the values given here are for the rate that is currently thought to be most likely.

The duration of the central helium burning phase in our $5\, M_\odot$ star (E–H) is about 22 Myr, i.e. approximately $0.27 \times \tau_{\text{MS}}$. This seems surprisingly long given that the energy gain per gram of He burning is only 10 % of that of H burning, while the luminosity of the star is (on average) somewhat larger than during the main sequence. The reason can be discerned from Fig. 10.3a: most of the luminosity during helium burning still comes from the H-burning shell surrounding the core, although the luminosity contribution of He burning (L_{He}) increases with time and becomes comparable towards the end of this phase.

We can understand the behaviour of L_{He} by considering that the properties of the helium core essentially depend only on the core mass M_c and are hardly affected by the surrounding envelope. Because the envelope is very extended the pressure it exerts on the core (eq. 9.14) is negligible compared

to the pressure inside the dense helium core. In fact L_{He} is a steep function of M_c, analogous to the main-sequence M-L relation – indeed, if the envelope were stripped away, the bare helium core would lie on a *helium main sequence*. The mass-luminosity relation for such helium main-sequence stars can be approximately described by the homology relation (7.32) if the appropriate value of μ is used. As a result of H-shell burning, M_c grows with time during the He-burning phase and L_{He} increases accordingly. Another consequence is that in models computed with convective overshooting L_{He} is larger on account of the larger core mass left after the main sequence (see Sect. 9.3.4). Therefore the duration of the He burning phase (i.e. the appropriate nuclear timescale, $\tau_{nuc} \propto M_c/L_{He}$) is *shorter* in models with overshooting. A $5\,M_\odot$ star of the same composition computed with overshooting has a main-sequence lifetime $\tau_{MS} = 100\,$Myr and a helium-burning lifetime of $16\,$Myr.

During helium burning intermediate-mass stars describe a loop in the H-R diagram (E–H in Fig. 10.2). After He ignition at the tip of the giant branch, the envelope contracts (on the nuclear timescale for helium burning) and the stellar radius decreases. Initially the luminosity also decreases while the envelope is mostly convective (E–F) and the star is forced to move along its Hayashi line. When most of the envelope has become radiative at point F, the star leaves the red giant branch and the effective temperature increases. This is the start of a so-called *blue loop*, the hottest point of which is reached at G when $X_{He} \approx 0.3$. This also corresponds to a minimum in the stellar radius, after which the envelope starts expanding and the star again approaches the giant branch when $X_{He} \approx 0.05$. By the end of core helium burning (H) the star is back on the Hayashi line, very close to its starting point (E). If we consider stars of different masses, the blue extension of the loops in the HRD increases (the loops extend to larger T_{eff} values) for increasing mass, up to $M \approx 12\,M_\odot$. (The behaviour of stars of larger masses can be more complicated, one of the reasons being strong mass loss, and we defer a discussion of this until Chapter 12.) On the other hand, for $M \lesssim 4\,M_\odot$ the loops always stay close to the red giant branch and do not become 'blue'.

The occurrence of blue loops is another example of a well-established result of detailed stellar evolution calculations, that is difficult to explain in terms of basic physics. The detailed models indicate that the occurrence and extension of blue loops depends quite sensitively on a number of factors: the chemical composition (mainly Z), the mass of the helium core relative to the envelope, and the shape of the hydrogen abundance profile above the core. It therefore also depends on whether convective overshooting was assumed to take place during the main sequence: this produces a larger core mass, which in turn has the effect of decreasing the blue-ward extension of the loops while increasing their luminosity.

The blue loops are important because they correspond to a slow, nuclear timescale phase of evolution. One therefore expects the corresponding region of the H-R diagram to be well populated. More precisely, since intermediate-mass stars spend part of their He-burning phase as red giants and part of it in a blue loop, one expects such stars to fill a wedge-shaped region in the HRD. Indeed one finds many stars in the corresponding region, both in the solar neighbourhood (Fig. 1.1, although this is dominated by *low-mass* stars) and in open clusters with ages less than $\sim 1\,$Gyr. The dependence of the loops on overshooting also makes observational tests of overshooting using He-burning stars possible. Another significant aspect of blue loops is that they are necessary for explaining Cepheid variables (see Sect. 10.4), which are important extragalactic distance indicators.

10.3.2 Helium burning in low-mass stars

In low-mass stars (with $M \lesssim 2\,M_\odot$) the helium burning phase differs from more massive stars in two important aspects: (1) helium ignition occurs under degenerate conditions, giving rise to a *helium flash*, and (2) all low-mass stars start helium burning with essentially the same core mass $M_c \approx 0.45\,M_\odot$ (Sect. 10.2.3). The luminosity of low-mass He-burning stars is therefore almost independent
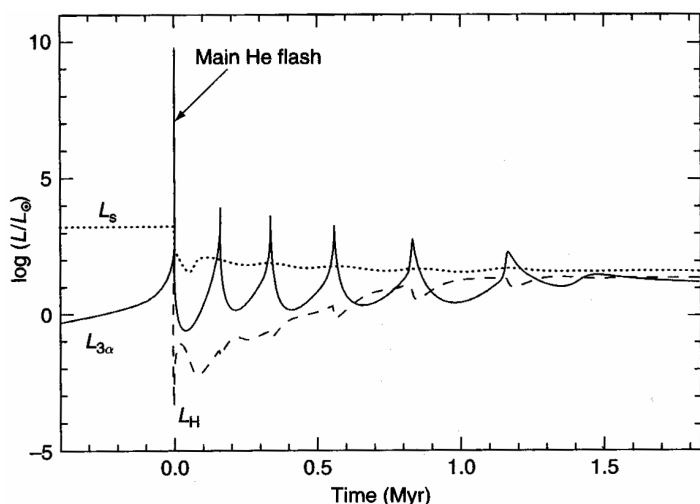
Figure 10.7. The helium flash. Evolution with time of the surface luminosity (L_s), the He-burning luminosity ($L_{3\alpha}$) and the H-burning luminosity (L_H) during the onset of He burning at the tip of the RGB in a low-mass star. Time $t = 0$ corresponds to the start of the main helium flash. Figure from SALARIS & CASSISI.

of their mass, giving rise to a *horizontal branch* in the HRD.

The helium flash

We again take a star of $1\,M_\odot$ as a typical example of all low-mass stars. Helium ignition occurs when $T_c \approx 10^8\,\mathrm{K}$ and $\rho_c \approx 10^6\,\mathrm{g/cm^3}$, so the helium core is strongly degenerate (see Fig. 10.1). We have seen in Sect. 7.5.2 that helium burning under these conditions is thermally unstable: the energy generated by the 3α reaction causes a temperature increase, rather than a decrease, and helium ignition thus initiates a *thermonuclear runaway*. The reason is that the degenerate pressure is basically independent of T, so that the energy released by fusion does not increase the pressure and therefore leads to negligible expansion and negligible work done. All nuclear energy released therefore goes into raising the internal energy. Since the internal energy of the degenerate *electrons* is a function of ρ and hence remains almost unchanged, it is the internal energy of the non-degenerate *ions* that increases and thus raises the temperature. As a result, the evolution is vertically upward in the ρ_c-T_c diagram.[3]

The thermonuclear runaway leads to an enormous overproduction of energy: at maximum, the local luminosity in the helium core is $l \approx 10^{10}\,L_\odot$ – similar to a small galaxy! However, this only lasts for a few seconds. Since the temperature increases at almost constant density, degeneracy is eventually lifted when $T \approx 3 \times 10^8\,\mathrm{K}$. Further energy release increases the pressure when the gas starts behaving like an ideal gas and thus causes expansion and cooling. All the energy released by the thermonuclear runaway is absorbed in the expansion of the core, and none of this nuclear power reaches the surface. The expansion and cooling results in a decrease of the energy generation rate, until it balances the energy loss rate and the core settles in thermal equilibrium at $T_c \approx 10^8\,\mathrm{K}$ and $\rho_c \approx 2 \times 10^4\,\mathrm{g/cm^3}$ (see Fig. 10.1). Further nuclear burning of helium is thermally stable.

Detailed numerical calculations of the helium flash indicate that this sequence of events indeed takes place, but helium is not ignited in the centre but in a spherical shell at $m \approx 0.1\,M_\odot$ where T has a maximum. This off-centre temperature maximum is due to *neutrino losses* during the preceding red giant phase. These neutrinos are not released by nuclear reactions, but by spontaneous weak interaction processes occurring at high density and temperature (see Section 6.5). Since neutrinos thus created escape without interacting with the stellar gas, this energy loss leads to effective cooling

---

[3]This part of the evolution is skipped in the $1\,M_\odot$ model shown in Fig. 10.1, which is why a gap appears in the evolution track. The evolution during the He flash is shown schematically as a dashed line for the $1\,M_\odot$ model in Fig. 8.4.
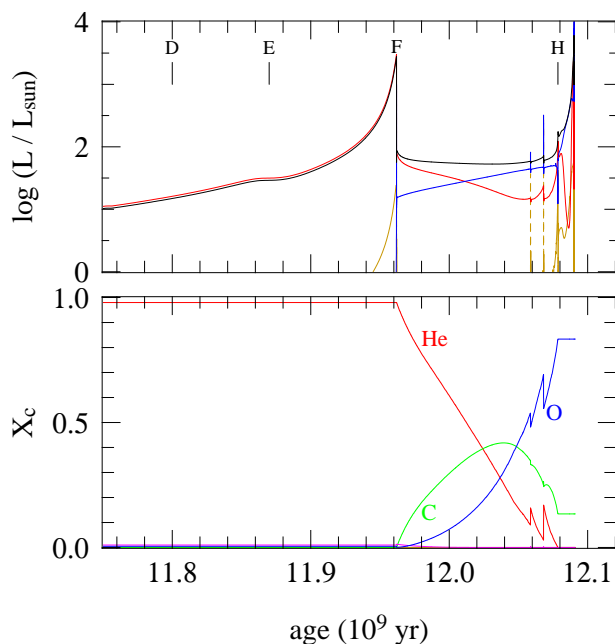
Figure 10.8. Evolution with time of the luminosities and central abundances in a $1\,M_\odot$ star during the late part of the red giant branch and during helium burning. Letters D...H correspond to the same evolution phases as in Fig. 10.5.

of the central region of the degenerate helium core. The mass coordinate at which $T_{\max}$ occurs (and where helium ignites) decreases somewhat with stellar mass.

The high local luminosity causes almost the entire region between the ignition point (at $m \approx 0.1\,M_\odot$) up to the bottom of the H-burning shell (at $0.45\,M_\odot$) to become convective. The energy released in the He flash is thus transported efficiently to the edge of the core, where it is absorbed by expansion of the surrounding non-degenerate layers. Convection also mixes the product of the He flash (^{12}C produced in the 3α reaction) throughout the core. About 3 % of the helium in the core is converted into carbon during the flash. Because the convective shell containing this carbon never overlaps with the convective envelope surrounding the H-burning shell, this carbon does not reach the surface. (However, this may be different at very low metallicity.)

After the He flash, the whole core expands somewhat but remains partially degenerate. In detailed models a series of smaller flashes follows the main He flash (see Fig. 10.7) during $\approx 1.5\,\mathrm{Myr}$, before degeneracy in the centre is completely lifted and further He burning proceeds stably in a convective core, as for intermediate-mass stars.

