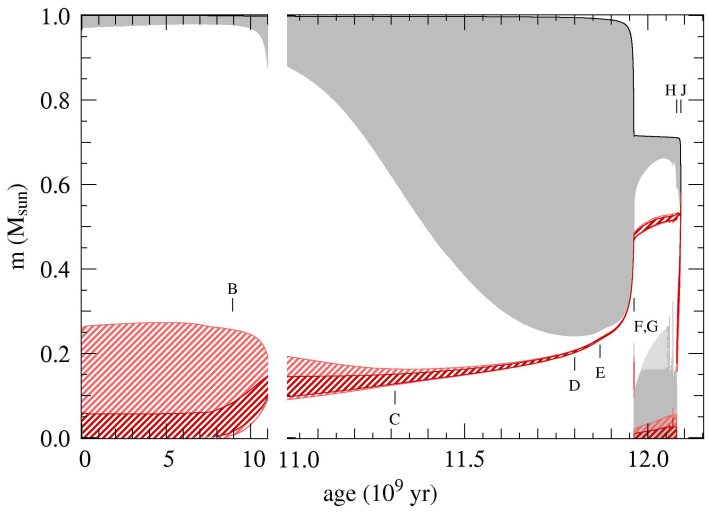
STELLAR STRUCTURE AND EVOLUTION



O.R. Pols

Astronomical Institute Utrecht
September 2011

Preface

These lecture notes are intended for an advanced astrophysics course on Stellar Structure and Evolution given at Utrecht University (NS-AP434M). Their goal is to provide an overview of the physics of stellar interiors and its application to the theory of stellar structure and evolution, at a level appropriate for a third-year Bachelor student or beginning Master student in astronomy. To a large extent these notes draw on the classical textbook by Kippenhahn & Weigert (1990; see below), but leaving out unnecessary detail while incorporating recent astrophysical insights and up-to-date results. At the same time I have aimed to concentrate on physical insight rather than rigorous derivations, and to present the material in a logical order, following in part the very lucid but somewhat more basic textbook by Prialnik (2000). Finally, I have borrowed some ideas from the textbooks by Hansen, Kawaler & Trimble (2004), Salaris & Cassissi (2005) and the recent book by Maeder (2009).

These lecture notes are evolving and I try to keep them up to date. If you find any errors or inconsistencies, I would be grateful if you could notify me by email (`O.R.Pols@uu.nl`).

Onno Pols
Utrecht, September 2011

Literature

- C.J. Hansen, S.D. Kawaler & V. Trimble, *Stellar Interiors*, 2004, Springer-Verlag, ISBN 0-387-20089-4 (HANSEN)
- R. Kippenhahn & A. Weigert, *Stellar Structure and Evolution*, 1990, Springer-Verlag, ISBN 3-540-50211-4 (KIPPENHAHN; K&W)
- A. Maeder, *Physics, Formation and Evolution of Rotating Stars*, 2009, Springer-Verlag, ISBN 978-3-540-76948-4 (MAEDER)
- D. Prialnik, *An Introduction to the Theory of Stellar Structure and Evolution*, 2nd edition, 2009, Cambridge University Press, ISBN 0-521-86604-9 (PRIALNIK)
- M. Salaris & S. Cassisi, *Evolution of Stars and Stellar Populations*, 2005, John Wiley & Sons, ISBN 0-470-09220-3 (SALARIS)

Physical and astronomical constants

Table 1. Physical constants in cgs units (CODATA 2006).

gravitational constant	G	$6.674\,3 \times 10^{-8}$ cm^3 g^{-1} s^{-2}
speed of light in vacuum	c	$2.997\,924\,58 \times 10^{10}$ cm s^{-1}
Planck constant	h	$6.626\,069 \times 10^{-27}$ erg s
radiation density constant	a	$7.565\,78 \times 10^{-15}$ erg cm^{-3} K^{-4}
Stefan-Boltzmann constant	$\sigma = \frac{1}{4}ac$	$5.670\,40 \times 10^{-5}$ erg cm^{-2} s^{-1} K^{-4}
Boltzmann constant	k	$1.380\,650 \times 10^{-16}$ erg K^{-1}
Avogadro's number	$N_\mathrm{A} = 1/m_\mathrm{u}$	$6.022\,142 \times 10^{23}$ g^{-1}
gas constant	$\mathcal{R} = kN_\mathrm{A}$	$8.314\,47 \times 10^{7}$ erg g^{-1} K^{-1}
electron volt	eV	$1.602\,176\,5 \times 10^{-12}$ erg
electron charge	e	$4.803\,26 \times 10^{-10}$ esu
	e^2	$1.440\,00 \times 10^{-7}$ eV cm
electron mass	m_e	$9.109\,382 \times 10^{-28}$ g
atomic mass unit	m_u	$1.660\,538\,8 \times 10^{-24}$ g
proton mass	m_p	$1.672\,621\,6 \times 10^{-24}$ g
neutron mass	m_n	$1.674\,927\,2 \times 10^{-24}$ g
α-particle mass	m_α	$6.644\,656\,2 \times 10^{-24}$ g

Table 2. Astronomical constants, mostly from the Astronomical Almanac (2008).

Solar mass	$M_\odot$	$1.988\,4 \times 10^{33}$ g
	$GM_\odot$	$1.327\,124\,42 \times 10^{26}$ cm^3 s^{-2}
Solar radius	$R_\odot$	6.957×10^{10} cm
Solar luminosity	$L_\odot$	3.842×10^{33} erg s^{-1}
year	yr	$3.155\,76 \times 10^{7}$ s
astronomical unit	AU	$1.495\,978\,71 \times 10^{13}$ cm
parsec	pc	$3.085\,678 \times 10^{18}$ cm

Chapter 1

Introduction

This introductory chapter sets the stage for the course, and briefly repeats some concepts from earlier courses on stellar astrophysics (e.g. the Utrecht first-year course *Introduction to stellar structure and evolution* by F. Verbunt).

The *goal* of this course on stellar evolution can be formulated as follows:

> to understand the structure and evolution of stars, and their observational properties, using known laws of physics

This involves applying and combining 'familiar' physics from many different areas (e.g. thermodynamics, nuclear physics) under extreme circumstances (high temperature, high density), which is part of what makes studying stellar evolution so fascinating.

What exactly do we mean by a 'star'? A useful definition for the purpose of this course is as follows: a star is an object that (1) radiates energy from an internal source and (2) is bound by its own gravity. This definition excludes objects like planets and comets, because they do not comply with the first criterion. In the strictest sense it also excludes brown dwarfs, which are not hot enough for nuclear fusion, although we will briefly discuss these objects. (The second criterion excludes trivial objects that radiate, e.g. glowing coals).

An important implication of this definition is that stars must *evolve* (why?). A star is born out of an interstellar (molecular) gas cloud, lives for a certain amount of time on its internal energy supply, and eventually dies when this supply is exhausted. As we shall see, a second implication of the definition is that stars can have only a limited range of masses, between ∼0.1 and ∼100 times the mass of the Sun. The *life and death* of stars forms the subject matter of this course. We will only briefly touch on the topic of *star formation*, a complex and much less understood process in which the problems to be solved are mostly very different than in the study of stellar evolution.

1.1 Observational constraints

Fundamental properties of a star include the *mass M* (usually expressed in units of the solar mass, $M_\odot = 1.99 \times 10^{33}$ g), the *radius R* (often expressed in $R_\odot = 6.96 \times 10^{10}$ cm) and the *luminosity L*, the rate at which the star radiates energy into space (often expressed in $L_\odot = 3.84 \times 10^{33}$ erg/s). The *effective temperature* $T_{\rm eff}$ is defined as the temperature of a black body with the same energy flux at the surface of the star, and is a good measure for the temperature of the photosphere. From the definition of effective temperature it follows that

$$L = 4\pi R^2 \sigma T_{\rm eff}^4. \tag{1.1}$$

In addition, we would like to know the *chemical composition* of a star. Stellar compositions are usually expressed as mass fractions X_i, where i denotes a certain element. This is often simplified to specifying the mass fractions X (of hydrogen), Y (of helium) and Z (of all heavier elements or 'metals'), which add up to unity. Another fundamental property is the *rotation rate* of a star, expressed either in terms of the rotation period $P_{\rm rot}$ or the equatorial rotation velocity $\upsilon_{\rm eq}$.

Astronomical observations can yield information about these fundamental stellar quantities:

- *Photometric measurements* yield the apparent brightness of a star, i.e. the energy flux received on Earth, in different wavelength bands. These are usually expressed as magnitudes, e.g. B, V, I, etc. Flux ratios or colour indices ($B-V$, $V-I$, etc.) give a measure of the effective temperature, using theoretical stellar atmosphere models and/or empirical relations. Applying a bolometric correction (which also depends on $T_{\rm eff}$) yields the apparent bolometric flux, $f_{\rm bol}$ (in erg s^{-1} cm^{-2}).

- In some cases the *distance* d to a star can be measured, e.g. from the parallax. The Hipparcos satellite has measured parallaxes with 1 milliarcsec accuracy of more than 10^5 stars. The luminosity then follows from $L = 4\pi d^2 f_{\rm bol}$, and the radius from eq. (1.1) if we have a measure of $T_{\rm eff}$.

- An independent measure of the effective temperature can be obtained from *interferometry*. This technique yields the angular diameter of a star if it is sufficiently extended on the sky, i.e. the ratio $\theta = R/d$. Together with a measurement of $f_{\rm bol}$ this can be seen from eq. (1.1) to yield $\sigma T_{\rm eff}^4 = f_{\rm bol}/\theta^2$. This technique is applied to red giants and supergiants. If the distance is also known, a direct measurement of the radius is possible.

- *Spectroscopy* at sufficiently high resolution yields detailed information about the physical conditions in the atmosphere. With detailed spectral-line analysis using stellar atmosphere models one can determine the photospheric properties of a star: the effective temperature and surface gravity ($g = GM/R^2$, usually expressed as $\log g$), surface abundances of various elements (usually in terms of number density relative to hydrogen) and a measure of the rotation velocity ($\upsilon_{\rm eq} \sin i$, where i is the unknown inclination angle of the equatorial plane). In addition, for some stars the properties of the *stellar wind* can be determined (wind velocities, mass loss rates). All this is treated in more detail in the Master course on *Stellar Atmospheres*.

- The most important fundamental property, the mass, cannot be measured directly for a single star. To measure stellar masses one needs *binary stars* showing radial velocity variations (spectroscopic binaries). Radial velocities alone can only yield masses up to a factor $\sin i$, where i is the inclination angle of the binary orbit. To determine absolute mass values one needs information on i, either from a visual orbit (visual binaries) or from the presence of eclipses (eclipsing binaries). In particular for so called double-lined eclipsing binaries, in which the spectral lines of both stars vary, it is possible to accurately measure both the masses and radii (with 1–2 % accuracy in some cases) by fitting the radial-velocity curves and the eclipse lightcurve. Together with a photometric or, better, spectroscopic determination of $T_{\rm eff}$ also the luminosity of such binaries can be measured with high accuracy, independent of the distance. For more details see the Master course on *Binary Stars*.

All observed properties mentioned above are surface properties. Therefore we need a *theory of stellar structure* to derive the internal properties of a star. However, some direct windows on the interior of a star exist:

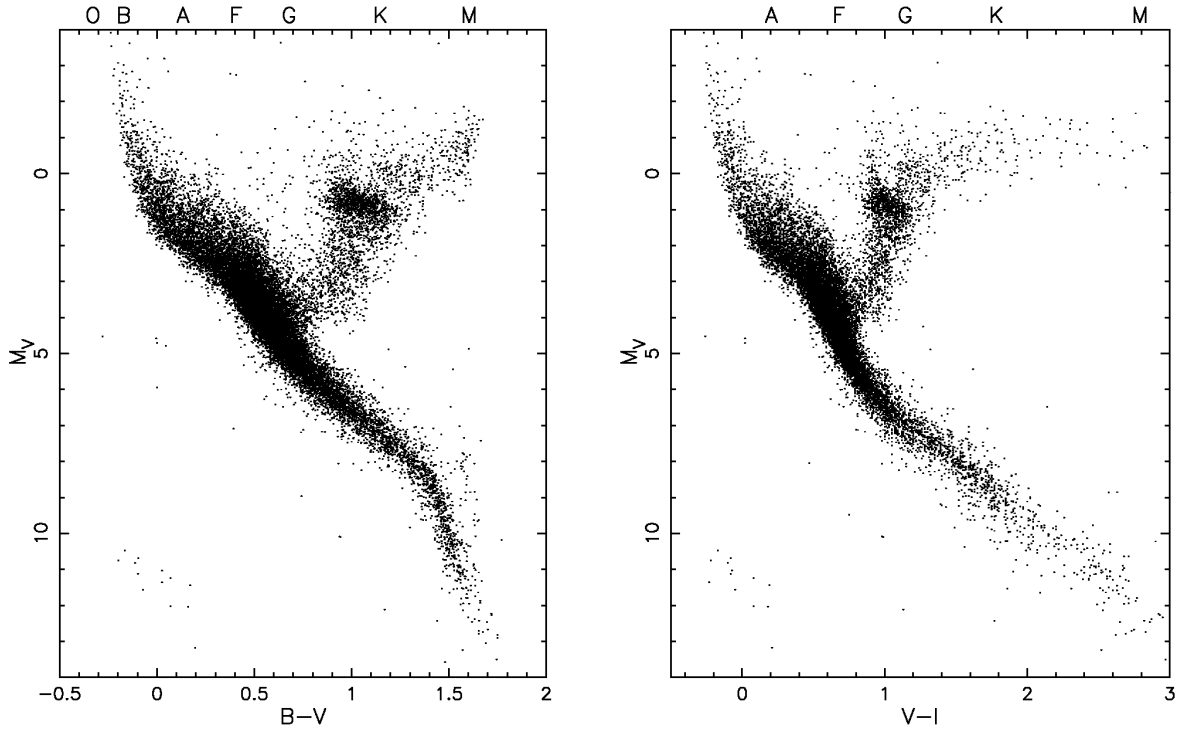


Figure 1.1. H-R diagram of solar neighbourhood. Source: Hipparcos, stars with d measured to $< 10\,\%$ accuracy.

- *neutrinos*, which escape from the interior without interaction. So far, the Sun is the only (non-exploding) star from which neutrinos have been detected.
- *oscillations*, i.e. stellar seismology. Many stars oscillate, and their frequency spectrum contains information about the speed of sound waves inside the star, and therefore about the interior density and temperature profiles. This technique has provided accurate constraints on detailed structure models for the Sun, and is now also being applied to other stars.

The timespan of any observations is much smaller than a stellar lifetime: observations are like snapshots in the life of a star. The observed properties of an individual star contain no (direct) information about its evolution. The diversity of stellar properties (radii, luminosities, surface abundances) does, however, depend on how stars evolve, as well as on intrinsic properties (mass, initial composition). Properties that are common to a large number of stars must correspond to long-lived evolution phases, and vice versa. By studying samples of stars statistically we can infer the (relative) lifetimes of certain phases, which provides another important constraint on the theory of stellar evolution.

Furthermore, observations of samples of stars reveal certain correlations between stellar properties that the theory of stellar evolution must explain. Most important are relations between luminosity and effective temperature, as revealed by the *Hertzsprung-Russell diagram*, and relations between mass, luminosity and radius.

1.1.1 The Hertzsprung-Russell diagram

The Hertzsprung-Russell diagram (HRD) is an important tool to test the theory of stellar evolution. Fig. 1.1 shows the colour-magnitude diagram (CMD) of stars in the vicinity of the Sun, for which the Hipparcos satellite has measured accurate distances. This is an example of a *volume-limited* sample

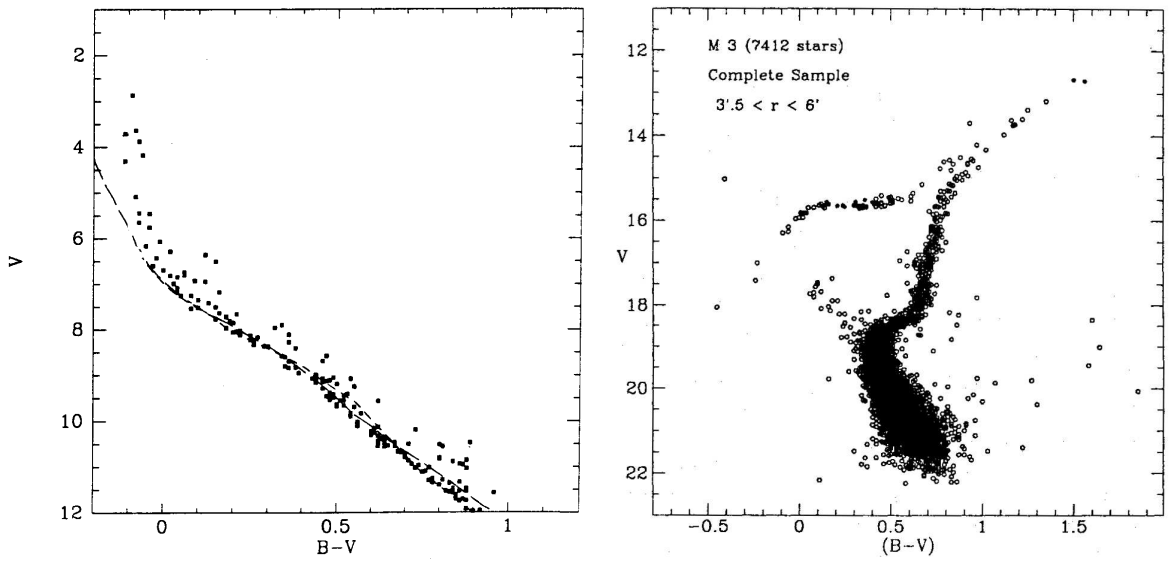


Figure 1.2. Colour-magnitude diagrams of a young open cluster, M45 (the Pleiades, left panel), and a globular cluster, M3 (right panel).

of stars. In this observers' HRD, the absolute visual magnitude M_V is used as a measure of the luminosity and a colour index ($B-V$ or $V-I$) as a measure for the effective temperature. It is left as an exercise to identify various types of stars and evolution phases in this HRD, such as the main sequence, red giants, the horizontal branch, white dwarfs, etc.

Star clusters provide an even cleaner test of stellar evolution. The stars in a cluster were formed within a short period of time (a few Myr) out of the same molecular cloud and therefore share the same age and (initial) chemical composition.[1] Therefore, to first-order approximation only the mass varies from star to star. A few examples of cluster CMDs are given in Fig. 1.2, for a young open cluster (the Pleiades) and an old globular cluster (M3). As the cluster age increases, the most luminous main-sequence stars disappear and a prominent red giant branch and horizontal branch appear. To explain the morphology of cluster HRDs at different ages is one of the goals of studying stellar evolution.

1.1.2 The mass-luminosity and mass-radius relations

For stars with measured masses, radii and luminosities (i.e. binary stars) we can plot these quantities against each other. This is done in Fig. 1.3 for the components of double-lined eclipsing binaries for which M, R and L are all measured with $\lesssim 2\,\%$ accuracy. These quantities are clearly correlated, and especially the relation between mass and luminosity is very tight. Most of the stars in Fig. 1.3 are long-lived main-sequence stars; the spread in radii for masses between 1 and 2 $M_\odot$ results from the fact that several more evolved stars in this mass range also satisfy the 2 % accuracy criterion. The observed relations can be approximated reasonably well by power laws:

$$L \propto M^{3.8} \qquad \text{and} \qquad R \propto M^{0.7}. \tag{1.2}$$

Again, the theory of stellar evolution must explain the existence and slopes of these relations.

[1]The stars in a cluster thus consitute a so-called *simple stellar population*. Recently, this simple picture has changed somewhat after the discovery of multiple populations in many star clusters.

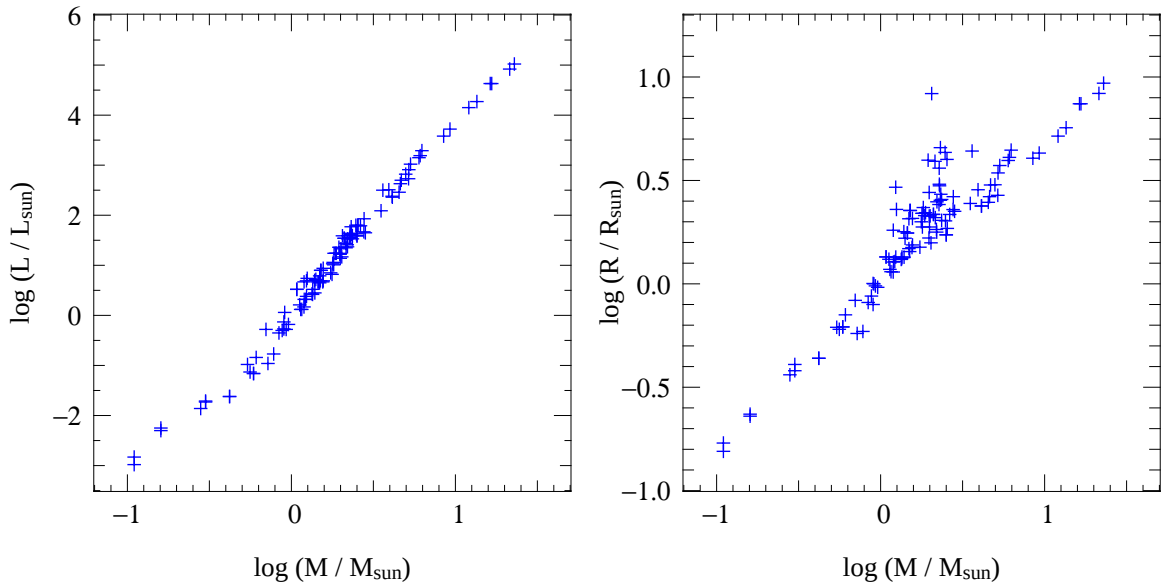


Figure 1.3. Mass-luminosity (left) and mass-radius (right) relations for components of double-lined eclipsing binaries with accurately measured M, R and L.

1.2 Stellar populations

Stars in the Galaxy are divided into different populations:

- Population I: stars in the galactic disk, in spiral arms and in (relatively young) open clusters. These stars have ages $\lesssim 10^9$ yr and are relatively metal-rich ($Z \sim 0.5 - 1\,Z_\odot$)
- Population II: stars in the galactic halo and in globular clusters, with ages $\sim 10^{10}$ yr. These stars are observed to be metal-poor ($Z \sim 0.01 - 0.1\,Z_\odot$).

An intermediate population (with intermediate ages and metallicities) is also seen in the disk of the Galaxy. Together they provide evidence for the *chemical evolution* of the Galaxy: the abundance of heavy elements (Z) apparently increases with time. This is the result of chemical enrichment by subsequent stellar generations.

The study of chemical evolution has led to the hypothesis of a 'Population III' consisting of the first generation of stars formed after the Big Bang, containing only hydrogen and helium and no heavier elements ('metal-free', $Z = 0$). No metal-free stars have ever been observed, probably due to the fact that they were massive and had short lifetimes and quickly enriched the Universe with metals. However, a quest for finding their remnants has turned up many very metal-poor stars in the halo, with the current record-holder having an iron abundance $X_{\rm Fe} = 4 \times 10^{-6} X_{\rm Fe,\odot}$.

1.3 Basic assumptions

We wish to build a theory of stellar evolution to explain the observational constraints highlighted above. In order to do so we must make some basic assumptions:

- stars are considered to be *isolated* in space, so that their structure and evolution depend only on *intrinsic* properties (mass and composition). For most single stars in the Galaxy this condition is satisfied to a high degree (compare for instance the radius of the Sun with the distance to its

nearest neighbour Proxima Centauri, see exercise 1.2). However, for stars in dense clusters, or in binary systems, the evolution can be influenced by interaction with neighbouring stars. In this course we will mostly ignore these complicating effects (many of which are treated in the Master course on *Binary Stars*).

- stars are formed with a *homogeneous composition*, a reasonable assumption since the molecular clouds out of which they form are well-mixed. We will often assume a so-called 'quasi-solar' composition ($X = 0.70$, $Y = 0.28$ and $Z = 0.02$), even though recent determinations of solar abundances have revised the solar metallicity down to $Z = 0.014$. In practice there is relatively little variation in composition from star to star, so that the initial mass is the most important parameter that determines the evolution of a star. The composition, in particular the metallicity Z, is of secondary influence but can have important effects especially in very metal-poor stars (see § 1.2).
- *spherical symmetry*, which is promoted by self-gravity and is a good approximation for most stars. Deviations from spherical symmetry can arise if non-central forces become important relative to gravity, in particular rotation and magnetic fields. Although many stars are observed to have magnetic fields, the field strength (even in highly magnetized neutron stars) is always negligible compared to gravity. Rotation can be more important, and the *rotation rate* can be considered an additional parameter (besides mass and composition) determining the structure and evolution of a star. For the majority of stars (e.g. the Sun) the forces involved are small compared to gravity. However, some rapidly rotating stars are seen (by means of interferometry) to be substantially flattened.

1.4 Aims and overview of the course

In the remainder of this course we will:

- understand the global properties of stars: energetics and timescales
- study the micro-physics relevant for stars: the equation of state, nuclear reactions, energy transport and opacity
- derive the equations necessary to model the internal structure of stars
- examine (quantitatively) the properties of simplified stellar models
- survey (mostly qualitatively) how stars of different masses evolve, and the endpoints of stellar evolution (white dwarfs, neutron stars)
- discuss a few ongoing research areas in stellar evolution

Suggestions for further reading

The contents of this introductory chapter are also largely covered by Chapter 1 of Prialnik, which provides nice reading.

Exercises

1.1 Evolutionary stages

In this course we use many concepts introduced in the introductory astronomy classes. In this exercise we recapitulate the names of evolutionary phases. During the lectures you are assumed to be familiar with these terms, in the sense that you are able to explain them in general terms.

We encourage you to use CARROLL & OSTLIE, *Introduction to Modern Astrophysics*, or the book of the first year course (VERBUNT, *Het leven van sterren*) to make a list of the concepts printed in *italic* with a brief explanation in your own words.

(a) Figure 1.1 shows the location of stars in the solar neighborhood in the Hertzsprung-Russel diagram. Indicate in Figure 1.1 where you would find:

main-sequence stars,	*neutron stars,*
the Sun,	*black holes,*
red giants,	*binary stars,*
horizontal branch stars,	*planets,*
asymptotic giant branch (AGB) stars,	*pre-main sequence stars,*
centrals star of planetary nebulae,	*hydrogen burning stars,*
white dwarfs,	*helium burning stars.*

(b) Through which stages listed above will the Sun evolve? Put them in chronological order. Through which stages will a massive star evolve?

(c) Describe the following concepts briefly in your own words. You will need the concepts indicated with * in the coming lectures.

ideal gas,*	*Jeans mass,*
black body,	*Schwarzschild criterion,*
virial theorem,*	*energy transport by radiation,*
first law of thermodynamics,*	*energy transport by convection,*
equation of state,	*pp-chain,*
binary stars,	*CNO cycle,*
star cluster,	*nuclear timescale*,*
interstellar medium,	*thermal or Kelvin-Helmholtz timescale*,*
giant molecular clouds,	*dynamical timescale**

1.2 Basic assumptions

Let us examine the three basic assumptions made in the theory of stellar evolution:

(a) *Stars are assumed to be isolated in space.* The star closest to the sun, Proxima Centauri, is 4.3 light-years away. How many solar radii is that? By what factors are the gravitational field and the radiation flux diminished? Many stars are formed in clusters and binaries. How could that influence the life of a star?

(b) *Stars are assumed to form with a uniform composition.* What elements is the Sun made of? Just after the Big Bang the Universe consisted almost purely of hydrogen and helium. Where do all the heavier elements come from?

(c) *Stars are assumed to be spherically symmetric.* Why are stars spherically symmetric to a good approximation? How would rotation affect the structure and evolution of a star? The Sun rotates around its axis every 27 days. Calculate the ratio of is the centrifugal acceleration a over the gravitational acceleration g for a mass element on the surface of the Sun. Does rotation influence the structure of the Sun?

1.3 Mass-luminosity and mass-radius relation

(a) The masses of stars are approximately in the range $0.08\,M_\odot \lesssim M \lesssim 100\,M_\odot$. Why is there an upper limit? Why is there a lower limit?

(b) Can you think of methods to measure (1) the mass, (2) the radius, and (3) the luminosity of a star? Can your methods be applied for any star or do they require special conditions. Discuss your methods with your fellow students.

(c) Figure 1.3 shows the luminosity versus the mass (left) and the radius versus the mass (right) for observed main sequence stars. We can approximate a mass-luminosity and mass-radius relation by fitting functions of the form

$$\frac{L}{L_\odot} = \left(\frac{M}{M_\odot}\right)^x, \qquad \frac{R}{L_\odot} = \left(\frac{M}{M_\odot}\right)^y \tag{1.3}$$

Estimate x and y from Figure 1.3.

(d) Which stars live longer, high mass stars (which have more fuel) or low mass stars? Derive an expression for the lifetime of a star as a function of its mass. (!)

[Hints: Stars spend almost all their life on the main sequence burning hydrogen until they run out of fuel. First try to estimate the life time as function of the mass (amount of fuel) and the luminosity (rate at which the fuel is burned).]

1.4 The ages of star clusters

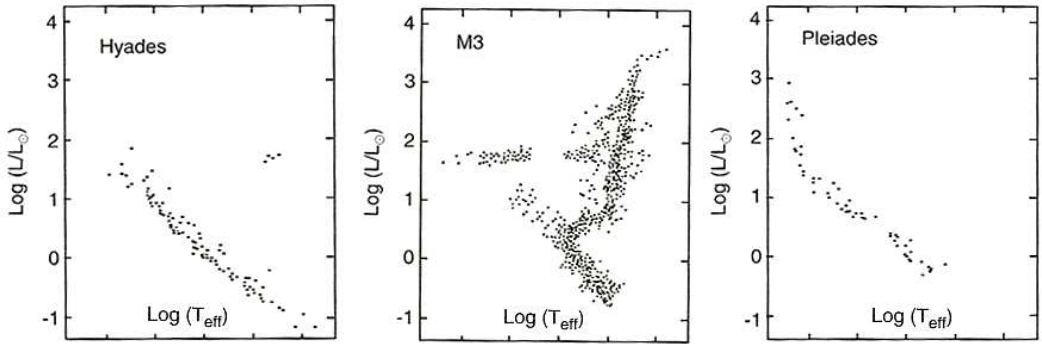


Figure 1.4. H-R diagrams of three star clusters (from PRIALNIK).

The stars in a star cluster are formed more or less simultaneously by fragmentation of a large molecular gas cloud.

(a) In Fig. 1.4 the H-R diagrams are plotted of the stars in three different clusters. Which cluster is the youngest?

(b) Think of a method to estimate the age of the clusters, discuss with your fellow students. Estimate the ages and compare with the results of your fellow students.

(c) (*) Can you give an error range on your age estimates?

Chapter 2

Mechanical and thermal equilibrium

In this chapter we apply the physical principles of mass conservation and momentum conservation to derive two of the fundamental stellar structure equations. We shall see that stars are generally in a state of almost complete *mechanical equilibrium*, which allows us to derive and apply the important *virial theorem*. We consider the basic stellar timescales and see that most (but not all) stars are also in a state of energy balance called *thermal equilibrium*.

2.1 Coordinate systems and the mass distribution

The assumption of spherical symmetry implies that all interior physical quantities (such as density ρ, pressure P, temperature T, etc) depend only on one radial coordinate. The obvious coordinate to use in a Eulerian coordinate system is the radius of a spherical shell, r ($\in 0 \ldots R$). In an evolving star, all quantities also depend on time t. When constructing the differential equations for stellar structure one should thus generally consider partial derivatives of physical quantities with respect to radius and time, $\partial/\partial r$ and $\partial/\partial t$, taken at constant t and r, respectively.

The principle of mass conservation applied to the mass dm of a spherical shell of thickness dr at radius r (see Fig. 2.1) gives

$$\mathrm{d}m(r,t) = 4\pi r^2 \rho \,\mathrm{d}r - 4\pi r^2 \rho \upsilon \,\mathrm{d}t, \tag{2.1}$$

where υ is the radial velocity of the mass shell. Therefore one has

$$\frac{\partial m}{\partial r} = 4\pi r^2 \rho \qquad \text{and} \qquad \frac{\partial m}{\partial t} = -4\pi r^2 \rho \upsilon. \tag{2.2}$$

The first of these partial differential equations relates the radial mass distribution in the star to the local density: it constitutes the first fundamental equation of stellar structure. Note that $\rho = \rho(r,t)$ is not known a priori, and must follow from other conditions and equations. The second equation of (2.2) represents the change of mass inside a sphere of radius r due to the motion of matter through its surface; at the stellar surface this gives the mass-loss rate (if there is a stellar wind with $\upsilon > 0$) or mass-accretion rate (if there is inflow with $\upsilon < 0$). In a static situation, where the velocity is zero, the first equation of (2.2) becomes an ordinary differential equation,

$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \rho. \tag{2.3}$$

This is almost always a good approximation for stellar interiors, as we shall see. Integration yields the mass $m(r)$ inside a spherical shell of radius r:

$$m(r) = \int_0^r 4\pi r'^2 \rho \,\mathrm{d}r'.$$

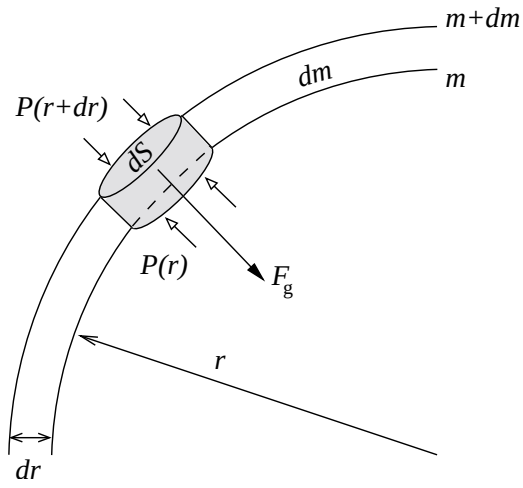


Figure 2.1. Mass shell inside a spherically symmetric star, at radius r and with thickness dr. The mass of the shell is d$m = 4\pi r^2 \rho \, \mathrm{d}r$. The pressure and the gravitational force acting on a cylindrical mass element are also indicated.

Since $m(r)$ increases monotonically outward, we can also use $m(r)$ as our radial coordinate, instead of r. This *mass coordinate*, often denoted as m_r or simply m, is a Lagrangian coordinate that moves with the mass shells:

$$m := m_r = \int_0^r 4\pi r'^2 \rho \, \mathrm{d}r' \qquad (m \in 0 \ldots M). \tag{2.4}$$

It is often more convenient to use a Lagrangian coordinate instead of a Eulerian coordinate. The mass coordinate is defined on a fixed interval, $m \in 0 \ldots M$, as long as the star does not lose mass. On the other hand r depends on the time-varying stellar radius R. Furthermore the mass coordinate follows the mass elements in the star, which simplifies many of the time derivatives that appear in the stellar evolution equations (e.g. equations for the composition). We can thus write all quantities as functions of m, i.e. $r = r(m)$, $\rho = \rho(m)$, $P = P(m)$, etc.