The horizontal branch

In our $1\,M_\odot$ example star, the helium flash occurs at point F in Fig. 10.5. Evolution through the helium flash was not calculated for the model shown in this figure. Instead, the evolution of the star is resumed at point G when the helium core has become non-degenerate and has settled into TE with stable He burning in the centre and H-shell burning around the core. (Models constructed in this way turn out to be very similar to models that are computed all the way through the He flash, such as shown in Fig. 10.7.) At this stage the luminosity and radius of the star have decreased by more than an order of magnitude from the situation just before the He flash. Here we again see the mirror principle at work: in this case the core has expanded (from a degenerate to a non-degenerate state) and the envelope has simultaneously contracted, with the H-burning shell acting as a 'mirror'.

In the $1\,M_\odot$ star of solar composition shown in Fig. 10.5, helium burning occurs between G and H. The position of the star in the H-R diagram does not change very much during this period, always staying close (but somewhat to the left of) the red giant branch. The luminosity is $\approx 50\,L_\odot$ for most
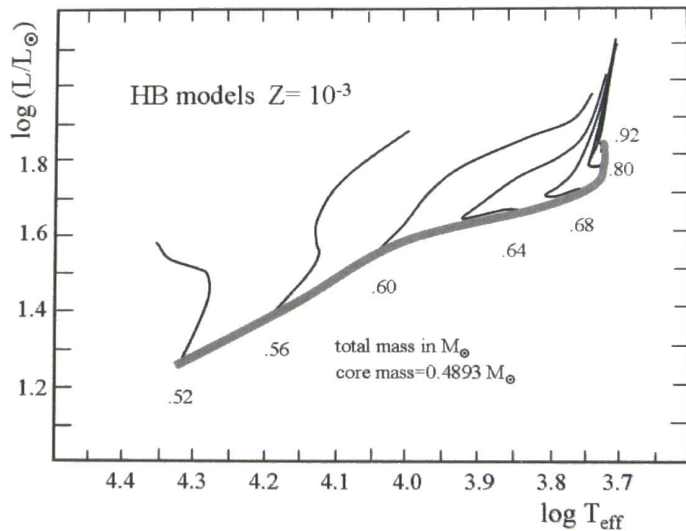
Figure 10.9. Location of the zero-age horizontal branch (think gray line) for a metallicity $Z = 0.001$ typical of globular clusters. These models have the same core mass (0.489 $M_\odot$) but varying total (i.e. envelope) mass, which determines their position in the H-R diagram. Evolution tracks during the HB for several total mass values are shown as thin solid lines. Figure from MAEDER.

of the time; this value is determined mainly by the core mass. Since the core mass at the start of helium burning is $\approx 0.45 \, M_\odot$ for all low-mass stars, independent of stellar mass, the luminosity at which He burning occurs is also almost independent of mass. If we consider He-burning stars of a given composition (e.g. in a star cluster), only the envelope mass may vary from star to star. At solar metallicity, all such stars occupy about the same position in the HRD. This gives rise to a so-called *red clump* in observed colour-magnitude diagrams of low-mass stellar populations (visible for instance in Fig. 1.1). However, the radius and effective temperature of He-burning stars depends on their envelope mass. Stars with a small envelope mass (either because of a smaller initial mass, or because they suffered a larger amount of mass loss on the RGB) can be substantially hotter than the one shown in Fig. 10.5. Furthermore, at low metallicity the critical envelope mass, below which He-burning stars become small and hot, is larger. Stars with different amounts of mass remaining in their envelopes can then form a *horizontal branch* in the HRD (Fig. 10.9). Horizontal branches are found in old stellar populations, especially in globular clusters of low metallicity (an example is the globular cluster M3 shown in Fig. 1.2). The observed distribution of stars along the HB varies greatly from cluster to cluster, and the origin of these different *HB morphologies* is not fully understood.

The duration of the core helium burning phase is about 120 Myr, again independent of stellar mass. While this is longer than in intermediate-mass stars, it is a much shorter fraction of the main-sequence lifetime because of the much higher luminosity of the He-burning phase. The evolution of the stellar structure during helium burning is qualitatively similar to that of intermediate-mass stars; see Figs. 10.5a and 10.8. The most striking differences are:

- The contribution of He-burning to the stellar luminosity is larger, especially towards the end of the phase. This is due to the relatively small envelope mass.

- The development of a substantial *semi-convective* region on top of the convective core. This is related to a difference in opacity between the C-rich convective core and the He-rich zone surrounding it, and gives rise to partial (non-homogeneous) mixing in this region.

- The occurrence of 'breathing pulses', giving rise to the sudden jumps in the central composition and in the luminosity. Whether these are real or simply a numerical artifact of one-dimensional stellar models is not clear.[4]

---

[4]For details about the latter two effects, see either SALARIS & CASSISI or John Lattanzio's tutorial at http://www.maths.monash.edu.au/~johnl/StellarEvolnDemo/.
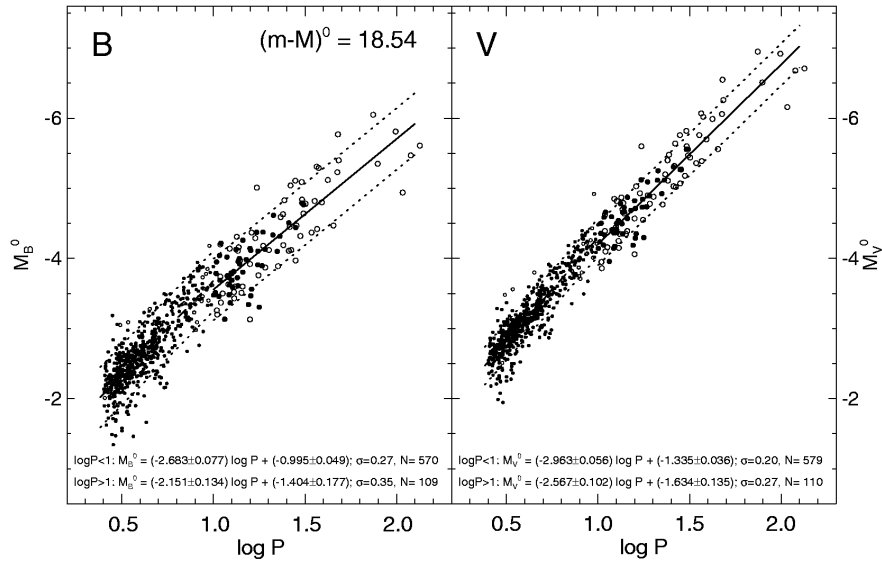
Figure 10.10. The period-luminosity relation for classical Cepheids in the Large Magellanic Cloud. Luminosity is expressed as absolute magnitude in the B band (left) and in the V band. Figure from Sandage et al. (2004, A&A 424, 43).

10.4 Pulsational instability during helium burning

During their post-main sequence evolution, stars may undergo one or more episodes during which they are unstable to radial pulsations. The most important manifestation of these pulsations are the *Cepheid* variables, luminous pulsating stars with periods between about 2 and 100 days. It turns out that there is a well-defined correlation between the pulsation period and the luminosity of these stars, first discovered for Cepheids in the Small Magellanic Cloud. A modern version of this empirical relation is shown in Fig. 10.10. Their importance for astronomy lies in the fact that the period can be easily determined, even for stars in other galaxies, and thus provides an estimate of the absolute luminosity of such a star, making Cepheids important *standard candles* for the extragalactic distance scale.

Cepheids lie along a pulsational instability strip in the H-R diagram (see Fig. 10.11). During the evolution of an intermediate-mass star, this instability strip is crossed up to three times. The first crossing occurs during H-shell burning (C–D in Fig. 10.2) but this is such a rapid phase that the probability of catching a star in this phase is very small. In stars with sufficiently extended blue loops, another two crossings occur (F-G and G–H) during a much slower evolution phase. Cepheids must thus be helium-burning stars undergoing a blue loop. Equivalently, the *RR Lyrae* variables seen in old stellar populations lie along the intersection of the instability strip and the horizontal branch.

Since pulsation is a dynamical phenomenon, the pulsation period is closely related to the dynamical timescale (eq. 2.18). Therefore the pulsation period Π is related the mean density: to first approximation once can write $\Pi \propto \bar{\rho}^{-1/2} \propto M^{-1/2}R^{3/2}$. Each passage of the instability strip yields a fairly well-defined radius and luminosity. Passage at a larger L corresponds to a larger R and therefore to a larger Π, because the variation in mass is smaller than that in radius and enters the relation with a smaller power. This provides a qualitative explanation of the period-luminosity relation. The minimum observed period should correspond to the the lowest-mass star undergoing a blue loop. Also the number of Cepheids as a function of period must correspond to the time it takes for a star of the corresponding mass to cross the instability strip. Thus Cepheids provide a potential test of stellar evolution models.

10.4.1 Physics of radial stellar pulsations

The radial oscillations of a pulsating star result from pressure waves, i.e sound waves that resonate in the stellar interior. These radial oscillation modes are essentially standing waves, with a node at the centre and an open end at the stellar surface – not unlike the sound waves in an organ pipe. Similarly, there are several possible modes of radial pulsation, the *fundamental mode* having just one node at the centre, while the *first* and *second overtone* modes have one or two additional nodes between the centre and surface, etc. Most radially pulsating stars, such as Cepheids, are oscillating in their fundamental mode.

In order to understand what powers the pulsations of stars in the instability strip, let us first reconsider the dynamical stability of stars. We have seen in Sec. 7.5.1 that overall dynamical stability requires $\gamma_{ad} > \frac{4}{3}$. In this situation a perturbation of pressure equilibrium will be restored, the restoring force being larger the more γ_{ad} exceeds the critical value of $\frac{4}{3}$. In practice, due to the inertia of the layers under consideration, this will give rise to an *oscillation* around the equilibrium structure. A linear perturbation analysis of the equation of motion (2.11) shows that a layer at mass coordinate m having equilibrium radius r_0 will undergo radial oscillations with a frequency

$$\omega^2 = (3\gamma_{ad} - 4)\frac{Gm}{r_0^3}, \tag{10.4}$$

if we assume the oscillations are adiabatic. Note that $\omega^2 > 0$ as long as $\gamma_{ad} > \frac{4}{3}$, consistent with dynamical stability. On the other hand, for $\gamma_{ad} < \frac{4}{3}$ the frequency becomes imaginary, which indicates
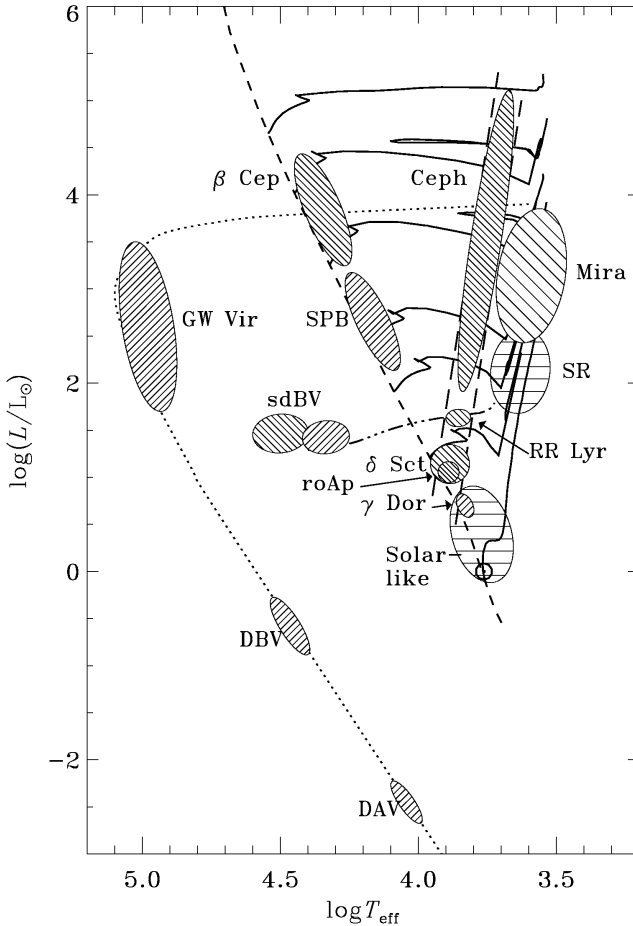

Figure 10.11. Occurrence of various classes of pulsating stars in the H-R diagram, overlaid on stellar evolution tracks (solid lines). Cepheid variables are indicated with 'Ceph', they lie within the pulsational instability strip in the HRD (long-dashed lines). Their equivalents are the RR Lyrae variables among HB stars (the horizontal branch is shown as a dash-dotted line), and the δ Scuti stars (δ Sct) among main-sequence stars. Pulsational instability is also found among luminous red giants (Mira variables), among massive main-sequence stars – β Cep variables and slowly pulsating B (SPB) stars, among extreme HB stars known as subdwarf B stars (sdBV) and among white dwarfs. Figure from Christensen-Dalsgaard (2004).

156

an exponential growth of the perturbation, i.e. dynamical instability. A proper average of ω over the star yields the pulsation frequency of the fundamental mode. We can obtain an approximate expression by replacing m with the total mass M and r_0 by the radius R, and taking γ_{ad} constant throughout the star. This yields

$$\Pi_0 = \frac{2\pi}{\sqrt{(3\gamma_{ad} - 4)\, GM/R^3}} = \left(\frac{3\pi}{(3\gamma_{ad} - 4)\, G\bar{\rho}}\right)^{1/2}. \tag{10.5}$$

This is indeed the same expression as for the dynamical timescale, to within a factor of unity. One can write

$$\Pi = Q\left(\frac{\bar{\rho}}{\bar{\rho}_\odot}\right)^{-1/2}, \tag{10.6}$$

where the pulsation constant Q depends on the structure of the star and is different for different modes of pulsation. For the fundamental mode, $Q \approx 0.04$ days and Q is smaller for higher overtones.

Driving and damping of pulsations

In an exactly adiabatic situation the oscillations will maintain the same (small) amplitude. In reality the situation is never exactly adiabatic, which means that the oscillations will generally be damped, unless there is an instability that drives the oscillation, i.e. that makes their amplitude grow.

The requirement for growth of an oscillation is that the net work done by a mass element in the star on its surroundings during an oscillation cycle must be positive, $\oint P\, dV > 0$. By the first law of thermodynamics, this work is provided by a net amount of heat being absorbed by the element during the cycle,

$$\oint dQ = \oint P\, dV > 0.$$

The change in entropy of the mass element is $dS = dQ/T$. Since entropy is a state variable, $\oint dQ/T = 0$ during a pulsation cycle. A mass element maintaining constant T during a cycle therefore cannot absorb any heat. Suppose that the temperature undergoes a small variation $T(t) = T_0 + \delta T(t)$ around an average value T_0. Then

$$0 = \oint \frac{dQ}{T} = \oint \frac{dQ}{T_0 + \delta T} \approx \oint \frac{dQ}{T_0}\left(1 - \frac{\delta T}{T_0}\right), \tag{10.7}$$

or

$$\oint dQ \approx \oint dQ\,\frac{\delta T}{T_0}. \tag{10.8}$$

Eq. (10.8) means that heat must enter the element ($dQ > 0$) when the temperature is high ($\delta T > 0$), i.e. when the layer is compressed, and/or heat must leave the layer ($dQ < 0$) during the low-temperature part of the cycle ($\delta T < 0$), i.e. during expansion. This is known in thermodynamics as a *heat engine*, and is analogous to what happens in a normal combustion motor, such as a car engine. In a pulsating star, some layers may absorb heat and do work to drive the pulsation, while other layers may lose heat and thereby damp the pulsation (if $\oint dQ = \oint P\, dV < 0$). To determine the overall effect, the contributions $\oint P\, dV$ must be integrated over all mass layers in the star.

In stars there are two possible mechanisms that can drive pulsations:

- If nuclear reactions occur in a region that is compressed during a pulsation, then the increase in T will lead to an increase in the energy generation rate ϵ_{nuc}. This satisfies the criterion (10.8) and is known as the ϵ-*mechanism*. Although this is always present, the amplitudes of the oscillations induced by this mechanism in the core of a star are usually so small that it cannot drive any significant pulsations. It may have important effects in very massive stars, but it is certainly not relevant for explaining Cepheid pulsations.

- If during the compression of a layer it becomes more *opaque*, then the energy flowing through this layer will be 'trapped'. The resulting increase in temperature and pressure pushes the layer outward. During the resulting expansion, the gas will become more transparent again and release the trapped heat. This so-called κ-*mechanism* can thus maintain the oscillation cycle and drive radial pulsations.

The condition for the κ-mechanism to work is therefore that the opacity must increase when the gas is compressed. The compression during a pulsation cycle is not exactly adiabatic, otherwise the mechanism would not work, but it is very close to adiabatic. Then the condition can be written as $(\mathrm{d}\ln\kappa/\mathrm{d}\ln P)_{\text{ad}} > 0$. We can write this as

$$\left(\frac{\mathrm{d}\ln\kappa}{\mathrm{d}\ln P}\right)_{\text{ad}} = \left(\frac{\partial\ln\kappa}{\partial\ln P}\right)_T + \left(\frac{\partial\ln\kappa}{\partial\ln T}\right)_P \left(\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right)_{\text{ad}} \equiv \kappa_P + \kappa_T\,\nabla_{\text{ad}}, \tag{10.9}$$

where κ_P and κ_T are shorthand notation for the partial derivatives of $\ln\kappa$ with respect to $\ln P$ and $\ln T$, respectively. For successful pulsations we must therefore have

$$\kappa_P + \kappa_T\,\nabla_{\text{ad}} > 0. \tag{10.10}$$

The instability strip and the period-luminosity relation

In stellar envelopes the opacity can be roughly described by a Kramers law, $\kappa \propto \rho T^{-3.5}$, which when combined with the ideal-gas law implies $\kappa_P \approx 1$ and $\kappa_T \approx -4.5$. Since for an ionized ideal gas $\nabla_{\text{ad}} = 0.4$, we normally have $\kappa_P + \kappa_T\,\nabla_{\text{ad}} < 0$, i.e. κ decreases upon compression and the star will not pulsate. In order to satisfy (10.10) one must have either:

- $\kappa_T > 0$, which is the case when the H^- opacity dominates, at $T < 10^4$ K. This may contribute to the driving of pulsations in very cool stars, such as Mira variables (Fig. 10.11), but the Cepheid instability strip is located at too high T_{eff} for this to be important.

- In case of a Kramers-like opacity, a small value of ∇_{ad} can lead to pulsation instability. For $\kappa_P \approx 1$ and $\kappa_T \approx -4.5$, eq. (10.10) implies $\nabla_{\text{ad}} \lesssim 0.22$. Such small values of ∇_{ad} can be found in *partial ionization zones*, as we have seen in Sec. 3.5 (e.g. see Fig. 3.5).

Stars generally have two important partial ionization zones, one at $T \approx 1.5 \times 10^4$ K where both $H \leftrightarrow H^+ + e^-$ and $He \leftrightarrow He^+ + e^-$ occur, and one at $T \approx 4 \times 10^4$ K where helium becomes twice ionized ($He^+ \leftrightarrow He^{++} + e^-$). These partial ionization zones can explain the location of the instability strip in the H-R diagram, as follows.

- At large T_{eff} (for $T_{\text{eff}} \gtrsim 7500$ K, the 'blue edge' of the instability strip) both ionization zones lie near the surface, where the density is very low. Although this region is indeed non-adiabatic, the mass and heat capacity of these zones is too small to drive pulsations effectively.

- As T_{eff} decreases, the ionization zones lie deeper into the stellar envelope. The mass and heat capacity in the partial ionization zones increase, while remaining non-adiabatic enough to absorb sufficient heat to drive pulsations.

- At still smaller T_{eff} (for $T_{\mathrm{eff}} \lesssim 5500\,\mathrm{K}$, the 'red edge' of the instability strip) the partial ionization zones lie at such high density that the gas behaves almost adiabatically. Although these zones still have a destabilizing effect, they cannot absorb enough heat to make the star as a whole unstable.

Thus the instability strip occupies a narrow region in the H-R diagram, as indicated in Fig. 10.11. Its location is related to the depth of the partial ionization zones. Since these zones occur in a specific temperature range, the instability strip also occurs for a narrow range of T_{eff} values, and is almost vertical in the H-R diagram (and parallel to the Hayashi line).

We can understand the period-luminosity relation from the dependence of the pulsation period on mass and radius (eq. 10.6). Since Cepheids follow a mass-luminosity relation, $M \propto L^{\alpha}$, and since $L \propto R^2 T_{\mathrm{eff}}^4$, we can write

$$\Pi \propto Q\, \frac{R^{3/2}}{M^{1/2}} \propto Q\, \frac{L^{(3/4)-(1/2\alpha)}}{T_{\mathrm{eff}}^3}.$$

With $\alpha \approx 3$ and $T_{\mathrm{eff}} \approx$ constant, we find $\Pi \propto L^{0.6}$ or $\log L \approx 1.7 \log \Pi +$ const. Detailed numerical models give

$$\log L = 1.270 \log \Pi + 2.570 \tag{10.11}$$

for the blue edge, and a slope of 1.244 and a constant 2.326 for the red edge. The smaller slope than in the simple estimate is mainly due to the fact that the effective temperature of the instability strip is not constant, but slightly decreases with increasing L.

Suggestions for further reading

The contents of this chapter are also covered by Chapters 25.3.2 and 26.1–26.5 of MAEDER, while stellar pulsations and Cepheids are treated in detail in Chapter 15. See also KIPPENHAHN & WEIGERT, Chapters 31 and 32.

Exercises

10.1 Conceptual questions

(a) Why does the luminosity of a star increase on the main sequence? Why do low-mass stars, like the Sun, expand less during the main sequence than higher-mass stars?

(b) Explain what happens during the 'hook' at the end of the main sequence of stars more massive than the Sun.

(c) What is *convective overshooting*? Think of at least three effects of overshooting on the evolution of a star.

(d) Explain the existence of a *Hertzsprung gap* in the HRD for high-mass stars. Why is there no Hertzsprung gap for low-mass stars?

(e) What do we mean by the *mirror principle*?

(f) Why does the envelope become convective on the red giant branch? What is the link with the *Hayashi line*?

10.2 Evolution of the abundance profiles

(a) Use Fig. 10.3 to sketch the profiles of the hydrogen and helium abundances as a function of the mass coordinate in a $5\,M_\odot$ star, at the ages corresponding to points C, E, G and H. Try to be as quantitative as possible, using the information provided in the figure.

(b) Do the same for a $1\,M_\odot$ star, using Figs. 10.5 and 10.8, at points B, D, F and H.

(c) The abundances plotted in Figs. 10.3 and 10.8 are central abundances. What happens to the abundances at the surface?

10.3 Red giant branch stars

(a) Calculate the total energy of the Sun assuming that the density is constant, i.e. using the equation for potential energy $E_{\mathrm{gr}} = -\frac{3}{5}GM^2/R$. In later phases, stars like the Sun become red giants, with $R \approx 100R_\odot$. What would be the total energy, if the giant had constant density. Assume that the mass did not change either. Is there something wrong? If so, why is it?