Using the coordinate transformation $r \to m$, i.e.

$$\frac{\partial}{\partial m} = \frac{\partial}{\partial r} \cdot \frac{\partial r}{\partial m}, \tag{2.5}$$

the first equation of stellar structure becomes in terms of the coordinate m:

$$\boxed{\frac{\partial r}{\partial m} = \frac{1}{4\pi r^2 \rho}} \tag{2.6}$$

which again becomes an ordinary differential equation in a static situation.

2.1.1 The gravitational field

Recall that a star is a self-gravitating body of gas, which implies that gravity is the driving force behind stellar evolution. In the general, non-spherical case, the gravitational acceleration $\boldsymbol{g}$ can be written as the gradient of the gravitational potential, $\boldsymbol{g} = -\boldsymbol{\nabla}\Phi$, where Φ is the solution of the Poisson equation

$$\nabla^2 \Phi = 4\pi G \rho.$$

Inside a spherically symmetric body, this reduces to $g := |\boldsymbol{g}| = \mathrm{d}\Phi/\mathrm{d}r$. The gravitational acceleration at radius r and equivalent mass coordinate m is then given by

$$g = \frac{Gm}{r^2}. \tag{2.7}$$

Spherical shells outside r apply no net force, so that g only depends on the mass distribution inside the shell at radius r. Note that g is the magnitude of the vector $\boldsymbol{g}$ which points inward (toward smaller r or m).

2.2 The equation of motion and hydrostatic equilibrium

We next consider conservation of momentum inside a star, i.e. Newton's second law of mechanics. The net acceleration on a gas element is determined by the sum of all forces acting on it. In addition to the gravitational force considered above, forces result from the pressure exerted by the gas surrounding the element. Due to spherical symmetry, the pressure forces acting horizontally (perpendicular to the radial direction) balance each other and only the pressure forces acting along the radial direction need to be considered. By assumption we ignore other forces that might act inside a star (Sect. 1.3).

Hence the net acceleration $\ddot{r} = \partial^2 r/\partial t^2$ of a (cylindrical) gas element with mass

$$\mathrm{d}m = \rho \, \mathrm{d}r \, \mathrm{d}S \tag{2.8}$$

(where $\mathrm{d}r$ is its radial extent and $\mathrm{d}S$ is its horizontal surface area, see Fig. 2.1) is given by

$$\ddot{r} \, \mathrm{d}m = -g \, \mathrm{d}m + P(r) \, \mathrm{d}S - P(r + \mathrm{d}r) \, \mathrm{d}S. \tag{2.9}$$

We can write $P(r + \mathrm{d}r) = P(r) + (\partial P/\partial r) \cdot \mathrm{d}r$, hence after substituting eqs. (2.7) and (2.8) we obtain the *equation of motion* for a gas element inside the star:

$$\frac{\partial^2 r}{\partial t^2} = -\frac{Gm}{r^2} - \frac{1}{\rho}\frac{\partial P}{\partial r}. \tag{2.10}$$

This is a simplified from of the Navier-Stokes equation of hydrodynamics, applied to spherical symmetry (see Maeder). Writing the pressure gradient $\partial P/\partial r$ in terms of the mass coordinate m by substituting eq. (2.6), the equation of motion is

$$\frac{\partial^2 r}{\partial t^2} = -\frac{Gm}{r^2} - 4\pi r^2 \frac{\partial P}{\partial m}. \tag{2.11}$$

Hydrostatic equilibrium The great majority of stars are obviously in such long-lived phases of evolution that no change can be observed over human lifetimes. This means there is no noticeable acceleration, and all forces acting on a gas element inside the star almost exactly balance each other. Thus most stars are in a state of mechanical equilibrium which is more commonly called *hydrostatic equilibrium* (HE).

The state of hydrostatic equilibrium, setting $\ddot{r} = 0$ in eq. (2.10), yields the second differential equation of stellar structure:

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{Gm}{r^2}\rho, \tag{2.12}$$

or with eq. (2.6)

$$\boxed{\frac{\mathrm{d}P}{\mathrm{d}m} = -\frac{Gm}{4\pi r^4}} \tag{2.13}$$

A direct consequence is that inside a star in hydrostatic equilibrium, the pressure always decreases outwards.

Eqs. (2.6) and (2.13) together determine the *mechanical structure* of a star in HE. These are two equations for three unknown functions of m (r, P and ρ), so they cannot be solved without a

third condition. This condition is usually a relation between P and ρ called the *equation of state* (see Chapter 3). In general the equation of state depends on the temperature T as well, so that the mechanical structure depends also on the temperature distribution inside the star, i.e. on its thermal structure. In special cases the equation of state is independent of T, and can be written as $P = P(\rho)$. In such cases (known as barotropes or polytropes) the mechanical structure of a star becomes independent of its thermal structure. This is the case for white dwarfs, as we shall see later.

Estimates of the central pressure A rough order-of-magnitude estimate of the central pressure can be obtained from eq. (2.13) by setting

$$\frac{\mathrm{d}P}{\mathrm{d}m} \sim \frac{P_{\rm surf} - P_c}{M} \approx -\frac{P_c}{M}, \quad m \sim \tfrac{1}{2}M, \quad r \sim \tfrac{1}{2}R$$

which yields

$$P_c \sim \frac{2}{\pi}\frac{GM^2}{R^4} \tag{2.14}$$

For the Sun we obtain from this estimate $P_c \sim 7 \times 10^{15}\,\mathrm{dyn/cm^2} = 7 \times 10^9\,\mathrm{atm}$.

A lower limit on the central pressure may be derived by writing eq. (2.13) as

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{Gm}{4\pi r^4}\frac{\mathrm{d}m}{\mathrm{d}r} = -\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{Gm^2}{8\pi r^4}\right) - \frac{Gm^2}{2\pi r^5},$$

and thus

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(P + \frac{Gm^2}{8\pi r^4}\right) = -\frac{Gm^2}{2\pi r^5} < 0. \tag{2.15}$$

The quantity $\Psi(r) = P + Gm^2/(8\pi r^4)$ is therefore a decreasing function of r. At the centre, the second term vanishes because $m \propto r^3$ for small r, and hence $\Psi(0) = P_c$. At the surface, the pressure is essentially zero. From the fact that Ψ must decrease with r it thus follows that

$$P_c > \frac{1}{8\pi}\frac{GM^2}{R^4}. \tag{2.16}$$

In contrast to eq. (2.14), this is a strict mathematical result, valid for any star in hydrostatic equilibrium regardless of its other properties (in particular, regardless of its density distribution). For the Sun we obtain $P_c > 4.4 \times 10^{14}\,\mathrm{dyn/cm^2}$. Both estimates indicate that an extremely high central pressure is required to keep the Sun in hydrostatic equilibrium. Realistic solar models show the central density to be $2.4 \times 10^{17}\,\mathrm{dyn/cm^2}$.

2.2.1 The dynamical timescale

We can ask what happens if the state of hydrostatic equilibrium is violated: how fast do changes to the structure of a star occur? The answer is provided by the equation of motion, eq. (2.10). For example, suppose that the pressure gradient that supports the star against gravity suddenly drops. All mass shells are then accelerated inwards by gravity: the star starts to collapse in "free fall". We can approximate the resulting (inward) acceleration by

$$|\ddot{r}| \approx \frac{R}{{\tau_{\rm ff}}^2} \quad \Rightarrow \quad \tau_{\rm ff} \approx \sqrt{\frac{R}{|\ddot{r}|}}$$

where $\tau_{\rm ff}$ is the free-fall timescale that we want to determine. Since $-\ddot{r} = g \approx GM/R^2$ for the entire star, we obtain

$$\tau_{\rm ff} \approx \sqrt{\frac{R}{g}} \approx \sqrt{\frac{R^3}{GM}}. \tag{2.17}$$

Of course each mass shell is accelerated at a different rate, so this estimate should be seen as an average value for the star to collapse over a distance R. This provides one possible estimate for the *dynamical timescale* of the star. Another estimate can be obtained in a similar way by assuming that gravity suddenly disappears: this gives the timescale for the outward pressure gradient to explode the star, which is similar to the time it takes for a sound wave to travel from the centre to the surface of the star. If the star is close to HE, all these timescales have about the same value given by eq. (2.17). Since the average density $\bar{\rho} = 3M/(4\pi R^3)$, we can also write this (hydro)dynamical timescale as

$$\tau_{\rm dyn} \approx \sqrt{\frac{R^3}{GM}} \approx \tfrac{1}{2}\,(G\bar{\rho})^{-1/2}. \tag{2.18}$$

For the Sun we obtain a very small value of $\tau_{\rm dyn} \approx 1600\,$sec or about half an hour (0.02 days). This is very much smaller than the age of the Sun, which is 4.6 Gyr or $\sim 1.5 \times 10^{17}$ sec, by 14 orders of magnitude. This result has several important consequences for the Sun and other stars:

- Any significant departure from hydrostatic equilibrium should very quickly lead to observable phenomena: either contraction or expansion on the dynamical timescale. If the star cannot recover from this disequilibrium by restoring HE, it should lead to a collapse or an explosion.
- Normally hydrostatic equilibrium can be restored after a disturbance (we will consider this *dynamical stability* of stars later). However a perturbation of HE may lead to small-scale oscillations on the dynamical timescale. These are indeed observed in the Sun and many other stars, with a period of minutes in the case of the Sun. Eq. (2.18) tells us that the pulsation period is a (rough) measure of the average density of the star.
- Apart from possible oscillations, stars are extremely close to hydrostatic equilibrium, since any disturbance is immediately quenched. We can therefore be confident that eq. (2.13) holds throughout most of their lifetimes. Stars do evolve and are therefore not completely static, but changes occur very slowly compared to their dynamical timescale. Stars can be said to evolve *quasi-statically*, i.e. through a series of near-perfect HE states.

2.3 The virial theorem

An important consequence of hydrostatic equilibrium is the *virial theorem*, which is of vital importance for the understanding of stars. It connects two important energy reservoirs of a star and allows predictions and interpretations of important phases in the evolution of stars.

To derive the virial theorem we start with the equation for hydrostatic equilibrium eq. (2.13). We multiply both sides by the enclosed volume $V = \frac{4}{3}\pi r^3$ and integrate over m:

$$\int_0^M \tfrac{4}{3}\pi r^3 \frac{{\rm d}P}{{\rm d}m}\,{\rm d}m = -\tfrac{1}{3}\int_0^M \frac{Gm}{r}\,{\rm d}m \tag{2.19}$$

The integral on the right-hand side has a straightforward physical interpretation: it is the *gravitational potential energy* of the star. To see this, consider the work done by the gravitational force $\boldsymbol{F}$ to bring a mass element δm from infinity to radius r:

$$\delta W = \int_\infty^r \boldsymbol{F}\cdot{\rm d}\boldsymbol{r} = \int_\infty^r \frac{Gm\,\delta m}{r^2}\,{\rm d}r = -\frac{GM}{r}\delta m.$$

The gravitational potential energy of the star is the work performed by the gravitational force to bring *all* mass elements from infinity to their current radius, i.e.

$$E_{\rm gr} = -\int_0^M \frac{Gm}{r}\,{\rm d}m \tag{2.20}$$

The left-hand side of eq. (2.19) can be integrated by parts:

$$\int_{P_c}^{P_s} V\,{\rm d}P = [V\cdot P]_{\rm c}^{\rm s} - \int_0^{V_s} P\,{\rm d}V \tag{2.21}$$

where c and s denote central and surface values. Combining the above expressions in eq. (2.19) we obtain

$$\tfrac{4}{3}\pi R^3\,P(R) - \int_0^{V_s} P\,{\rm d}V = \tfrac{1}{3}E_{\rm gr}, \tag{2.22}$$

with $P(R)$ the pressure at the surface of the volume. This expression is useful when the pressure from the surrounding layers is substantial, e.g. when we consider only the core of a star. If we consider the star as a whole, however, the first term vanishes because the pressure at the stellar surface is negligible. In that case

$$-3\int_0^{V_s} P\,{\rm d}V = E_{\rm gr}, \tag{2.23}$$

or, since ${\rm d}V = {\rm d}m/\rho$,

$$-3\int_0^{M} \frac{P}{\rho}\,{\rm d}m = E_{\rm gr}. \tag{2.24}$$

This is the general form of the virial theorem, which will prove valuable later. It tells us that that the average pressure needed to support a star in HE is equal to $-\frac{1}{3}E_{\rm gr}/V$. In particular it tells us that a star that contracts quasi-statically (that is, slowly enough to remain in HE) must increase its internal pressure, since $|E_{\rm gr}|$ increases while its volume decreases.

The virial theorem for an ideal gas The pressure of a gas is related to its internal energy. We will show this in Ch. 3, but for the particular case of an ideal monatomic gas it is easy to see. The pressure of an ideal gas is given by

$$P = nkT = \frac{\rho}{\mu m_{\rm u}}kT, \tag{2.25}$$

where $n = N/V$ is the number of particles per unit volume, and μ is mass of a gas particle in atomic mass units. The kinetic energy per particle is $\epsilon_k = \frac{3}{2}kT$, and the internal energy of an ideal monatomic gas is equal to the kinetic energy of its particles. The internal energy per unit mass is then

$$u = \frac{3}{2}\frac{kT}{\mu m_{\rm u}} = \frac{3}{2}\frac{P}{\rho}. \tag{2.26}$$

We can now interpret the left-hand side of the virial theorem (eq. 2.24) as $\int (P/\rho)\,{\rm d}m = \frac{2}{3}\int u\,{\rm d}m = \frac{2}{3}E_{\rm int}$, where $E_{\rm int}$ is the total internal energy of the star. The virial theorem for an ideal gas is therefore

$$\boxed{E_{\rm int} = -\tfrac{1}{2}E_{\rm gr}} \tag{2.27}$$

This important relation establishes a link between the gravitational potential energy and the internal energy of a star in hydrostatic equilibrium that consists of an ideal gas. (We shall see later that the ideal gas law indeed holds for most stars, at least on the main sequence.) The virial theorem tells us that a more tightly bound star must have a higher internal energy, i.e. it must be *hotter*. In other words, a star that contracts quasi-statically must get hotter in the process. The full implications of this result will become clear when we consider the total energy of a star in a short while.

Estimate of the central temperature Using the virial theorem we can obtain an estimate of the average temperature inside a star composed of ideal gas. The gravitational energy of the star is found from eq. (2.20) and can be written as

$$E_{\rm gr} = -\alpha \frac{GM^2}{R}, \tag{2.28}$$

where α is a constant of order unity (determined by the distribution of matter in the star, i.e. by the density profile). Using eq. (2.26), the internal energy of the star is $E_{\rm int} = \frac{3}{2}k/(\mu m_{\rm u}) \int T {\rm d}m = \frac{3}{2}k/(\mu m_{\rm u})\bar{T}M$, where $\bar{T}$ is the temperature averaged over all mass shells. By the virial theorem we then obtain

$$\bar{T} = \frac{\alpha}{3}\frac{\mu m_{\rm u}}{k}\frac{GM}{R}. \tag{2.29}$$

Taking $\alpha \approx 1$ and $\mu = 0.5$ for ionized hydrogen, we obtain for the Sun $\bar{T} \sim 4 \times 10^6$ K. This is the average temperature required to provide the pressure that is needed to keep the Sun in hydrostatic equilibrium. Since the temperature in a star normally decreases outwards, it is also an approximate lower limit on the central temperature of the Sun. At these temperatures, hydrogen and helium are indeed completely ionized. We shall see that $T_c \approx 10^7$ K is high enough for hydrogen fusion to take place in the central regions of the Sun.

The virial theorem for a general equation of state Also for equations of state other than an ideal gas a relation between pressure and internal energy exists, which we can write generally as

$$u = \phi \frac{P}{\rho}. \tag{2.30}$$

We have seen above that $\phi = \frac{3}{2}$ for an ideal gas, but it will turn out (see Ch. 3) that this is valid not only for an ideal gas, but for all non-relativistic particles. On the other hand, if we consider a gas of relativistic particles, in particular photons (i.e. radiation pressure), $\phi = 3$. If ϕ is constant throughout the star we can integrate the left-hand side of eq. (2.23) to obtain a more general form of the virial theorem:

$$E_{\rm int} = -\tfrac{1}{3}\phi E_{\rm gr} \tag{2.31}$$

2.3.1 The total energy of a star

The total energy of a star is the sum of its gravitational potential energy, its internal energy and its kinetic energy $E_{\rm kin}$ (due to bulk motions of gas inside the star, not the thermal motions of the gas particles):

$$E_{\rm tot} = E_{\rm gr} + E_{\rm int} + E_{\rm kin}. \tag{2.32}$$

The star is bound as long as its total energy is negative.

For a star in hydrostatic equilibrium we can set $E_{\rm kin} = 0$. Furthermore for a star in HE the virial theorem holds, so that $E_{\rm gr}$ and $E_{\rm int}$ are tightly related by eq. (2.31). Combining eqs. (2.31) and (2.32) we obtain the following relations:

$$E_{\rm tot} \;=\; E_{\rm int} + E_{\rm gr} \;=\; \frac{\phi - 3}{\phi} E_{\rm int} \;=\; (1 - \tfrac{1}{3}\phi) E_{\rm gr} \tag{2.33}$$

As long as $\phi < 3$ the star is bound. This is true in particular for the important case of a star consisting of an ideal gas (eq. 2.27), for which we obtain

$$\boxed{E_{\rm tot} \;=\; E_{\rm int} + E_{\rm gr} \;=\; -\,E_{\rm int} \;=\; \tfrac{1}{2}E_{\rm gr} \;<\; 0} \tag{2.34}$$

In other words, its total energy of such a star equals half of its gravitational potential energy.

From eq. (2.34) we can see that the virial theorem has the following important consequences:

- Gravitationally bound gas spheres must be *hot* to maintain hydrostatic equilibrium: heat provides the pressure required to balance gravity. The more compact such a sphere, the more strongly bound, and therefore the hotter it must be.
- A hot sphere of gas radiates into surrounding space, therefore a star must lose energy from its surface. The rate at which energy is radiated from the surface is the *luminosity* of the star. In the absence of an internal energy source, this energy loss must equal the decrease of the total energy of the star: $L = -\mathrm{d}E_{\rm tot}/\mathrm{d}t > 0$, since L is positive by convention.
- Taking the time derivative of eq. (2.34), we find that as a consequence of losing energy:

  $$\dot{E}_{\rm gr} = -2L < 0,$$

  meaning that the star *contracts* (becomes more strongly bound), and

  $$\dot{E}_{\rm int} = L > 0,$$

  meaning that the star *gets hotter* – unlike familiar objects which cool when they lose energy. Therefore a star can be said to have a *negative heat capacity*. Half the energy liberated by contraction is used for heating the star, the other half is radiated away.

For the case of a star that is dominated by radiation pressure, we find that $E_{\rm int} = -E_{\rm gr}$, and therefore the total energy $E_{\rm tot} = 0$. Therefore a star dominated by radiation pressure (or more generally, by the pressure of relativistic particles) is only marginally bound. No energy is required to expand or contract such a star, and a small perturbation would be enough to render it unstable and to trigger its collapse or complete dispersion.

2.3.2 Thermal equilibrium

If internal energy sources are present in a star due to nuclear reactions taking place in the interior, then the energy loss from the surface can be compensated: $L = L_{\rm nuc} \equiv -\mathrm{d}E_{\rm nuc}/\mathrm{d}t$. In that case the total energy is conserved and eq. (2.34) tells us that $\dot{E}_{\rm tot} = \dot{E}_{\rm int} = \dot{E}_{\rm gr} = 0$. The virial theorem therefore states that both $E_{\rm int}$ and $E_{\rm gr}$ are conserved as well: the star cannot, for example, contract and cool while keeping its total energy constant.

In this state, known as *thermal equilibrium* (TE), the star is in a stationary state. Energy is radiated away at the surface at the same rate at which it is produced by nuclear reactions in the interior. The star neither expands nor contracts, and it maintains a constant interior temperature. We shall see later that this temperature is regulated by the nuclear reactions themselves, which in combination with the virial theorem act like a stellar thermostat. Main-sequence stars like the Sun are in thermal equilibrium, and a star can remain in this state as long as nuclear reactions can supply the necessary energy.

Note that the arguments given above imply that both hydrostatic equilibrium and thermal equilibrium are *stable* equilibria, an assumption that we have yet to prove (see Ch. 7). It is relatively easy to understand why TE is stable, at least as long as the ideal-gas pressure dominates ($\phi < 3$ in eq. 2.31). Consider what happens when TE is disturbed, e.g. when $L_{\rm nuc} > L$ temporarily. The total energy then increases, and the virial theorem states that as a consequence the star must expand and cool. Since the nuclear reaction rates typically increase strongly with temperature, the rate of nuclear burning and thus $L_{\rm nuc}$ will decrease as a result of this cooling, until TE is restored when $L = L_{\rm nuc}$.

2.4 The timescales of stellar evolution

Three important timescales are relevant for stellar evolution, associated with changes to the mechanical structure of a star (described by the equation of motion, eq. 2.11), changes to its thermal structure (as follows from the virial theorem, see also Sect. 5.1) and changes in its composition, which will be discussed in Ch. 6.

The first timescale was already treated in Sec. 2.2.1: it is the *dynamical timescale* given by eq. (2.18),

$$\tau_{\rm dyn} \approx \sqrt{\frac{R^3}{GM}} \approx 0.02 \left(\frac{R}{R_\odot}\right)^{3/2} \left(\frac{M_\odot}{M}\right)^{1/2} \text{days} \tag{2.35}$$

The dynamical timescale is the timescale on which a star reacts to a perturbation of hydrostatic equilibrium. We saw that this timescale is typically of the order of hours or less, which means that stars are extremely close to hydrostatic equilibrium.

2.4.1 The thermal timescale

The second timescale describes how fast changes in the thermal structure of a star can occur. It is therefore also the timescale on which a star in thermal equilibrium reacts when its TE is perturbed. To obtain an estimate, we turn to the virial theorem: we saw in Sec. 2.3.1 that a star without a nuclear energy source contracts by radiating away its internal energy content: $L = \dot{E}_{\rm int} \approx -2\dot{E}_{\rm gr}$, where the last equality applies strictly only for an ideal gas. We can thus define the *thermal* or *Kelvin-Helmholtz timescale* as the timescale on which this gravitational contraction would occur:

$$\tau_{\rm KH} = \frac{E_{\rm int}}{L} \approx \frac{|E_{\rm gr}|}{2L} \approx \frac{GM^2}{2RL} \approx 1.5 \times 10^7 \left(\frac{M}{M_\odot}\right)^2 \frac{R_\odot}{R} \frac{L_\odot}{L} \text{ yr} \tag{2.36}$$

Here we have used eq. (2.28) for $E_{\rm gr}$ with $\alpha \approx 1$.

The thermal timescale for the Sun is about 1.5×10^7 years, which is many orders of magnitude larger than the dynamical timescale. There is therefore no direct observational evidence that any star is in thermal equilibrium. In the late 19th century gravitational contraction was proposed as the energy source of the Sun by Lord Kelvin and, independently, by Hermann von Helmholtz. This led to an age of the Sun and an upper limit to the age the Earth that was in conflict with emerging geological evidence, which required the Earth to be much older. Nuclear reactions have since turned out to be a much more powerful energy source than gravitational contraction, allowing stars to be in thermal equilibrium for most ($> 99\,\%$) of their lifetimes. However, several phases of stellar evolution, during which the nuclear power source is absent or inefficient, do occur on the thermal timescale.

2.4.2 The nuclear timescale

A star can remain in thermal equilibrium for as long as its nuclear fuel supply lasts. The associated timescale is called the *nuclear timescale*, and since nuclear fuel (say hydrogen) is burned into 'ash' (say helium), it is also the timescale on which composition changes in the stellar interior occur.

The energy source of nuclear fusion is the direct conversion of a small fraction ϕ of the rest mass of the reacting nuclei into energy. For hydrogen fusion, $\phi \approx 0.007$; for fusion of helium and heavier elements ϕ is smaller by a factor 10 or more. The total nuclear energy supply can therefore be written as $E_{\rm nuc} = \phi M_{\rm nuc} c^2 = \phi f_{\rm nuc} M c^2$, where $f_{\rm nuc}$ is that fraction of the mass of the star which may serve as nuclear fuel. In thermal equilibrium $L = L_{\rm nuc} = \dot{E}_{\rm nuc}$, so we can estimate the nuclear timescale as

$$\tau_{\rm nuc} = \frac{E_{\rm nuc}}{L} = \phi f_{\rm nuc} \frac{Mc^2}{L} \approx 10^{10} \, \frac{M}{M_\odot} \frac{L_\odot}{L} \, \text{yr}. \tag{2.37}$$

The last approximate equality holds for hydrogen fusion in a star like the Sun, with has 70 % of its initial mass in hydrogen and fusion occurring only in the inner $\approx$ 10 % of its mass (the latter result comes from detailed stellar models). This long timescale is consistent with the geological evidence for the age of the Earth.

We see that, despite only a small fraction of the mass being available for fusion, the nuclear timescale is indeed two to three orders of magnitude larger than the thermal timescale. Therefore the assumption that stars can reach a state of thermal equilibrium is justified. To summarize, we have found:

$$\tau_{\rm nuc} \gg \tau_{\rm KH} \gg \tau_{\rm dyn}.$$

As a consequence, the rates of nuclear reactions determine the pace of stellar evolution, and stars may be assumed to be in hydrostatic and thermal equilibrium throughout most of their lives.

Suggestions for further reading

The contents of this chapter are covered more extensively by Chapter 1 of Maeder and by Chapters 1 to 4 of Kippenhahn & Weigert.

Exercises

2.1 Density profile

In a star with mass M, assume that the density decreases from the center to the surface as a function of radial distance r, according to

$$\rho = \rho_c \left[1 - \left(\frac{r}{R} \right)^2 \right], \tag{2.38}$$

where ρ_c is a given constant and R is the radius of the star.

(a) Find $m(r)$.

(b) Derive the relation between M and R.

(c) Show that the average density of the star is $0.4\rho_c$.

2.2 Hydrostatic equilibrium

(a) Consider an infinitesimal mass element dm inside a star, see Fig. 2.1. What forces act on this mass element?

(b) Newton's second law of mechanics, or the equation of motion, states that the net force acting on a body is equal to its acceleration times it mass. Write down the equation of motion for the gas element.

(c) In hydrostatic equilibrium the net force is zero and the gas element is not accelerated. Find an expression of the pressure gradient in hydrostatic equilibrium.

(d) Find an expression for the central pressure P_c by integrating the pressure gradient. Use this to derive the lower limit on the central pressure of a star in hydrostatic equilibrium, eq. (2.16).

(e) Verify the validity of this lower limit for the case of a star with the density profile of eq. (2.38).

2.3 The virial theorem

An important consequence of hydrostatic equilibrium is that it links the gravitational potential energy $E_{\rm gr}$ and the internal thermal energy $E_{\rm int}$.

(a) Estimate the gravitational energy $E_{\rm gr}$ for a star with mass M and radius R, assuming (1) a constant density distribution and (2) the density distribution of eq. (2.38).

(b) Assume that a star is made of an ideal gas. What is the kinetic internal energy per particle for an ideal gas? Show that the total internal energy, $E_{\rm int}$ is given by:

$$E_{\rm int} = \int_0^R \left(\frac{3}{2}\frac{k}{\mu m_{\rm u}}\rho(r)T(r)\right) 4\pi r^2\,{\rm d}r. \tag{2.39}$$

(c) Estimate the internal energy of the Sun by assuming constant density and $T(r) \approx \langle T\rangle \approx \frac{1}{2}T_{\rm c} \approx 5\times10^6 K$ and compare your answer to your answer for a). What is the total energy of the Sun? Is the Sun bound according to your estimates?

It is no coincidence that the order of magnitude for $E_{\rm gr}$ and $E_{\rm int}$ are the same[1]. This follows from hydrostatic equilibrium and the relation is known as the virial theorem. In the next steps we will derive the virial theorem starting from the pressure gradient in the form of eq. (2.12).

(d) Multiply by both sides of eq. (2.12) by $4\pi r^3$ and integrate over the whole star. Use integration by parts to show that

$$\int_0^R 3\,P\;4\pi r^2\,{\rm d}r = \int_0^R \frac{Gm(r)}{r}\rho 4\pi r^2\,{\rm d}r. \tag{2.40}$$

(e) Now derive a relation between $E_{\rm gr}$ and $E_{\rm int}$, the virial theorem for an ideal gas.

(f) (*) Also show that for the average pressure of the star

$$\langle P\rangle = \frac{1}{V}\int_0^{R_*} P\;4\pi r^2\,{\rm d}r = -\frac{1}{3}\frac{E_{\rm gr}}{V}, \tag{2.41}$$

where V is the volume of the star.

As the Sun evolved towards the main sequence, it contracted under gravity while remaining close to hydrostatic equilibrium. Its internal temperature changed from about 30 000 K to about 6×10^6K.

(g) Find the total energy radiated during away this contraction. Assume that the luminosity during this contraction is comparable to $L_\odot$ and estimate the time taken to reach the main sequence.

2.4 Conceptual questions

[1] In reality $E_{\rm gr}$ is larger than estimated above because the mass distribution is more concentrated to the centre.

(a) Use the virial theorem to explain why stars are hot, i.e. have a high internal temperature and therefore radiate energy.

(b) What are the consequences of energy loss for the star, especially for its temperature?

(c) Most stars are in thermal equilibrium. What is compensating for the energy loss?

(d) What happens to a star in thermal equilibrium (and in hydrostatic equilibrium) if the energy production by nuclear reactions in a star drops (slowly enough to maintain hydrostatic equilibrium)?

(e) Why does this have a stabilizing effect? On what time scale does the change take place?

(f) What happens if hydrostatic equilibrium is violated, e.g. by a sudden increase of the pressure.

(g) On which timescale does the change take place? Can you give examples of processes in stars that take place on this timescale.

2.5 Three important timescales in stellar evolution

(a) The nuclear timescale $\tau_{\rm nuc}$.

i. Calculate the total mass of hydrogen available for fusion over the lifetime of the Sun, if 70% of its mass was hydrogen when the Sun was formed, and only 13% of all hydrogen is in the layers where the temperature is high enough for fusion.

ii. Calculate the fractional amount of mass converted into energy by hydrogen fusion. (Refer to Table 1 for the mass of a proton and of a helium nucleus.)

iii. Derive an expression for the nuclear timescale in solar units, i.e. expressed in terms of $R/R_\odot$, $M/M_\odot$ and $L/L_\odot$.

iv. Use the mass-radius and mass-luminosity relations for main-sequence stars to express the nuclear timescale of main-sequence stars as a function of the mass of the star only.

v. Describe in your own words the meaning of the nuclear timescale.

(b) The thermal timescale $\tau_{\rm KH}$.

i-iii. Answer question (a) iii, iv and v for the thermal timescale and calculate the age of the Sun according to Kelvin.

iv. Why are most stars observed to be main-sequence stars and why is the Hertzsprung-gap called a gap?

(c) The dynamical timescale $\tau_{\rm dyn}$.

i-iii. Answer question (a) iii, iv and v for the dynamical timescale.

iv. In stellar evolution models one often assumes that stars evolve *quasi-statically*, i.e. that the star remains in hydrostatic equilibrium throughout. Why can we make this assumption?

v. Rapid changes that are sometimes observed in stars may indicate that dynamical processes are taking place. From the timescales of such changes - usually oscillations with a characteristic period - we may roughly estimate the average density of the Star. The sun has been observed to oscillate with a period of minutes, white dwarfs with periods of a few tens of seconds. Estimate the average density for the Sun and for white dwarfs.

(d) Comparison.

i. Summarize your results for the questions above by computing the nuclear, thermal and dynamical timescales for a 1, 10 and 25 $M_\odot$ main-sequence star. Put your answers in tabular form.

ii. For each of the following evolutionary stages indicate on which timescale they occur: *pre-main sequence contraction, supernova explosion, core hydrogen burning, core helium burning*.

iii. When the Sun becomes a red giant (RG), its radius will increase to $200R_\odot$ and its luminosity to $3000L_\odot$. Estimate $\tau_{\rm dyn}$ and $\tau_{\rm KH}$ for such a RG.

iv. How large would such a RG have to become for $\tau_{\rm dyn} > \tau_{\rm KH}$? Assume both R and L increase at constant effective temperature.

Chapter 3

Equation of state of stellar interiors

3.1 Local thermodynamic equilibrium

Empirical evidence shows that in a part of space isolated from the rest of the Universe, matter and radiation tend towards a state of *thermodynamic equilibrium*. This equilibrium state is achieved when sufficient interactions take place between the material particles ('collisions') and between the photons and mass particles (scatterings and absorptions). In such a state of thermodynamic equilibrium the radiation field becomes isotropic and the photon energy distribution is described by the Planck function (blackbody radiation). The statistical distribution functions of both the mass particles and the photons are then characterized by a single temperature T.

We know that stars are not isolated systems, because they emit radiation and generate (nuclear) energy in their interiors. Indeed, the surface temperature of the Sun is about 6000 K, while we have estimated from the virial theorem (Sec. 2.3) that the interior temperature must of the order of 10^7 K. Therefore stars are *not* in global thermodynamic equilibrium. However, it turns out that locally within a star, a state of thermodynamic equilibrium *is* achieved. This means that within a region much smaller than the dimensions of a star ($\ll R_*$), but larger than the average distance between interactions of the particles (both gas particles and photons), i.e. larger than the mean free path, there is a well-defined *local temperature* that describes the particle statistical distributions.

We can make this plausible by considering the mean free path for photons:

$$\ell_{\rm ph} = 1/\kappa\rho$$

where κ is the opacity coefficient, i.e. the effective cross section per unit mass. For fully ionized matter, a minimum is given by the electron scattering cross section, which is $\kappa_{\rm es} = 0.4\,{\rm cm}^2/{\rm g}$ (see Ch. 5). The average density in the Sun is $\bar{\rho} = 1.4\,{\rm g/cm}^3$, which gives a mean free path of the order of $\ell_{\rm ph} \sim 1\,{\rm cm}$. In other words, stellar matter is very opaque to radiation. The temperature difference over a distance $\ell_{\rm ph}$, i.e. between emission and absorption, can be estimated as

$$\Delta T \approx \frac{dT}{dr}\ell_{\rm ph} \approx \frac{T_c}{R}\ell_{\rm ph} \approx \frac{10^7}{10^{11}} \approx 10^{-4}\,{\rm K}$$

which is a tiny fraction (10^{-11}) of the typical interior temperature of 10^7 K. Using a similar estimate, it can be shown that the mean free path for interactions between ionized gas particles (ions and electrons) is several orders of magnitude smaller than $\ell_{\rm ph}$. Hence a small region can be defined (a 'point' for all practical purposes) which is $> \ell_{\rm ph}$ but much smaller than the length scale over which significant changes of thermodynamic quantities occur. This is called *local thermodynamic equilibrium* (LTE). We can therefore assume a well-defined temperature distribution inside the star.