(b) What really happens is that red giants have a dense, degenerate, pure helium cores which grow to $\sim 0.45M_\odot$ at the end of the red giant branch (RGB). What is the maximum radius the core can have for the total energy to be smaller than the energy of the Sun? (N.B. Ignore the envelope – why are you allowed to do this?)

(c) For completely degenerate stars, one has

$$R = 2.6 \times 10^9 \, \mu_e^{-5/3} \left(\frac{M}{M_\odot}\right)^{-1/3} \mathrm{cm}, \qquad (10.12)$$

where μ_e is the molecular weight per electron and $\mu_e = 2$ for pure helium. Is the radius one finds from this equation consistent with upper limit derived in (b)?

10.4 Core mass-luminosity relation for RGB stars

Low-mass stars on the RGB obey a core mass-luminosity relation, which is approximately given by eq. (10.2). The luminosity is provided by hydrogen shell burning.

(a) Derive relation between luminosity L and the rate at which the core grows dM_c/dt. Use the energy released per gram in hydrogen shell burning.

(b) Derive how the core mass evolves in time, i.e, $M_c = M_c(t)$.

(c) Assume that a star arrives to the RGB when its core mass is 15% of the total mass, and that it leaves the RGB when the core mass is $0.45\,M_\odot$. Calculate the total time a $1\,M_\odot$ star spends on the RGB and do the same for a $2\,M_\odot$ star. Compare these to the main sequence (MS) lifetimes of these stars.

(d) What happens when the core mass reaches $0.45\,M_\odot$? Describe the following evolution of the star (both its interior and the corresponding evolution in the HRD).

(e) What is the difference in evolution with stars more massive than $2\,M_\odot$?

10.5 Jump in composition

Consider a star with the following distribution of hydrogen:

$$X(m) = \begin{cases} 0.1 & \text{for } m < m_c \\ 0.7 & \text{for } m \le m_c \end{cases} \qquad (10.13)$$

(a) In this star a discontinuous jump in the composition profile occurs at $m = m_c$. What could have caused such a chemical profile? Explain why P and T must be continuous functions.

(b) Calculate the jump in density $\Delta\rho/\rho$.

(c) Also calculate the jump in opacity, $\Delta\kappa/\kappa$, if the opacity is given as:
- Kramers: $\kappa_{bf} \sim Z(1 + X)\rho T^{-3.5}$
- Electron scattering: $\kappa_e = 0.2(1 + X)$

160

Chapter 11

Late evolution of low- and intermediate-mass stars

After the central helium burning phase a central core composed of carbon and oxygen is formed. As discussed before, the further evolution of a star differs greatly between massive stars on the one hand, and low- and intermediate-mass stars on the other hand. The evolution of massive stars, in which the core avoids degeneracy and undergoes further nuclear burning cycles, will be discussed in the next chapter.

In low- and intermediate-mass stars, up to about $8\,M_\odot$, the C-O core becomes degenerate and their late evolution is qualitatively similar. These stars evolve along the so-called *asymptotic giant branch* (AGB) in the H-R diagram. The AGB is a brief but interesting and important phase of evolution, among other things because it is the site of rich nucleosynthesis. AGB stars also suffer from strong mass loss, which eventually removes their envelope and leaves the degenerate C-O core, which after a brief transition stage as the central star of a planetary nebula, becomes a long-lived cooling *white dwarf*.

11.1   The asymptotic giant branch

The AGB phase starts at the exhaustion of helium in the centre. In the examples of the 5 and $1\,M_\odot$ stars discussed in the previous chapter, this occurs at point H in the evolution tracks (Figs. 10.2 and 10.5). The star resumes its climb along the giant branch, which was interrupted by central helium burning, towards higher luminosity. In low-mass stars the AGB lies at similar luminosities but somewhat higher effective temperature than the preceding RGB phase. This is the origin of the name 'asymptotic' giant branch. For stars more massive than about $2.5\,M_\odot$ the AGB lies at higher luminosities than the RGB and the name has no morphological meaning.

One can distinguish two or three phases during the evolution of a star along the AGB. These are highlighted in Fig. 11.1 for our $5\,M_\odot$ example star, but the evolution of lower-mass stars is qualitatively similar.

The early AGB phase

After central He exhaustion the carbon-oxygen core contracts. During a brief transition all layers below the H-burning shell contract (shortly after point H), until He burning shifts to a shell around the CO core. The star now has two active burning shells and a double mirror effect operates: the core contracts, the He-rich layers above expand, and the outer envelope starts contracting. However, due
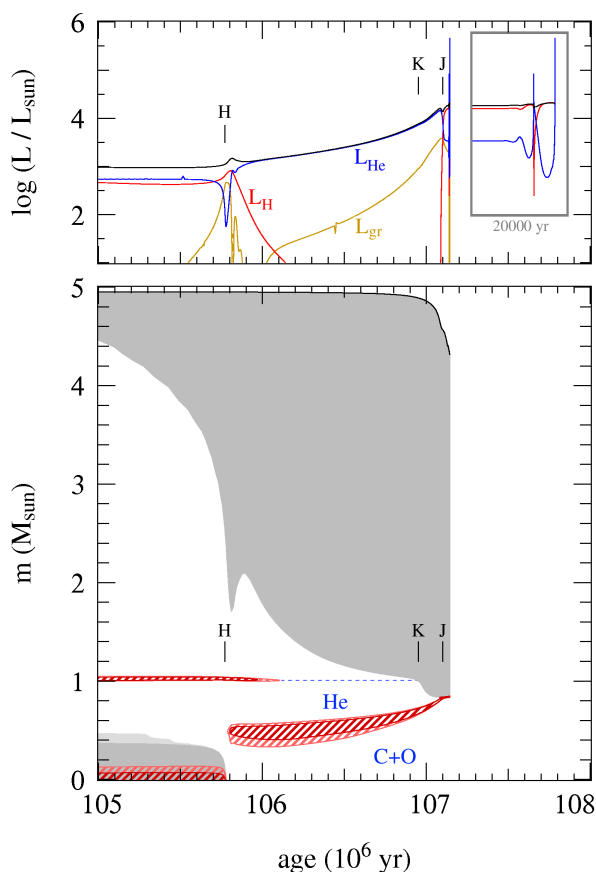
Figure 11.1. Evolution of luminosities (upper panel) and internal structure (lower panel) with time in a 5 $M_\odot$ star (with composition $X = 0.70$, $Z = 0.02$) during the last stages of helium burning and on the AGB. Compare with Fig. 10.3 for the same star. The early AGB starts at point H, when He burning shifts quite suddenly from the centre to a shell around the former convective core. The H-burning shell extinguishes and at point K second dredge-up occurs. The H-burning shell is re-ignited some time later at point J. This is the start of the double shell-burning phase, which soon afterward leads to thermal pulses of the He-burning shell (and break-down of this particular model). The first thermal pulses can be seen in the inset of the upper panel which shows the last 20 000 yr of this model calculation. Strong mass loss is then expected to remove the stellar envelope within $\lesssim 10^6$ yr, leaving the degenerate CO core as a cooling white dwarf.

to expansion of the He-rich zone the temperature in the H-shell decreases and the H-burning shell is extinguished. Thus only one 'mirror' remains and now the entire envelope – He-rich layer plus H-rich outer envelope – starts expanding in response to core contraction. A fairly long-lived phase follows in which the stellar luminosity is provided almost entirely by He-shell burning (phase H-K in Fig. 11.1). This is called the *early AGB* phase.

The He-burning shell gradually adds mass to the growing CO core, which becomes degenerate due to its increasing density. As the envelope expands and cools the convective envelope penetrates deeper until it reaches the composition discontinuity left by the extinct H-shell at point K.

Second dredge-up

In stars of sufficiently high mass, $M \gtrsim 4\,M_\odot$ (depending somewhat on the initial composition and on whether overshooting is included) a convective dredge-up episode can occur, called the *second dredge-up*. At point K in Fig. 11.1 the convective envelope is seen to penetrate down into the helium-rich layers. This is due to a combination of the continuing expansion and cooling of these layers, which increases their opacity, and the growing energy flux produced by the He-burning shell – note that the luminosity has been steadily growing. For lower-mass stars the H-burning shell remains active at a low level, which prevents the convective envelope from penetrating deeper into the star. Consequently, the second dredge-up does not occur in lower-mass stars.

In the material that is dredged up hydrogen has been burned into helium, while ^{12}C and ^{16}O have been almost completely converted into ^{14}N by the CNO-cycle. The amount of He- and N-rich material dredged up is about 0.2 $M_\odot$ in the example shown, and can be as much as 1 $M_\odot$ in the most massive AGB stars. This material is mixed with the outer convective envelope and appears at the surface.

Thus the second dredge-up has a qualitatively similar, but much more dramatic effect that the first dredge-up phase that occurred on the RGB.

An additional important effect of the second dredge-up is the reduction of the mass of the H-exhausted core, thus limiting the mass of the white dwarf that remains. Effectively, the occurrence of second dredge-up thus increases the upper initial mass limit, M_{up}, of stars that produce white dwarfs.

The thermally pulsing AGB phase

As the He-burning shell approaches the H-He discontinuity, its luminosity decreases as it runs out of fuel. The layers above then contract somewhat in response, thus heating the extinguished H-burning shell until it is re-ignited. Both shells now provide energy and a phase of *double shell burning* begins. However, the shells do not burn at the same pace: the He-burning shell becomes thermally unstable and undergoes periodic *thermal pulses*, discussed in detail in Sec. 11.1.1. This phase is thus referred to as the *thermally pulsing AGB* (TP-AGB).

The structure of a star during the TP-AGB phase is schematically depicted in Fig. 11.2. The thermally pulsing phase of the AGB has a number of salient properties:

- The periodic thermal pulses alternate with mixing episodes and give rise to a unique *nucleosynthesis* of (among others) ^{12}C, ^{14}N, and elements heavier than iron (Sec. 11.1.2). This process gradually makes the stellar envelope and atmosphere more carbon-rich.

- Similar to the RGB, the stellar properties mainly depend on the size of the degenerate CO core. In particular there is a tight *core mass-luminosity* relation,

$$L = 5.9 \times 10^4 L_\odot \left(\frac{M_c}{M_\odot} - 0.52 \right), \tag{11.1}$$

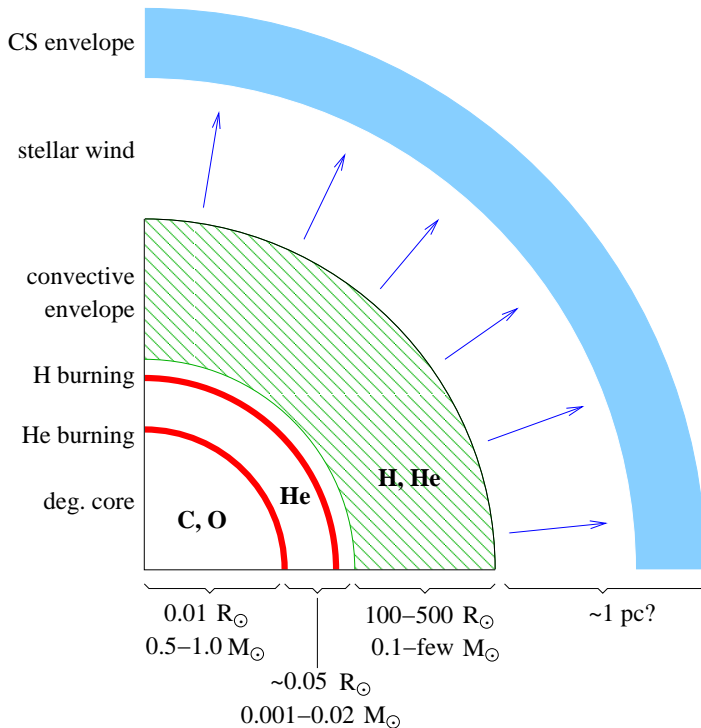which is not as steep as the RGB relation (10.2).

Figure 11.2. Schematic structure of an AGB star during its thermally pulsing phase. The CO core is degenerate and very compact, and is surrounded by two burning shells very close together in mass coordinate. The convective envelope by contrast is very extended and tenuous, having a radius 10^4–10^5 times the size of the core. This loosely bound envelope is gradually eroded by the strong stellar wind, which forms a dusty circumstellar envelope out to several hundreds of stellar radii. The convective envelope, stellar atmosphere and circumstellar envelope have a rich and changing chemical composition driven by nucleosynthesis processes in the burning shells in the deep interior.

- Strong *mass loss* ($10^{-7} - 10^{-4} \, M_\odot$/yr), probably driven by dynamical (Mira) pulsations combined with radiation pressure on dust particles formed in the cool atmosphere (Sec. 11.1.3), gradually removes the envelope and replenishes the interstellar medium with the synthesized elements.

- The extended stellar atmosphere and circumstellar envelope, formed by the outflow, have a rich molecular and dust chemistry. This is mainly revealed in their infra-red spectra, which have been observed by space telescope missions such as ISO and Spitzer.

11.1.1 Thermal pulses and dredge-up

After the H-burning shell is reignited, the He-burning shell that lies underneath it becomes geometrically thin. Nuclear burning in such a thin shell is thermally unstable, for the reasons discussed in Sect. 7.5.2. This gives rise to periodic *thermal pulses* of the He-burning shell. What happens during a thermal pulse cycle is depicted schematically in Fig. 11.3.

- For most of the time, the He-burning shell is inactive. The H-burning shell adds mass to the He-rich region between the burning shells (the intershell region), which increases the pressure and temperature at the bottom of this region.

- When the mass of the intershell region reaches a critical value, helium is ignited in an unstable manner, giving rise to a thermonuclear runaway called a *helium shell flash*. (Note the difference with the *core* He flash in low-mass red giants, where electron degeneracy causes the thermonuclear runaway.) Values of $L_{He} \approx 10^8 \, L_\odot$ are reached during ~ 1 year. The large energy flux drives convection in the whole intershell region (producing an *intershell convection zone*, ICZ).
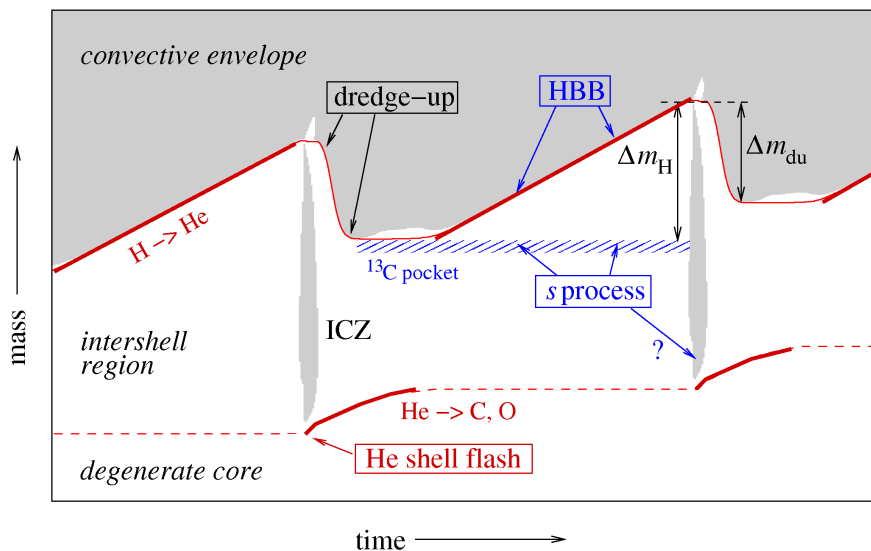

Figure 11.3. Schematic evolution of an AGB star through two thermal-pulse cycles. Convective regions are shown as gray shaded areas, where 'ICZ' stands for the intershell convection zone driven by the He-shell flash. The H-exhausted core mass is shown as a thin red solid line and the He-exhausted core mass as a dashed line. Thick red lines indicate when nuclear burning is active in these shells. Only the region around the two burning shells is shown, comprising $\sim 0.01 \, M_\odot$. The hatched region indicates a shell or 'pocket' rich in ^{13}C that is formed at the interface of the H-rich envelope and the C-rich intershell region, following a dredge-up episode. Note that the time axis is highly non-linear: the He shell-flash and dredge-up phases (lasting ~ 100 years) are expanded relative to the interpulse phase ($10^4 - 10^5$ years).
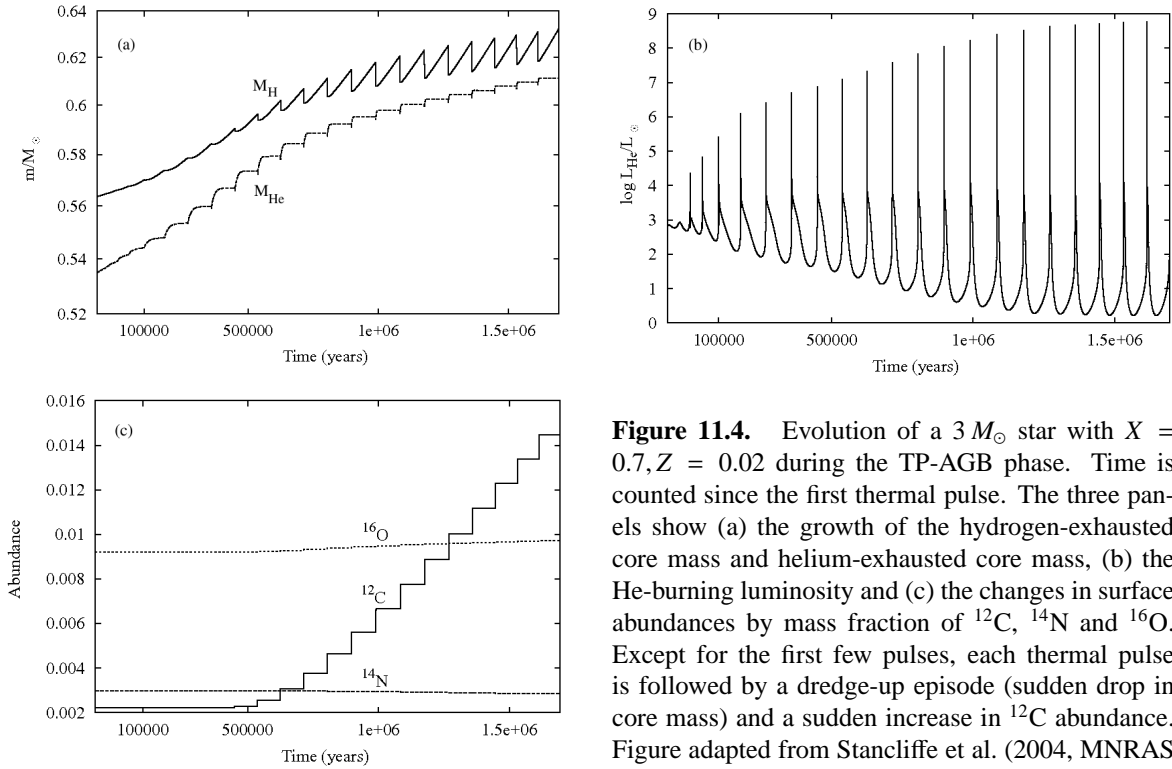
Figure 11.4. Evolution of a 3 $M_\odot$ star with $X = 0.7, Z = 0.02$ during the TP-AGB phase. Time is counted since the first thermal pulse. The three panels show (a) the growth of the hydrogen-exhausted core mass and helium-exhausted core mass, (b) the He-burning luminosity and (c) the changes in surface abundances by mass fraction of ^{12}C, ^{14}N and ^{16}O. Except for the first few pulses, each thermal pulse is followed by a dredge-up episode (sudden drop in core mass) and a sudden increase in ^{12}C abundance. Figure adapted from Stancliffe et al. (2004, MNRAS 352, 984).

This mixes ^{12}C produced by the 3α reaction, as well as other elements produced during He burning, throughout the intershell region.

- The large energy release by the He-shell flash mostly goes into expansion of the intershell region against the gravitational potential. This eventually allows the He-burning shell to expand and cool as well, so that the He-shell flash dies down after several years. A phase of stable He-shell burning follows which lasts up to a few hundred years. As a result of the expansion and cooling of the intershell region after the He-shell flash, the H-burning shell extinguishes.

- Expansion and cooling of the intershell region can also lead to a deeper penetration of the outer convective envelope. In some cases convection can penetrate beyond the now extinct H-burning shell, such that material from the intershell region is mixed into the outer envelope. This phenomenon is called *third dredge-up*. Note that this term is used even for stars that do not experience the second dredge-up, and is used for all subsequent dredge-up events following further thermal pulses. Helium as well as the products of He burning, in particular ^{12}C, can thus appear at the surface.

- Following third dredge-up, the H-burning shell is reignited and the He-burning shell becomes inactive again. A long phase of stable H-shell burning follows in which the mass of the intershell region grows until the next thermal pulse occurs. The duration of this *interpulse period* depends on the core mass, lasting between 50,000 yrs (for low-mass AGB stars with CO cores of $\sim 0.5\,M_\odot$) to < 1000 yrs for the most massive AGB stars.

This thermal pulse cycle can repeat many times, as shown for a 3 $M_\odot$ AGB star in Fig. 11.4. The pulse amplitude (the maximum helium-burning luminosity) increases with each pulse, which facilitates dredge-up after several thermal pulses. In the example shown, third dredge-up first occurs after the 7th thermal pulse ($\sim 5 \times 10^5$ yr after the start of the TP-AGB phase) and then follows after

every subsequent pulse. The efficiency of dredge-up is often measured by a parameter λ, which is defined as the ratio of the mass dredged up into the envelope over the mass by which the H-exhausted core has grown during the preceding interpulse period (see Fig. 11.3),

$$\lambda = \frac{\Delta M_{du}}{\Delta M_H}. \tag{11.2}$$

Third dredge-up has two important consequences. First, unlike the first and second dredge-up which only mix up H-burning products, the third dredge-up bring products of *helium burning* to the surface. This leads to important nucleosynthesis (see Sec. 11.1.2). Second, third dredge-up limits the growth of the CO core mass. Efficient dredge-up with $\lambda \approx 1$ means that in the long run, the core mass does not increase.

11.1.2 Nucleosynthesis and abundance changes on the AGB

The main effect of thermal pulses and third dredge-up operating in AGB stars is the appearance of helium-burning products at the surface, in particular a large production of carbon. In the $3\,M_\odot$ model shown in Fig. 11.4, the surface ^{12}C abundance increases after every dredge-up episode and thus gradually increases, until it exceeds the ^{16}O abundance after 1.3×10^6 yr.