Furthermore, the average time between particle interactions (the mean free time) is much shorter than the timescale for changes of the macroscopic properties. Therefore a state of LTE is secured at all times in the stellar interior. The assumption of LTE[1] constitutes a great simplification. It enables the calculation of all thermodynamic properties of the stellar gas in terms of the local values of temperature, density and composition, as they change from the centre to the surface.

3.2 The equation of state

The equation of state (EOS) describes the microscopic properties of stellar matter, for given density ρ, temperature T and composition X_i. It is usually expressed as the relation between the pressure and these quantities:

$$\boxed{P = P(\rho, T, X_i)} \tag{3.1}$$

Using the laws of thermodynamics, and a similar equation for the internal energy $U(\rho, T, X_i)$, we can derive from the EOS the thermodynamic properties that are needed to describe the structure of a star, such as the specific heats c_V and c_P, the adiabatic exponent $\gamma_{\rm ad}$ and the adiabatic temperature gradient $\nabla_{\rm ad}$.

An example is the ideal-gas equation of state, which in the previous chapters we have tacitly assumed to hold for stars like the Sun:

$$P = nkT \quad \text{or} \quad P = \frac{k}{\mu m_{\rm u}}\rho T.$$

In this chapter we will see whether this assumption was justified, and how the EOS can be extended to cover all physical conditions that may prevail inside a star. The ideal-gas law pertains to particles that behave according to classical physics. However, both quantum-mechanical and special relativistic effects may be important under the extreme physical conditions in stellar interiors. In addition, photons (which can be described as extremely relativistic particles) can be an important source of pressure.

We can define an ideal or *perfect* gas as a mixture of free, non-interacting particles. Of course the particles in such a gas do interact, so more precisely we require that their interaction energies are small compared to their kinetic energies. In that case the internal energy of the gas is just the sum of all kinetic energies. From statistical mechanics we can derive the properties of such a perfect gas, both in the classical limit (recovering the ideal-gas law) and in the quantum-mechanical limit (leading to electron degeneracy), and both in the non-relativistic and in the relativistic limit (e.g. valid for radiation). This is done in Sect. 3.3.

In addition, various *non-ideal* effects may become important. The high temperatures ($> 10^6$ K) in stellar interiors ensure that the gas will be fully ionized, but at lower temperatures (in the outer layers) partial ionization has to be considered, with important effects on the thermodynamic properties (see Sect. 3.5). Furthermore, in an ionized gas *electrostatic interactions* between the ions and electrons may be important under certain circumstances (Sect. 3.6).

3.3 Equation of state for a gas of free particles

We shall derive the equation of state for a perfect gas from the principles of statistical mechanics. This provides a description of the ions, the electrons, as well as the photons in the deep stellar interior.

[1] N.B. note the difference between (local) *thermodynamic equilibrium* ($T_{\rm gas}(r) = T_{\rm rad}(r) = T(r)$) and the earlier defined, global property of *thermal equilibrium* ($E_{\rm tot}$ = const, or $L = L_{\rm nuc}$).

Let $n(p)$ be the distribution of momenta of the gas particles, i.e. $n(p)\,\mathrm{d}p$ represents the number of particles per unit volume with momenta $p \in [p \ldots p + \mathrm{d}p]$. If $n(p)$ is known then the total number density (number of particles per unit volume), the internal energy density (internal energy per unit volume) and the pressure can be obtained from the following integrals:

$$\text{number density} \qquad n = \int_0^\infty n(p)\,\mathrm{d}p \tag{3.2}$$

$$\text{internal energy density} \qquad U = \int_0^\infty \epsilon_p n(p)\,\mathrm{d}p = n\langle\epsilon_p\rangle \tag{3.3}$$

$$\text{pressure} \qquad P = \tfrac{1}{3}\int_0^\infty p v_p n(p)\,\mathrm{d}p = \tfrac{1}{3} n\langle p v_p\rangle \tag{3.4}$$

Here ϵ_p is the kinetic energy of a particle with momentum p, and v_p is its velocity. Eq. (3.2) is trivial, and eq. (3.3) follows from the perfect-gas assumption. The pressure integral eq. (3.4) requires some explanation.

Consider a gas of n particles in a cubical box with sides of length $L = 1\,\mathrm{cm}$. Each particle bounces around in the box, and the pressure on one side of the box results from the momentum imparted by all the particles colliding with it. Consider a particle with momentum p and corresponding velocity v coming in at an angle θ with the normal to the surface, as depicted in Fig. 3.1. The time between two collisions with the same side is

$$\Delta t = \frac{2L}{v\cos\theta} = \frac{2}{v\cos\theta}.$$

The collisions are elastic, so the momentum transfer is twice the momentum component perpendicular to the surface,

$$\Delta p = 2p\cos\theta. \tag{3.5}$$

The momentum transferred per particle per second and per cm^2 is therefore

$$\frac{\Delta p}{\Delta t} = vp\,\cos^2\theta. \tag{3.6}$$

The number of particles in the box with $p \in [p \ldots p + \mathrm{d}p]$ and $\theta \in [\theta \ldots \theta + \mathrm{d}\theta]$ is denoted as $n(\theta, p)\,\mathrm{d}\theta\,\mathrm{d}p$. The contribution to the pressure from these particles is then

$$\mathrm{d}P = vp\,\cos^2\theta\, n(\theta, p)\,\mathrm{d}\theta\,\mathrm{d}p. \tag{3.7}$$

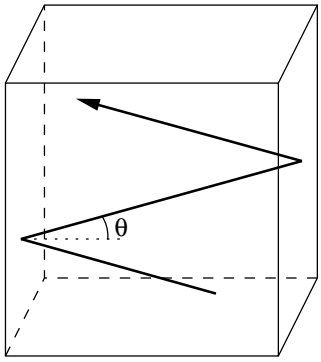


Figure 3.1. Gas particle in a cubical box with a volume of $1\,\mathrm{cm}^3$. Each collision with the side of the box results in a transfer of momentum; the pressure inside the box is the result of the collective momentum transfers of all n particles in the box.

Since the momenta are distributed isotropically over all directions within a solid angle 2π, and the solid angle $\mathrm{d}\omega$ subtended by those particles with $\theta \in [\theta \ldots \theta + \mathrm{d}\theta]$ equals $2\pi \sin\theta \,\mathrm{d}\theta$, we have $n(\theta, p)\,\mathrm{d}\theta = n(p)\, sin\theta\, \mathrm{d}\theta$ and

$$\mathrm{d}P = vp\, n(p)\, \cos^2\theta \sin\theta \,\mathrm{d}\theta \,\mathrm{d}p. \tag{3.8}$$

The total pressure is obtained by integrating over all angles ($0 \leq \theta \leq \pi/2$) and momenta. This results in eq. (3.4) since $\int_0^{\pi/2} \cos^2\theta \sin\theta \,\mathrm{d}\theta = \int_0^1 \cos^2\theta \,\mathrm{d}\cos\theta = \frac{1}{3}$.

3.3.1 Relation between pressure and internal energy

In general, the particle energies and velocities are related to their momenta according to special relativity:

$$\epsilon^2 = p^2c^2 + m^2c^4, \qquad \epsilon_p = \epsilon - mc^2 \tag{3.9}$$

and

$$v_p = \frac{\partial\epsilon}{\partial p} = \frac{pc^2}{\epsilon}. \tag{3.10}$$

We can obtain generally valid relations between the pressure and the internal energy of a perfect gas in the non-relativistic (NR) limit and the extremely relativistic (ER) limit:

NR limit: in this case the momenta $p \ll mc$, so that $\epsilon_p = \epsilon - mc^2 = \frac{1}{2}p^2/m$ and $v = p/m$. Therefore $\langle pv\rangle = \langle p^2/m\rangle = 2\langle\epsilon_p\rangle$ so that eq. (3.4) yields

$$\boxed{P = \tfrac{2}{3}U} \tag{3.11}$$

ER limit: in this case $p \gg mc$, so that $\epsilon_p = pc$ and $v = c$. Therefore $\langle pv\rangle = \langle pc\rangle = \langle\epsilon_p\rangle$, and eq. (3.4) yields

$$\boxed{P = \tfrac{1}{3}U} \tag{3.12}$$

These relations are generally true, for *any particle* (electrons, ions and photons). We will apply this in the coming sections. As we saw in the previous Chapter, the change from $\frac{2}{3}$ to $\frac{1}{3}$ in the relation has important consequences for the virial theorem, and for the stability of stars.

3.3.2 The classical ideal gas

Using the tools of statistical mechanics, we can address the origin of the ideal-gas law. The momentum distribution $n(p)$ for classical, non-relativistic particles of mass m in LTE is given by the *Maxwell-Boltzmann* distribution:

$$n(p)\,\mathrm{d}p = \frac{n}{(2\pi mkT)^{3/2}}\, e^{-p^2/2mkT}\, 4\pi p^2\,\mathrm{d}p. \tag{3.13}$$

Here the exponential factor ($e^{-\epsilon_p/kT}$) represents the equilibrium distribution of kinetic energies, the factor $4\pi p^2\,\mathrm{d}p$ is the volume in momentum space (p_x, p_y, p_z) for $p \in [p \ldots p + \mathrm{d}p]$, and the factor $n/(2\pi mkT)^{3/2}$ comes from the normalization of the total number density n imposed by eq. (3.2). (You can verify this by starting from the standard integral $\int_0^\infty e^{-ax^2}\,\mathrm{d}x = \frac{1}{2}\sqrt{\pi/a}$, and differentiating once with respect to a to obtain the integral $\int_0^\infty e^{-ax^2}x^2\,\mathrm{d}x$.)

The pressure is calculated by using $v = p/m$ for the velocity in eq. (3.4):

$$P = \tfrac{1}{3}\frac{n}{(2\pi mkT)^{3/2}}\int_0^\infty \frac{p^2}{m}\, e^{-p^2/2mkT}\, 4\pi p^2\, \mathrm{d}p. \tag{3.14}$$

By performing the integration (for this you need to differentiate $\int_0^\infty e^{-ax^2}x^2\,\mathrm{d}x$ once more with respect to a) you can verify that this indeed yields the ideal gas law

$$\boxed{P = nkT}\,. \tag{3.15}$$

(N.B. This derivation is for a gas of *non-relativistic* classical particles, but it can be shown that the same relation $P = nkT$ is also valid for *relativistic* classical particles.)

3.3.3 Mixture of ideal gases, and the mean molecular weight

The ideal gas relation was derived for identical particles of mass m. It should be obvious that for a mixture of free particles of different species, it holds for the partial pressures of each of the constituents of the gas separately. In particular, it holds for both the ions and the electrons, as long as quantum-mechanical effects can be ignored. The total gas pressure is then just the sum of partial pressures

$$P_{\rm gas} = P_{\rm ion} + P_{\rm e} = \textstyle\sum_i P_i + P_{\rm e} = (\sum_i n_i + n_{\rm e})kT = nkT$$

where n_i is the number density of ions of element i, with mass $m_i = A_i m_{\rm u}$ and charge $Z_i e$. Then n_i is related to the density and the mass fraction X_i of this element as

$$n_i = \frac{X_i\,\rho}{A_i\,m_{\rm u}} \qquad \text{and} \qquad n_{\rm ion} = \sum_i \frac{X_i}{A_i}\frac{\rho}{m_{\rm u}} \equiv \frac{1}{\mu_{\rm ion}}\frac{\rho}{m_{\rm u}}, \tag{3.16}$$

which defines the mean atomic mass per ion $\mu_{\rm ion}$. The partial pressure due to all ions is then

$$P_{\rm ion} = \frac{1}{\mu_{\rm ion}}\frac{\rho}{m_{\rm u}}kT = \frac{\mathcal{R}}{\mu_{\rm ion}}\rho T. \tag{3.17}$$

We have used here the universal gas constant $\mathcal{R} = k/m_{\rm u} = 8.31447 \times 10^7\,\mathrm{erg\,g^{-1}\,K^{-1}}$. The number density of electrons is given by

$$n_{\rm e} = \sum_i Z_i n_i = \sum_i \frac{Z_i X_i}{A_i}\frac{\rho}{m_{\rm u}} \equiv \frac{1}{\mu_{\rm e}}\frac{\rho}{m_{\rm u}}, \tag{3.18}$$

which defines the *mean molecular weight per free electron* $\mu_{\rm e}$. As long as the electrons behave like classical particles, the electron pressure is thus given by

$$P_{\rm e} = \frac{1}{\mu_{\rm e}}\frac{\rho}{m_{\rm u}}kT = \frac{\mathcal{R}}{\mu_{\rm e}}\rho T. \tag{3.19}$$

When the gas is fully ionized, we have for hydrogen $Z_i = A_i = 1$ while for helium and the most abundant heavier elements, $Z_i/A_i \approx \frac{1}{2}$. In terms of the hydrogen mass fraction X we then get

$$\mu_{\rm e} \approx \frac{2}{1+X}, \tag{3.20}$$

which for the Sun ($X = 0.7$) amounts to $\mu_{\rm e} \approx 1.18$, and for hydrogen-depleted gas gives $\mu_{\rm e} \approx 2$.

The total gas pressure is then given by

$$P_{\rm gas} = P_{\rm ion} + P_{\rm e} = \Big(\frac{1}{\mu_{\rm ion}} + \frac{1}{\mu_{\rm e}}\Big)\mathcal{R}\rho T = \frac{\mathcal{R}}{\mu}\rho T \tag{3.21}$$

where the *mean molecular weight* μ is given by

$$\frac{1}{\mu} = \frac{1}{\mu_{\text{ion}}} + \frac{1}{\mu_{\text{e}}} = \sum_i \frac{(Z_i + 1)X_i}{A_i}. \tag{3.22}$$

It is left as an exercise to show that for a fully ionized gas, μ can be expressed in terms of the mass fractions X, Y and Z as

$$\mu \approx \frac{1}{2X + \frac{3}{4}Y + \frac{1}{2}Z} \tag{3.23}$$

if we assume that for elements heavier than helium, $A_i \approx 2Z_i \approx 2(Z_i + 1)$.

3.3.4 Quantum-mechanical description of the gas

According to quantum mechanics, the accuracy with which a particle's location and momentum can be known simultaneously is limited by Heisenberg's uncertainty principle, i.e. $\Delta x \Delta p \geq h$. In three dimensions, this means that if a particle is located within a volume element ΔV then its localization within three-dimensional momentum space $\Delta^3 p$ is constrained by

$$\Delta V \, \Delta^3 p \geq h^3. \tag{3.24}$$

The quantity h^3 defines the volume in six-dimensional phase space of one quantum cell. The *number of quantum states* in a spatial volume V and with momenta $p \in [p \ldots p + \mathrm{d}p]$ is therefore given by

$$g(p)\,\mathrm{d}p = g_s \frac{V}{h^3} 4\pi p^2 \,\mathrm{d}p, \tag{3.25}$$

where g_s is the number of intrinsic quantum states of the particle, e.g. spin or polarization.

The relative occupation of the available quantum states for particles in thermodynamic equilibrium depends on the type of particle:

- *fermions* (e.g. electrons or nucleons) obey the Pauli exclusion principle, which postulates that no two such particles can occupy the same quantum state. The fraction of states with energy ϵ_p that will be occupied at temperature T is given by

  $$f_{\text{FD}}(\epsilon_p) = \frac{1}{e^{(\epsilon_p - \mu)/kT} + 1}, \tag{3.26}$$

  which is always ≤ 1.

- *bosons* (e.g. photons) have no restriction on the number of particles per quantum state, and the fraction of states with energy ϵ_p that is occupied is

  $$f_{\text{BE}}(\epsilon_p) = \frac{1}{e^{(\epsilon_p - \mu)/kT} - 1}, \tag{3.27}$$

  which can be > 1.

The actual distribution of momenta for particles in LTE is given by the product of the occupation fraction $f(\epsilon_p)$ and the number of quantum states, given by eq. (3.25). The quantity μ appearing in eqs. (3.26) and (3.27) is the so-called *chemical potential*. It can be seen as a normalization constant, determined by the total number of particles in the volume considered (i.e., by the constraint imposed by eq. 3.2).

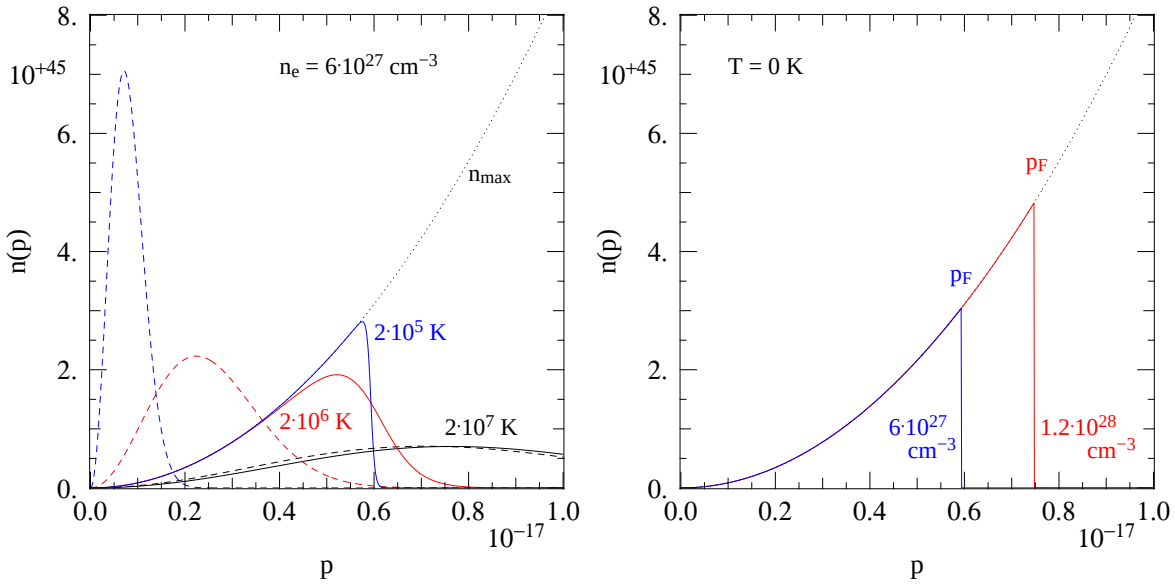


Figure 3.2. *Left:* Electron momentum distributions $n(p)$ for an electron density of $n_e = 6 \times 10^{27}$ cm^{-3} (corresponding to $\rho = 2 \times 10^4$ g/cm^{-3} if $\mu_e = 2$), and for three different temperatures: $T = 2 \times 10^7$ K (black lines), 2×10^6 K (red lines) and 2×10^5 K (blue lines). The actual distributions, governed by quantum mechanics, are shown as solid lines while the Maxwell-Boltzmann distributions for the same n_e and T values are shown as dashed lines. The dotted line n_{max} is the maximum possible number distribution if all quantum states with momentum p are occupied. *Right:* Distributions in the limit $T = 0$, when all lowest available momenta are fully occupied. The blue line is for the same density as in the left panel, while the red line is for a density two times as high.

3.3.5 Electron degeneracy

Electrons are fermions with two spin states, i.e. $g_e = 2$. According to eq. (3.25), the maximum number density of electrons with momentum p allowed by quantum mechanics is therefore

$$n_{max}(p)\,dp = \frac{g_e}{h^3}\,4\pi p^2\,dp = \frac{8\pi}{h^3}\,p^2\,dp. \tag{3.28}$$

This is shown as the dotted line in Fig. 3.2. The actual momentum distribution of electrons $n_e(p)$ is given by the product of eq. (3.28) and eq. (3.26). In the non-relativistic limit we have $\epsilon_p = p^2/2m_e$, giving

$$n_e(p)\,dp = \frac{2}{h^3}\,\frac{1}{e^{(p^2/2m_e kT)-\psi}+1}\,4\pi p^2\,dp, \tag{3.29}$$

where we have replaced the chemical potential by the *degeneracy parameter* $\psi = \mu/kT$. The value of ψ is determined by the constraint that $\int_0^\infty n_e(p)\,dp = n_e$ (eq. 3.2).

The limitation imposed by the Pauli exclusion principle means that electrons can exert a higher pressure than predicted by classical physics (eq. 3.19). To illustrate this, in Fig. 3.2 the momentum distribution eq. (3.29) is compared to the Maxwell-Boltzmann distribution for electrons, eq. (3.13),

$$n_{MB}(p)\,dp = \frac{n_e}{(2\pi m_e kT)^{3/2}}\,e^{-p^2/2m_e kT}\,4\pi p^2\,dp. \tag{3.30}$$

The situation shown is for an electron density $n_e = 6 \times 10^{27}$ cm^{-3}, which corresponds to a mass density of 2×10^4 g/cm^{-3} (assuming a hydrogen-depleted gas with $\mu_e = 2$). At high temperatures, $T = 2 \times 10^7$ K, the momentum distribution (solid line) nearly coincides with the M-B distribution

(dashed line): none of the quantum states are fully occupied ($n_e(p) < n_{max}(p)$ for all values of p) and the electrons behave like classical particles. As the temperature is decreased, e.g. at $T = 2 \times 10^6$ K (red lines), the peak in the M-B distribution shifts to smaller p and is higher (since the integral over the distribution must equal n_e). The number of electrons with small values of p expected from classical physics, $n_{MB}(p)$, then exceeds the maximum allowed by the Pauli exclusion principle, $n_{max}(p)$. These electrons are forced to assume quantum states with higher p: the peak in the distribution $n_e(p)$ occurs at higher p. Due to the higher momenta and velocities of these electrons, the electron gas exerts a higher pressure than inferred from classical physics. This is called *degeneracy pressure*. If the temperature is decreased even more, e.g. at $T = 2 \times 10^5$ K (blue lines), the lowest momentum states become nearly all filled and $n_e(p)$ follows $n_{max}(p)$ until it drops sharply. In this state of strong degeneracy, further decrease of T hardly changes the momentum distribution, so that the electron pressure becomes nearly *independent of temperature*.

Complete electron degeneracy

In the limit that $T \to 0$, all available momentum states are occupied up to a maximum value, while all higher states are empty, as illustrated in the right panel of Fig. 3.2. This is known as *complete degeneracy*, and the maximum momentum is called the *Fermi momentum* p_F. Then we have

$$n_e(p) = \frac{8\pi p^2}{h^3} \qquad \text{for } p \le p_F, \tag{3.31}$$

$$n_e(p) = 0 \qquad \text{for } p > p_F. \tag{3.32}$$

The Fermi momentum is determined by the electron density through eq. (3.2), i.e. $\int_0^{p_F} n_e(p)\,\mathrm{d}p = n_e$, which yields

$$p_F = h\Big(\frac{3}{8\pi} n_e\Big)^{1/3}. \tag{3.33}$$

The pressure of a completely degenerate electron gas is now easy to compute using the pressure integral eq. (3.4). It depends on whether the electrons are relativistic or not. In the *non-relativistic* limit we have $v = p/m$ and hence

$$P_e = \tfrac{1}{3}\int_0^{p_F} \frac{8\pi p^4}{h^3 m_e}\,\mathrm{d}p = \frac{8\pi}{15 h^3 m_e}\, {p_F}^5 = \frac{h^2}{20 m_e}\left(\frac{3}{\pi}\right)^{2/3} {n_e}^{5/3}. \tag{3.34}$$

Using eq. (3.18) for n_e this can be written as

$$\boxed{P_e = K_{NR}\left(\frac{\rho}{\mu_e}\right)^{5/3}} \qquad \text{with} \qquad K_{NR} = \frac{h^2}{20 m_e\, m_u^{5/3}}\left(\frac{3}{\pi}\right)^{2/3} = 1.0036 \times 10^{13}\ \text{[cgs]}. \tag{3.35}$$

As more electrons are squeezed into the same volume, they have to occupy states with larger momenta, as illustrated in Fig. 3.2. Therefore the electron pressure increases with density, as expressed by eq. (3.35).

If the electron density is increased further, at some point the velocity of the most energetic electrons, p_F/m_e, approaches the speed of light. We then have to replace $v = p/m$ by the relativistic kinematics relation (3.10). In the *extremely relativistic* limit when the majority of electrons move at relativistic speeds, we can take $v = c$ and

$$P_e = \tfrac{1}{3}\int_0^{p_F} \frac{8\pi c p^3}{h^3}\,\mathrm{d}p = \frac{8\pi c}{12 h^3}\, p_F^4 = \frac{hc}{8}\left(\frac{3}{\pi}\right)^{1/3} {n_e}^{4/3}, \tag{3.36}$$

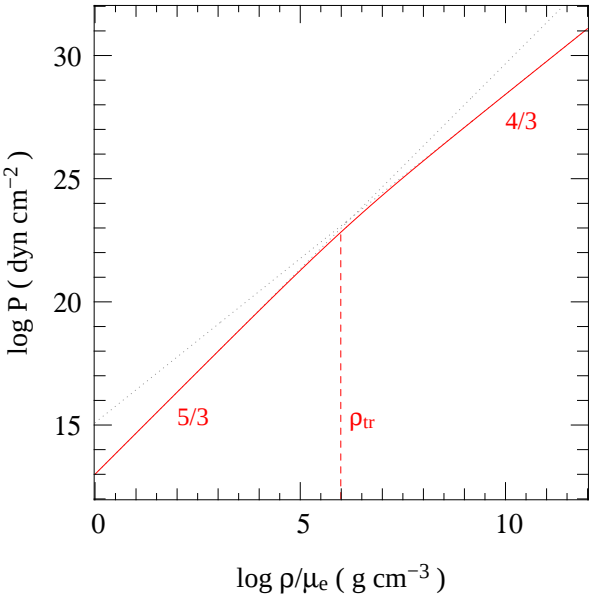


Figure 3.3. The equation of state for completely degenerate electrons. The slope of the $\log P$-$\log \rho$ relation changes from 5/3 at relatively low densities, where the electrons are non-relativistic, to 4/3 at high density when the electrons are extremely relativistic. The transition is smooth, but takes place at densities around $\rho_{\rm tr} \approx 10^6 \mu_{\rm e}$ g cm^{-3}.

which gives

$$\boxed{P_{\rm e} = K_{\rm ER} \left(\frac{\rho}{\mu_{\rm e}}\right)^{4/3}} \qquad \text{with} \qquad K_{\rm ER} = \frac{hc}{8 m_{\rm u}^{4/3}} \left(\frac{3}{\pi}\right)^{1/3} = 1.2435 \times 10^{15} \text{ [cgs]}. \tag{3.37}$$

In the ER limit the pressure still increases with density, but with a smaller exponent ($\frac{4}{3}$ instead of $\frac{5}{3}$). The transition between the NR regime, eq. (3.35), and the ER regime, eq. (3.37), is smooth and can be expressed as a function of $x = p_{\rm F}/m_{\rm e}c$, see Maeder Sec. 7.7. Roughly, the transition occurs at a density $\rho_{\rm tr}$ given by the condition $p_{\rm F} \approx m_{\rm e}c$, which can be expressed as

$$\rho_{\rm tr} \approx \mu_{\rm e}\, m_{\rm u} \frac{8\pi}{3} \left(\frac{m_{\rm e}c}{h}\right)^3. \tag{3.38}$$

The relation between $P_{\rm e}$ and ρ for a completely degenerate electron gas is shown in Fig. 3.3.

Partial degeneracy

Although the situation of complete degeneracy is only achieved at $T = 0$, it is a very good approximation whenever the degeneracy is strong, i.e. when the temperature is sufficiently low, as illustrated by Fig. 3.2. It corresponds to the situation when the degeneracy parameter $\psi \gg 0$ in eq. (3.29). In that case eqs. (3.35) and (3.37) can still be used to calculate the pressure to good approximation.

The transition between the classical ideal gas situation and a state of strong degeneracy occurs smoothly, and is known as *partial degeneracy*. To calculate the pressure the full expression eq. (3.29) has to be used in the pressure integral, which becomes rather complicated. The integral then depends on ψ, and can be expressed as one of the so-called *Fermi-Dirac* integrals, see Maeder Sec. 7.7 for details (the other Fermi-Dirac integral relates to the internal energy density U). The situation of partial degeneracy corresponds to $\psi \sim 0$.

When $\psi \ll 0$ the classical description is recovered, i.e. eq. (3.29) becomes the Maxwell-Boltzmann distribution. In that case $1/(e^{(p^2/2m_{\rm e}kT)-\psi} + 1) = e^{-(p^2/2m_{\rm e}kT)+\psi}$ and therefore

$$\frac{2}{h^3} e^{\psi} = \frac{n_{\rm e}}{(2\pi m_{\rm e} kT)^{3/2}} \qquad \text{or} \qquad \psi = \ln \frac{h^3 n_{\rm e}}{2(2\pi m_{\rm e} kT)^{3/2}}.$$

This only holds for $\psi \ll 0$, but more generally it can be shown that $\psi = \psi(n_e/T^{2/3})$. We have to consider (partial) degeneracy if $\psi \gtrsim 0$, i.e. if

$$n_e \gtrsim \frac{2(2\pi m_e kT)^{3/2}}{h^3}. \tag{3.39}$$

The limit of strong (almost complete) degeneracy is reached when n_e is roughly a factor 10 higher.

Importance of electron degeneracy in stars

As a star, or its core, contracts the density may become so high that the electrons become degenerate and exert a (much) higher pressure than they would if they behaved classically. Since in the limit of strong degeneracy the pressure no longer depends on the temperature, this degeneracy pressure can hold the star up against gravity, regardless of the temperature. Therefore a degenerate star does not have to be hot to be in hydrostatic equilibrium, and it can remain in this state forever even when it cools down. This is the situation in *white dwarfs*.

The importance of relativity is that, when a degenerate star becomes more compact and the density increases further, the pressure increases less steeply with density. This has important consequences for massive white dwarfs, and we shall see that it implies that there is a maximum mass for which white dwarfs can exist (the Chandrasekhar mass).

We note that although electron degeneracy can be (very) important in stars, degeneracy of the *ions* is not. Since the ions have masses $\gtrsim 2000$ larger than electrons, their momenta ($p = \sqrt{2m\epsilon}$) are much larger at energy equipartition, and the condition (3.39) above (with m_e replaced by m_{ion}) implies that much higher densities are required at a particular temperature. In practice this never occurs: before such densities are reached the protons in the atomic nuclei will capture free electrons, and the composition becomes one of (mostly) neutrons. Degeneracy of *neutrons* does become important when we consider neutron stars.

3.3.6 Radiation pressure

Photons can be treated as quantum-mechanical particles that carry momentum and therefore exert pressure when they interact with matter. In particular photons are *bosons* with $g_s = 2$ (two polarization states), so they can be described by the Bose-Einstein statistics, eq. (3.27). The number of photons is not conserved, they can be destroyed and created until thermodynamic equilibrium is achieved. This means that $\mu = 0$ in eq. (3.27) and hence

$$n(p)\,\mathrm{d}p = \frac{2}{h^3}\,\frac{1}{e^{\epsilon_p/kT} - 1}\,4\pi p^2\,\mathrm{d}p \tag{3.40}$$

Photons are completely relativistic with $\epsilon_p = pc = h\nu$, so in terms of frequency ν their distribution in LTE becomes the *Planck function* for blackbody radiation:

$$n(\nu)\,\mathrm{d}\nu = \frac{8\pi}{c^3}\,\frac{\nu^2\,\mathrm{d}\nu}{e^{h\nu/kT} - 1} \tag{3.41}$$

Applying eqs. (3.2) and (3.3) one can show that the photon number density and the energy density of radiation are

$$n_{ph} = \int_0^\infty n(p)\,\mathrm{d}p = b\,T^3 \tag{3.42}$$

$$U_{rad} = \int_0^\infty pc\,n(p)\,\mathrm{d}p = a\,T^4 \tag{3.43}$$

where $b = 20.3\,\mathrm{cm^{-3}\,K^{-3}}$ and a is the radiation constant

$$a = \frac{8\pi^5 k^4}{15h^3c^3} = 7.56 \times 10^{-15}\,\mathrm{erg\,cm^{-3}\,K^{-4}}.$$

Since photons are always extremely relativistic, $P = \frac{1}{3}U$ by eq. (3.12) and the *radiation pressure* is given by

$$\boxed{P_{\mathrm{rad}} = \tfrac{1}{3}aT^4} \tag{3.44}$$

Pressure of a mixture of gas and radiation

The pressure inside a star is the sum of the gas pressure and radiation pressure,

$$P = P_{\mathrm{rad}} + P_{\mathrm{gas}} = P_{\mathrm{rad}} + P_{\mathrm{ion}} + P_{\mathrm{e}}.$$

where P_{rad} is given by eq. (3.44) and P_{ion} by eq. (3.17). In general P_{e} must be calculated as described in Sect. 3.3.5. In the classical limit it is given by eq. (3.19), and in the limits of non-relativistic and extremely relativistic degeneracy by eqs. (3.35) and (3.37), respectively. If the electrons are non-degenerate then the pressure can be written as

$$P = \tfrac{1}{3}aT^4 + \frac{\mathcal{R}}{\mu}\rho T. \tag{3.45}$$

If the electrons are strongly degenerate their pressure dominates over that of the (classical) ions, so in that case P_{ion} can be neglected in the total pressure.

The fraction of the pressure contributed by the gas is customarily expressed as β, i.e.

$$P_{\mathrm{gas}} = \beta P \qquad \text{and} \qquad P_{\mathrm{rad}} = (1-\beta)P. \tag{3.46}$$

3.3.7 Equation of state regimes

The different sources of pressure we have discussed so far dominate the equation of state at different temperatures and densities. In Fig. 3.4 the boundaries between these regimes are plotted schematically in the $\log T$, $\log\rho$ plane.

- The boundary between regions where radiation and ideal-gas pressure dominate is defined by $P_{\mathrm{rad}} = P_{\mathrm{gas}}$, giving $T/\rho^{1/3} = 3.2 \times 10^7 \mu^{-1/3}$ when T and ρ are expressed in cgs units. (Verify this by comparing eqs. 3.21 and 3.44.) This is a line with slope $\frac{1}{3}$ in the $\log T$ vs $\log\rho$ plane.
- Similarly, the boundary between the regions dominated by ideal-gas pressure and non-relativistic degenerate electron pressure can be defined by $P_{\mathrm{gas,ideal}} = P_{e,\mathrm{NR}}$ as given by eq. (3.35), giving $T/\rho^{2/3} = 1.21 \times 10^5 \mu\,\mu_{\mathrm{e}}^{-5/3}$ (again with T and ρ in cgs units). This is a line with slope $\frac{2}{3}$ in the $\log T$-$\log\rho$ plane.
- The approximate boundary between non-relativistic and relativistic degeneracy is given by eq. (3.38), $\rho = 9.7 \times 10^5 \mu_{\mathrm{e}}\,\mathrm{g/cm^3}$.
- At high densities the boundary between ideal gas pressure and extremely relativistic degeneracy is found by equating eqs. (3.21) and (3.37), giving $T/\rho^{1/3} = 1.50 \times 10^7 \mu\,\mu_{\mathrm{e}}^{-4/3}$ (with T and ρ in cgs units), again a line with slope $\frac{1}{3}$.