At the low temperatures in the stellar atmosphere, most of the C and O atoms are bound into CO molecules, such that the spectral features of AGB stars strongly depend on the C/O number ratio. If $n(C)/n(O) < 1$ (simply written as 'C/O < 1'), then the remaining O atoms form *oxygen-rich* molecules and dust particles, such as TiO, H_2O and silicate grains. The spectra of such O-rich AGB stars are classified as type M or S. As a result of repeated dredge-ups, at some point the C/O ratio can exceed unity. If C/O > 1 then all O is locked into CO molecules and the remaining C forms *carbon-rich* molecules and dust grains, e.g. C_2, CN and carbonaceous grains like graphite. Such more evolved AGB stars are classified as *carbon stars* with spectral type C.

Besides carbon, the surface abundances of many other elements and isotopes change during the TP-AGB phase. The direct evidence for active nucleosynthesis in AGB stars was the detection in 1953 of technetium, an element with only radioactive isotopes of which the longest-lived one (^{99}Tc) decays on a timescale of 2×10^5 yrs. AGB stars are nowadays considered to be major producers in the Universe of carbon, nitrogen and of elements heavier than iron by the *s-process*. They also make an important contribution to the production of ^{19}F, ^{25}Mg, ^{26}Mg and other isotopes.

Production of heavy elements: the s-process

Spectroscopic observations show that many AGB stars are enriched in elements heavier than iron, such as Zr, Y, Sr, Tc, Ba, La and Pb. These elements are produced via slow neutron capture reactions on Fe nuclei, the so-called *s-process*. In this context 'slow' means that the time between successive neutron captures is long compared to the β-decay timescale of unstable, neutron-rich isotopes.

The synthesis of s-process elements requires a source of free neutrons, which can be produced in the He-rich intershell region by either of two He-burning reactions: ^{13}C$(\alpha, n)^{16}$O and ^{22}Ne$(\alpha, n)^{25}$Mg. The latter reaction can take place during the He-shell flash if the temperature exceeds 3.5×10^8 K, which is only reached in rather massive AGB stars. The ^{22}Ne required for this reaction is abundant in the intershell region, because the ^{14}N that is left by the CNO-cycle is all converted into ^{22}Ne by He-burning: ^{14}N$(\alpha, \gamma)^{18}$F$(\beta^+)^{18}$O$(\alpha, \gamma)^{22}$Ne.

The main neutron source in low-mass stars (up to $3\,M_\odot$) is probably the ^{13}C$(\alpha, n)^{16}$O reaction. The current idea is that a thin shell or 'pocket' of ^{13}C is formed (shown as a hatched region in Fig. 11.3) by partial mixing of protons and ^{12}C at the interface between the H-rich envelope and the C-rich intershell region, which produces ^{13}C by the first step of the CN-cycle. The ^{13}C subsequently reacts

with helium when the temperature reaches 10^8 K, releasing the required neutrons. The s-enriched pocket is ingested into the ICZ during the next pulse, and mixed throughout the intershell region, together with carbon produced by He burning. The carbon and s-process material from the intershell region is subsequently mixed to the surface in the next dredge-up phase (see Fig. 11.3).

Hot bottom burning

In stars with $M \gtrsim 4-5\,M_\odot$, the temperature at the base of the convective envelope during the interpulse period becomes so high ($T_{\mathrm{BCE}} \gtrsim 3 \times 10^7$ K) that H-burning reactions take place. The CNO cycle then operates on material in the convective envelope, a process known as *hot bottom burning*. Its main effects are: (1) an increase in the surface luminosity, which breaks the core mass-luminosity relation; (2) the conversion of dredged-up ^{12}C into ^{14}N, besides many other changes in the surface composition. Hot bottom burning thus prevents massive AGB stars from becoming carbon stars, and turns such stars into efficient producers of *nitrogen*. Other nuclei produced during hot bottom burning are ^{7}Li, ^{23}Na, and ^{25,26}Mg.

11.1.3 Mass loss and termination of the AGB phase

Once a star enters the TP-AGB phase it can experience a large number of thermal pulses. The number of thermal pulses and the duration of the TP-AGB phase is limited by (1) the decreasing mass of the H-rich envelope and (2) the growing mass of the degenerate CO core. If the CO core mass is able to grow close to the *Chandrasekhar mass*, $M_{\mathrm{Ch}} \approx 1.46\,M_\odot$, carbon will be ignited in the centre in a so-called 'carbon flash' that has the power to disrupt the whole star (see Chapter 13). However, white dwarfs are observed in rather young open clusters that still contain massive main-sequence stars. This tells us that the carbon flash probably never happens in AGB stars, even when the total mass is $8\,M_\odot$, much larger than M_{Ch}. The reason is that *mass loss* becomes so strong on the AGB that the entire H-rich envelope can be removed before the core has had time to grow significantly. The lifetime of the TP-AGB phase, $1 - 2 \times 10^6$ yr, is essentially determined by the mass-loss rate.
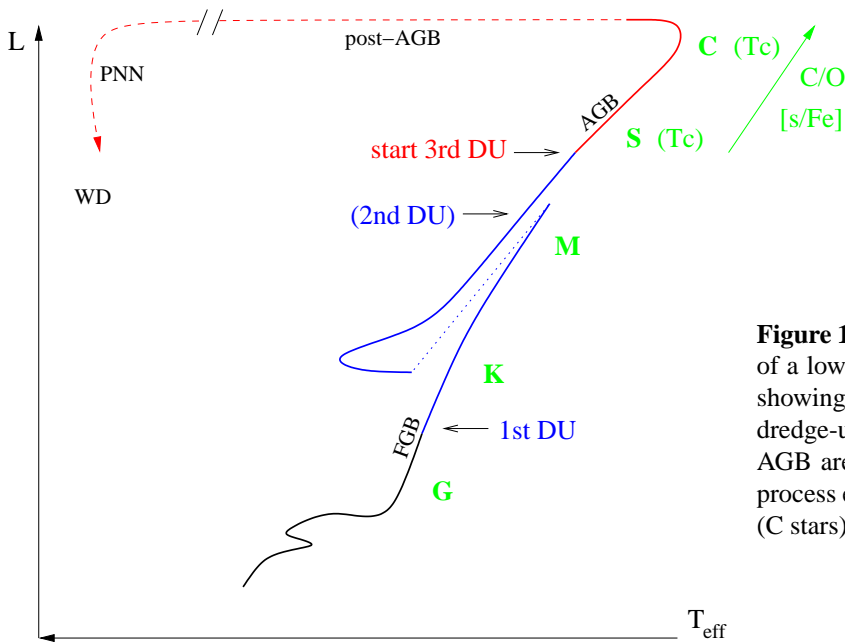

Figure 11.5. Schematic evolution track of a low-mass star in the H-R diagram, showing the occurrence of the various dredge-up episodes. Stars on the upper AGB are observed to be enriched in s-process elements (S stars) and in carbon (C stars).
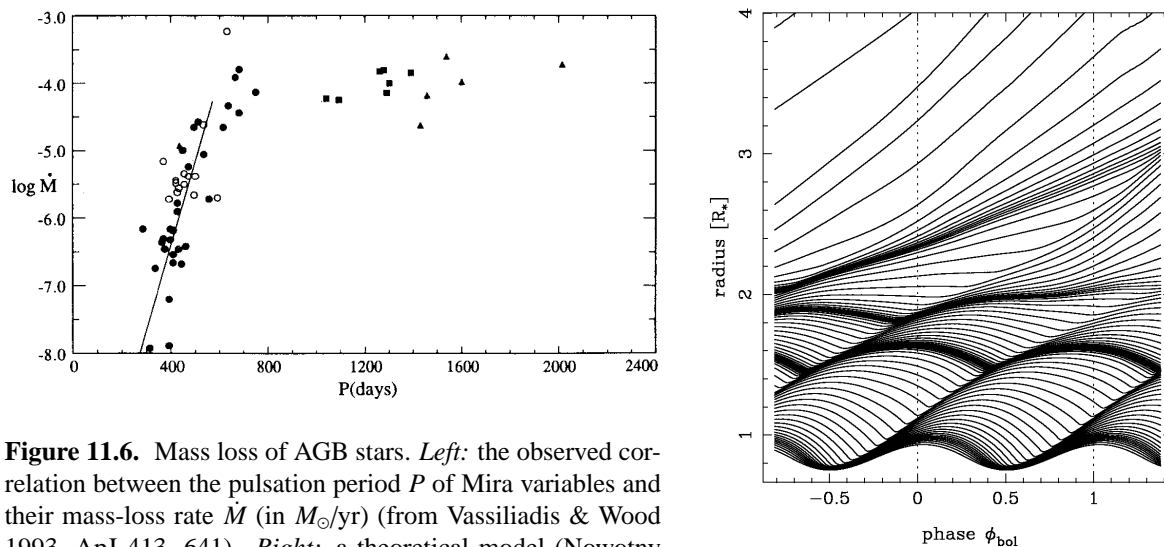
167

Figure 11.6. Mass loss of AGB stars. *Left:* the observed correlation between the pulsation period P of Mira variables and their mass-loss rate $\dot{M}$ (in $M_\odot$/yr) (from Vassiliadis & Wood 1993, ApJ 413, 641). *Right:* a theoretical model (Nowotny et al. 2005, A&A 437, 273) showing streamlines in the outer atmosphere of an AGB star undergoing radial pulsations. At $r \gtrsim 2 R_*$ dust particles form in the dense shocked regions and radiation pressure on the dust then pushes the mass out.

AGB mass loss

That AGB stars have strong stellar winds is clear from their spectral energy distributions, which show a large excess at infrared wavelengths. Many AGB stars (known as OH/IR stars) are even completely enshrouded in a dusty circumstellar envelope and are invisible at optical wavelengths. The mechanisms driving such strong mass loss are not yet completely understood, but a combination of dynamical *pulsations* and *radiation pressure* on dust particles formed in the atmosphere probably plays an essential role. Stars located on the AGB in the H-R diagram are found to undergo strong radial pulsations, they are known as *Mira* variables (see Fig. 10.11). An observational correlation exists between the pulsation period and the mass-loss rate, shown in Fig. 11.6a. As a star evolves towards larger radii along the AGB, the pulsation period increases and so does the mass-loss rate, from $\sim 10^{-8}\,M_\odot$/yr to $\sim 10^{-4}\,M_\odot$/yr for pulsation periods in excess of about 600 days.

The basic physical picture is illustrated in Fig. 11.6b. The pulsations induce shock waves in the stellar atmosphere, which brings gas out to larger radii and thus increases the gas density in the outer atmosphere. At about $1.5 - 2$ stellar radii, the temperature is low enough ($\sim 1500\,\mathrm{K}$) that dust particles can condense. The dust particles are very opaque and, once they have formed, can easily be accelerated by the radiation pressure that results from the high stellar luminosity. In the absence of pulsations, the gas density at such a distance from the star would be too low to form dust. Even though the gas in the atmosphere is mostly in molecular form (H_2, CO, etc.) and the dust fraction is only about 1%, the molecular gas is dragged along by the accelerated dust particles resulting in a large-scale outflow.

Observationally, the mass-loss rate levels off at a maximum value of $\sim 10^{-4}\,M_\odot$/yr (this is the value inferred for dust-enshrouded OH/IR stars, the stars with the largest pulsation periods in Fig. 11.6). This phase of very strong mass loss is sometimes called a 'superwind'. Once an AGB star enters this superwind phase, the H-rich envelope is rapidly removed. This marks the end of the AGB phase. The high mass-loss rate during the superwind phase therefore determines both the maximum luminosity that a star can reach on the AGB, and its final mass, i.e. the mass of the white-dwarf remnant (Fig. 11.7).