As shown in Fig. 3.4, detailed models of zero-age (that is, homogeneous) main-sequence stars with masses between 0.1 and 100 $M_\odot$ cover the region where ideal-gas pressure dominates the equation of state. This justifies the assumptions made in Ch. 2 when discussing the virial theorem and its consequences for stars, and when estimating temperatures in the stellar interior.

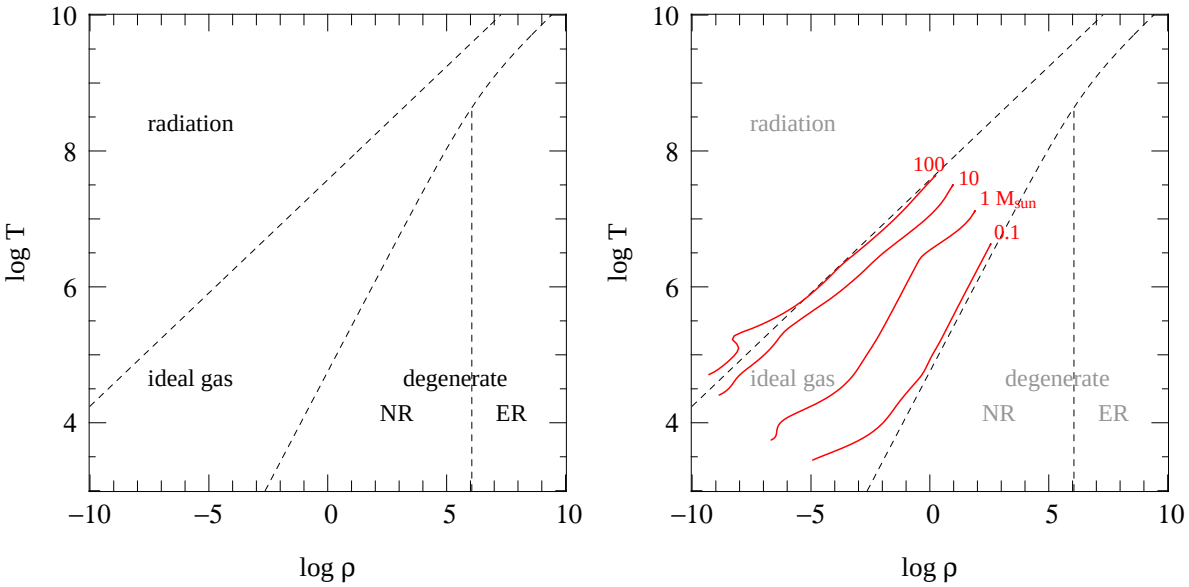


Figure 3.4. *Left:* The equation of state for a gas of free particles in the $\log T$, $\log \rho$ plane. The dashed lines are approximate boundaries between regions where radiation pressure, ideal gas pressure, non-relativistic electron degeneracy and extremely relativistic electron degeneracy dominate, for a composition $X = 0.7$ and $Z = 0.02$. *Right:* Detailed structure models for homogeneous main-sequence stars of 0.1...100 $M_\odot$ have been added (solid lines). The 1 $M_\odot$ model is well within the ideal-gas region of the equation of state. In the 0.1 $M_\odot$ star electron degeneracy pressure is important, except in the outer layers (at low ρ and T). In stars more massive than 10 $M_\odot$, radiation pressure becomes important, and it dominates in the surface layers of the 100 $M_\odot$ model.

3.4 Adiabatic processes

It is often important to consider processes that occur on such a short (e.g. hydrodynamical) timescale that there is no heat exchange with the environment; such processes are *adiabatic*. To derive the properties of stellar interiors under adiabatic conditions we need several thermodynamic derivatives. We therefore start from the laws of thermodynamics.

The *first law* of thermodynamics states that the amount of heat absorbed by a system (δQ) is the sum of the change in its internal energy (δU) and the work done on the system ($\delta W = P\delta V$). The *second law* of thermodynamics states that, for a reversible process, the change in entropy equals the change in the heat content divided by the temperature. Entropy is a state variable, unlike the heat content. For a unit mass (1 gram) of matter the combination of these laws can be expressed as

$$\mathrm{d}q = T\,\mathrm{d}s = \mathrm{d}u + P\,\mathrm{d}v = \mathrm{d}u - \frac{P}{\rho^2}\,\mathrm{d}\rho. \tag{3.47}$$

Here $\mathrm{d}q$ is the change in heat content, $\mathrm{d}u$ is the change in internal energy ($u = U/\rho$ is the *specific* internal energy, i.e. per gram), s is the specific entropy (i.e. the entropy per unit mass) and $v = 1/\rho$ is the volume of a unit mass. Note that $\mathrm{d}u$ and $\mathrm{d}s$ are exact differentials, whereas $\mathrm{d}q$ is not.

Differential form of the equation of state To compute general expressions for thermodynamic derivatives such as the specific heats and the adiabatic derivatives it is useful to write the equation of state in differential form, i.e.

$$\frac{\mathrm{d}P}{P} = \chi_T \frac{\mathrm{d}T}{T} + \chi_\rho \frac{\mathrm{d}\rho}{\rho}, \tag{3.48}$$

where χ_T and χ_ρ are defined as

$$\chi_T = \left(\frac{\partial \log P}{\partial \log T}\right)_{\rho,X_i} = \frac{T}{P}\left(\frac{\partial P}{\partial T}\right)_{\rho,X_i}, \tag{3.49}$$

$$\chi_\rho = \left(\frac{\partial \log P}{\partial \log \rho}\right)_{T,X_i} = \frac{\rho}{P}\left(\frac{\partial P}{\partial \rho}\right)_{T,X_i}. \tag{3.50}$$

The subscript X_i means that the composition is held constant as well. In a general equation of state χ_T and χ_ρ can depend on T and ρ themselves, but if they are (approximately) constant then we can write the equation of state in power-law form:

$$P = P_0\, \rho^{\chi_\rho}\, T^{\chi_T}.$$

For example, for an ideal gas without radiation we have $\chi_T = \chi_\rho = 1$, while for a radiation-dominated gas $\chi_T = 4$ and $\chi_\rho = 0$.

3.4.1 Specific heats

The specific heats at constant volume c_V and at constant pressure c_P for a unit mass of gas follow from eq. (3.47):

$$c_V = \left(\frac{\mathrm{d}q}{\mathrm{d}T}\right)_v = \left(\frac{\partial u}{\partial T}\right)_v, \tag{3.51}$$

$$c_P = \left(\frac{\mathrm{d}q}{\mathrm{d}T}\right)_P = \left(\frac{\partial u}{\partial T}\right)_P - \frac{P}{\rho^2}\left(\frac{\partial \rho}{\partial T}\right)_P, \tag{3.52}$$

where a partial derivative taken at constant v is the same as one taken at constant ρ. For an ideal gas, with $u = U/\rho = \frac{3}{2}P/\rho$, we obtain from eq. (3.21) the familiar result $c_V = \frac{3}{2}\mathcal{R}/\mu$. For a radiation-dominated gas, eq. (3.43) yields $c_V = 4aT^3/\rho$. Using thermodynamic transformations and some algebraic manipulation (see Appendix 3.A), it follows quite generally that the specific heats are related by

$$c_P - c_V = \frac{P}{\rho T}\frac{{\chi_T}^2}{\chi_\rho}. \tag{3.53}$$

For an ideal gas this amounts to $c_P - c_V = \mathcal{R}/\mu$, and therefore $c_P = \frac{5}{2}\mathcal{R}/\mu$. For a radiation-dominated gas $\chi_\rho = 0$ and hence $c_P \to \infty$: indeed, since $P_{\rm rad}$ only depends on T, a change in temperature cannot be performed at constant pressure.

The ratio of specific heats is denoted as γ:

$$\gamma = \frac{c_P}{c_V} = 1 + \frac{P}{\rho T c_V}\frac{{\chi_T}^2}{\chi_\rho}, \tag{3.54}$$

so that $\gamma = \frac{5}{3}$ for an ideal gas.

Expressions for dq It is often useful to have expressions for the change in heat content dq (eq. 3.47) in terms of variations of T and ρ or T and P. Making use of the specific heats one can derive (see Appendix 3.A)

$$\mathrm{d}q \;=\; T\,\mathrm{d}s \;=\; c_V\,\mathrm{d}T - \chi_T\,\frac{P}{\rho^2}\,\mathrm{d}\rho \;=\; c_P\,\mathrm{d}T - \frac{\chi_T}{\chi_\rho}\,\frac{\mathrm{d}P}{\rho}. \tag{3.55}$$

3.4.2 Adiabatic derivatives

The thermodynamic response of a system to adiabatic changes is measured by the so-called *adiabatic derivatives*. Two of these have special importance for stellar structure:

- The *adiabatic exponent*[2] $\gamma_{\rm ad}$ measures the response of the pressure to adiabatic compression or expansion, i.e. to a change in the density. It is defined as

$$\gamma_{\rm ad} = \left(\frac{\partial \log P}{\partial \log \rho}\right)_{\rm ad} \tag{3.56}$$

  where the subscript 'ad' means that the change is performed adiabatically, that is, at constant entropy. If $\gamma_{\rm ad}$ is constant then $P \propto \rho^{\gamma_{\rm ad}}$ for adiabatic changes. As we shall see later, $\gamma_{\rm ad}$ is related to the *dynamical stability* of stars.

- The *adiabatic temperature gradient* is defined as

$$\nabla_{\rm ad} = \left(\frac{\partial \log T}{\partial \log P}\right)_{\rm ad} \tag{3.57}$$

  It is in fact another exponent that describes the behaviour of the temperature under adiabatic compression or expansion ($T \propto P^{\nabla_{\rm ad}}$ if $\nabla_{\rm ad}$ is constant), which turns out to be important for stability against *convection*.

The adiabatic exponent For an adiabatic process $dq = 0$ in eq. (3.47) and therefore

$$\mathrm{d}u = \frac{P}{\rho^2}\,\mathrm{d}\rho. \tag{3.58}$$

We have seen in Sect. 3.3.1 that for a perfect gas of free particles the internal energy density U is proportional to P, in both the NR and ER limits. For such a simple system we can therefore write, as we did in Sect. 2.3,

$$u = \phi\,\frac{P}{\rho} \tag{3.59}$$

with ϕ a constant (between $\frac{3}{2}$ and 3). If we differentiate this and substitute into eq. (3.58) we obtain for an adiabatic change

$$\frac{\mathrm{d}P}{P} = \frac{\phi+1}{\phi}\frac{\mathrm{d}\rho}{\rho}. \tag{3.60}$$

Therefore, according to the definition of $\gamma_{\rm ad}$ (eq. 3.56),

$$\gamma_{\rm ad} = \frac{\phi+1}{\phi} \qquad \text{(for a simple, perfect gas).} \tag{3.61}$$

[2] In many textbooks one finds instead the adiabatic exponents Γ_1, Γ_2, and Γ_3 introduced by Chandrasekhar. They are defined, and related to $\gamma_{\rm ad}$ and $\nabla_{\rm ad}$, as follows:

$$\Gamma_1 = \left(\frac{\partial \log P}{\partial \log \rho}\right)_{\rm ad} = \gamma_{\rm ad}, \qquad \frac{\Gamma_2}{\Gamma_2 - 1} = \left(\frac{\partial \log P}{\partial \log T}\right)_{\rm ad} = \frac{1}{\nabla_{\rm ad}}, \qquad \Gamma_3 = \left(\frac{\partial \log T}{\partial \log \rho}\right)_{\rm ad} + 1.$$

They obey the relation

$$\frac{\Gamma_1}{\Gamma_3 - 1} = \frac{\Gamma_2}{\Gamma_2 - 1}.$$

- for *non-relativistic* particles (e.g. a classical ideal gas, NR degenerate electrons) $\phi = \frac{3}{2}$ and therefore $\gamma_{\rm ad} = \frac{5}{3}$
- for *extremely relativistic* particles (e.g. photons, ER degenerate electrons) $\phi = 3$ and therefore $\gamma_{\rm ad} = \frac{4}{3}$
- for a mixture of gas and radiation ($0 \leq \beta \leq 1$) and/or moderately relativistic degenerate electrons, $\frac{4}{3} \leq \gamma_{\rm ad} \leq \frac{5}{3}$

For a general equation of state, described by eq. (3.48), one can derive (see Appendix 3.A)

$$\gamma_{\rm ad} = \chi_\rho + \frac{P}{\rho T c_V} {\chi_T}^2. \tag{3.62}$$

Therefore $\gamma_{\rm ad}$ is related to the ratio of specific heats (eq. 3.54), $\gamma_{\rm ad} = \gamma \chi_\rho$. The γ's are equal if $\chi_\rho = 1$ (as in the case of an ideal gas).

The adiabatic temperature gradient By writing eq. (3.56) as $\mathrm{d}P/P = \gamma_{\rm ad}\, \mathrm{d}\rho/\rho$ for an adiabatic change, and eliminating $d\rho$ with the help of eq. (3.48), we obtain a general relation between the adiabatic temperature gradient $\nabla_{\rm ad}$ and the adiabatic exponent $\gamma_{\rm ad}$:

$$\nabla_{\rm ad} = \frac{\gamma_{\rm ad} - \chi_\rho}{\gamma_{\rm ad} \chi_T}, \tag{3.63}$$

This gives the following limiting cases:

- for an ideal gas without radiation ($\beta = 1$) we have $\chi_T = \chi_\rho = 1$, which together with $\gamma_{\rm ad} = \frac{5}{3}$ gives $\nabla_{\rm ad} = \frac{2}{5} = 0.4$.
- for a radiation-dominated gas ($\beta = 0$) $\chi_T = 4$ and $\chi_\rho = 0$ so that $\nabla_{\rm ad} = \frac{1}{4} = 0.25$.

For a general equation of state one has to consider the general expression for $\gamma_{\rm ad}$ (eq. 3.62) in eq. (3.63). From the expression of $\mathrm{d}q$ in terms of $\mathrm{d}T$ and $\mathrm{d}P$ (3.55) it follows that

$$\nabla_{\rm ad} = \frac{P}{\rho T c_P} \frac{\chi_T}{\chi_\rho}. \tag{3.64}$$

This means that for a general *non*-adiabatic process we can write eq. (3.55) as

$$\mathrm{d}q = c_P \left(\mathrm{d}T - \nabla_{\rm ad} \frac{T}{P} \mathrm{d}P \right), \tag{3.65}$$

which will prove to be a useful relation later on.

We give some important results without derivations, which can be found in K&W Chapters 13.2 and 16.3 or in Hansen Chapter 3.7:

- for a mixture of gas and radiation with $0 < \beta < 1$, $\nabla_{\rm ad}$ and $\gamma_{\rm ad}$ both depend on β and take on intermediate values, i.e. $0.25 < \nabla_{\rm ad} < 0.4$.
- for a non-relativistic degenerate gas, we have to consider that although electrons dominate the pressure, there is a (tiny) temperature dependence due to the ion gas which must be taken into account in calculating χ_T and therefore $\nabla_{\rm ad}$. After some manipulation it can be shown that in this case $\nabla_{\rm ad} = 0.4$, as for the ideal classical gas.
- for an extremely relativistic degenerate gas one also has to consider that while the electrons are relativistic, the ions are still non-relativistic. It turns out that in this limit $\nabla_{\rm ad} = 0.5$.

3.5 Ionization

We have so far implicitly assumed complete ionization of the gas, i.e. that it consists of bare atomic nuclei and free electrons. This is a good approximation in hot stellar interiors, where $T > 10^6\,\mathrm{K}$ so that typical energies kT are much larger than the energy needed to ionize an atom, i.e. to knock off a bound electron. In the cooler outer layers of a star, however, we need to consider the *partial ionization* of the elements. In this case quasi-static changes of the state variables (ρ and T) will lead to changes in the degree of ionization. This can have a large effect on the thermodynamic properties of the gas, e.g. on γ_{ad} and ∇_{ad}.

In LTE the number densities of ionized and neutral species are determined by the *Saha equation*

$$\frac{n_{r+1}}{n_r} n_{\mathrm{e}} = \frac{u_{r+1}}{u_r} \frac{2(2\pi m_{\mathrm{e}} kT)^{3/2}}{h^3} e^{-\chi_r/kT} \tag{3.66}$$

where n_r and n_{r+1} indicate the number densities of r and $r+1$ times ionized nuclei, χ_r is the ionization potential, i.e. the energy required to remove the r-th bound electron, and u_r and u_{r+1} are the partition functions. The partition functions depend on T but can in most cases be approximated by the statistical weights of the ground states of the bound species. (This equation can be derived from statistical mechanics, e.g. see K&W Chapter 14.1.)

3.5.1 Ionization of hydrogen

As an example, we consider the simple case where the gas consists only of hydrogen. Then there are just three types of particle, electrons and neutral and ionized hydrogen, with $u_{\mathrm{H}} = u_0 = 2$ and $u_{\mathrm{H}+} = u_1 = 1$. We write their number densities as n_+ and n_0 so that

$$\frac{n_+}{n_0} n_{\mathrm{e}} = \frac{(2\pi m_{\mathrm{e}} kT)^{3/2}}{h^3} e^{-\chi_{\mathrm{H}}/kT} \tag{3.67}$$

where $\chi_{\mathrm{H}} = 13.6\,\mathrm{eV}$. The gas pressure is given by $P_{\mathrm{gas}} = (n_0 + n_+ + n_{\mathrm{e}})\,kT$ and the density is $\rho = (n_0 + n_+)\,m_u$. The *degree of ionization* is defined as

$$x = \frac{n_+}{n_0 + n_+} \tag{3.68}$$

so that P_{gas} can be written in terms of the degree of ionization

$$P_{\mathrm{gas}} = (1 + x)\,\mathcal{R}\rho T \tag{3.69}$$

We can then rewrite Saha's equation as

$$\frac{x^2}{1 - x^2} = \frac{(2\pi m_{\mathrm{e}})^{3/2}}{h^3} \frac{(kT)^{5/2}}{P_{\mathrm{gas}}} e^{-\chi_{\mathrm{H}}/kT} \tag{3.70}$$

We see that the degree of ionization increases with T, as expected since more atoms are broken up by the energetic photons. However, x decreases with gas pressure (or density) when T is kept constant, because this increases the probability of recombination which is proportional to n_{e}. From eq. (3.69) we see that the mean molecular weight $\mu = 1/(1 + x)$ decreases as hydrogen becomes ionized (one atomic mass is divided over two particles).

To estimate the effect on the thermodynamic properties of the gas, we note that in the case of partial ionization the internal energy has a contribution from the available potential energy of recombination. Per unit volume this contribution is equal to $n_+\chi_{\mathrm{H}}$, so per unit mass it equals $n_+\chi_{\mathrm{H}}/\rho = x\chi_{\mathrm{H}}/m_{\mathrm{u}}$. Thus

$$u = \frac{3}{2}\frac{P_{\mathrm{gas}}}{\rho} + x\frac{\chi_{\mathrm{H}}}{m_{\mathrm{u}}} = \tfrac{3}{2}(1 + x)\,\mathcal{R}T + x\frac{\chi_{\mathrm{H}}}{m_{\mathrm{u}}}. \tag{3.71}$$

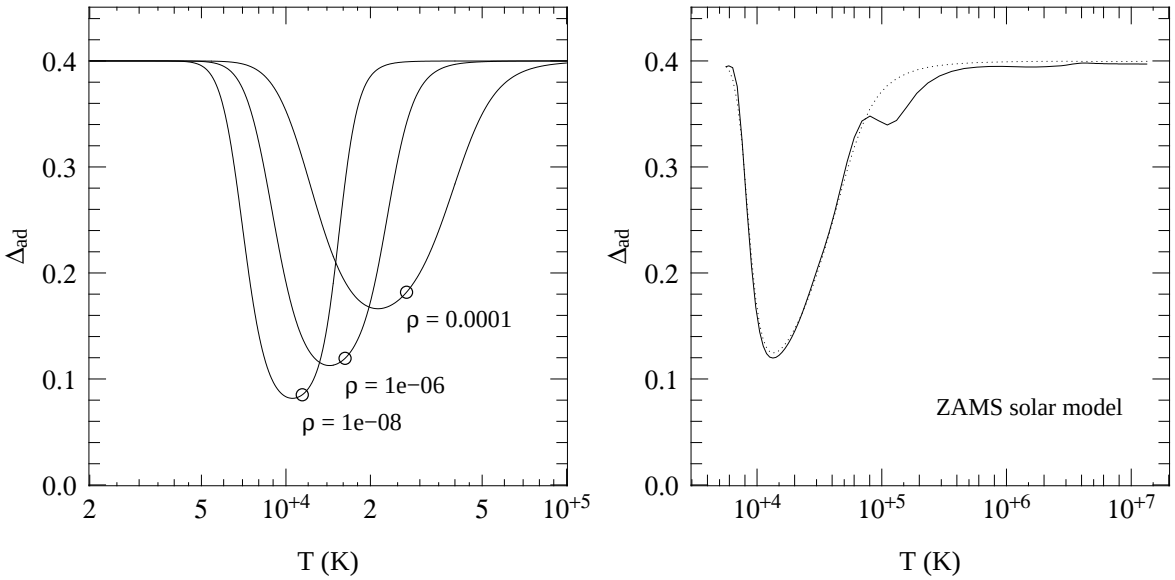


Figure 3.5. The adiabatic temperature gradient ∇_{ad} plotted against temperature. The left panel shows the effect of partial ionization for the simple case of a pure hydrogen gas, for three values of the density (10^{-4}, 10^{-6} and 10^{-8} g/cm^3). When hydrogen is partially ionized, ∇_{ad} is decreased below its ideal-gas value of 0.4. The circles indicate the points where the degree of ionization $x = 0.5$, close to the minimum of ∇_{ad}. As the density increases, a higher temperature is needed to reach the same ionization degree. The right panel shows how ∇_{ad} varies with temperature in a detailed stellar model of 1 $M_\odot$, between the surface (at $T \approx 6000$ K) and the centre (at $T \approx 1.5 \times 10^7$ K). Apart from the hydrogen ionization zone around 10^4 K, a second depression of ∇_{ad} around 10^5 K is seen which is due to the first ^{4}He ionization zone. The second He ionization zone is merged with H ionization because it occurs at similar temperatures and densities. Note that the region where $T < 10^6$ K comprises only the outer 1 % of the mass of the Sun. (The dotted line shows how ∇_{ad} would vary with T in this model if the composition were pure hydrogen, as was assumed in the left panel.)

A small increase in temperature increases the degree of ionization, which results in a large amount of energy being absorbed by the gas. In other words, the *specific heat* of a partially ionized gas will be much larger than for an unionized gas, or for a completely ionized gas (in the latter case $x = 1$ so that the second term in eq. (3.71) becomes a constant and therefore irrelevant).

Now consider what happens if the gas is adiabatically compressed. Starting from neutral hydrogen, for which $\nabla_{ad} = 0.4$, the temperature initially increases as $T \propto P^{0.4}$. Further compression (work done on the gas) increases u, but when partial ionization sets in most of this energy goes into raising the degree of ionization (second term of eq. 3.71) and only little into raising the temperature (first term). In other words, T increases less strongly with with P, and therefore $\nabla_{ad} < 0.4$. A detailed calculation (e.g. see K&W Chapter 14.3) shows that under typical conditions ∇_{ad} reaches a minimum value of ≈ 0.1 when $x \approx 0.5$. As the gas becomes almost fully ionized, ∇_{ad} rises back to 0.4. The variation of ∇_{ad} with temperature for a pure hydrogen gas is shown in the left panel of Fig. 3.5 for different values of the density.

The decrease of ∇_{ad} in partial ionization zones can induce *convection* in the outer layers of stars, as we shall see in Ch. 5. Similarly it can be shown that γ_{ad} decreases in partial ionization zones, from $\frac{5}{3}$ to $\gamma_{ad} \approx 1.2$ when $x \approx 0.5$. This has consequences for the stability of stars, as we shall also see.

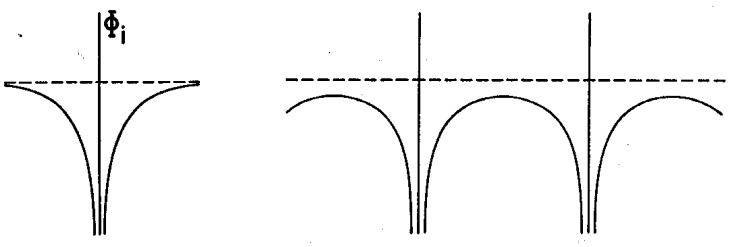


Figure 3.6. Schematic depiction of the electrostatic potential of an isolated ion (left) and the superposition of the potentials of neighbouring ions (right). Figure reproduced from KIPPENHANHN & WEIGERT.

3.5.2 Ionization of a mixture of gases

In a mixture of gases the situation becomes more complicated because many, partly ionized species have to be considered, the densities of which all depend on each other (see e.g. K&W Chapter 14.4-14.5). However the basic physics remains the same as considered above for the simple case of pure hydrogen. The effect on the thermodynamic properties is that e.g. $\nabla_{\rm ad}$ can show additional deviations below 0.4 at different temperatures, especially where helium (the second-most abundant element in stars) is partially ionized. This is illustrated in Fig. 3.5b which shows the variation of $\nabla_{\rm ad}$ with temperature in a homogeneous model for the initial Sun.

3.5.3 Pressure ionization

As ρ increases indefinitely, the Saha equation gives $x \to 0$, i.e. ionized gas recombines to form atoms. This is obviously nonsense at very high density, and becomes incorrect when the average distance d between ions becomes less than an atomic radius. In this situation the ionization energy is suppressed (there are fewer bound excited states; see Fig. 3.6), a situation known as *pressure ionization*.

Consider the case of hydrogen: the volume per H atom is $1/n_{\rm H}$ so that $d = (\frac{4\pi}{3} n_{\rm H})^{-1/3}$. Pressure ionization sets in when $d \lesssim a_0 = 5 \times 10^{-9}$ cm (the Bohr radius). This implies

$$n_{\rm H} \gtrsim \frac{1}{\frac{4\pi}{3} {a_0}^3}$$

or $\rho = n_{\rm H} m_{\rm H} \gtrsim 3\,{\rm g\,cm^{-3}}$. Other elements are pressure-ionized at similar values of the density, within an order of magnitude. At densities $\gtrsim 10\,{\rm g\,cm^{-3}}$, therefore, we can again assume complete ionization.

Fig. 3.7 shows the approximate boundary in the density-temperature diagram between neutral and ionized hydrogen according the Saha equation for $\rho < 1\,{\rm g\,cm^{-3}}$, and as a result of pressure ionization at higher densities.

3.6 Other effects on the equation of state

3.6.1 Coulomb interactions and crystallization

We have so far ignored the effect of electrostatic or Coulomb interactions between the ions and electrons in the gas. Is this a reasonable approximation, i.e. are the interaction energies indeed small compared to the kinetic energies, as we have assumed in Sect. 3.3?

The average distance between gas particles (with mass $Am_{\rm u}$) is $d \approx (\frac{4\pi}{3} n)^{-1/3}$ where n is the number density, $n = \rho/(Am_{\rm u})$. The typical Coulomb energy per particle (with charge Ze) is $\epsilon_{\rm C} \approx Z^2 e^2/d$, while the average kinetic energy is $\epsilon_{\rm kin} = \frac{3}{2} kT$. The ratio of Coulomb energy to kinetic energy is usually called the Coulomb parameter $\Gamma_{\rm C}$, defined as

$$\Gamma_{\rm C} = \frac{Z^2 e^2}{d\,kT} = \frac{Z^2 e^2}{kT} \left(\frac{4\pi\rho}{3Am_{\rm u}} \right)^{1/3} = 2.275 \times 10^5 \, \frac{Z^2}{A^{1/3}} \, \frac{\rho^{1/3}}{T}, \tag{3.72}$$

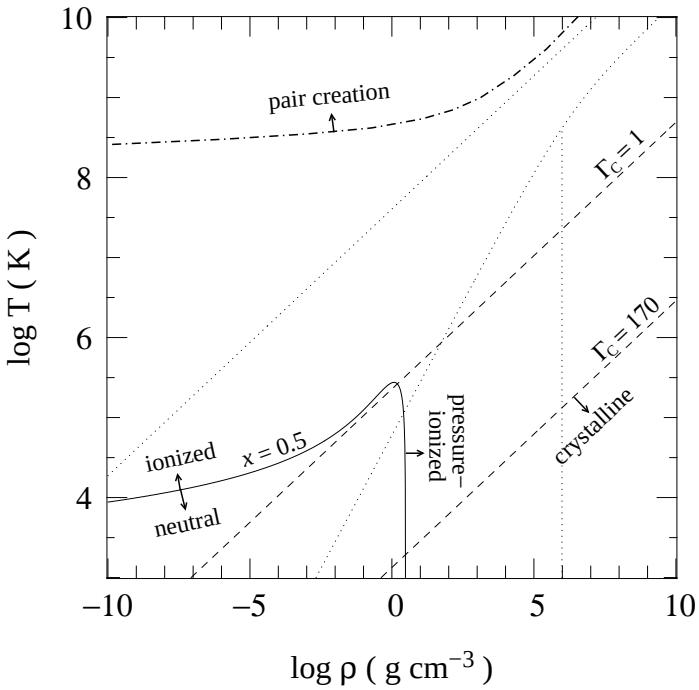


Figure 3.7. The equation of state in the ρ, T plane for a pure hydrogen gas. The dotted lines are the borders, also shown in Fig. 3.4, between regions where radiation, ideal gas and degenerate electrons dominate the pressure. The solid line shows where the ionization fraction of hydrogen is 0.5 according to the Saha equation, and where hydrogen becomes pressure-ionized at high density. The dashed lines show where the Coulomb interaction parameter Γ_C equals 1, above which Coulomb interactions become important, and where $\Gamma_C = 170$, above which the ions form a crystalline lattice. Above the dash-dotted line e^+e^- pairs play an important role in stellar interiors.

where in the last equality the numerical factor is in cgs units. We see that Coulomb interactions increase in importance at high densities or low temperatures. Roughly, Coulomb interactions start to become important in stellar interiors when $\Gamma_C \gtrsim 1$.

To estimate the typical value of Γ_C in stellar interiors we approximate $\rho \approx \bar{\rho} = M/(\frac{4\pi}{3}R^3)$, and we approximate T by the average temperature estimated from the virial theorem, $T \approx \bar{T} \approx \frac{1}{3}\frac{Am_u}{k}\frac{GM}{R}$ (eq. 2.29). Ignoring factors of order unity, we get

$$\Gamma_C \approx 0.01 \frac{Z^2}{A^{4/3}} \left(\frac{M}{M_\odot}\right)^{-2/3} . \tag{3.73}$$

The ratio $Z^2/A^{4/3}$ depends on the composition, and represents an average over the constituents of the gas. In stars mostly composed of hydrogen, $A \approx 1$ and $Z \approx 1$, and we find that in the Sun the Coulomb energy contributes of the order of 1 % to the particle energies (and hence has a similar effect on the pressure). We are therefore justified in ignoring Coulomb interactions in stars similar to or more massive than the Sun. However, eq. (3.73) shows that in low-mass stars Coulomb interactions can start to contribute significantly. This can also be seen by comparing Fig. 3.4 and Fig. 3.7, where the location of the condition $\Gamma_C = 1$ is indicated in the ρ-T diagram. Detailed models of low-mass stars need to take this effect into account. For $M \lesssim 10^{-3}\, M_\odot$ the Coulomb energies dominate. Such objects are not stars but planets (Jupiter's mass is about $10^{-3}\, M_\odot$). Calculations of the structure of planets requires a much more complicated equation of state than for stars.

Crystallization

If $\Gamma_C \gg 1$ the thermal motions of the ions are overwhelmed by the Coulomb interactions. In this situation the ions will tend to settle down into a conglomerate with a lower energy, in other words they will form a crystalline lattice. Detailed estimates indicate that this transition takes place at a critical value of $\Gamma_C \approx 170$. This condition is also indicated in Fig. 3.4 for a pure hydrogen gas. In reality, this situation will never occur in hydrogen-rich stellar interiors, but it can take place in cooling

white dwarfs (in which the temperature gradually decreases with time while the density remains constant). White dwarfs are usually composed of carbon and oxygen, so in this case we have to take into account the composition which raises the temperature at which the transition occurs (the 'melting' temperature) by a factor $Z^2/A^{1/3}$ according to eq. (3.72).

Finally we note that crystallization only occurs in the region where the electrons are strongly degenerate. You may verify that the Coulomb interaction energy between electrons and ions (Ze^2/d) is always smaller than the typical electron energy ($p_F^2/2m_e$). The electrons therefore behave as a free degenerate gas, even if the ions form a crystalline structure.

3.6.2 Pair production

A very different process can take place at very high temperatures and relatively low densities. A photon may turn into an electron-positron pair if its energy $h\nu$ exceeds the rest-mass energy of the pair, $h\nu > 2m_e c^2$. This must take place during the interaction with a nucleus, since otherwise momentum and energy cannot both be conserved. Pair production takes place at a typical temperature $kT \approx h\nu \approx 2m_e c^2$, or $T \approx 1.2 \times 10^{10}$ K. However, even at $T \sim 10^9$ K the number of energetic photons in the tail of the Planck distribution (eq. 3.41) is large enough to produce a large number of e^+e^- pairs. The newly created positrons tend to be annihilated quickly by the inverse reaction ($e^+ + e^- \rightarrow 2\gamma$), as a result of which the number of positrons reaches equilibrium. At a few times 10^9 K, depending on the electron density, the number of positrons is a significant fraction of the number of electrons.

Pair production is similar to an ionization process: an increase in temperature leads to an increase in the number of particles at the expense of the photon energy (and pressure). Therefore pair production gives rise to a decrease of the adiabatic gradient γ_{ad} and of ∇_{ad}, similar to partial ionization. This is the main importance of pair production for stellar evolution: it affects the stability of very massive stars in advanced stages of evolution (when their temperature may reach values in excess of 10^9 K) and can trigger their collapse.

Suggestions for further reading

The contents of this chapter are also covered by Chapter 7 of Maeder and by Chapters 13 to 16 of Kippenhahn & Weigert. However, a more elegant derivation of the equation of state, which is also more consistent with the way it is derived in these lecture notes, is given in Chapter 3 of Hansen, Kawaler & Trimble. Explicit expressions for many of the results that are only mentioned here can be found in this book.

Exercises

3.1 Conceptual questions

These questions are intended to test your understanding of the lectures. Try to answer them without referring to the lecture notes.

(a) What do we mean by *local thermodynamic equilibrium* (LTE)? Why is this a good assumption for stellar interiors? What is the difference between LTE and *thermal equilibrium* (as treated in Ch. 2)?

(b) In what type of stars does degeneracy become important? Is it important in main-sequence stars? Is it more important in high mass or low mass MS stars?

(c) Explain qualitatively why for degenerate matter, the pressure increases with the density.

(d) Why do electrons become relativistic when they are compressed into a smaller volume? Why does the pressure increase less steeply with the density in this case?

(e) In the central region of a star we find free electrons and ions. Why do the electrons become degenerate first? Why do the ions never become degenerate in practice?

3.2 Mean molecular weight

Derive a general expression for the mean molecular weight of an ionized gas, as a function of composition X, Y, Z. Assume that, for elements heavier than H, nuclei are composed of equal number of protons and neutrons, so that the nuclear charge Z_i is half of the mass number A_i.

3.3 The $\rho - T$ plane

Consider a gas of ionized hydrogen. In the $\rho - T$ plane compute the approximate boundary lines between the regions where:

(a) radiation pressure dominates,

(b) the electrons behave like a classical ideal gas,

(c) the electrons behave like a degenerate gas,

(d) the electrons are relativistically degenerate.

3.4 The pressure of a gas of free particles

In this exercise you will derive some important relations from this chapter for yourself.

(a) Suppose that the particles in a gas have momenta distributed as $n(p)\,\mathrm{d}p$. Show that the pressure can be expressed by eq. (3.4).

(b) For classical particles in LTE, the momentum distribution is given by the Maxwell-Boltzmann distribution, eq. (3.13). Calculate the pressure using eq. (3.4). Does the result look familiar?

(c) Show that for a gas of free, non-relativistic particles $P = \frac{2}{3}U$ (eq. 3.11), where U is the internal energy density. Show that in the extremely relativistic limit $P = \frac{1}{3}U$ (eq. 3.12).

(d) Electrons are fermions with 2 spin states. Explain why the maximum number of electrons per volume with momentum p can be written as eq. (3.28).

(e) In the extreme case of complete degeneracy, $T \to 0$, the electrons fill up all available quantum states up to a maximum p_F, the Fermi momentum. Show that

$$p_\mathrm{F} = h\left(\frac{3n_\mathrm{e}}{8\pi}\right)^{\frac{1}{3}}$$

(f) Show that the pressure as function of the density for a non-relativistic degenerate electron gas can be written as

$$P = K_\mathrm{NR}\left(\frac{\rho}{\mu_\mathrm{e}}\right)^{x}$$

and derive an expression for K_NR and x.

(g) Show that the pressure as function of the density for an extremely relativistic degenerate electron gas can be written as

$$P = K_\mathrm{ER}\left(\frac{\rho}{\mu_\mathrm{e}}\right)^{y}$$

and derive an expression for K_ER and y.

(h) Photons are bosons, and the distribution of their momenta is given by the Planck function (eq. 3.27). Show that in this case

$$U \propto T^4$$

(Hint: to derive an expression for the proportionality constant a, you might want to use Mathematica or a list of standard integrals.)

(i) Now use (c) to show that the radiation pressure is given by $P_{\rm rad} = \frac{1}{3}aT^4$.

3.5 Adiabatic derivatives

(a) Use the first law of thermodynamics to show that, for an ideal gas in an adiabatic process,

$$P \propto \rho^{\gamma_{\rm ad}} \tag{3.74}$$

and give a value for the adiabatic exponent $\gamma_{\rm ad}$.

(b) Use the ideal gas law in combination with eq. (3.74) to show that

$$\nabla_{\rm ad} = \left(\frac{{\rm d}\ln T}{{\rm d}\ln P}\right)_{\rm ad,id} = 0.4.$$

(c) The quantity $\nabla_{\rm ad}$ is referred to as the *adiabatic temperature gradient*. Normally you would use the term 'gradient of a quantity A' for ${\rm d}A/{\rm d}r$, or if you use mass coordinates instead of radius coordinates, ${\rm d}X/{\rm d}m$. Do you understand why ∇_{ad} can be referred to as a temperature 'gradient'?

(d) (*) Show that for a mixture of an ideal gas plus radiation, the adiabatic exponent is given by

$$\gamma_{\rm ad} = \frac{32 - 24\beta - 3\beta^2}{24 - 21\beta},$$

where $\beta = P_{\rm gas}/P$.

(Hints: write down the equation of state for the mixture in differential form as in eq. (3.48), and express χ_T and χ_ρ in terms of β. Then apply the first law of therrmodynamics for an adiabatic process.)

(e) (*) What is the value of $\gamma_{\rm ad}$ in the limit where radiation dominates and where pressure dominates? Does this look familiar?

3.6 Ionization effects

(a) The particles in an ionized gas are charged and therefore undergo electrostatic (Coulomb) interactions. Why can can we nevertheless make the ideal-gas assumption in most stars (i.e. that the internal energy of the gas is just the sum of the kinetic energies of the particles)? For which stars do Coulomb interactions have a significant effect?

(b) Why does the gas in the interior of a star become pressure-ionized at high densities?

(c) Explain qualitatively why partial ionization leads to $\nabla_{\rm ad} < \nabla_{\rm ad,ideal} = 0.4$, in other words: why does adiabatic compression lead to a smaller temperature increase when the gas is partly ionized, compared to a completely ionized (or unionized) gas?

3.A Appendix: Themodynamic relations

In this Appendix we derive some of the thermodynamic relations that were given without proof in Chapter 3.

The first law of thermodynamics states that the heat added to a mass element of gas is the sum of the change in its internal energy and the work done by the mass element. Taking the element to be of unit mass, we can wite this as

$$\mathrm{d}q = \mathrm{d}u + P\,\mathrm{d}v = \mathrm{d}u - \frac{P}{\rho^2}\,\mathrm{d}\rho, \tag{3.75}$$

because the volume of a unit mass is $v = 1/\rho$. We can write the change in the internal energy of a unit mass in terms of the changes in the state variables (T and ρ) as

$$\mathrm{d}u = \left(\frac{\partial u}{\partial T}\right)_\rho \mathrm{d}T + \left(\frac{\partial u}{\partial \rho}\right)_T \mathrm{d}\rho. \tag{3.76}$$

The change in the entropy per unit mass, $\mathrm{d}s = \mathrm{d}q/T$, is therefore

$$\mathrm{d}s = \frac{\mathrm{d}q}{T} = \frac{1}{T}\left(\frac{\partial u}{\partial T}\right)_\rho \mathrm{d}T + \frac{1}{T}\left[\left(\frac{\partial u}{\partial \rho}\right)_T - \frac{P}{\rho^2}\right]\mathrm{d}\rho. \tag{3.77}$$

Because s is a function of state, $\partial^2 s/\partial\rho\partial T = \partial^2 s/\partial T\partial\rho$, which means that

$$\frac{1}{T}\,\frac{\partial^2 u}{\partial T\partial\rho} = \frac{\partial}{\partial T}\left[\frac{1}{T}\left(\frac{\partial u}{\partial \rho}\right)_T - \frac{P}{\rho^2 T}\right], \tag{3.78}$$

where the $\partial/\partial T$ on the right-hand side should be taken at constant ρ. Working out the right-hand side allows us to eliminate the second derivative of u, giving

$$\frac{1}{T^2}\left(\frac{\partial u}{\partial \rho}\right)_T = \frac{P}{\rho^2 T^2} - \frac{1}{\rho^2 T}\left(\frac{\partial P}{\partial T}\right)_\rho.$$

With the definition of χ_T (eq. 3.49) we can write $(\partial P/\partial T)_\rho = \chi_T\, P/T$, and thus

$$\left(\frac{\partial u}{\partial \rho}\right)_T = (1-\chi_T)\frac{P}{\rho^2}. \tag{3.79}$$

Specific heats

The definitions of the specfic heats at constant volume and at constant pressure are

$$c_V \equiv \left(\frac{\mathrm{d}q}{\mathrm{d}T}\right)_v = \left(\frac{\partial u}{\partial T}\right)_\rho, \tag{3.80}$$

$$c_P \equiv \left(\frac{\mathrm{d}q}{\mathrm{d}T}\right)_P = \left(\frac{\partial u}{\partial T}\right)_P - \frac{P}{\rho^2}\left(\frac{\partial \rho}{\partial T}\right)_P. \tag{3.81}$$

To work out an expression for c_P, we need $(\partial u/\partial T)_P$ and $(\partial\rho/\partial T)_P$. To start with the latter, we use the differential form of the equation of state (3.48). At constant pressure $\mathrm{d}P = 0$ this gives

$$\chi_\rho\frac{\mathrm{d}\rho}{\rho} = -\chi_T\frac{\mathrm{d}T}{T} \qquad \Rightarrow \qquad \left(\frac{\partial\rho}{\partial T}\right)_P = -\frac{\rho}{T}\,\frac{\chi_T}{\chi_\rho}. \tag{3.82}$$

To obtain an expression for $(\partial u/\partial T)_P$ we use eq. (3.76), which we can write as

$$\frac{\mathrm{d}u}{\mathrm{d}T} = \left(\frac{\partial u}{\partial T}\right)_\rho + \left(\frac{\partial u}{\partial \rho}\right)_T \frac{\mathrm{d}\rho}{\mathrm{d}T}$$

and therefore

$$\left(\frac{\partial u}{\partial T}\right)_P = \left(\frac{\partial u}{\partial T}\right)_\rho + \left(\frac{\partial u}{\partial \rho}\right)_T \left(\frac{\partial \rho}{\partial T}\right)_P = \left(\frac{\partial u}{\partial T}\right)_\rho + (\chi_T - 1)\frac{\chi_T}{\chi_\rho}\frac{P}{\rho T}\,. \tag{3.83}$$

To obtain the last equality we used eqs. (3.79) and (3.82). From the definitions (3.80) and (3.81) we thus arrive at the following relation between c_P and c_V:

$$\boxed{c_P - c_V = \frac{{\chi_T}^2}{\chi_\rho}\frac{P}{\rho T}} \tag{3.84}$$

which is eq. (3.53).

Expressions for dq

It is useful to be able to write the change in heat content of a unit mass in terms of the changes in the state variables. Eq. (3.77) already shows how dq is written in terms of T and ρ, i.e.

$$\mathrm{d}q = T\,\mathrm{d}s = c_V\,\mathrm{d}T - \chi_T\frac{P}{\rho^2}\,\mathrm{d}\rho, \tag{3.85}$$

making use of (3.79) and (3.80). It is often useful to express dq is terms of T and P, rather than ρ. To do this we write dρ with the help of eq. (3.48),

$$\mathrm{d}\rho = \frac{\rho}{\chi_\rho}\left(\frac{\mathrm{d}P}{P} - \chi_T\frac{\mathrm{d}T}{T}\right), \tag{3.86}$$

so that

$$\mathrm{d}q = c_V\,\mathrm{d}T - \frac{\chi_T}{\chi_\rho}\frac{P}{\rho}\left(\frac{\mathrm{d}P}{P} - \chi_T\frac{\mathrm{d}T}{T}\right) = \left(c_V + \frac{{\chi_T}^2}{\chi_\rho}\frac{P}{\rho T}\right)\mathrm{d}T - \frac{\chi_T}{\chi_\rho}\frac{1}{\rho}\,\mathrm{d}P\,. \tag{3.87}$$

The terms with parentheses in the last equality are simple c_P, according to (3.84), and therefore

$$\mathrm{d}q = T\,\mathrm{d}s = c_P\,\mathrm{d}T - \frac{\chi_T}{\chi_\rho}\frac{\mathrm{d}P}{\rho}. \tag{3.88}$$

Adiabatic derivatives

Eq. (3.88) makes it easy to derive an expression for the adiabatic temperature gradient (3.57),

$$\nabla_{\mathrm{ad}} \equiv \left(\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right)_{\mathrm{ad}}. \tag{3.89}$$

An adiabatic change in T and P means the changes take place at constant s, or with dq = 0. Hence (3.88) shows that

$$\left(\frac{\mathrm{d}T}{\mathrm{d}P}\right)_s = \frac{1}{\rho c_P}\frac{\chi_T}{\chi_\rho} \qquad \Rightarrow \qquad \left(\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right)_s = \frac{P}{T}\left(\frac{\mathrm{d}T}{\mathrm{d}P}\right)_s = \frac{P}{\rho T c_P}\frac{\chi_T}{\chi_\rho}. \tag{3.90}$$

This means

$$\boxed{\nabla_{\mathrm{ad}} = \frac{P}{\rho T c_P}\frac{\chi_T}{\chi_\rho}}\,, \tag{3.91}$$

which is eq. (3.64). With the help of this expression we can also write (3.88) as

$$\mathrm{d}q = c_P\left(\mathrm{d}T - \nabla_{\mathrm{ad}}\frac{T}{P}\,\mathrm{d}P\right). \tag{3.92}$$

To derive an expression for the adiabatic exponent (3.56),

$$\gamma_{\rm ad} \equiv \left(\frac{{\rm d}\ln P}{{\rm d}\ln\rho}\right)_{\rm ad}, \tag{3.93}$$

we use (3.85) and (3.88) and set ${\rm d}q = 0$ in both expressions. This gives

$${\rm d}T = \frac{P}{\rho^2 c_V}\chi_T\,{\rm d}\rho \qquad \text{and} \qquad {\rm d}T = \frac{1}{\rho c_P}\frac{\chi_T}{\chi_\rho}\,{\rm d}P.$$

Eliminating ${\rm d}T$ from both expressions gives

$$\frac{{\rm d}P}{P} = \frac{c_P}{c_V}\chi_\rho\frac{{\rm d}\rho}{\rho} \qquad \Rightarrow \qquad \left(\frac{{\rm d}\ln P}{{\rm d}\ln\rho}\right)_s = \frac{c_P}{c_V}\chi_\rho.$$

This means

$$\boxed{\gamma_{\rm ad} = \frac{c_P}{c_V}\chi_\rho = \gamma\chi_\rho}, \tag{3.94}$$

where $\gamma = c_P/c_V$ is the ratio of specific heats. Using eq. (3.84) this can also be written as

$$\gamma_{\rm ad} = \chi_\rho + \frac{P}{\rho T c_V}{\chi_T}^2, \tag{3.95}$$

which is eq. (3.62).

Chapter 4

Polytropic stellar models

As mentioned in Sec. 2.2, the equation of hydrostatic equilibrium can be solved if the pressure is a known function of the density, $P = P(\rho)$. In this situation the mechanical structure of the star is completely determined. A special case of such a relation between P and ρ is the *polytropic relation*,

$$P = K\rho^{\gamma} \tag{4.1}$$

where K and γ are both constants. The resulting stellar models are known as *polytropic stellar models* or simply polytropes. Polytropic models have played an important role in the historical development of stellar structure theory. Although nowadays their practical use has mostly been superseded by more realistic stellar models, due to their simplicity polytropic models still give useful insight into several important properties of stars. Moreover, in some cases the polytropic relation is a good approximation to the real equation of state. We have encountered a few examples of polytropic equations of state in Chapter 3, e.g. the pressure of degenerate electrons, and the case where pressure and density are related adiabatically.

In this brief chapter – and the accompanying computer practicum – we will derive the analytic theory of polytropes and construct polytropic models, and study to which kind of stars they correspond, at least approximately.

4.1 Polytropes and the Lane-Emden equation

If the equation of state can be written in polytropic form, the equations for mass continuity ($\mathrm{d}m/\mathrm{d}r$, eq. 2.3) and for hydrostatic equilibrium ($\mathrm{d}P/\mathrm{d}r$, eq. 2.12) can be combined with eq. (4.1) to give a second-order differential equation for the density:

$$\frac{1}{\rho r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\rho^{\gamma-2}\frac{\mathrm{d}\rho}{\mathrm{d}r}\right) = -\frac{4\pi G}{K\gamma} \tag{4.2}$$

The exponent γ is often replaced by the so-called polytropic index n, which is defined by

$$n = \frac{1}{\gamma - 1} \qquad \text{or} \qquad \gamma = 1 + \frac{1}{n} \tag{4.3}$$

In order to construct a polytropic stellar model we have to solve eq. (4.2), together with two boundary conditions which are set in the centre, $r = 0$:

$$\rho(0) = \rho_c \qquad \text{and} \qquad \left(\frac{\mathrm{d}\rho}{\mathrm{d}r}\right)_{r=0} = 0, \tag{4.4}$$

where ρ_c is a parameter to be chosen, or determined from other constraints.

Table 4.1. Numerical values for polytropic models with index n.

n	z_n	Θ_n	$\rho_c/\overline{\rho}$	N_n	W_n
0	2.44949	4.89898	1.00000	...	0.119366
1	3.14159	3.14159	3.28987	0.63662	0.392699
1.5	3.65375	2.71406	5.99071	0.42422	0.770140
2	4.35287	2.41105	11.40254	0.36475	1.638183
3	6.89685	2.01824	54.1825	0.36394	11.05068
4	14.97155	1.79723	622.408	0.47720	247.559
4.5	31.8365	1.73780	6189.47	0.65798	4921.84
5	∞	1.73205	∞	∞	∞

In order to simplify eq. (4.2), we define two new dimensionless variables w (related to the density) and z (related to the radius) by writing

$$\rho = \rho_c w^n, \tag{4.5}$$

$$r = \alpha z, \qquad \text{with} \quad \alpha = \left(\frac{n+1}{4\pi G} K \rho_c^{1/n-1} \right)^{1/2}. \tag{4.6}$$

This choice of α ensures that the constants K and $4\pi G$ are eliminated after substituting r and ρ into eq. (4.2). The resulting second-order differential equation is called the *Lane-Emden equation*:

$$\frac{1}{z^2} \frac{\mathrm{d}}{\mathrm{d}z} \left(z^2 \frac{\mathrm{d}w}{\mathrm{d}z} \right) + w^n = 0. \tag{4.7}$$

A polytropic stellar model can be constructed by integrating this equation outwards from the centre. The boundary conditions (4.4) imply that in the centre ($z = 0$) we have $w = 1$ and $\mathrm{d}w/\mathrm{d}z = 0$. For $n < 5$ the solution $w(z)$ is found to decrease monotonically and to reach zero at finite $z = z_n$, which corresponds to the surface of the model.

No general analytical solution of the Lane-Emden equation exists. The only exceptions are $n = 0$, 1 and 5, for which the solutions are:

$$n = 0: \qquad w(z) = 1 - \frac{z^2}{6} \qquad z_0 = \sqrt{6}, \tag{4.8}$$

$$n = 1: \qquad w(z) = \frac{\sin z}{z} \qquad z_1 = \pi, \tag{4.9}$$

$$n = 5: \qquad w(z) = \left(1 + \frac{z^2}{3} \right)^{-1/2} \qquad z_5 = \infty. \tag{4.10}$$

The case $n = 0$ ($\gamma = \infty$) corresponds to a homogeneous gas sphere with constant density ρ_c, following eq. (4.5). The solution for $n = 5$ is peculiar in that it has infinite radius; this is the case for all $n \geq 5$, while for $n < 5$ z_n grows monotonically with n. For values of n other than 0, 1 or 5 the solution must be found by numerical integration (this is quite straightforward, see the accompanying computer practicum). Table 4.1 lists the value of z_n for different values of n, as well as several other properties of the solution that will be discussed below.

4.1.1 Physical properties of the solutions

Once the solution $w(z)$ of the Lane-Emden equation is found, eq. (4.5) fixes the relative density distribution of the model, which is thus uniquely determined by the polytropic index n. Given the solution for a certain n, the physical properties of a polytropic stellar model, such as its mass and radius, are then determined by the parameters K and ρ_c, as follows.

The radius of a polytropic model follows from eq. (4.6):

$$R = \alpha z_n = \left[\frac{(n+1)K}{4\pi G}\right]^{1/2} \rho_c^{(1-n)/2n}\, z_n. \tag{4.11}$$

The mass $m(z)$ interior to z can be obtained from integrating eq. (2.3), using eqs. (4.5), (4.6) and (4.7):

$$m(z) = \int_0^{\alpha z} 4\pi r^2 \rho\, dr = -4\pi\alpha^3 \rho_c\, z^2 \frac{\mathrm{d}w}{\mathrm{d}z}. \tag{4.12}$$

Hence the total mass of a polytropic model is

$$M = 4\pi\alpha^3 \rho_c\, \Theta_n = 4\pi \left[\frac{(n+1)K}{4\pi G}\right]^{3/2} \rho_c^{(3-n)/2n}\, \Theta_n, \tag{4.13}$$

where we have defined Θ_n as

$$\Theta_n \equiv \left(- z^2 \frac{\mathrm{d}w}{\mathrm{d}z}\right)_{z=z_n}. \tag{4.14}$$

By eliminating ρ_c from eqs. (4.11) and (4.13) we can find a relation between M, R and K,

$$K = N_n\, G M^{(n-1)/n} R^{(3-n)/n} \qquad \text{with} \qquad N_n = \frac{(4\pi)^{1/n}}{n+1}\, \Theta_n^{(1-n)/n}\, z_n^{(n-3)/n}. \tag{4.15}$$

Numerical values of Θ_n and N_n are given in Table 4.1. From the expressions above we see that $n = 1$ and $n = 3$ are special cases. For $n = 1$ the radius is independent of the mass, and is uniquely determined by the value of K. Conversely, for $n = 3$ the mass is independent of the radius and is uniquely determined by K. For a given K there is only one value of M for which hydrostatic equilibrium can be satisfied if $n = 3$.

The average density $\bar{\rho} = M/(\frac{4}{3}\pi R^3)$ of a polytropic star is related to the central density by eqs. (4.11) and (4.13) as

$$\bar{\rho} = \left(- \frac{3}{z}\frac{\mathrm{d}w}{\mathrm{d}z}\right)_{z=z_n} \rho_c = \frac{3\Theta_n}{z_n^3}\, \rho_c \tag{4.16}$$

Hence the ratio $\rho_c/\bar{\rho}$, i.e. the degree of central concentration of a polytrope, only depends on the polytropic index n. This dependence is also tabulated in Table 4.1. One may invert this relation to find the central density of a polytropic star of a given mass and radius.

The central pressure of a polytropic star follows from eq. (4.1), which can be written as

$$P_c = K \rho_c^{(n+1)/n}.$$

In combination with (4.15) and (4.16) this gives

$$P_c = W_n \frac{GM^2}{R^4} \qquad \text{with} \qquad W_n = \frac{z_n^4}{4\pi(n+1)\Theta_n^2}. \tag{4.17}$$

Note that in our simple scaling estimate, eq. (2.14), we found the same proportionality $P_c \propto GM^2/R^4$, where the proportionality constant W_n is now determined by the polytropic index n (see Table 4.1). We can eliminate R in favour of ρ_c to obtain the very useful relation

$$P_c = C_n\, GM^{2/3}\rho_c^{4/3} \qquad \text{with} \qquad C_n = \frac{(4\pi)^{1/3}}{n+1}\Theta_n^{-2/3}, \tag{4.18}$$

where you may verify that the constant C_n is only weakly dependent on n, unlike W_n in (4.17).

We give without derivation an expression for the gravitational potential energy of a polytrope of index n:

$$E_{\rm gr} = -\frac{3}{5-n}\,\frac{GM^2}{R}. \tag{4.19}$$

(The derivation can be found in K&W Sec. 19.9 and Maeder Sec. 24.5.1.)

4.2 Application to stars

Eq. (4.15) expresses a relation between the constant K in eq. (4.1) and the mass and radius of a polytropic model. This relation can be interpreted in two very different ways:

- The constant K may be given in terms of physical constants. This is the case, for example, for a star dominated by the pressure of degenerate electrons, in either the non-relativistic limit or the extremely relativistic limit. In that case eq. (4.15) defines a unique relation between the mass and radius of a star.

- In other cases the constant K merely expresses proportionality in eq. (4.1), i.e. K is a free parameter that is constant in a particular star, but may vary from star to star. In this case there are many different possible values of M and R. For a star with a given mass and radius, the corresponding value of K for this star can be determined from eq. (4.15).

In this section we briefly discuss examples for each of these two interpretations.

4.2.1 White dwarfs and the Chandrasekhar mass

Stars that are so compact and dense that their interior pressure is dominated by degenerate electrons are known observationally as *white dwarfs*. They are the remnants of stellar cores in which hydrogen has been completely converted into helium and, in most cases, also helium has been fused into carbon and oxygen. Since the pressure of a completely degenerate electron gas is a function of density only (Sec. 3.3.5), the mechanical structure if a white dwarf is fixed and is independent of temperature. We can thus understand some of the structural properties of white dwarfs by means of polytropic models.

We start by considering the equation of state for a degenerate, non-relativistic electron gas. From eq. (3.35) this can be described by a polytropic relation with $n = 1.5$. Since the corresponding K is determined by physical constants, eq. (4.15) shows that such a polytrope follows a mass-radius relation of the from

$$R \propto M^{-1/3}. \tag{4.20}$$

More massive white dwarfs are thus more compact, and therefore have a higher density. Above a certain density the electrons will become relativistic as they are pushed up to higher momenta by the Pauli exclusion principle. The degree of relativity increases with density, and therefore with the mass of the white dwarf, until at a certain mass all the electrons become extremely relativistic, i.e., their

speed $v_e \to c$. In this limit the equation of state has changed from eq. (3.35) to eq. (3.37), which is also a polytropic relation but with $n = 3$. We have already seen above that an $n = 3$ polytrope is special in the sense that it has a unique mass, which is determined by K and is independent of the radius:

$$M = 4\pi\,\Theta_3 \left(\frac{K}{\pi G}\right)^{3/2}. \tag{4.21}$$

This value corresponds to an upper limit to the mass of a gas sphere in hydrostatic equilibrium that can be supported by degenerate electrons, and thus to the maximum possible mass for a white dwarf. Its existence was first found by Chandrasekhar in 1931, after whom this limiting mass was named. Substituting the proper numerical values into eq. (4.21), with K corresponding to eq. (3.37), we obtain the *Chandrasekhar mass*

$$M_{\rm Ch} = 5.836\,\mu_e^{-2}\,M_\odot. \tag{4.22}$$

White dwarfs are typically formed of helium, carbon or oxygen, for which $\mu_e = 2$ and therefore $M_{\rm Ch} = 1.46\,M_\odot$. Indeed no white dwarf with a mass exceeding this limit is known to exist.

4.2.2 Eddington's standard model

As an example of a situation where K is not fixed by physical constants but is essentially a free parameter, we consider a star in which the pressure is given by a mixture of ideal gas pressure and radiation pressure, eq. (3.45). In particular we make the assumption that the ratio β of gas pressure to total pressure is constant, i.e. has the same value in each layer of the star. Since $P_{\rm gas} = \beta P$ we can write

$$P = \frac{1}{\beta}\frac{\mathcal{R}}{\mu}\rho T, \tag{4.23}$$

while also

$$1 - \beta = \frac{P_{\rm rad}}{P} = \frac{aT^4}{3P}. \tag{4.24}$$

Thus the assumption of constant β means that $T^4 \propto P$ throughout the star. If we substitute the complete expression for T^4 into eq. (4.24) we obtain

$$P = \left(\frac{3\mathcal{R}^4}{a\mu^4}\,\frac{1-\beta}{\beta^4}\right)^{1/3} \rho^{4/3}, \tag{4.25}$$

which is a polytropic relation with $n = 3$ for constant β. Since we are free to choose β between 0 and 1, the constant K is indeed a free parameter dependent on β.

The relation (4.25) was derived by Arthur Eddington in the 1920s for his famous 'standard model'. He found that in regions with a high opacity κ (see Ch. 5) the ratio of local luminosity to mass coordinate l/m is usually small, and vice versa. Making the assumption that $\kappa l/m$ is constant throughout the star is equivalent to assuming that β is constant (again, see Ch. 5). Indeed, for stars in which radiation is the main energy transport mechanism this turns out to be approximately true, even though it is a very rough approximation to the real situation. Nevertheless, the structure of stars on the main sequence with $M \gtrsim M_\odot$ is reasonably well approximated by that of a $n = 3$ polytrope. Since the mass of a $n = 3$ polytrope is given by eq. (4.21), we see from eq. (4.25) that there is a unique relation between the mass M of a star and β. The relative contribution of radiation pressure increases with the mass of a star. This was also noted by Eddington, who pointed out that the limited range of known stellar masses corresponds to values of β that are significantly different from 0 or 1.

Suggestions for further reading

Polytropic stellar models are briefly covered in Chapter 24.5 of Maeder and treated more extensively in Chapter 19 of Kippenhahn & Weigert and Chapter 7.2 of Hansen.

Exercises

4.1 The Lane-Emden equation

(a) Derive eq. (4.2) from the stellar structure equations for mass continuity and hydrostatic equilibrium. (Hint: multiply the hydrostatic equation by r^2/ρ and take the derivative with respect to r).

(b) What determines the second boundary condition of eq. (4.4), i.e., why does the density gradient have to vanish at the center?

(c) By making the substitutions (4.3), (4.5) and (4.6), derive the Lane-Emden equation (4.7).

(d) Solve the Lane-Emden equation analytically for the cases $n = 0$ and $n = 1$.

4.2 Polytropic models

(a) Derive K and γ for the equation of state of an ideal gas at a fixed temperature T, of a non-relativistic degenerate gas and of a relativistic degenerate gas.

(b) Using the Lane-Emden equation, show that the mass distribution in a polytropic star is given by eq. (4.12), and show that this yields eq. (4.13) for the total mass of a polytrope.

(c) Derive the expressions for the central density ρ_c and the central pressure P_c as function of mass and radius, eqs. (4.16) and (4.17).

(d) Derive eq. (4.18) and compute the constant C_n for several values of n.

4.3 White dwarfs

To understand some of the properties of white dwarfs (WDs) we start by considering the equation of state for a degenerate, non-relativistic electron gas.

(a) What is the value of K for such a star? Remember to consider an appropriate value of the mean molecular weight per free electron μ_e.

(b) Derive how the central density ρ_c depends on the mass of a non-relativistic WD. Using this with the result of Exercise 4.2(b), derive a radius-mass relation $R = R(M)$. Interpret this physically.

(c) Use the result of (b) to estimate for which WD masses the relativistic effects would become important.

(d) Show that the derivation of a $R = R(M)$ relation for the extreme relativistic case leads to a unique mass, the so-called *Chandrasekhar mass*. Calculate its value, i.e. derive eq. (4.22).

4.4 Eddington's standard model

(a) Show that for constant β the virial theorem leads to

$$E_{\rm tot} = \frac{\beta}{2} E_{\rm gr} = -\frac{\beta}{2-\beta} E_{\rm int}, \tag{4.26}$$

for a classical, non-relativistic gas. What happens in the limits $\beta \to 1$ and $\beta \to 0$?

(b) Verify eq. (4.25), and show that the corresponding constant K depends on β and the mean molecular weight μ as

$$K = \frac{2.67 \times 10^{15}}{\mu^{4/3}} \left(\frac{1-\beta}{\beta^4} \right)^{1/3} . \tag{4.27}$$

(c) Use the results from above and the fact that the mass of an $n = 3$ polytrope is uniquely determined by K, to derive the relation $M = M(\beta, \mu)$. This is useful for numerically solving the amount of radiation pressure for a star with a given mass.

Chapter 5

Energy transport in stellar interiors

The energy that a star radiates from its surface is generally replenished from sources or reservoirs located in its hot central region. This represents an outward energy flux at every layer in the star, and it requires an effective means of transporting energy through the stellar material. This transfer of energy is possible owing to a non-zero temperature gradient in the star. While radiation is often the most important means of energy transport, and its is always present, it is not the only means. In stellar interiors, where matter and radiation are always in local thermodynamic equilibrium (Chapter 3) and the mean free paths of both photons and gas particles are extremely small, energy (heat) can be transported from hot to cool regions in two basic ways:

- Random thermal motions of the particles – either photons or gas particles – by a process that can be called *heat diffusion*. In the case of photons, the process is known as *radiative diffusion*. In the case of gas particles (atoms, ions, electrons) it is usually called *heat conduction*.
- Collective (bulk) motions of the gas particles, which is known as *convection*. This is an important process in stellar interiors, not only because it can transport energy very efficiently, it also results in rapid mixing. Unfortunately, convection is one of the least understood ingredients of stellar physics.

The transport of energy in stars is the subject of this chapter, which will lead us to two additional differential equations for the stellar structure.

5.1 Local energy conservation

In Chapter 2 we considered the global energy budget of a star, regulated by the virial theorem. We have still to take into account the conservation of energy on a local scale in the stellar interior. To do this we turn to the first law of thermodynamics (Sect. 3.4), which states that the internal energy of a system can be changed by two forms of energy transfer: heat and work. By δf we denote a change in a quantity f occurring in a small time interval δt. For a gas element of unit mass the first law can be written as (see eq. 3.47)

$$\delta u = \delta q + \frac{P}{\rho^2}\,\delta\rho. \tag{5.1}$$

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

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

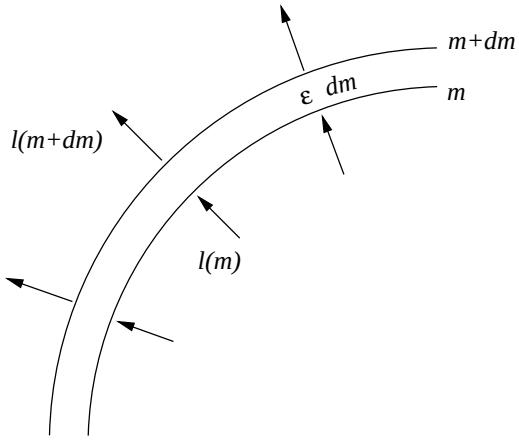


Figure 5.1. Energy generation and heat flow into and out of a spherical mass shell.

- Heat is added by the release of nuclear energy, if available. The rate at which nuclear energy is produced per unit mass and per second is written as $\epsilon_{\rm nuc}$. The details of nuclear energy generation will be treated in Chapter 6.