Post-AGB evolution

When the mass of the H-rich envelope becomes very small, $10^{-2} - 10^{-3}\,M_\odot$ depending on the core mass, the envelope shrinks and the star leaves the AGB. The resulting decrease in stellar radius occurs at almost constant luminosity, because the H-burning shell is still fully active and the star keeps following the core mass-luminosity relation. The star thus follows a horizontal track in the H-R diagram towards higher effective temperatures. This is the *post-AGB* phase of evolution. Note that the star remains in complete equilibrium during this phase: the evolution towards higher $T_{\rm eff}$ is caused by the decreasing mass of the envelope, which is eroded at the bottom by H-shell burning and at the top by continuing mass loss. The typical timescale for this phase is $\sim 10^4$ yrs.

As the star gets hotter and $T_{\rm eff}$ exceeds 30,000 K, two effects start happening: (1) the star develops a weak but fast wind, driven by radiation pressure in UV absorption lines (similar to the winds of massive OB-type stars, see Sec. 12.1); and (2) the strong UV flux destroys the dust grains in the circumstellar envelope, dissociates the molecules and finally ionizes the gas. Part of the circumstellar envelope thus becomes ionized (an HII region) and starts radiating in recombination lines, appearing as a *planetary nebula*. Current ideas about the formation of planetary nebulae are that they result from the interaction between the slow AGB wind and the fast wind from the central star, which forms a compressed optically thin shell from which the radiation is emitted.

When the envelope mass has decreased to $10^{-5}\,M_\odot$, the H-burning shell is finally extinguished. This happens when $T_{\rm eff} \approx 10^5$ K and from this point the luminosity starts decreasing. The remnant now cools as a white dwarf. In some cases the star can still experience a final thermal pulse during its post-AGB phase (a *late thermal pulse*), or even during the initial phase of white dwarf cooling (a
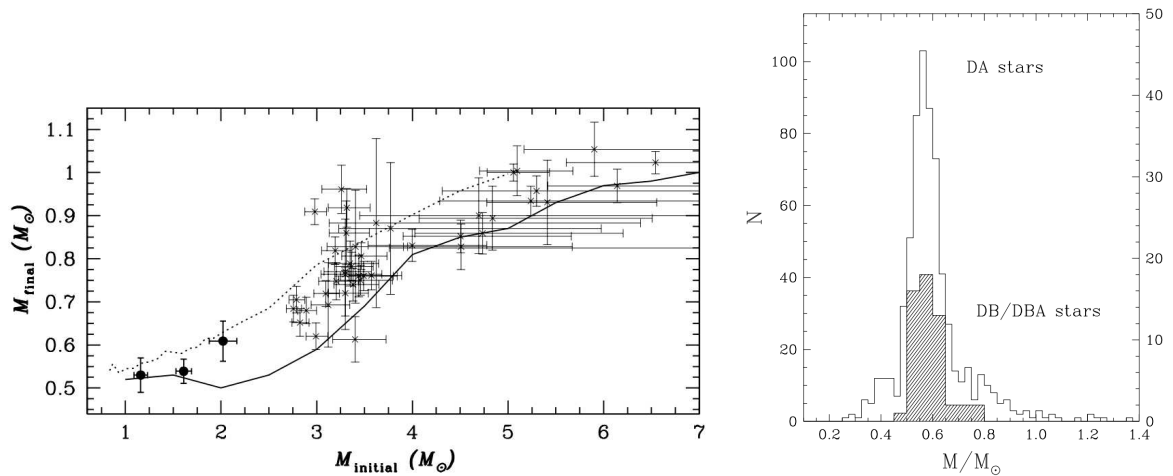

Figure 11.7. *Left:* Relation between between the initial and final mass of low- and intermediate-mass stars, from Kalirai et al. (2008, ApJ 676, 594). The data points represent white dwarfs observed in open clusters, for which the mass has been determined from their spectra. The age of the cluster $t_{\rm cl}$ and the cooling time of the white dwarf $t_{\rm wd}$ have been used to estimate the initial mass, because $t_{\rm cl} - t_{\rm wd}$ corresponds to the lifetime of the progenitor star. The solid line shows model predictions for the core mass of a star at the start of the TP-AGB phase (from Marigo 2001, A&A 370, 194) for solar metallicity. The dotted line shows the final mass of these models, which is reasonably consistent with the data points. The growth of the core mass on the AGB is severely limited by dredge-up and strong mass loss.

Right: Observed mass distribution of white dwarfs, for a large sample of DA white dwarfs and a smaller sample of DB white dwarfs (from Bergeron et al. 2007). There is a sharp peak between 0.55 and $0.6\,M_\odot$, as can be expected from the initial-final mass relation because most white dwarfs come from low-mass stars with $M < 2\,M_\odot$.

very late thermal pulse). This can temporarily bring the star back to the AGB (sometimes referred to as the 'born-again AGB' scenario).

11.2 White dwarfs

All stars with initial masses up to about $8\,M_\odot$ develop electron-degenerate cores and lose their envelopes during the AGB phase, and thus end their lives as white dwarfs. Nuclear fusion no longer provides energy and white dwarfs shine by radiating the thermal energy stored in their interiors, cooling at almost constant radius and decreasing luminosities. The faintest white dwarfs detected have $L \approx 10^{-4.5}\,L_\odot$. Observed WD masses are mostly in a narrow range around $0.6\,M_\odot$, see Fig. 11.7b, which corresponds to the CO core mass of low-mass ($\lesssim 2\,M_\odot$) AGB progenitors. This sharply peaked mass distribution, along with the observationally induced initial-to-final mass relation (Fig. 11.7a), are further evidence that AGB mass loss is very efficient at removing the stellar envelope.

The great majority of white dwarfs are indeed composed of C and O. Those with $M < 0.45\,M_\odot$ are usually He white dwarfs, formed by a low-mass star that lost its envelope already on the RGB. This is not expected to happen in single stars, but can result from binary interaction and indeed most low-mass WDs are found in binary systems. White dwarfs with $M > 1.2\,M_\odot$, on the other hand, are mostly ONe white dwarfs. They result from stars that underwent carbon burning in the core but developed degenerate ONe cores, which is expected to happen in a small initial mass range around $8\,M_\odot$.

The *surface* composition of white dwarfs is usually completely different than their interior composition. The strong surface gravity has resulted in separation of the elements, such that any hydrogen left is found as the surface layer while all heavier elements have settled at deeper layers. Most white dwarfs, regardless of their interior composition, therefore show spectra completely dominated by H lines and are classified as DA white dwarfs. A minority of white dwarfs show only helium lines and have spectroscopic classification DB. These have lost all hydrogen from the outer layers during their formation process, probably as a result of a late or very late thermal pulse.

11.2.1 Structure of white dwarfs

As discussed earlier, the equation of state of degenerate matter is independent of temperature, which means that the mechanical structure of a white dwarf is independent of its thermal properties. As a white dwarf cools, its radius therefore remains constant. As long as the electrons are non-relativistic the structure of a white dwarf can be described as a $n = \frac{3}{2}$ polytrope with constant K. Such stars follow a mass-radius relation of the form $R \propto M^{-1/3}$, depicted in Fig. 11.8 as a dashed line. A proper theory for WDs should take into account that the most energetic electrons in the Fermi sea can move with relativistic speeds, even in fairly low-mass white dwarfs. This means that the equation of state is generally not of polytropic form, but the relation $P(\rho)$ has a gradually changing exponent between $\frac{5}{3}$ and $\frac{4}{3}$, as shown in Fig. 3.3. The pressure in the central region is therefore somewhat smaller than that of a purely non-relativistic electron gas. Thus WD radii are smaller than given by the polytropic relation, the difference growing with increasing mass (and increasing central density). The relativistic theory, worked out by Chandrasekhar, predicts the mass-radius relation shown as a solid line in Fig. 11.8. As the mass approaches the Chandrasekhar mass, given by eq. (4.22),

$$M_{\mathrm{Ch}} = 1.459 \left(\frac{2}{\mu_{\mathrm{e}}}\right)^2 M_\odot, \qquad (11.3)$$

the radius goes to zero as all electrons become extremely relativistic. White dwarfs more massive than M_{Ch} must collapse as the relativistic degeneracy pressure is insufficient to balance gravity.
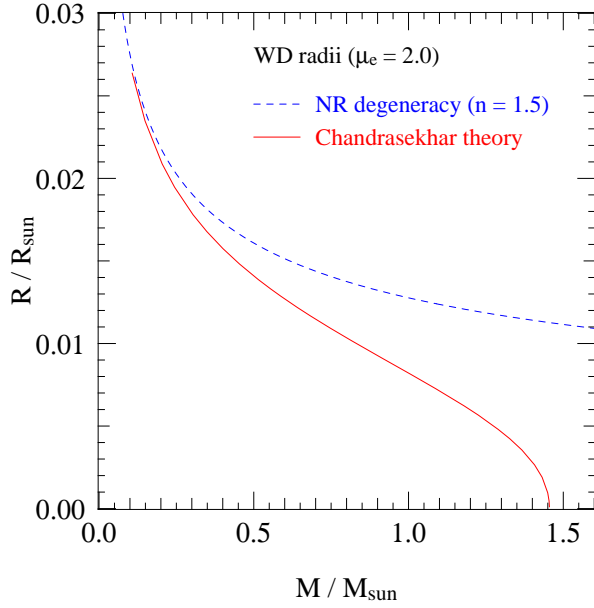
Figure 11.8. Comparison of the radius-mass relation of a completely degenerate star computed using Chandrasekhar's theory for white dwarfs (taking into account the partly relativistic velocities of the electrons in the Fermi sea) and an approximation based on non-relativistic degeneracy.

Chandrasekhar's white dwarf theory assumes the electrons are fully degenerate and non-interacting. In reality, certain corrections have to be made to the structure, in particular *electrostatic interactions* between the electrons and ions (see Sec. 3.6.1). These give a negative correction to the electron pressure, leading to a somewhat smaller radius at a particular mass. Furthermore, at high densities *inverse β-decays* become important. Examples are the reactions

$$^{24}\text{Mg} + e^- \rightarrow {}^{24}\text{Na} + \nu, \qquad {}^{24}\text{Na} + e^- \rightarrow {}^{24}\text{Ne} + \nu.$$

A neutron-rich nucleus such as ^{24}Na is normally unstable to β-decay ($^{24}\text{Na} \rightarrow {}^{24}\text{Mg} + e^- + \bar{\nu}$), but at high density is stabilized by the Fermi sea of energetic electrons: the decay is prevented because the energy of the released electron is lower than the Fermi energy. Reactions such as these (also called *electron captures*) decrease the electron pressure at high density. Their main effect is a lowering of the effective Chandrasekhar mass, from the 'ideal' value of 1.459 $M_\odot$ for a CO white dwarf to 1.4 $M_\odot$.

11.2.2 Thermal properties and evolution of white dwarfs

In the interior of a white dwarf, the degenerate electrons provide a high thermal conductivity (Sec. 5.2.4). This leads to a very small temperature gradient, especially because L is also very low. The degenerate interior can thus be considered to a have a constant temperature. However, the outermost layers have much lower density and are non-degenerate, and here energy transport is provided by radiation. Due to the high opacity in these layers, radiation transport is much less effective than electron conduction in the interior. The non-degenerate outer layers thus act to insulate the interior from outer space, and here a substantial temperature gradient is present.