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

- Finally, heat is absorbed or emitted according to the balance of heat fluxes flowing into and out of the shell. We define a new variable, the *local luminosity* l, as the rate at which energy in the form of heat flows outward through a sphere of radius r (see Fig. 5.1). In spherical symmetry, l is related to the radial energy flux F (in erg s^{-1} cm^{-2}) as

$$l = 4\pi r^2 F. \tag{5.2}$$

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

We can therefore write:

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

with $l(m+\Delta m) = l(m) + (\partial l/\partial m)\cdot\Delta m$, so that after dividing by Δm,

$$\delta q = \left(\epsilon_{\rm nuc} - \epsilon_\nu - \frac{\partial l}{\partial m}\right)\delta t. \tag{5.3}$$

Combining eqs. (5.3) and (5.1) and taking the limit $\delta t \to 0$ yields:

$$\boxed{\frac{\partial l}{\partial m} = \epsilon_{\rm nuc} - \epsilon_\nu - \frac{\partial u}{\partial t} + \frac{P}{\rho^2}\frac{\partial \rho}{\partial t}} \tag{5.4}$$

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

$$\begin{aligned} \epsilon_{\rm gr} &= -\frac{\partial u}{\partial t} + \frac{P}{\rho^2}\frac{\partial \rho}{\partial t} \\ &= -T\frac{\partial s}{\partial t} \end{aligned} \tag{5.5}$$

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

$$\frac{\partial l}{\partial m} = \epsilon_{\rm nuc} - \epsilon_\nu + \epsilon_{\rm gr} \tag{5.6}$$

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

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

$$\frac{{\rm d}l}{{\rm d}m} = \epsilon_{\rm nuc} - \epsilon_\nu. \tag{5.7}$$

If we integrate this equation over the mass we obtain

$$L = \int_0^M \epsilon_{\rm nuc}\,{\rm d}m - \int_0^M \epsilon_\nu\,{\rm d}m \equiv L_{\rm nuc} - L_\nu \tag{5.8}$$

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

5.2 Energy transport by radiation and conduction

We have seen that most stars are in a long-lived state of thermal equilibrium, in which energy generation in the stellar centre exactly balances the radiative loss from the surface. What would happen if the nuclear energy source in the centre is suddenly quenched? The answer is: very little, at least initially. Photons that carry the energy are continually scattered, absorbed and re-emitted in random directions. Because stellar matter is very *opaque* to radiation, the photon mean free path $\ell_{\rm ph}$ is very small (typically $\ell_{\rm ph} \sim 1\,{\rm cm} \ll R_\odot$, see Sect. 3.1). As a result, radiation is trapped within the stellar interior, and photons diffuse outwards very slowly by a 'random walk' process. The time it takes radiation to escape from the centre of the Sun by this random walk process is roughly 10^7 years, despite the fact that photons travel at the speed of light (see Exercise 5.1). Changes in the Sun's luminosity would only occur after millions of years, on the timescale for radiative energy transport, which you may recognise as the Kelvin-Helmholtz timescale for thermal readjustment.

We also estimated in Sec. 3.1 that the temperature difference over a distance $\ell_{\rm ph}$ is only $\Delta T \sim 10^{-4}\,{\rm K}$. This means that the radiation field is extremely close to black-body radiation with $U = u\rho = aT^4$ (Sec. 3.3.6). Black-body radiation is isotropic and as a result no net energy transport would take place. However, a small anisotropy is still present due to the tiny relative temperature difference $\Delta T/T \sim 10^{-11}$. This small anisotropy is enough to carry the entire energy flux in the Sun (see Exercise 5.1). These estimates show that radiative energy transport in stellar interiors can be described as a diffusion process. This yields a great simplification of the physical description.

5.2.1 Heat diffusion by random motions

Fick's law of diffusion states that, when there is a gradient $\boldsymbol{\nabla} n$ in the density of particles of a certain type, the diffusive flux $\boldsymbol{J}$ – i.e. the net flux of such particles per unit area per second – is given by

$$\boldsymbol{J} = -D\,\boldsymbol{\nabla} n \qquad \text{with} \qquad D = \tfrac{1}{3}\bar{v}\ell. \tag{5.9}$$

Here D is the *diffusion coefficient*, which depends on the average particle velocity $\bar{v}$ and their mean free path ℓ. The origin of this equation can be understood as follows.

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

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \tfrac{1}{6} n \bar{v},$$

The factor $\frac{1}{6}$ comes from the fact that half of the particles cross the surface in one direction, and because their motions are isotropic the average velocity perpendicular to the surface is $\frac{1}{3}\bar{v}$ (this can be proven in the same way as the factor $\frac{1}{3}$ appearing in the pressure integral eq. 3.4). If there is a gradient in the particle density along the z direction, $\partial n/\partial z$, then the particles moving upwards with mean free path ℓ on average have a density $n(z-\ell)$ and those moving down on average have a density $n(z+\ell)$. Therefore the net particle flux across the surface is

$$J \;=\; \tfrac{1}{6}\bar{v}\,n(z-\ell) - \tfrac{1}{6}\bar{v}\,n(z+\ell) \;=\; \tfrac{1}{6}\bar{v}\cdot\Big(-2\ell\,\frac{\partial n}{\partial z}\Big) \;=\; -\tfrac{1}{3}\bar{v}\ell\,\frac{\partial n}{\partial z}.$$

Eq. (5.9) is the generalisation of this expression to three dimensions.

Suppose now that, in addition to a gradient in particle density, there is a gradient in the energy density U carried by these particles (e.g. photons or gas particles). Then by analogy, there is a net flux of energy across the surface, since the particles moving 'up' on average carry more energy than those moving 'down'. Therefore a gradient in the energy density $\boldsymbol{\nabla} U$ gives rise to a net energy flux

$$\boldsymbol{F} = -D\,\boldsymbol{\nabla} U, \tag{5.10}$$

Since a gradient in energy density is associated with a temperature gradient, $\boldsymbol{\nabla} U = (\partial U/\partial T)_V\,\boldsymbol{\nabla} T = C_V\,\boldsymbol{\nabla} T$, we can write this as an equation for heat conduction,

$$\boldsymbol{F} = -K\,\boldsymbol{\nabla} T \qquad \text{with} \qquad K = \tfrac{1}{3}\bar{v}\ell\,C_V, \tag{5.11}$$

where K is the *conductivity*. This description is valid for all particles in LTE, photons as well as gas particles.

5.2.2 Radiative diffusion of energy

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

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = -\kappa_\nu \rho\, I_\nu, \tag{5.12}$$

where κ_ν is the mass absorption coefficient or opacity coefficient (in $\mathrm{cm}^2\ \mathrm{g}^{-1}$) at frequency ν. The mean free path is the distance over which the intensity decreases by a factor of e, which obviously

depends on the frequency. If we make a proper average over frequencies (see Sec. 5.2.3), we can write

$$\ell_{\rm ph} = \frac{1}{\kappa\rho}. \tag{5.13}$$

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

$$K_{\rm rad} = \tfrac{4}{3}\frac{acT^3}{\kappa\rho}, \tag{5.14}$$

such that the radiative energy flux is

$$\boldsymbol{F}_{\rm rad} = -K_{\rm rad}\,\boldsymbol{\nabla}T = -\tfrac{4}{3}\frac{acT^3}{\kappa\rho}\boldsymbol{\nabla}T. \tag{5.15}$$

In spherical symmetric star the flux is related to the local luminosity, $F_{\rm rad} = l/4\pi r^2$ (eq. 5.2). We can thus rearrange the equation to obtain the temperature gradient

$$\frac{\partial T}{\partial r} = -\frac{3\kappa\rho}{16\pi acT^3}\frac{l}{r^2} \tag{5.16}$$

or when combined with eq. (2.6) for $\partial r/\partial m$,

$$\boxed{\frac{\partial T}{\partial m} = -\frac{3}{64\pi^2 ac}\frac{\kappa l}{r^4T^3}} \tag{5.17}$$

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

Eq. (5.17) is valid as long as $\ell_{\rm ph} \ll R$, i.e. as long as the LTE conditions hold. This breaks down when the stellar surface, the photosphere, is approached: this is where the photons escape, i.e. $\ell_{\rm ph} \gtrsim R$. Near the photosphere the diffusion approximation is no longer valid and we need to solve the full, and much more complicated, equations of radiative transfer. This is the subject of the study of *stellar atmospheres*. Fortunately, the LTE conditions and the diffusion approximation hold over almost the entire stellar interior.

In hydrostatic equilibrium, we can combine eqs. (5.17) and (2.13) as follows

$$\frac{{\rm d}T}{{\rm d}m} = \frac{{\rm d}P}{{\rm d}m}\cdot\frac{{\rm d}T}{{\rm d}P} = -\frac{Gm}{4\pi r^4}\,\frac{T}{P}\cdot\frac{{\rm d}\log T}{{\rm d}\log P}$$

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

$$\boxed{\nabla_{\rm rad} = \left(\frac{{\rm d}\log T}{{\rm d}\log P}\right)_{\rm rad} = \frac{3}{16\pi acG}\frac{\kappa lP}{mT^4}} \tag{5.18}$$

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

5.2.3 The Rosseland mean opacity

The radiative diffusion equations derived above are independent of frequency ν, since the flux F is integrated over all frequencies. However, in general the opacity coefficient κ_ν depends on frequency, such that the κ appearing in eq. (5.16) or (5.17) must represent a proper average over frequency. This average must be taken in a particular way.

If $\boldsymbol{F}_\nu\, d\nu$ represents the radiative flux in the frequency interval $[\nu, \nu + d\nu]$, then eq. (5.10) must be replaced by

$$\boldsymbol{F}_\nu = -D_\nu\, \boldsymbol{\nabla} U_\nu = -D_\nu\, \frac{\partial U_\nu}{\partial T}\, \boldsymbol{\nabla} T \tag{5.19}$$

where

$$D_\nu = \tfrac{1}{3} c \ell_\nu = \frac{c}{3\kappa_\nu \rho}. \tag{5.20}$$

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

$$U_\nu = \frac{8\pi h}{c^3}\, \frac{\nu^3}{e^{h\nu/kT} - 1} \tag{5.21}$$

which is proportional to the Planck function for the intensity of black-body radiation. The total flux is obtained by integrating eq. (5.19) over all frequencies,

$$\boldsymbol{F} = -\Big[\frac{c}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, \mathrm{d}\nu\Big]\, \boldsymbol{\nabla} T. \tag{5.22}$$

This is eq. (5.11) but with conductivity

$$K_{\rm rad} = \frac{c}{3\rho} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, \mathrm{d}\nu. \tag{5.23}$$

Comparing with eq. (5.14) shows that the proper average of opacity as it appears in eq. (5.16) or (5.17) is

$$\frac{1}{\kappa} = \frac{1}{4aT^3} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial U_\nu}{\partial T}\, \mathrm{d}\nu. \tag{5.24}$$

This is the so-called *Rosseland mean opacity*. The factor $4aT^3$ appearing in eq. (5.24) is equal to $\int_0^\infty (\partial U_\nu/\partial T)\, \mathrm{d}\nu$, so that the Rosseland mean can be seen as the harmonic mean of κ_ν with weighting function $\partial U_\nu/\partial T$. (The weighting function has a maximum at $h\nu \approx 4kT$, as can be verified by differentiating eq. (5.21) with respect to T, and subsequently with respect to ν.)

We can interpret the Rosseland mean in another way. The integrand of eq. (5.24) also appears in the expression (5.19) for the monochromatic flux, $\boldsymbol{F}_\nu$, when combined with (5.20). The Rosseland mean therefore favours the frequency range where the flux is large. In other words, $1/\kappa$ represents the average *transparency* of the stellar gas.

5.2.4 Conductive transport of energy

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

However, the situation can be quite different when the electrons become degenerate. In that case both their velocities increase (their momenta approach the Fermi momentum, see Sec. 3.3.5) and, more importantly, their mean free paths increase (most of the quantum cells of phase space are occupied, so an electron has to travel further to find an empty cell and transfer its momentum). At very high densities, when $\ell_e \gg \ell_{\rm ph}$, electron conduction becomes a much more efficient way of transporting energy than radiative diffusion (see Sec. 5.3). This is important for stars in late stages of evolution with dense degenerate cores and for white dwarfs, in which efficient electron conduction results in almost isothermal cores.

The energy flux due to heat conduction can be written as

$$\boldsymbol{F}_{\rm cd} = -K_{\rm cd}\,\boldsymbol{\nabla} T \tag{5.25}$$

such that the sum of radiative and conductive fluxes is

$$\boldsymbol{F} = \boldsymbol{F}_{\rm rad} + \boldsymbol{F}_{\rm cd} = -(K_{\rm rad} + K_{\rm cd})\,\boldsymbol{\nabla} T. \tag{5.26}$$

We can define a *conductive opacity* $\kappa_{\rm cd}$ by analogy with the radiative opacity, if we write the conductivity in the same form as eq. (5.14),

$$K_{\rm cd} = \frac{4acT^3}{3\kappa_{\rm cd}\,\rho}. \tag{5.27}$$

Then we can write the combined flux due to radiation and conduction in the same form as the radiative flux, eq. (5.15),

$$\boldsymbol{F} = -\frac{4acT^3}{3\kappa\rho}\boldsymbol{\nabla} T \qquad \text{with} \qquad \frac{1}{\kappa} = \frac{1}{\kappa_{\rm rad}} + \frac{1}{\kappa_{\rm cd}} \tag{5.28}$$

This result simply means that the transport mechanism with the largest flux will dominate, that is, the mechanism for which the stellar matter has the highest transparency. With κ defined as in eq. (5.28), the stellar structure equation (5.17) also accounts for the effects of conduction, if present.

5.3 Opacity

The opacity coefficient κ appearing in eq. (5.17) determines the flux that can be transported by radiation for a certain temperature gradient, or more to the point, how large the temperature gradient must be in order to carry a given luminosity l by radiation. Therefore κ is an important quantity that has a large effect on the structure of a star.

5.3.1 Sources of opacity

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

Electron scattering

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

$$\sigma_{\rm e} = \frac{8\pi}{3}\left(\frac{e^2}{m_{\rm e}c^2}\right)^2 = 6.652\times10^{-25}\,{\rm cm}^2 \tag{5.29}$$

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

$$\kappa_{\rm es} = \frac{\sigma_{\rm e}}{\mu_{\rm e} m_{\rm u}} = 0.20\,(1+X)\ \ {\rm cm}^2/{\rm g} \tag{5.30}$$

Since the electron scattering opacity is independent of frequency, this expression also gives the Rosseland mean. In the last equality we have assumed that the gas is completely ionized so that $\mu_{\rm e} = 2/(1+X)$ (eq. 3.20). Electron scattering is an important opacity source in an ionized gas that is not too dense. When the degree of ionization drops (typically when $T \lesssim 10^4$ K in hydrogen-rich gas) the electron density becomes so small that the electron scattering opacity is strongly reduced below eq. (5.30).

When the photon energy becomes a significant fraction of the rest mass of the electron, $h\nu \gtrsim 0.1 m_e c^2$, the exchange of momentum between photon and electron must be taken into account (Compton scattering). This occurs at high temperature, since the Planck function has a maximum at $h\nu = 4.965\,kT$ (Wien's law), i.e. when $kT \gtrsim 0.02 m_e c^2$ or $T \gtrsim 10^8$ K. At such temperatures the electron scattering opacity is smaller than given by eq. (5.30).

Free-free absorption

A free electron cannot absorb a photon because this would violate momentum and energy conservation. However, if a charged ion is in its vicinity, absorption is possible because of the electromagnetic coupling between the ion and electron. This *free-free absorption* is the inverse process of bremsstrahlung, where an electron emits a photon when it passes by and interacts with an ion.

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

$$\kappa_{\nu,{\rm ff}} \propto \frac{n_{\rm e}}{\rho} \sum_i n_i\, Z_i{}^2\, T^{-1/2}\, \nu^{-3}.$$

In a completely ionized gas, $n_{\rm e}/\rho = 1/(\mu_{\rm e} m_{\rm u}) = (1+X)/2m_{\rm u}$. Following Sec. 3.3.3, the sum over ions can be written as $\sum_i n_i Z_i{}^2 = (\rho/m_{\rm u}) \sum_i (X_i Z_i{}^2/A_i) = (\rho/m_{\rm u})\,(X+Y+B)$, where $B = \sum_{i>2} (X_i Z_i{}^2/A_i)$ is the contribution of elements heavier than helium. As long as their abundance is small, one can take $X + Y + B \approx 1$ to a reasonable approximation.

When we take the Rosseland mean, the factor ν^{-3} becomes a factor T^{-3} (this can be verified by performing the integration of eq. 5.24 with $\kappa_\nu \propto \nu^{-\alpha}$, see Exercise 5.2). We thus obtain

$$\kappa_{\rm ff} \propto \rho\, T^{-7/2}. \tag{5.31}$$

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

$$\kappa_{\rm ff} \approx 3.8 \times 10^{22}\,(1+X)\,\rho\, T^{-7/2}\ \ {\rm cm}^2/{\rm g}. \tag{5.32}$$

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

Bound-free and bound-bound absorption

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

$$\kappa_{\rm bf} \approx 4.3 \times 10^{25}\,(1+X)Z\,\rho\,T^{-7/2}\ \ {\rm cm}^2/{\rm g}. \tag{5.33}$$

Again one should use this formula with caution. It should certainly not be applied for $T \lesssim 10^4\,{\rm K}$ because most of the photons are not energetic enough to ionize the electrons, while at very high T most species are fully ionized so the bound-free opacity is smaller than eq. (5.33) suggests. Keeping these limitations in mind, the bound-free opacity is seen to depend directly on the metallicity Z. One thus has, very approximately, $\kappa_{\rm bf} \approx 10^3 Z \times \kappa_{\rm ff}$. We may thus expect bound-free absorption to dominate over free-free absorption for $Z \gtrsim 10^{-3}$.

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

The negative hydrogen ion

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

$$\kappa_{\rm H^-} \approx 2.5 \times 10^{-31} \left(\frac{Z}{0.02}\right) \rho^{1/2}\,T^9\ \ {\rm cm}^2/{\rm g} \tag{5.34}$$

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

Molecules and dust

In cool stars with $T_{\rm eff} \lesssim 4000\,{\rm K}$ opacity sources arising from molecules and (at even lower temperatures) dust grains become dominant. Here one has to deal with complicated molecular chemistry and dust formation processes, which still contains a lot of uncertainty. When dust grains form, at $T \lesssim 1500\,{\rm K}$, they are very effective absorbers in the outer atmospheres of very cool stars.

Conductive opacities

As we saw in Sec. 5.2.4, energy transport by means of heat conduction can also be described by means of a conductive opacity coefficient $\kappa_{\rm cd}$. Under ideal gas conditions, conduction is very inefficient

compared to radiative transport of energy ($\kappa_{\rm cd} \gg \kappa_{\rm rad}$). Therefore we only need to consider the case of a degenerate electron gas. In this case the following approximation holds

$$\kappa_{\rm cd} \approx 4.4 \times 10^{-3} \, \frac{\sum_i Z_i^{5/3} X_i / A_i}{(1+X)^2} \, \frac{(T/10^7\,{\rm K})^2}{(\rho/10^5\,{\rm g/cm^3})^2} \;\; {\rm cm^2/g}. \tag{5.35}$$

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

5.3.2 A detailed view of stellar opacities

In general, $\kappa = \kappa(\rho, T, X_i)$ is a complicated function of density, temperature and composition. While certain approximations can be made, as in the examples shown above, these are usually too simplified and inaccurate to apply in detailed stellar models. An additional complication is that the Rosseland mean opacity (eq. 5.24) is not additive: the opacity of a mixture of gases is not simply equal to the sum of the opacities of its components. Instead, one first has to add the frequency-dependent opacities, $\kappa_\nu = \sum_i X_i \kappa_{\nu,i}$ and then integrate over ν to calculate the Rosseland mean.

In practical stellar structure calculations one usually interpolates in pre-computed opacity tables, e.g. as calculated in the 1990s by the OPAL project. An example is shown in Fig. 5.2 for a quasi-solar mixture of elements. One may recognize the various regions in the density-temperature plane where one of the processes discussed above dominates. At low density and high temperature, κ has a constant value given by electron scattering. Opacity increases towards higher ρ and lower T due to free-free and bound-free absorptions. For $T < 10^4$ K opacity decreases drastically due to recombination of hydrogen, the main opacity source here is the H^- ion. At lower temperatures still, κ rises again

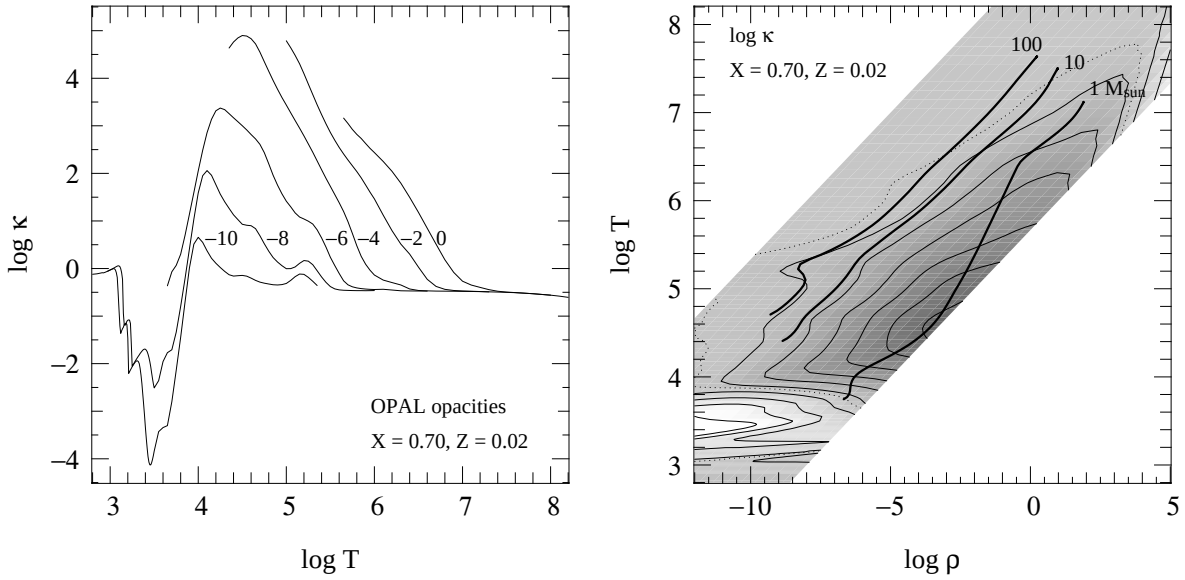


Figure 5.2. Rosseland mean opacities as a function of T and ρ, for a mixture of elements representative of solar abundances ($X = 0.7, Z = 0.02$), calculated by the OPAL project for high temperatures and by J. Ferguson for low temperatures ($\log T < 3.8$). The left panel shows curves of $\log\kappa$ (in cm^2/g) versus temperature for several values of the density, labelled by the value of $\log\rho$ (in g/cm^3). The right panel shows contour lines of constant $\log\kappa$ in the ρ-T plane, in steps of 1.0 between -4 and 5, over the region in temperature and density for which the radiative opacity has been calculated. The thick lines are detailed structure models for main-sequence stars of 1, 10 and 100 $M_\odot$, as in Fig. 3.4.

due to molecules and dust formation. Finally, at very high density the opacity is dominated by the conductivity of degenerate electrons and decreases strongly with increasing ρ (just visible in the upper right corner of Fig. 5.2). It should be clear that there is much more structure in the function $\kappa(\rho, T)$ than in the simple power-law approximations, such as the Kramers law. The many ridges and bumps show that the Kramers law is a rather poor approximation of the actual opacity.

For comparison, interior structure models for main-sequence stars of different masses are also shown. The opacity in the interior of a 1 $M_\odot$ star is dominated by free-free and bound-free absorption, and is very high (up to $10^5\,\mathrm{cm}^2/\mathrm{g}$) in the envelope, at temperatures between 10^4 and $10^5\,\mathrm{K}$. In the surface layers the opacity rapidly decreases due to the H^- opacity. More massive stars are located at lower densities than the Sun, and generally have much lower opacities in their envelopes. In the most massive stars the opacity is dominated by electron scattering, at low ρ and high T. However, even here one has to deal with additional opacity bumps, most prominently the one due to bound-free transitions of Fe at $\log T \approx 5.3$.

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

5.4 The Eddington luminosity

We have seen that radiative transport of energy inside a star requires a temperature gradient $\mathrm{d}T/\mathrm{d}r$, the magnitude of which is given by eq. (5.16). Since $P_\mathrm{rad} = \frac{1}{3}aT^4$, this means there is also a gradient in the radiation pressure:

$$\frac{\mathrm{d}P_\mathrm{rad}}{\mathrm{d}r} = -\frac{4}{3}aT^3\frac{\mathrm{d}T}{\mathrm{d}r} = -\frac{\kappa\rho}{4\pi c}\frac{l}{r^2}. \tag{5.36}$$

This radiation pressure gradient represents an outward force due to the net flux of photons outwards. Of course, for a star in hydrostatic equilibrium this outward radiation force must be smaller than the inward force of gravity, as given by the pressure gradient necessary for HE, eq. (2.12). In other words,

$$\left|\frac{\mathrm{d}P_\mathrm{rad}}{\mathrm{d}r}\right| < \left|\left(\frac{\mathrm{d}P}{\mathrm{d}r}\right)_\mathrm{HE}\right| \quad \Rightarrow \quad \frac{\kappa\rho}{4\pi c}\frac{l}{r^2} < \frac{Gm\rho}{r^2}.$$

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

$$l < \frac{4\pi cGm}{\kappa} = l_\mathrm{Edd}. \tag{5.37}$$

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

The inequality expressed by eq. (5.37) can be violated in the case of a very large heat flux (large l), which may result from intense nuclear burning, or in the case of a very high opacity κ. As we saw in Sec. 5.3, high opacities are encountered at relatively low temperatures, near the ionization temperature of hydrogen and helium (and for example in the outer layers of the Sun). In such cases hydrostatic equilibrium (eq. 2.13) and radiative equilibrium (eq. 5.17) cannot hold simultaneously. Therefore, if the star is to remain in HE, energy must be transported by a different means than radiative diffusion. This means of transport is *convection*, the collective motion of gas bubbles that carry heat and can distribute it efficiently. We shall consider convection in detail in Sec. 5.5. It will turn out that eq. (5.37) is a necessary, but not a sufficient condition for a region of a star to be stable against convection.

The surface layer of a star is always radiative, since it is from here that energy escapes the star in the form of photons. Applying eq. (5.37) at the surface of the star ($m = M$) we get

$$L < L_\mathrm{Edd} = \frac{4\pi cGM}{\kappa}, \tag{5.38}$$

where κ is the opacity in the photosphere. Violation of this condition now means violation of hydrostatic equilibrium: matter is accelerated away from the star by the photon pressure, giving rise to violent mass loss. The Eddington luminosity expressed by eq. (5.38) is a critical stellar luminosity that cannot be exceeded by a star in hydrostatic equilibrium. If we assume κ to be approximately constant (in very luminous main-sequence stars the opacity is dominated by electron scattering, so this is not a bad assumption) then $L_{\rm Edd}$ is only dependent on M. It can be expressed as follows

$$L_{\rm Edd} = 3.8 \times 10^4 \left(\frac{M}{M_\odot}\right)\left(\frac{0.34\,{\rm cm}^2/{\rm g}}{\kappa}\right) L_\odot. \tag{5.39}$$

The value of $0.34\,{\rm cm}^2/{\rm g}$ corresponds to the electron scattering opacity for $X = 0.7$.

Since $L_{\rm Edd}$ is proportional to M, while stars (at least on the main sequence) follow a mass-luminosity relation $L \propto M^x$ with $x > 1$ (Sec. 1.1.2), this implies that for stars of increasing mass L will at some point exceed $L_{\rm Edd}$. Hence, we can expect a *maximum mass* to exist for main-sequence stars. Note that the existence of a steep mass-luminosity relation (with $x \approx 3$) can be derived directly for stars in which energy transport occurs by radiation (see Exercise 5.3, and also Sec. 7.4), without having to assume anything about how energy is generated.

5.5 Convection

For radiative diffusion to transport energy outwards, a certain temperature gradient is needed, given by eq. (5.16) or eq. (5.17). The larger the luminosity that has to be carried, the larger the temperature gradient required. There is, however, an upper limit to the temperature gradient inside a star – if this limit is exceeded an instability in the gas sets in. This instability leads to cyclic macroscopic motions of the gas, known as *convection*. Convection can be regarded as a type of dynamical instability, although (as we shall see later in this section) it does not have disruptive consequences. In particular, it does not lead to an overall violation of hydrostatic equilibrium. Convection affects the structure of a star only as an efficient means of heat transport and as an efficient mixing mechanism.

In Sec. 5.4 we already derived an upper limit to the luminosity that can be transported by radiation. We will now derive a more stringent criterion for convection to occur, based on considerations of dynamical stability.

5.5.1 Criteria for stability against convection

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

Consider a mass element that, due to a small perturbation, is displaced upwards by a small distance as depicted in Fig. 5.3. At its original position (at radius r) the density and pressure are ρ_1 and P_1, and at its new position ($r + \Delta r$) the ambient density and pressure are ρ_2 and P_2. Since pressure decreases outwards, $P_2 < P_1$ and the gas element will expand to restore pressure equilibrium with its surroundings. Hence the pressure of the gas element at position 2 is $P_e = P_2$, but its new density after expansion ρ_e is not necessarily equal to ρ_2. If $\rho_e > \rho_2$, the gas element will experience a net buoyancy force downwards (by Archimedes' law), which pushes it back towards its original position. Then the small perturbation is quenched, and the situation is stable. On the other hand, if $\rho_e < \rho_2$ then there is a net buoyancy force upwards and we have an *unstable* situation that leads to convection.

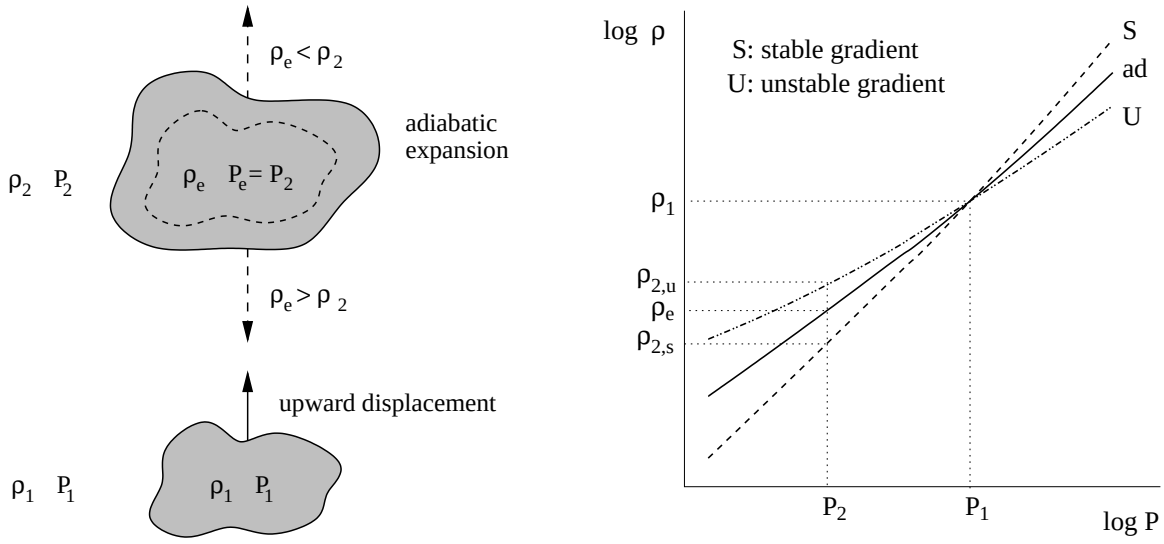


Figure 5.3. Schematic illustration of the Schwarzschild criterion for stability against convection. A gas element is perturbed and displaced upwards from position 1 to position 2, where it expands adiabatically to maintain pressure equilibrium with its surroundings. If its density is larger than the ambient density, it will sink back to its original position. If its density is smaller, however, buoyancy forces will accelerate it upwards: convection occurs. On the right the situation is shown in a density-pressure diagram. A layer is stable against convection if the density varies more steeply with pressure than for an adiabatic change.

The expansion of the gas element as it rises over Δr occurs on the local dynamical timescale (i.e. with the speed of sound), which is typically much shorter than the local timescale for heat exchange, at least in the deep interior of the star. The displacement and expansion of the gas element will therefore be very close to adiabatic. We have seen in Sec. 3.4 that the adiabatic exponent $\gamma_{\rm ad}$ defined by eq. (3.56) describes the logarithmic response of the pressure to an adiabatic change in the density. Writing as $\delta\rho_e$ and δP_e the changes in the density and pressure of the element when it is displaced over a small distance Δr, we therefore have

$$\frac{\delta P_e}{P_e} = \gamma_{\rm ad} \frac{\delta\rho_e}{\rho_e}. \tag{5.40}$$

Here δP_e is determined by the pressure gradient $\mathrm{d}P/\mathrm{d}r$ inside the star because $P_e = P_2$, i.e. $\delta P_e = P_2 - P_1 = (\mathrm{d}P/\mathrm{d}r)\,\Delta r$. Therefore the change in density $\delta\rho_e$ follows from eq. (5.40)

$$\delta\rho_e = \frac{\rho_e}{P_e} \frac{1}{\gamma_{\rm ad}} \frac{\mathrm{d}P}{\mathrm{d}r} \Delta r. \tag{5.41}$$

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

$$\delta\rho_e > \frac{\mathrm{d}\rho}{\mathrm{d}r} \Delta r, \tag{5.42}$$

which combined with eq. (5.41) yields an upper limit to the density gradient for which a layer inside the star is stable against convection,

$$\frac{1}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}r} < \frac{1}{P}\frac{\mathrm{d}P}{\mathrm{d}r}\frac{1}{\gamma_{\rm ad}}, \tag{5.43}$$

where we have replaced P_e and ρ_e by P and ρ, since the perturbations are assumed to be very small. Remember, however, that both $\mathrm{d}\rho/\mathrm{d}r$ and $\mathrm{d}P/\mathrm{d}r$ are negative. Therefore, in absolute value the sign

of eq. (5.43) reverses, which means that the density gradient must be *steeper* than a critical value, determined by $\gamma_{\rm ad}$. If we divide (5.43) by dP/dr we obtain the general criterion for stability against convection, which is depicted on the right-hand side in Fig. 5.3,

$$\frac{\mathrm{d}\log\rho}{\mathrm{d}\log P} > \frac{1}{\gamma_{\rm ad}}. \tag{5.44}$$

If condition (5.44) is violated then convective motions will develop. Gas bubbles that, due to a small perturbation, are slightly hotter than their surroundings will move up, transporting their heat content upwards until they are dissolved. Other bubbles may be slightly cooler than their environment, these will move down and have a smaller heat content than their surroundings. When these bubbles finally dissolve, they absorb heat from their surroundings. Therefore, both the upward and downward moving convective bubbles effectively transport heat in the upward direction. Hence there is a *net upward heat flux*, even though there is no net mass flux, since upward and downward moving bubbles carry equal amounts of mass. This is the principle behind convective heat transport.

The Schwarzschild and Ledoux criteria

The stability criterion (5.44) is not of much practical use, because it involves computation of a density gradient which is not part of the stellar structure equations. We would rather have a criterion for the temperature gradient, because this also appears in the equation for radiative transport. We can rewrite eq. (5.44) in terms of temperature by using the equation of state. We write the equation of state in its general, differential form (eq. 3.48) but now also take into account a possible variation in composition. If we characterize the composition by the mean molecular weight μ then $P = P(\rho, T, \mu)$ and we can write

$$\frac{\mathrm{d}P}{P} = \chi_T \frac{\mathrm{d}T}{T} + \chi_\rho \frac{\mathrm{d}\rho}{\rho} + \chi_\mu \frac{\mathrm{d}\mu}{\mu}, \tag{5.45}$$

with χ_T and χ_ρ defined by eqs. (3.49) and (3.50), and χ_μ is defined as

$$\chi_\mu = \left(\frac{\partial \log P}{\partial \log \mu}\right)_{\rho,T}. \tag{5.46}$$

For an ideal gas $\chi_\mu = -1$. With the help of eq. (5.45) we can write the variation of density with pressure through the star as

$$\frac{\mathrm{d}\log\rho}{\mathrm{d}\log P} = \frac{1}{\chi_\rho}\left(1 - \chi_T \frac{\mathrm{d}\log T}{\mathrm{d}\log P} - \chi_\mu \frac{\mathrm{d}\log\mu}{\mathrm{d}\log P}\right) = \frac{1}{\chi_\rho}(1 - \chi_T\nabla - \chi_\mu\nabla_\mu). \tag{5.47}$$

Here we have introduced, by analogy with eq. (5.18), the symbols $\nabla \equiv \mathrm{d}\log T/\mathrm{d}\log P$ and $\nabla_\mu \equiv \mathrm{d}\log\mu/\mathrm{d}\log P$. These quantities represent the actual gradients of temperature and of mean molecular weight through the star, regarding P as the variable that measures depth. In the displaced gas element the composition does not change, and from eq. (3.63) we can write

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

so that the stability criterion (5.44) becomes

$$\nabla < \nabla_{\rm ad} - \frac{\chi_\mu}{\chi_T}\nabla_\mu. \tag{5.48}$$

If all the energy is transported by radiation then $\nabla = \nabla_{\rm rad}$ as given by eq. (5.18). Hence we can replace ∇ by $\nabla_{\rm rad}$ in eq. (5.48) and thus arrive at the *Ledoux criterion* which states that a layer is stable against convection if

$$\boxed{\nabla_{\rm rad} < \nabla_{\rm ad} - \frac{\chi_\mu}{\chi_T}\nabla_\mu} \qquad \text{(Ledoux)} \tag{5.49}$$

In chemically homogeneous layers $\nabla_\mu = 0$ and eq. (5.49) reduces to the simple *Schwarzschild criterion* for stability against convection[1]

$$\boxed{\nabla_{\rm rad} < \nabla_{\rm ad}} \qquad \text{(Schwarzschild)} \tag{5.50}$$

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

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

$$\nabla_{\rm rad} < \nabla_{\rm ad} + \nabla_\mu. \tag{5.51}$$

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

Occurrence of convection

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

$$\nabla_{\rm rad} = \frac{3}{16\pi acG}\frac{P}{T^4}\frac{\kappa l}{m} > \nabla_{\rm ad}. \tag{5.52}$$

This requires one of following:

- A large value of κ, that is, convection occurs in opaque regions of a star. Examples are the outer envelope of the Sun (see Fig. 5.2) and of other cool stars, because opacity increases with decreasing temperature. Since low-mass stars are cooler than high-mass stars, we may expect low-mass stars to have convective envelopes.

- A large value of l/m, i.e. regions with a large energy flux. We note that towards the centre of a star $l/m \approx \epsilon_{\rm nuc}$ by eq. (5.4), so that stars with nuclear energy production that is strongly peaked towards the centre can be expected to have convective cores. We shall see that this is the case for relatively massive stars.

[1] We can relate the convection criterion to the Eddington limit derived in Sec. 5.4. By writing $\nabla_{\rm rad}$ in terms of l, $l_{\rm Edd}$ (defined in eq. 5.37) and $P_{\rm rad} = (1-\beta)P$ we can rewrite the Schwarzschild criterion for stability as

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

(see Exercise 5.6). For $\beta > 0$ and $\nabla_{\rm ad} > 0.25$ we see that convection already sets in before the Eddington limit is reached.

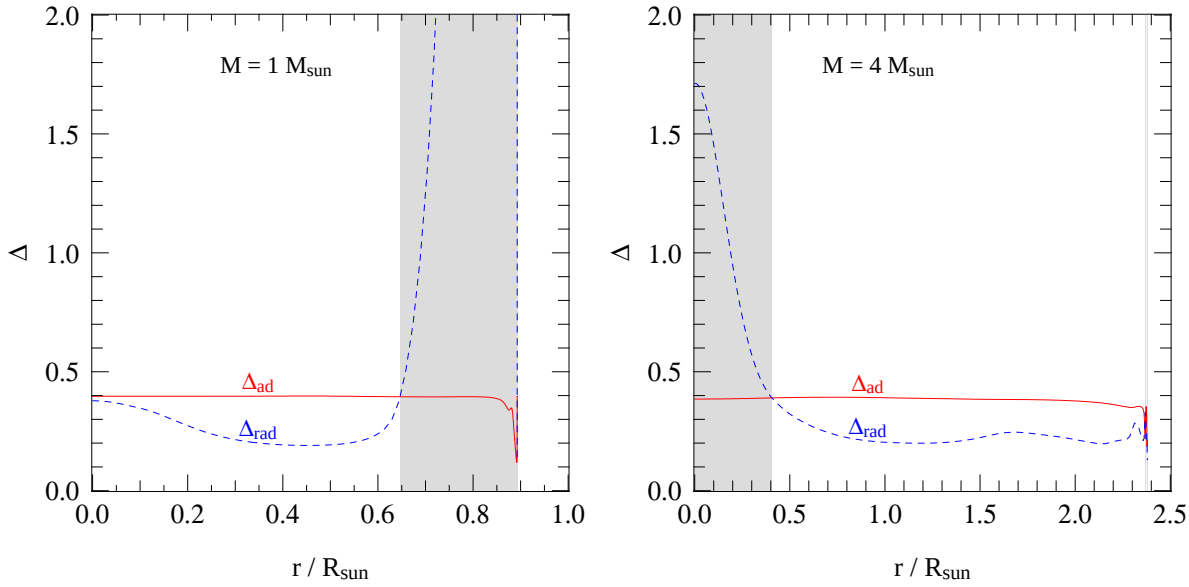


Figure 5.4. The variation of $\nabla_{\rm ad}$ (red, solid line) and $\nabla_{\rm rad}$ (blue, dashed line) with radius in two detailed stellar models of $1\,M_\odot$ and $4\,M_\odot$ at the start of the main sequence. The solar-mass model has a very large opacity in its outer layers, resulting in a large value of $\nabla_{\rm rad}$ which gives rise to a convective envelope where $\nabla_{\rm rad} > \nabla_{\rm ad}$ (indicated by gray shading). On the other hand, the $4\,M_\odot$ model has a hotter outer envelope with lower opacity so that $\nabla_{\rm rad}$ stays small. The large energy generation rate in the centre now results in a large $\nabla_{\rm rad}$ and a convective core extending over the inner $0.4\,R_\odot$. In both models $\nabla_{\rm ad} \approx 0.4$ since the conditions are close to an ideal gas. In the surface ionization zones, however, $\nabla_{\rm ad} < 0.4$ and a thin convective layer appears in the $4\,M_\odot$ model.

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

These effects are illustrated in Fig. 5.4.

5.5.2 Convective energy transport

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

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

of the local pressure scale height, which is the radial distance over which the pressure changes by an e-folding factor,

$$H_P = \left| \frac{\mathrm{d}r}{\mathrm{d}\ln P} \right| = \frac{P}{\rho g}. \tag{5.53}$$

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

The convective energy flux

Within the framework of MLT we can calculate the convective energy flux, and the corresponding temperature gradient required to carry this flux, as follows. After rising over a radial distance ℓ_m the temperature difference between the gas element (e) and its surroundings (s) is

$$\Delta T = T_\mathrm{e} - T_\mathrm{s} = \left[\left(\frac{\mathrm{d}T}{\mathrm{d}r} \right)_\mathrm{e} - \frac{\mathrm{d}T}{\mathrm{d}r} \right] \ell_\mathrm{m} = \Delta \left(\frac{\mathrm{d}T}{\mathrm{d}r} \right) \ell_\mathrm{m}.$$

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

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

noting that the '$-$' sign appears because $\mathrm{d}P/\mathrm{d}r < 0$ in eq. (5.53). Hence

$$\Delta T = T \, \frac{\ell_\mathrm{m}}{H_P} \, (\nabla - \nabla_\mathrm{ad}). \tag{5.54}$$

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

$$F_\mathrm{conv} = v_\mathrm{c} \, \rho \Delta u = v_\mathrm{c} \, \rho c_P \Delta T \tag{5.55}$$

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

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

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

$$v_\mathrm{c} \approx \sqrt{\tfrac{1}{2} \ell_\mathrm{m} g \frac{\Delta T}{T}} \approx \sqrt{\frac{{\ell_\mathrm{m}}^2 g}{2 H_P} \, (\nabla - \nabla_\mathrm{ad})}. \tag{5.56}$$

Combining this with eq. (5.55) gives

$$F_\mathrm{conv} = \rho c_P T \left(\frac{\ell_\mathrm{m}}{H_P} \right)^2 \sqrt{\tfrac{1}{2} g H_P} \, (\nabla - \nabla_\mathrm{ad})^{3/2}. \tag{5.57}$$

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

Estimate of the convective temperature gradient

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

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

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

$$F_{\rm conv} \sim \frac{M}{R^3}\left(\frac{GM}{R}\right)^{3/2} (\nabla - \nabla_{\rm ad})^{3/2}. \tag{5.58}$$

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

$$\nabla - \nabla_{\rm ad} \sim \left(\frac{LR}{M}\right)^{2/3} \frac{R}{GM} \tag{5.59}$$

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

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

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

$$\frac{{\rm d}T}{{\rm d}m} = -\frac{Gm}{4\pi r^4}\frac{T}{P}\nabla \qquad \text{with} \qquad \nabla = \nabla_{\rm ad}. \tag{5.60}$$

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

As the surface is approached, convection becomes very inefficient at transporting energy. Then $F_{\rm conv} \ll F_{\rm rad}$ so that radiation effectively transports all the energy, and $\nabla \approx \nabla_{\rm rad}$ despite convection taking place. These effects are shown in Fig. 5.5 for a detailed solar model.

5.5.3 Convective mixing

Besides being an efficient means of transporting energy, convection is also a very efficient *mixing mechanism*. We can see this by considering the average velocity of convective cells, eq. (5.56), and taking $\ell_{\rm m} \approx H_P$ and $\sqrt{gH_P} \approx \upsilon_{\rm s}$, so that

$$\upsilon_{\rm c} \approx \upsilon_{\rm s}\sqrt{\nabla - \nabla_{\rm ad}}. \tag{5.61}$$

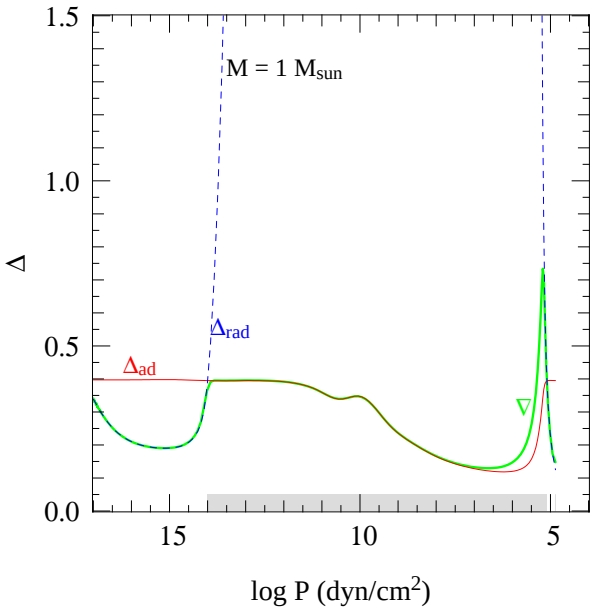


Figure 5.5. The variation of $\nabla_{\rm ad}$ (red, solid line) and $\nabla_{\rm rad}$ (blue, dashed line) in the same detailed model of 1 $M_\odot$ as shown in Fig. 5.4, but now plotted against $\log P$ rather than radius to focus on the outermost layers (where the pressure gradient is very large). The thick green line shows the actual temperature gradient ∇. The partial ionization zones are clearly visible as depressions in $\nabla_{\rm ad}$ (compare to Fig. 3.5b). The convection zone stretches from $\log P \approx 14$ to 5 (indicated by a gray bar along the bottom). In the deep interior (for $\log P > 8$) convection is very efficient and $\nabla = \nabla_{\rm ad}$. Higher up, at lower pressures, convection becomes less and less efficient at transporting energy and requires a larger T-gradient, $\nabla > \nabla_{\rm ad}$. In the very outer part of the convection zone convection is very inefficient and $\nabla \approx \nabla_{\rm rad}$.

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

By substituting into eq. (5.61) rough estimates for the interior of a star, i.e. $v_{\rm s} \sim \sqrt{GM/R}$ and eq. (5.59) for $\nabla - \nabla_{\rm ad}$, we obtain $v_{\rm c} \sim (LR/M)^{1/3} \approx 5 \times 10^3$ cm/s for a star like the Sun. These velocities are large enough to mix a convective region on a small timescale. We can estimate the timescale on which a region of radial size $d = qR$ is mixed as $\tau_{\rm mix} \approx d/v_{\rm c} \sim q(R^2 M/L)^{1/3}$, which is $\sim q \times 10^7$ sec for solar values. Depending on the fractional extent q of a convective region, the convective mixing timescale is of the order of weeks to months. Hence $\tau_{\rm mix} \ll \tau_{\rm KH} \ll \tau_{\rm nuc}$, so that over a thermal timescale, and certainly over a nuclear timescale, a convective region inside a star will be mixed homogeneously. (Note that convective mixing remains very efficient in the outer layers of a star, even though convection becomes inefficient at transporting energy.)

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

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

- A star with a deep *convective envelope*, such that it extends into regions where nuclear burning has taken place, will mix the burning products outwards towards the surface. This process

(often called ‘dredge-up’), which happens when stars become red giants, can therefore modify the surface composition, and in such a star measurements of the surface abundances provide a window into nuclear processes that have taken place deep inside the star.

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

5.5.4 Convective overshooting

To determine the extent of a region that is mixed by convection, we need to look more closely at what happens at the boundary of a convective zone. According to the Schwarzschild criterion derived in Sec. 5.5.1, in a chemically homogeneous layer this boundary is located at the surface where $\nabla_{\rm rad} = \nabla_{\rm ad}$. At this point the acceleration due to the buoyancy force, $a \approx g(\nabla - \nabla_{\rm ad})$, vanishes. Just outside this boundary, the acceleration changes sign and a convective bubble will be strongly braked – even more so when the non-mixed material outside the convective zone has a lower μ and hence a lower density. However, the convective eddies have (on average) a non-zero velocity when they cross the Schwarzschild boundary, and will *overshoot* by some distance due to their inertia. A simple estimate of this overshooting distance shows that it should be much smaller than a pressure scale height, so that the Schwarzschild criterion should determine the convective boundary quite accurately. However the convective elements also carry some heat and mix with their surroundings, so that both $|\nabla - \nabla_{\rm ad}|$ and the μ-gradient decrease. Thus also the effective buoyancy force that brakes the elements decreases, and a positive feedback loop can develop that causes overshooting elements to penetrate further and further. This is a highly non-linear effect, and as a result the actual overshooting distance is very uncertain and could be substantial.

Convective overshooting introduces a large uncertainty in the extent of mixed regions, with important consequences for stellar evolution. A convectively mixed core that is substantially larger will generate a larger fuel supply for nuclear burning, and thus affects both the hydrogen-burning lifetime and the further evolution of a star. In stellar evolution calculations one usually parametrizes the effect of overshooting by assuming that the distance $d_{\rm ov}$ by which convective elements penetrate beyond the Schwarzschild boundary is a fixed fraction of the local pressure scale height, $d_{\rm ov} = \alpha_{\rm ov} H_P$. Here $\alpha_{\rm ov}$ is a free parameter, that can be calibrated against observations (see Chapter 9).

Suggestions for further reading

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

Exercises

5.1 Radiation transport

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

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

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

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

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

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

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

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

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

5.2 Opacity

(a) Identify the various processes contributing to the opacity as shown in Fig. 5.2, and the T and ρ ranges where they are important.

(b) Compare the opacity curve for $\log\rho = -6$ in the left panel of Fig. 5.2 to the approximations given in Sec. 5.3.1 for (1) electron scattering, (2) free-free absorption, (3) bound-free absorption and (4) the H^- ion. How well do these approximations fit the realistic opacity curve?

(c) Calculate (up to an order of magnitude) the photon mean free path in a star of 1 $M_\odot$ at radii where the temperature is 10^7 K, 10^5 K and 10^4 K, using the right panel of Fig. 5.2.

(d) Suppose that the frequency-dependent opacity coefficient has the form $\kappa_\nu = \kappa_0 \nu^{-\alpha}$. Show that the Rosseland mean opacity depends on the temperature as $\kappa \propto T^{-\alpha}$.

5.3 Mass-luminosity relation for stars in radiative equilibrium

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

(a) Derive how the central temperature, T_c, scales with the mass, M, radius, R, and luminosity, L, for a star in which the energy transport is by radiation. To do this, use the stellar structure equation (5.16) for the temperature gradient in radiative equilibrium. Assume that $r \sim R$ and that the temperature is proportional to T_c, $l \sim L$ and estimating $\mathrm{d}T/\mathrm{d}r \sim -T_c/R$.

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

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

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

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

5.4 Conceptual questions: convection

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

(b) Explain the Schwarzschild criterion

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

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

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

5.5 Applying Schwarzschild's criterion

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

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

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

5.6 The Eddington luminosity

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

(a) Show that

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

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

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

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

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

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

Chapter 6

Nuclear processes in stars

For a star in thermal equilibrium, an internal energy source is required to balance the radiative energy loss from the surface. This energy source is provided by *nuclear reactions* that take place in the deep interior, where the temperature and density are sufficiently high. In ordinary stars, where the ideal-gas law holds, this stellar nuclear reactor is very stable: the rate of nuclear reactions adapts itself to produce exactly the amount of energy that the star radiates away from its surface. Nuclear reactions do not determine the luminosity of the star – this is set by how fast the energy can be transported, i.e. by the opacity of the stellar gas – but they do determine for how long the star is able to sustain its luminosity. In stars composed of degenerate gas, on the other hand, nuclear reactions are unstable and may give rise to flashes or even explosions.

Apart from energy generation, another important effect of nuclear reactions is that they change the composition by transmutations of chemical elements into other, usually heavier, elements. In this way stars produces all the elements in the Universe heavier than helium – a process called *stellar nucleosynthesis*.

6.1 Basic nuclear properties

Consider a reaction whereby a nucleus X reacts with a particle a, producing a nucleus Y and a particle b. This can be denoted as

$$X + a \rightarrow Y + b \qquad \text{or} \qquad X(a,b)Y \quad . \tag{6.1}$$

The particle a is generally another nucleus, while the particle b could also be a nucleus, a γ-photon or perhaps another kind of particle. Some reactions produce more than two particles (e.g. when a weak interaction is involved, an electron and anti-neutrino can be produced in addition to nucleus Y), but the general principles discussed here also apply to reactions involving different numbers of nuclei. Each nucleus is characterized by two integers, the charge Z_i (representing the number of protons in the nucleus) and the baryon number or mass number A_i (equal to the total number of protons plus neutrons). Charges and baryon numbers must be conserved during a reaction, i.e. for the example above:

$$Z_X + Z_a = Z_Y + Z_b \qquad \text{and} \qquad A_X + A_a = A_Y + A_b. \tag{6.2}$$

If a or b are non-nuclear particles then $A_i = 0$, while for reactions involving weak interactions the lepton number must also be conserved during the reaction. Therefore any three of the reactants uniquely determine the fourth.

6.1.1 Nuclear energy production

The masses of atomic nuclei are not exactly equal to the sum of the masses of the individual nucleons (protons and neutrons), because the nucleons are bound together by the strong nuclear force. If m_i denotes the mass of a nucleus i, then the *binding energy* of the nucleus can be defined as

$$E_{B,i} = [(A_i - Z_i)m_\mathrm{n} + Z_i m_\mathrm{p} - m_i]\, c^2, \tag{6.3}$$

where m_n and m_p are the masses of a free neutron and proton respectively. Therefore, although $\sum A_i$ is conserved during a nuclear reaction, the sum of the actual masses involved in the reaction is not. This mass difference Δm is converted into energy according to Einstein's formula $E = \Delta m c^2$. The energy released by a reaction of the kind $X(a,b)Y$ is therefore

$$Q = (m_X + m_a - m_Y - m_b)\, c^2. \tag{6.4}$$

Note that Q may be negative if energy is absorbed by the reaction; such reactions are called *endothermic*. Reactions that release energy ($Q > 0$) are called *exothermic*.

In practice, one often uses atomic masses rather than nuclear masses to calculate Q. This is allowed because the number of electrons is conserved during a reaction – despite the fact that the nuclei are completely ionized under the conditions where nuclear reactions take place. Atomic masses of a few important isotopes are given in Table 6.1. The energy release by a reaction is related to the so-called *mass defect* of nuclei, defined as

$$\Delta M_i = (m_i - A_i m_\mathrm{u})\, c^2. \tag{6.5}$$

Since nucleon number is conserved during a reaction, we can write (6.4) as

$$Q = \Delta M_X + \Delta M_a - \Delta M_Y - \Delta M_b. \tag{6.6}$$

Nuclear binding energies and reaction Q-values are usually expressed in MeV. Published tables of atomic masses often list the mass defects in MeV, rather than the masses themselves. Remember that m_u is defined as 1/12 times the mass of the ^{12}C atom; a useful identity is $m_\mathrm{u}c^2 = 931.494\,\mathrm{MeV}$.

When comparing different nuclei, the *binding energy per nucleon* E_B/A is a more informative quantity than E_B itself. In Fig. 6.1 this quantity is plotted against mass number A. With the exception of the lightest nuclei, typical values are around 8 MeV. This reflects the short range of the strong nuclear force: a nucleon only 'feels' the attraction of the nucleons in its immediate vicinity, so that E_B/A quickly saturates with increasing A. There is a slow increase with A up to a maximum at ^{56}Fe,

Table 6.1. Atomic masses of several important isotopes.

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

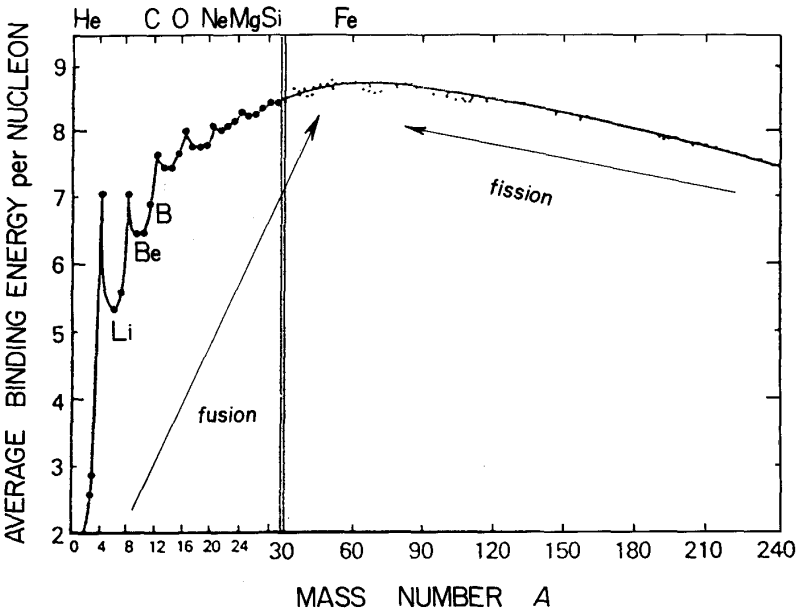


Figure 6.1. Binding energy of atomic nuclei per nucleon, E_B/A as a function of mass number A.

which has $E_B/A = 8.79\,\mathrm{MeV}$, beyond which the binding energy per nucleon decreases towards larger A. This decrease is due to the increase in the number of protons Z with A: the protons inside the nucleus experience a repulsive Coulomb force, which has a long range and does not saturate with increasing Z. There is additional structure in the curve, caused by the shell structure of nuclei and pairing effects.

The most tightly bound nuclei occur around the maximum at ^{56}Fe. Energy can be gained from the fusion of light nuclei into heavier ones as long as E_B/A increases; this is the main energy source in stars. Fusion of nuclei heavier than ^{56}Fe would be endothermic and does not occur in nature (but energy can be released by fission reactions that break up heavy nuclei into lighter ones). ^{56}Fe thus forms the natural endpoint of the stellar nuclear reaction cycles. In a star initially consisting mostly of hydrogen, each step in the transformation of H into Fe releases energy: a total of 8.8 MeV per nucleon, of which 7.0 MeV are already used up in the first step, the fusion of H into He.

6.2 Thermonuclear reaction rates

Consider again a reaction of the type $X(a,b)Y$. Let us first suppose that particles X are bombarded by particles a with a particular velocity υ. The rate at which they react then depends on the *cross-section*, i.e. the effective surface area of the particle X for interacting with particle a. The cross-section is defined as

$$\sigma = \frac{\text{number of reactions } X(a,b)Y \text{ per second}}{\text{flux of incident particles } a},$$

which indeed has a unit of area (cm^2). We denote the reacting particles X and a by indices i and j and their number densities as n_i and n_j, respectively. The incident flux of particles a is then $n_j\,\upsilon$, so that the number of reactions with a certain particle X taking place per second is $n_j\,\upsilon\,\sigma$. The number of reactions per second in a unit volume is therefore

$$\tilde{r}_{ij} = n_i n_j\,\upsilon\,\sigma,$$

which defines the reaction rate at a particular relative velocity υ. This expression applies if X and a are of a different kind. If the reacting particles are identical, then the number of possible reacting

pairs is not $n_i n_j$ but $\frac{1}{2}n_i(n_i-1) \approx \frac{1}{2}n_i^2$ for large particle numbers. Thus we can write more generally

$$\tilde{r}_{ij} = \frac{1}{1+\delta_{ij}} n_i n_j \, \upsilon \, \sigma, \tag{6.7}$$

since $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$.

In general, $\sigma = \sigma(\upsilon)$ depends on the relative velocity. In a stellar gas there is a distribution of velocities $\phi(\upsilon)$, normalized such that $\int_o^\infty \phi(\upsilon)\,\mathrm{d}\upsilon = 1$. The overall reaction rate, i.e the number of reactions taking place per second and per unit volume, is therefore

$$r_{ij} = \frac{1}{1+\delta_{ij}} \, n_i n_j \int_0^\infty \phi(\upsilon)\,\sigma(\upsilon)\upsilon \,\mathrm{d}\upsilon = \frac{1}{1+\delta_{ij}} \, n_i n_j \langle \sigma\upsilon \rangle. \tag{6.8}$$

In an ideal gas in LTE, the particle velocities are given by the Maxwell-Boltzmann distribution, eq. (3.13). If each particle velocity distribution is Maxwellian, then so is their *relative* velocity distribution,

$$\phi(\upsilon) = 4\pi\upsilon^2 \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{m\upsilon^2}{2kT}\right), \tag{6.9}$$

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

$$m = \frac{m_i m_j}{m_i + m_j}. \tag{6.10}$$

We replace the relative velocity υ by the kinetic energy in the centre-of-mass frame, $E = \frac{1}{2}m\upsilon^2$. Using the fact that $\phi(\upsilon)\,\mathrm{d}\upsilon = \phi(E)\,\mathrm{d}E$, we can write the average over $\sigma\upsilon$ in eq. (6.8) as

$$\langle \sigma\upsilon \rangle = \left(\frac{8}{\pi m}\right)^{1/2} (kT)^{-3/2} \int_0^\infty \sigma(E)\, E \, \exp\left(-\frac{E}{kT}\right) \mathrm{d}E. \tag{6.11}$$

This depends only on temperature, i.e. the dependence on velocity in eq. (6.7) turns into a dependence on the *temperature* in the overall reaction rate. The temperature dependence of a nuclear reaction is thus expressed by the factor $\langle \sigma\upsilon \rangle$. To understand this temperature dependence, we must consider in more detail the reaction cross sections and their dependence on energy.

6.2.1 Nuclear cross-sections

The cross-section σ appearing in the reaction rate equation (6.8) is a measure of the probability that a nuclear reaction occurs, given the number densities of the reacting nuclei. While the energy gain from a reaction can be simply calculated from the mass deficits of the nuclei, the cross-section is much more difficult to obtain. Classically, the geometrical cross-section for a reaction between nuclei i and j with radii R_i and R_j is $\sigma = \pi(R_i + R_j)^2$. A good approximation to the nuclear 'radius', or rather for the range of the nuclear force, is

$$R_i \approx R_0 A_i^{1/3} \qquad \text{with} \qquad R_0 = 1.44 \times 10^{-13}\,\mathrm{cm}. \tag{6.12}$$

This would yield typical cross-sections of the order of 10^{-25}–$10^{-24}\,\mathrm{cm}^2$. On the other hand, quantum-mechanically the particles 'see' each other as smeared out over a length equal to the de Broglie wavelength associated with their relative momentum p,

$$\lambda = \frac{\hbar}{p} = \frac{\hbar}{(2mE)^{1/2}}, \tag{6.13}$$

with m and E the reduced mass and relative kinetic energy as defined before. The last equality assumes non-relativistic particles. A better estimate of the geometrical cross-section is therefore $\sigma = \pi\lambda^2$. At typical conditions in the stellar gas, this is (much) larger than the classical estimate since $\lambda > R_i + R_j$. The real situation is much more complicated owing to a number of effects:

- Charged nuclei experience a repulsive Coulomb force which, although weaker than the strong nuclear force, has a much longer range. This Coulomb barrier would prevent any reaction to occur under stellar condition, were it not for the quantum-mechanical *tunnel effect*.
- The nature of the force involved in the reaction determines the strength of the interaction. For a reaction $X(a,b)Y$, the emitted particle may be either another nucleus, a γ-photon, or an $e^-\overline{\nu}$ or $e^+\nu$ pair. In the first case, only the strong force is involved and the cross-section may be close to the geometrical one. The second case also involves the electromagnetic force, which is weaker and gives a lower reaction probability, i.e. a smaller cross-section. In the last case, a weak interaction must occur which has an even lower probability and smaller cross-section.
- Nuclear structure effects can have a strong influence on the cross-section. This is particularly true in the case of *resonant interactions*.

Coulomb barrier and the tunnel effect

At distances r larger than the range of the nuclear force, two nuclei with charges Z_i and Z_j experience a repulsive Coulomb potential

$$V(r) = \frac{Z_i Z_j e^2}{r} = 1.44 \frac{Z_i Z_j}{r\,[\mathrm{fm}]}\,\mathrm{MeV}, \tag{6.14}$$

with r expressed in fm = 10^{-13} cm in the last equality. To experience the attractive nuclear force the particles have to approach each other within a typical distance $r_\mathrm{n} \sim A^{1/3} R_0$ as given by eq. (6.12). For $r < r_\mathrm{n}$ the nuclear attraction gives a potential drop to roughly $V_0 \approx -30$ MeV. The particles must therefore overcome a typical Coulomb barrier $E_\mathrm{C} = V(r_\mathrm{n}) \approx Z_1 Z_2$ MeV, see Fig. 6.2.

If an incoming particle has a kinetic energy E at infinity in the reference frame of the nucleus, it can classically only come within a distance r_c given by $E = V(r_\mathrm{c})$. In stellar interiors the kinetic energies of nuclei have a Maxwellian distribution, with an average value $\langle E \rangle = \frac{3}{2}kT \approx 1.3$ keV at 10^7 K, which is typical of the centre of the Sun and other main-sequence stars. This falls short of the Coulomb barrier by a factor of about 1000. Even considering the high-energy tail of the Maxwell-Boltzmann distribution, the fraction of particles with $E > E_\mathrm{C}$ is vanishingly small. With purely classical considerations nuclear reactions have no chance of happening at such temperatures.

We need to turn to quantum mechanics to see how nuclear reactions are possible at stellar temperatures. As was discovered by G. Gamow, there is a finite probability that the projectile penetrates the repulsive Coulomb barrier even if $E \ll E_\mathrm{C}$. The tunnelling probability can be estimated as

$$P \sim \exp\left(- \int_{r_\mathrm{n}}^{r_\mathrm{c}} \frac{\sqrt{2m[V(r) - E]}}{\hbar}\,\mathrm{d}r \right)$$

where

$$r_\mathrm{c} = \frac{Z_i Z_j e^2}{E}$$

is the classical distance of closest approach. The result is

$$P = P_0 \exp(-b\,E^{-1/2}) \qquad \text{with} \qquad b = 2\pi \frac{Z_i Z_j e^2}{\hbar} \left(\frac{m}{2}\right)^{1/2} = 31.29\, Z_i Z_j A^{1/2}\,[\mathrm{keV}]^{1/2}. \tag{6.15}$$

Here $A = A_i A_j/(A_i + A_j)$ is the reduced mass in units of m_u and P_0 is a constant. P increases steeply with E and decreases with $Z_i Z_j$, i.e., with the height of the Coulomb barrier. Therefore, at relative low temperatures only the lightest nuclei (with the smallest $Z_i Z_j$) have a non-negligible chance to react. Reactions with heavier nuclei, with larger $Z_i Z_j$, require larger energies and therefore higher temperatures to have a comparable penetration probability.

Nuclear structure effects on the cross-section

A typical thermonuclear reaction proceeds as follows. After penetrating the Coulomb barrier, the two nuclei can from an unstable, excited *compound nucleus* which after a short time decays into the product particles, e.g.

$$X + a \rightarrow C^* \rightarrow Y + b\,.$$

Although the lifetime of the compound nucleus C^* is very short, it is much longer than the crossing time of the nucleus at the speed of light ($\sim 10^{-21}$ s). Therefore by the time it decays, the compound nucleus has no 'memory' of how it was formed, and the decay depends only on the energy.