We can obtain a simple description by starting from the radiative envelope solutions discussed in Sec. 7.2.3, assuming an ideal gas and a Kramers opacity law $\kappa = \kappa_0 \rho T^{-7/2}$, and assuming P and T approach zero at the surface:

$$T^{17/2} = B P^2 \qquad \text{with} \qquad B = \frac{17}{4} \frac{3}{16\pi a c G} \frac{\kappa_0 \mu}{\mathcal{R}} \frac{L}{M}. \tag{11.4}$$

Replacing $P = (\mathcal{R}/\mu)\rho T$ and solving for ρ, we find that within the non-degenerate envelope

$$\rho = B^{-1/2} \frac{\mu}{\mathcal{R}} T^{13/4}. \tag{11.5}$$
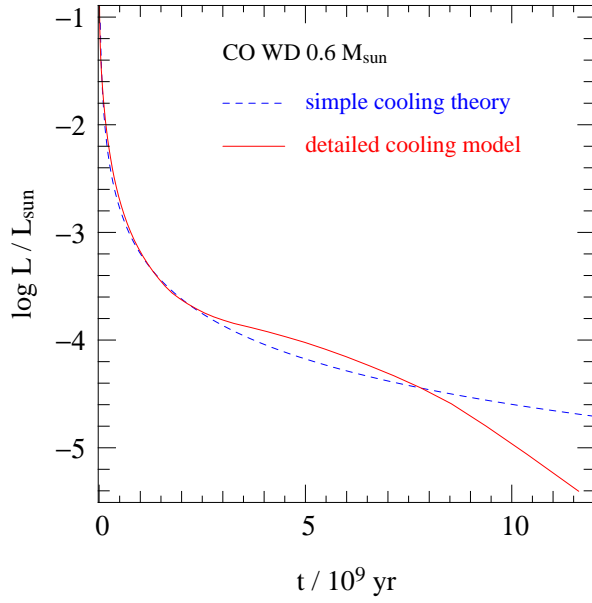
171

Figure 11.9. Theoretical cooling curves for a CO white dwarf with a typical mass of $0.6\,M_\odot$. The dashed (blue) line shows the evolution of luminosity with time based on the simple cooling theory by Mestel, which yields $L \propto t^{-7/5}$. The solid (red) line is a detailed cooling model for a DA white dwarf by M. Wood (1995, LNP 443, 41). This model takes into account (among other things) the effect of *crystallization*, a phase transition that releases an additional amount of energy, visible as the slowing down of the cooling after about 2 Gyr. When crystallization is almost complete after about 7 Gyr, the cooling speeds up again.

Let us assume that the transition point with the degenerate interior is located where the degenerate electron pressure equals the ideal-gas pressure of the *electrons* in the envelope, $P_e = (\mathcal{R}/\mu_e)\rho T$, since the ions are non-degenerate everywhere. At this point, denoted with subscript 'b', we have

$$\frac{\mathcal{R}}{\mu_e}\rho_b T_b = K_{NR}\left(\frac{\rho_b}{\mu_e}\right)^{5/3}.$$

T_b and ρ_b must match the value given by eq. (11.5) at the transition point. Eliminating ρ_b gives

$$T_b^{7/2} = \frac{\mathcal{R}^5\mu_e^2}{K_{NR}^3\mu^2}\,B = \frac{51\mathcal{R}^4}{64\pi acGK_{NR}^3}\kappa_0\frac{\mu_e^2}{\mu}\frac{L}{M}. \tag{11.6}$$

Since the degenerate interior is nearly isothermal, T_b is approximately the temperature of the entire interior or 'core' of the white dwarf. We can thus write (11.6) as $T_c^{7/2} = \alpha L/M$. To evaluate the proportionality constant α we have to substitute appropriate values for κ_0 and the composition (μ_e and μ), which is somewhat arbitrary. Assuming bound-free absorption (eq. 5.33) and $\mu_e = 2$ in the envelope, which is reasonable because the envelope is H-depleted except for the very surface layers, we get $\alpha \approx 1.38 \times 10^{29}\,Z/\mu$ in cgs units. In a typical DA white dwarf, most of the non-degenerate layers are helium-rich so assuming $Z = 0.02$ and $\mu \approx 1.34$ is reasonable. With these assumptions we obtain the following relation between the temperature in the interior and the luminosity and mass of the white dwarf,

$$T_c \approx 7.7 \times 10^7\,\text{K}\left(\frac{L/L_\odot}{M/M_\odot}\right)^{2/7}. \tag{11.7}$$

The typical masses and luminosities of white dwarfs, $M \approx 0.6\,M_\odot$ and $L < 10^{-2}\,L_\odot$, imply 'cold' interiors with $T < 2 \times 10^7$ K.

We can use these properties of white dwarfs to obtain a simple model for their cooling, i.e. the change in WD luminosity with time. Since there are no nuclear energy sources, the virial theorem applied to degenerate objects tells us that the luminosity radiated away comes from the decrease of internal energy. Since the electrons fill their lowest energy states up to the Fermi level, their internal energy cannot change and neither can energy be released by contraction. The only source of energy
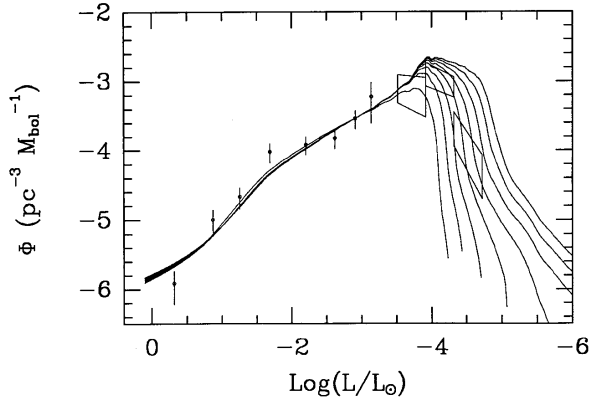
Figure 11.10. Observed and theoretical distributions of white dwarf luminosities in the Galactic disk, from Wood (1992, ApJ 386, 539), based on cooling models similar to the one shown in Fig. 11.9. The curves are for assumed ages of the Galactic disk between 6 and 13 Gyr. The paucity of observed white dwarfs with $\log(L/L_\odot) < -4.3$, shown as a slanted box, implies an age of the local Galactic disk of 8–11 Gyr.

available is the thermal energy stored in the non-degenerate ions, that make up the bulk of the mass of the white dwarf. Since the interior is isothermal at temperature T_c, the total thermal energy is

$$E_{in} = c_V M T_c, \tag{11.8}$$

where c_V is the specific heat per unit mass. For ions behaving as an ideal gas we have $c_V = \frac{3}{2}\mathcal{R}/\mu_{ion}$ which is a constant. The luminosity is thus given by

$$L = -\frac{dE_{in}}{dt} = -c_V M \frac{dT_c}{dt}, \tag{11.9}$$

where L is related to M and T_c by eq. (11.6). If we write this relation as $T_c^{7/2} = \alpha L/M$ we obtain

$$T_c^{7/2} = -\alpha c_V \frac{dT_c}{dt},$$

which can be easily integrated between an initial time t_0, when the white dwarf forms, and a generic time t to give

$$\tau \equiv t - t_0 = \frac{2}{5}\alpha c_V (T_c^{-5/2} - T_{c,0}^{-5/2}). \tag{11.10}$$

Once the white dwarf has cooled significantly, its core temperature is much smaller than the initial value so that $T_{c,0}^{-5/2}$ can be neglected. We thus obtain a simple relation between the cooling time τ of a white dwarf and its core temperature, and thus between τ and the luminosity,

$$\tau \approx \frac{2}{5}\alpha c_V T_c^{-5/2} = \frac{2}{5}c_V \alpha^{2/7} \left(\frac{L}{M}\right)^{-5/7}. \tag{11.11}$$

Making the same assumptions in calculating α as in eq. (11.7), and substituting $c_V = \frac{3}{2}\mathcal{R}/\mu_{ion}$, we can write this relation as

$$\tau \approx \frac{1.05 \times 10^8 \text{ yr}}{\mu_{ion}} \left(\frac{L/L_\odot}{M/M_\odot}\right)^{-5/7}. \tag{11.12}$$

This approximate cooling law was derived by Mestel. It shows that more massive white dwarfs evolve more slowly, because more ionic thermal energy is stored in their interior. Also, increasing the mean mass of the ions μ_{ion} in a white dwarf of the same total mass decreases the cooling time, because there are fewer ions per unit mass storing heat. For a CO white dwarf composed in equal parts of C and O, $\mu_{ion} \approx 14$.

173

This simple cooling law, depicted in Fig. 11.9 for a $0.6\,M_\odot$ CO white dwarf, predicts cooling times greater than 1 Gyr when $L < 10^{-3}\,L_\odot$, and greater than the age of the Universe when $L < 10^{-5}\,L_\odot$. More realistic models take into account the effect of contraction of the non-degenerate envelope, which provides some additional energy during the initial cooling phase, and more importantly, the effects of Coulomb interactions and of *crystallization* in particular. As the ion gas cools, electrostatic interactions become more important (Sec. 3.6.1) and the ions settle into a lattice structure. This releases latent heat (in other words, $c_V > \frac{3}{2}\mathcal{R}/\mu_{\text{ion}}$) and the cooling is correspondingly slower than given by the Mestel law. Once crystallization is almost complete, c_V decreases and cooling speeds up again. A more detailed WD cooling model that includes these effects is shown in Fig. 11.9. White dwarfs that have cooled for most of the age of the Universe cannot have reached luminosities much less than $10^{-5}\,L_\odot$ and should still be detectable. Observed white dwarf luminosities thus provide a way to derive the age of a stellar population (e.g. see Fig. 11.10).

Suggestions for further reading

The evolution of AGB stars is treated in Chapter 26.6–26.8 of MAEDER and Chapter 33.2–33.3 of KIPPENHAHN & WEIGERT. White dwarfs are discussed in more detail in Chapter 35 of KIPPENHAHN & WEIGERT and Chapter 7.4 of SALARIS & CASSISI.

Exercises

11.1 Core mass luminosity relation for AGB stars

The luminosity of an AGB star is related to its core mass via the Paczynski relation (11.1). The nuclear burning in the H- and He-burning shells add matter to the core at a rate of $\dot{M}_c/M_\odot = 1.0\times 10^{-11}(L_*/L_\odot)$. Assume that a star enters the AGB with a luminosity of $10^3\,L_\odot$ and a total mass of $2\,M_\odot$.

  (a) Derive an expression for the luminosity as a function of time after the star entered the AGB phase.

  (b) Assume that T_{eff} remains constant at 3000 and derive an expression for the radius as a function of time.

  (c) Derive an expression for the core-mass as a function of time.

11.2 Mass loss of AGB stars

The masses of white dwarfs and the luminosity on the tip of the AGB are completely determined by mass loss during the AGB phase. The mass loss rate is very uncertain, but for this exercise assume that the mass loss rate is given by the Reimers relation, eq. (10.3), with $\eta \approx 3$ for AGB stars. Now, also assume that a star entered the AGB phase with a mass of $2\,M_\odot$ and a luminosity of $10^3\,L_\odot$.

  (a) Derive an expression for the mass of the star as a function of time, using $L(t)$ and $R(t)$ from Exercise 11.1. (Hint: $-\dot{M}M = 0.5\,\mathrm{d}(M^2)/\mathrm{d}t$).

  (b) Use the expression from (a) and the one for $M_c(t)$ from Exercise 11.1 to derive:
  - the time when the star leaves the AGB ($M_{\text{env}} \simeq 0$).
  - the luminosity at the tip of the AGB.
  - the mass of the resulting white dwarf. (This requires a numerical solution of a simple equation).

  (c) Derive the same quantities in the cases when the mass loss rate on the AGB is three times larger, i.e., $\eta = 9$, and when it is three times smaller, i.e., $\eta = 1$.