The decay can proceed via different channels, e.g. $C^* \rightarrow X + a$, $\rightarrow Y_1 + b_1$, $\rightarrow Y_2 + b_2$, $\ldots$, $\rightarrow C + \gamma$. In the first case the original particles are reproduced, the last case is a decay to a stable energy level of C with γ-emission. In the other cases the particles b_1, b_2, etc. may be protons, neutrons or α-particles. (Reactions involving electron and neutrino emission do not proceed via a compound intermediate state, since the necessary β-decays would be prohibitively slow.) In order for a certain energy level of C^* to decay via a certain channel, the conservations laws of energy, momentum, angular momentum and nuclear symmetries must be fulfilled. The outgoing particles obtain a certain kinetic energy, which – with the exception of neutrinos that escape without interaction – is quickly thermalised, i.e. shared among the other gas particles owing to the short photon and particle mean free paths in the stellar gas.

The energy levels of the compound nucleus play a crucial role in determining the reaction cross-section, see Fig. 6.2. Let $E_{\rm min}$ be the minimum energy required to remove a nucleon from the ground state of C to infinity, analogous to the ionization energy of an atom. Energy levels below $E_{\rm min}$ correspond to bound states in an atom; these can only decay by γ-emission which is relatively improbable. These 'stationary' energy levels have long lifetimes τ and correspondingly small widths Γ, since according to Heisenberg's uncertainty relation

$$\Gamma = \frac{\hbar}{\tau}. \tag{6.16}$$

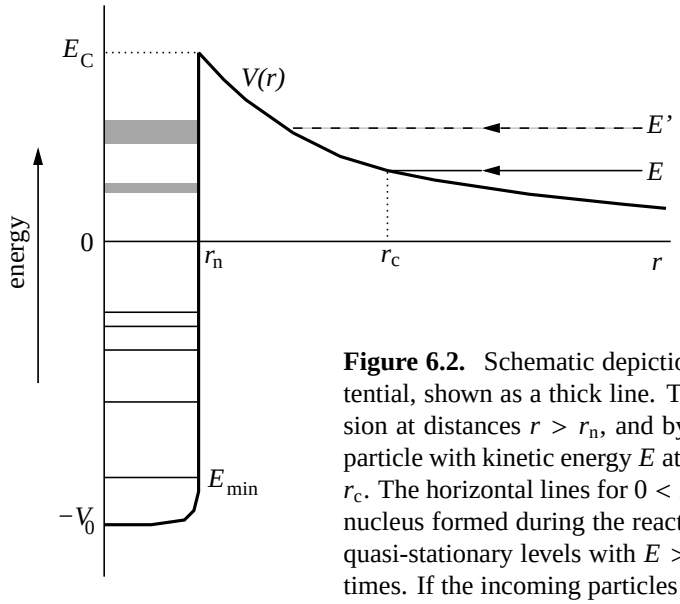


Figure 6.2. Schematic depiction of the combined nuclear and Coulomb potential, shown as a thick line. The potential is dominated by Coulomb repulsion at distances $r > r_{\rm n}$, and by nuclear attraction for $r < r_{\rm n}$. An incoming particle with kinetic energy E at infinity can classically approach to a distance $r_{\rm c}$. The horizontal lines for $0 < r < r_{\rm n}$ indicate energy levels in the compound nucleus formed during the reaction. The ground state is at energy $-E_{\rm min}$; the quasi-stationary levels with $E > 0$ are broadened due to their very short lifetimes. If the incoming particles have energy E' corresponding to such a level they can find a resonance in the compound nucleus (see text).

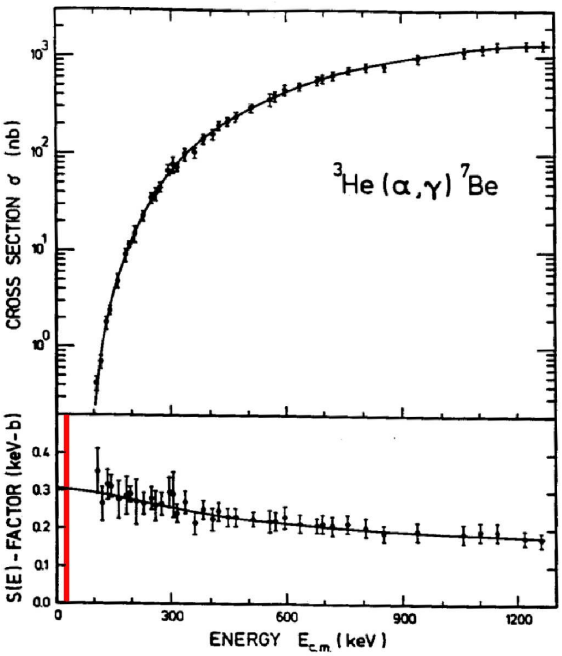


Figure 6.3. Example of the dependence of the reaction cross section on energy E for the ^{3}He + ^{4}He $\rightarrow$ ^{7}Be + γ reaction. Although σ varies very strongly with energy, and becomes immeasurably small at very low E, the factor $S(E)$ is only a very weak function of E and – at least for this reaction – can be safely extrapolated to the low energies that are relevant for nuclear reactions in stars (15–30 keV, vertical bar on the left).

Energy levels above $E_{\rm min}$ can also expel particles, which is much more probable than γ-emission. These levels also have finite lifetimes because of the sharp potential rise beyond $r_{\rm n}$, but eventually the particles can escape by the tunnel effect. These 'quasi-stationary' levels have much shorter lifetimes and correspondingly larger widths. The probability of escape increases with energy and so does the level width, until eventually Γ is larger than the distance between levels resulting in a continuum of energy states above a certain $E_{\rm max}$.

The possible existence of discrete energy levels above $E_{\rm min}$ can give rise to so-called 'resonances' with much increased reaction probabilities. Suppose we let X and a react with gradually increasing relative energy E (measured at large distance). As long as E is in a region without or in between quasi-stationary levels, the reaction probability will simply increase with the penetration probability (6.15). However, if E coincides with such a level (e.g. energy E' in Fig. 6.2), then the reaction probability can be enhanced by several orders of magnitude. For energies close to such a level $E_{\rm res}$ the cross-section has an energy dependence with a typical resonance form,

$$\xi(E) \propto \frac{1}{(E - E_{\rm res})^2 + (\Gamma/2)^2}. \tag{6.17}$$

At $E = E_{\rm res}$ the cross-section can be close to the geometrical cross-section, $\pi\lambda^2$, where λ is the de Broglie wavelength (6.13). We can thus expect the cross-section to depend on energy as

$$\sigma(E) \propto \pi\lambda^2\, P(E)\, \xi(E). \tag{6.18}$$

The astrophysical cross-section factor

Since $\lambda^2 \propto 1/E$ and $P(E) \propto \exp(-b\,E^{-1/2})$, one usually writes

$$\boxed{\sigma(E) = S(E)\,\frac{\exp(-b\,E^{-1/2})}{E}}\,. \tag{6.19}$$

This equation defines the 'astrophysical S-factor' $S(E)$, which contains all remaining effects, i.e. the intrinsic nuclear properties of the reaction including possible resonances.

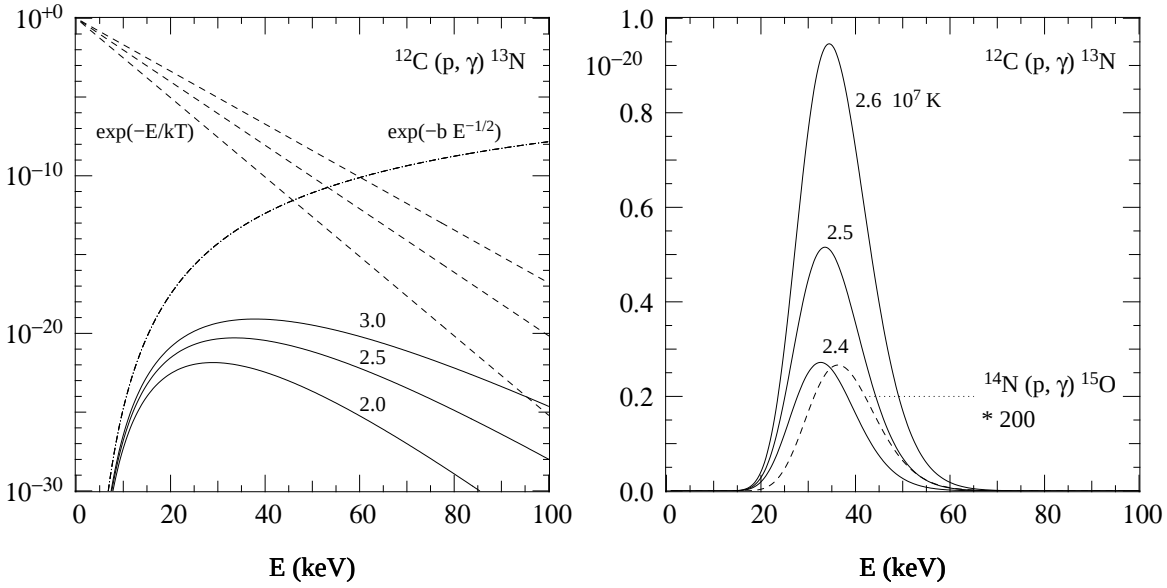


Figure 6.4. Example of the Gamow peak for the ^{12}C(p, γ)^{13}N reaction. The left panel shows as dash-dotted lines the tunnelling probability factor, $\exp(-b/E^{1/2})$, and as dashed lines the tail of the Maxwell distribution, $\exp(-E/kT)$, for three values of temperature: $T = 2.0{\times}10^7$ K (lower curve), $2.5{\times}10^7$ K (middle) and $3.0{\times}10^7$ K (upper). The solid lines show the product of these two factors, $f(E)$ as in eq. (6.21), labelled by $T_7 \equiv T/10^7$ K. Note the enormous range of the vertical log-scale. To appreciate the sharpness of the Gamow peak, and the enormous sensitivity to temperature, the right panel shows $f(E)$ on a linear scale for $T_7 = 2.4$, 2.5 and 2.6. The dashed line is the Gamow peak for the ^{14}N(p, γ)^{15}O reaction for $T_7 = 2.4$, multiplied by a factor 200.

The S-factor can in principle be calculated, but in practice one relies on laboratory measurements of the cross-section to obtain $S(E)$. The difficulty is that such measurements are only feasible at large E, typically > 0.1 MeV, because cross-sections quickly become unmeasurably small at lower energies. This lowest energy is still an order of magnitude larger than the energies at which reactions typically take place under stellar conditions. One therefore has to extrapolate $S(E)$ down over quite a large range of E to the relevant energies. In many cases $S(E)$ is nearly constant or varies slowly with E – unlike $\sigma(E)$! – and this procedure can be quite reliable (e.g. see Fig. 6.3). However, when resonances occur in the range of energies over which to extrapolate, the results can be very uncertain.

6.2.2 Temperature dependence of reaction rates

Combining eqs. (6.11) and (6.19), the cross-section factor $\langle\sigma\upsilon\rangle$ can be written as

$$\langle\sigma\upsilon\rangle \;=\; (8/\pi m)^{1/2}(kT)^{-3/2} \int_0^\infty S(E)\,\exp\!\left(-\frac{E}{kT}-\frac{b}{E^{1/2}}\right)\mathrm{d}E. \tag{6.20}$$

We will look at the case of *non-resonant* reactions, where we can assume that $S(E)$ varies slowly with E. The integrand is then dominated by the product of two exponential factors: $\exp(-E/kT)$, the tail of the Maxwell-Boltzmann distribution which decreases rapidly with E; and $\exp(-bE^{-1/2})$, the penetration probability due to the tunnel effect which increases rapidly with E. The product of these two exponentials,

$$f(E) = \exp\!\left(-\frac{E}{kT}-\frac{b}{E^{1/2}}\right), \tag{6.21}$$

is a sharply peaked function called the *Gamow peak*, which has appreciable values only around a maximum at energy E_0. Fig. 6.4 shows an example for the reaction $^{12}\mathrm{C} + \mathrm{p} \rightarrow {}^{13}\mathrm{N} + \gamma$. Since by assumption $S(E)$ varies slowly with E, we can take $S(E) \approx S(E_0)$ out of the integral (6.20) and obtain

$$\langle \sigma \upsilon \rangle \approx (8/\pi m)^{1/2} (kT)^{-3/2} S(E_0) \int_0^\infty \exp\left(-\frac{E}{kT} - \frac{b}{E^{1/2}} \right) \mathrm{d}E. \tag{6.22}$$

The reaction rate then only depends on the integral $\int_0^\infty f(E)\,\mathrm{d}E$.

Properties of the Gamow peak

The value of the Gamow peak energy E_0 can be found by taking $\mathrm{d}f/\mathrm{d}E = 0$, which gives

$$E_0 = (\tfrac{1}{2} bkT)^{2/3} = 5.665\,(Z_i^2 Z_j^2 A\, T_7^2)^{1/3}\,\mathrm{keV}. \tag{6.23}$$

To obtain the last equality we have substituted b as given by eq. (6.15) and we use the notation $T_n = T/(10^n\,\mathrm{K})$, while A is the reduced mass in m_u as before. For reactions between light nuclei at temperatures $T \sim 1\text{–}2 \times 10^7\,\mathrm{K}$, $E_0 \sim 10\text{–}30\,\mathrm{keV}$, while the average kinetic energies are 1–2 keV. The peak is quite narrow, having a width ΔE at half maximum that is always smaller than E_0. Thus, the nuclei that contribute to the reaction rate have energies in a narrow interval around 10 times the thermal energy, but about 2 orders of magnitude below the Coulomb barrier.

The right panel of Fig. 6.4 illustrates the strong dependence of the maximum value $f(E_0)$ of the Gamow peak on the temperature. In the case of the $^{12}\mathrm{C}(\mathrm{p},\gamma)^{13}\mathrm{N}$ reaction, an increase in temperature by 4% (from $T_7 = 2.4$ to 2.5, or from 2.5 to 2.6) almost doubles the maximum value of $f(E)$. The width of the peak also increases modestly, such that the area under the curve – which is the integral that appears in eq. (6.22) – increases enormously with increasing temperature. This is the reason why thermonuclear reaction rates are extremely sensitive to the temperature.

When we compare different reactions, the factor $b \propto Z_1 Z_2 A^{1/2}$ changes and thereby the penetration probability at a certain energy. A reaction between heavier nuclei (with larger A and Z) will therefore have a much lower rate at certain fixed temperature. This is illustrated in the right panel of Fig. 6.4 by the dashed curve, showing the Gamow peak for the $^{14}\mathrm{N}(\mathrm{p},\gamma)^{15}\mathrm{O}$ reaction at $T_7 = 2.4$, multiplied by a factor 200. Hence, the probability of this reaction is 200 times smaller than that of the $^{12}\mathrm{C}(\mathrm{p},\gamma)^{13}\mathrm{N}$ reaction at the same temperature. In other words, reactions between heavier nuclei will need a higher temperature to occur at an appreciable rate.

To summarize, the properties of the Gamow peak imply that

- the reaction rate $\langle \sigma \upsilon \rangle$ increases *very strongly with temperature*.
- $\langle \sigma \upsilon \rangle$ decreases strongly with increasing Coulomb barrier.

Analytic expressions for the temperature dependence

We can find an analytical expression for the reaction rate if we approximate the integrand $f(E)$ in eq. (6.22) by a Gaussian centred at E_0, i.e.,

$$f(E) \approx f(E_0) \exp\left[-\left(\frac{E - E_0}{\Delta E} \right)^2 \right], \tag{6.24}$$

Considering the shapes of the curves in Fig. 6.4, this is not a bad approximation. From eq. (6.21) we find $f(E_0) = \exp(-3E_0/kT) \equiv \exp(-\tau)$, which defines the often used quantity τ,

$$\tau = \frac{3E_0}{kT} = 19.72 \left(\frac{Z_i^2 Z_j^2 A}{T_7} \right)^{1/3}. \tag{6.25}$$

The width ΔE of the Gaussian can be obtained by expanding eq. (6.21) for $f(E)$ in a Taylor series around E_0,

$$f(E) = f(E_0) + f'(E_0)\,(E - E_0) + \tfrac{1}{2} f''(E_0)\,(E - E_0)^2 + \ldots,$$

in which the second term equals zero because $f'(E_0) = 0$. Comparing this with a similar expansion of the Gaussian approximation to $f(E)$ yields the same expression, to second order, if

$$\Delta E = \left(\frac{-2f}{f''}\right)^{1/2}_{E=E_0} = \left(\frac{4E_0 kT}{3}\right)^{1/2} \tag{6.26}$$

We can then approximate the integral in eq. (6.22) by

$$\int_0^\infty f(E)\,\mathrm{d}E \;\approx\; e^{-\tau} \int_0^\infty \exp\left[-\left(\frac{E - E_0}{\Delta E}\right)^2\right] \mathrm{d}E \;\approx\; e^{-\tau}\,\sqrt{\pi}\,\Delta E. \tag{6.27}$$

In the last step we have extended the integral from $-\infty$ to ∞ to obtain the result $\sqrt{\pi}\,\Delta E$, which introduces only a very small error because the exponential is negligibly small for $E < 0$. When we substitute (6.27) with the expression (6.26) for ΔE into (6.22), and we eliminate E_0 and kT in favour of τ and b using (6.23) and (6.25), then we find after some manipulation

$$\langle \sigma v \rangle \;\approx\; \frac{8}{9}\left(\frac{2}{3m}\right)^{1/2} \frac{S(E_0)}{b}\,\tau^2\,e^{-\tau} \;=\; \frac{7.21 \times 10^5}{Z_i Z_j A}\left(\frac{S(E_0)}{\mathrm{keV\,cm^2}}\right)\tau^2\,e^{-\tau}. \tag{6.28}$$

In the last equality we have substituted the explicit expression (6.15) for b. Since $\tau \propto T^{-1/3}$ this gives a temperature dependence of the form

$$\langle \sigma v \rangle \propto \frac{1}{T^{2/3}}\,\exp\left(-\frac{C}{T^{1/3}}\right), \tag{6.29}$$

where the constant C in the exponential factor depends on $Z_i Z_j$, i.e. on the height of the Coulomb barrier. This is indeed a strongly increasing function of temperature.

If we consider a small range of temperatures around some value T_0, we can write

$$\langle \sigma v \rangle = \langle \sigma v \rangle_0 \left(\frac{T}{T_0}\right)^\nu \qquad \text{with} \qquad \nu \equiv \frac{\partial \log\langle \sigma v \rangle}{\partial \log T} = \frac{\tau - 2}{3}. \tag{6.30}$$

The last equality follows from (6.28) and (6.25). Therefore the exponent ν is not a constant but depends on T itself – in fact ν decreases with T roughly as $T^{-1/3}$. In general, however, any particular reaction is only important in quite a limited range of temperatures, so that taking ν as constant in (6.30) is approximately correct. Values of the exponent ν are in all cases $\gg 1$. For example, at $T_7 = 1.5$ we find $\langle \sigma v \rangle \propto T^{3.9}$ for the p + p reaction for hydrogen fusion and $\langle \sigma v \rangle \propto T^{20}$ for the $^{14}\mathrm{N(p,\gamma)}$ reaction in the CNO cycle (see Sec. 6.4.1). Thus thermonuclear reaction rates are about the most strongly varying functions found in physics. This temperature sensitivity has a strong influence on stellar models, as we shall see.

6.2.3 Electron screening

We found that the repulsive Coulomb force between nuclei plays a crucial role in determining the rate of a thermonuclear reaction. In our derivation of the cross section we have ignored the influence of the surrounding free electrons, which provide overall charge neutrality in the gas. In a dense medium, the attractive Coulomb interactions between atomic nuclei and free electrons cause each nucleus to

be effectively surrounded by a cloud of electrons. This electron cloud reduces the Coulomb repulsion between the nuclei at large distances, and may thus increase the probability of tunneling through the Coulomb barrier. This effect is know as *electron screening* or *electron shielding*.

We simply give the main results, the derivation of which can be found in MAEDER Sec. 9.4 or KIPPENHAHN Sec. 18.4. The repulsive Coulomb potential (eq. 6.14) is reduced by a factor $\exp(-r/r_{\rm D})$, where $r_{\rm D}$, the so-called Debye-Hückel radius, represents the effective radius of the electron cloud. The stronger the Coulomb interactions between nuclei and electrons, the smaller $r_{\rm D}$. We have found (Sec. 3.6.1) that Coulomb interactions increase in strength with increasing density and decreasing temperature, and so does the magnitude of the electron screening effect. It turns out that the reaction rate $\langle\sigma\upsilon\rangle$ is enhanced by a factor

$$f = \exp\left(\frac{E_{\rm D}}{kT}\right), \tag{6.31}$$

where, for small values of $E_{\rm D}/kT \ll 1$,

$$\frac{E_{\rm D}}{kT} = \frac{Z_1 Z_2 e^2}{r_{\rm D} kT} \sim 0.006\, Z_1 Z_2\, \frac{\rho^{1/2}}{{T_7}^{3/2}}. \tag{6.32}$$

This is the *weak screening* approximation, which applies to relatively low densities and high temperatures such as found in the centre of the Sun and other main-sequence stars. Under these conditions, reaction rates are enhanced only by modest factors, $f \lesssim 1.1$.

The description of electron screening becomes complicated at high densities and relatively low temperatures, where the weak screening approximation is no longer valid. A general result is that with increasing strength of electron screening, the temperature sensitivity of the reaction rate diminishes and the density dependence becomes stronger. At very high densities, $\rho \gtrsim 10^6$ g/cm^3, the screening effect is so large that it becomes the dominating factor in the reaction rate. The shielding of the Coulomb barrier can be so effective that the reaction rate depends mainly on the density and no longer on temperature. Reactions between charged nuclei become possible even at low temperature, if the density exceeds a certain threshold. One then speaks of *pycnonuclear reactions*, which can play an important role in late stages of stellar evolution. In a very cool and dense medium one must also take into account the effect of crystallization, which decreases the mobility of the nuclei and thus the probability of collisions.

6.3 Energy generation rates and composition changes

Having obtained an expression for the cross-section factor $\langle\sigma\upsilon\rangle$, the reaction rate r_{ij} follows from eq. (6.8). We can then easily obtain the energy generation rate. Each reaction releases an amount of energy Q_{ij} according to eq. (6.4), so that $Q_{ij}\, r_{ij}$ is the energy generated per unit volume and per second. The energy generation rate per *unit mass* from the reaction between nuclei of type i and j is then

$$\epsilon_{ij} = \frac{Q_{ij}\, r_{ij}}{\rho}. \tag{6.33}$$

We can express the energy generation rate in terms of the mass fractions X_i and X_j and the density ρ using eq. (6.8). Replacing the number density n_i by the mass fraction X_i according to $n_i = X_i\,\rho/(A_i m_u)$, eq. (6.33) can be written as

$$\epsilon_{ij} = \frac{Q_{ij}}{(1+\delta_{ij})\, A_i A_j m_{\rm u}^2}\, \rho\, X_i X_j\, \langle\sigma\upsilon\rangle_{ij} = \frac{q_{ij}}{(1+\delta_{ij})\, A m_{\rm u}}\, \rho\, X_i X_j\, \langle\sigma\upsilon\rangle_{ij}, \tag{6.34}$$

In the last identity $A = A_i A_j/(A_i + A_j)$ is the reduced mass in units of m_u, and we have replaced Q_{ij} by the energy released per unit mass by this reaction

$$q_{ij} = \frac{Q_{ij}}{m_i + m_j} \approx \frac{Q_{ij}}{(A_i + A_j)m_\mathrm{u}}. \tag{6.35}$$

Remember that $\langle\sigma\upsilon\rangle$ contains the temperature dependence of the reaction rate. If we use the power-law approximation (6.30) we can write the energy generation rate of a reaction as

$$\epsilon_{ij} = \epsilon_{0,ij}\, X_i X_j\, \rho\, T^\nu. \tag{6.36}$$

The total nuclear energy generation rate results from all reactions taking place in a certain mass element in the star, i.e.

$$\epsilon_\mathrm{nuc} = \sum_{i,j} \epsilon_{ij}. \tag{6.37}$$

This is the quantity ϵ_nuc that appears in the stellar structure equation for the luminosity, eq. (5.4).

Composition changes

The reaction rates also determine the rate at which the composition changes. The rate of change in the number density n_i of nuclei of type i owing to reactions with nuclei of type j is

$$\left(\frac{\mathrm{d}n_i}{\mathrm{d}t}\right)_j = -(1+\delta_{ij})\, r_{ij} = -n_i n_j\, \langle\sigma\upsilon\rangle_{ij}. \tag{6.38}$$

The factor $1+\delta_{ij}$ takes into account that a reaction between identical nuclei consumes *two* such nuclei. One can define the *nuclear lifetime* of a species i owing to reactions with j as

$$\tau_{i,j} = \frac{n_i}{|(\mathrm{d}n_i/\mathrm{d}t)_j|} = \frac{1}{n_j\, \langle\sigma\upsilon\rangle_{ij}}, \tag{6.39}$$

which is the timescale on which the abundance of i changes as a result of this reaction.

The overall change in the number n_i of nuclei of type i in a unit volume can generally be the result of different nuclear reactions. Some reactions (with rate r_{ij} as defined above) consume i while other reactions, e.g. between nuclei k and l, may produce i. If we denote the rate of reactions of the latter type as $r_{kl,i}$, we can write for the total rate of change of n_i:

$$\frac{\mathrm{d}n_i}{\mathrm{d}t} = -\sum_j (1+\delta_{ij})\, r_{ij} + \sum_{k,l} r_{kl,i} \tag{6.40}$$

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

$$\frac{\mathrm{d}X_i}{\mathrm{d}t} = A_i \frac{m_\mathrm{u}}{\rho}\left(-\sum_j (1+\delta_{ij})\, r_{ij} + \sum_{k,l} r_{kl,i}\right) \tag{6.41}$$

For each nuclear species i one can write such an equation, describing the composition change at a particular mass shell inside the star (with density ρ and temperature T) resulting from nuclear reactions. In the presence of internal mixing (in particular of *convection*, Sec. 5.5.3) the redistribution of composition between different mass shells should also be taken into account.

Note the similarity between the expressions for the nuclear energy generation rate (6.37) and the equation for composition changes (6.41), both of which are proportional to r_{ij}. Using eq. (6.35) for the energy released per gram, we can write the reaction rate as

$$r_{ij} = \frac{\epsilon_{ij}}{q_{ij}(A_i + A_j)} \frac{\rho}{m_\mathrm{u}}. \tag{6.42}$$

If we substitute this expression into eq. (6.41) the factor ρ/m_u drops out. We obtain a useful expression in simple cases where only *one* reaction occurs, or a reaction chain in which one reaction determines the overall rate. An example is the fusion of $4\,^1\mathrm{H}$ into $^4\mathrm{He}$, which is the net result of a chain of reactions (see Sec. 6.4.1). In that case you may verify that (6.41) and (6.42) reduce to

$$\frac{\mathrm{d}Y}{\mathrm{d}t} = -\frac{\mathrm{d}X}{\mathrm{d}t} = \frac{\epsilon_\mathrm{H}}{q_\mathrm{H}}, \tag{6.43}$$

where ϵ_H is the energy generation rate by the complete chain of H-burning reactions, and q_H is amount of energy produced by converting 1 gram of $^1\mathrm{H}$ into $^4\mathrm{He}$.

6.4 The main nuclear burning cycles

In principle, many different nuclear reactions can occur simultaneously in a stellar interior. If one is interested in following the detailed isotopic abundances produced by all these reactions, or if structural changes occur on a very short timescale, a large network of reactions has to be calculated (as implied by eq. 6.41). However, for the calculation of the structure and evolution of a star usually a much simpler procedure is sufficient, for the following reasons:

- The very strong dependence of nuclear reaction rates on the temperature, combined with the sensitivity to the Coulomb barrier $Z_1 Z_2$, implies that nuclear fusions of different possible fuels – hydrogen, helium, carbon, etc. – are well separated by substantial temperature differences. The evolution of a star therefore proceeds through several distinct *nuclear burning cycles*.
- For each nuclear burning cycle, only a handful of reactions contribute significantly to energy production and/or cause major changes to the overall composition.
- In a chain of subsequent reactions, often one reaction is by far the slowest and determines the rate of the whole chain. Then only the rate of this bottleneck reaction needs to be taken into account.

6.4.1 Hydrogen burning

The net result of hydrogen burning is the fusion of four $^1\mathrm{H}$ nuclei into a $^4\mathrm{He}$ nucleus,

$$4\,^1\mathrm{H} \rightarrow {}^4\mathrm{He} + 2\,\mathrm{e}^+ + 2\,\nu\,. \tag{6.44}$$

You may verify using Sec. 6.1.1 that the total energy release is 26.734 MeV. In order to create a $^4\mathrm{He}$ nucleus two protons have to be converted into neutrons. Therefore two neutrinos are released by weak interactions ($\mathrm{p} \rightarrow \mathrm{n} + \mathrm{e}^+ + \nu$), which escape without interacting with the stellar matter. It is customary not to include the neutrino energies into the overall energy release Q, but to take into account only the energy that is used to heat the stellar gas. This includes energy released in the form of γ-rays (including the γ-rays resulting from pair annihilation after e^+ emission) and in the form of kinetic energies of the resulting nuclei. The effective Q-value of hydrogen burning is therefore somewhat smaller than 26.734 MeV and depends on the reaction in which the neutrinos are emitted.

Since a simultaneous reaction between four protons is extremely unlikely, a chain of reactions is always necessary for hydrogen burning. This can take place in two distinct ways: either direct fusion of protons via the *p-p chain*, or by using already present CNO-nuclei as catalysts in the *CNO cycle*. Hydrogen burning in stars takes place at temperatures ranging between 8×10^6 K and 5.0×10^7 K, depending on stellar mass and evolution stage.

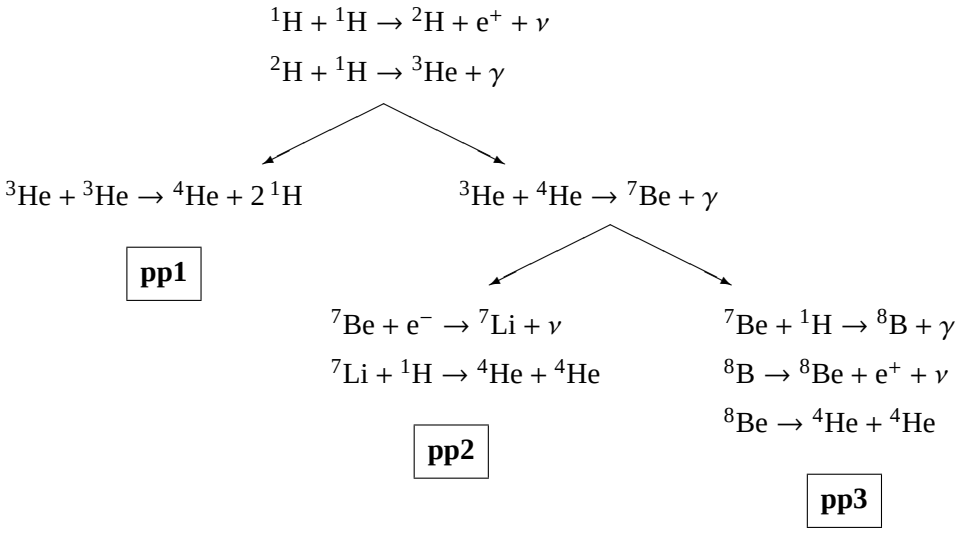
The p-p chains

The first reaction is the so-called p-p reaction:

$$^1\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^2\mathrm{H} + \mathrm{e}^+ + \nu \qquad \text{or} \qquad \mathrm{p} + \mathrm{p} \rightarrow \mathrm{D} + \mathrm{e}^+ + \nu\,. \tag{6.45}$$

This involves the simultaneous β-decay of one of the protons during the strong nuclear interaction. This is very unlikely and the p-p reaction therefore has an extremely small cross-section, about 10^{-20} times that of a typical reaction involving only strong interactions. The reaction rate cannot be measured in the laboratory and is only known from theory.

After some deuterium is produced, it rapidly reacts with another proton to from ^{3}He. Subsequently three different branches are possible to complete the chain towards ^{4}He:



(6.46)

The pp1 branch requires two ^{3}He nuclei, so the first two reactions in the chain have to take place twice. The alternative pp2 and pp3 branches require only one ^{3}He nucleus and an already existing ^{4}He nucleus (either present primordially, or produced previously by hydrogen burning). The resulting ^{7}Be nucleus can either capture an electron or fuse with another proton, giving rise to the second branching into pp2 and pp3. Three of the reactions in the chains are accompanied by neutrino emission, and the (average) neutrino energy is different in each case: $\langle E_\nu \rangle = 0.265$ MeV for the p-p reaction, 0.814 MeV for electron capture of ^{7}Be and 6.71 MeV for the β-decay of ^{8}B. Therefore the total energy release Q_H for the production of one ^{4}He nucleus is different for each chain: 26.20 MeV (pp1), 25.66 MeV (pp2) and only 19.76 MeV for pp3.

The relative frequency of the three chains depends on temperature and chemical composition. Because the ^{3}He + ^{4}He reaction is slightly more sensitive to temperature than the ^{3}He + ^{3}He reaction (it has a somewhat higher reduced mass and larger τ, eq. 6.25), the pp1 chain dominates over the other two at relatively low temperature ($T_7 \lesssim 1.5$). The pp1 chain is the main energy-producing reaction chain in the Sun. At increasing T, first the pp2 chain and then the pp3 chain become increasingly important.

At low temperatures ($T < 8 \times 10^6$ K) the rates of all reactions should be calculated separately to obtain the energy generation rate and the changes in abundances. In particular, the $^3\mathrm{He} + {}^3\mathrm{He}$ reaction is quite slow and a substantial abundance of ^{3}He can accumulate before further reactions occur. For $T \gtrsim 8 \times 10^6$ K all reactions in the chain are fast enough that they reach a steady state, where once a D nucleus is produced by the first, very slow reaction, all successive reactions proceed very quickly until ^{4}He is formed. The nuclear lifetimes (eq. 6.39) of the intermediate nuclei D, ^{3}He, ^{7}Li, etc, are very short compared to the overall nuclear timescale, and their abundances are very small. The overall rate of the whole reaction chain is then set by the rate of the bottleneck p-p reaction, r_{pp}. In this steady-state or 'equilibrium' situation the rate of each subsequent reaction adapts itself to the pp rate.[1] The energy generation rate (given by the sum of energies released by each reaction, eq. 6.37) can then be expressed in a single term of the form (6.33), i.e. $\epsilon_{\mathrm{nuc}} = Q_{\mathrm{H}} r_{\mathrm{pp}}/\rho$ where Q_{H} is the total energy released in the whole chain (6.44). The above expression applies to the pp2 and pp3 chains, which each require one p-p reaction to complete. For the pp1 chain two p-p reactions are needed and therefore in that case $\epsilon_{\mathrm{nuc}} = \frac{1}{2} Q_{\mathrm{H}} r_{\mathrm{pp}}/\rho$. Expressing r_{pp} in terms of the cross section factor $\langle \sigma \upsilon \rangle_{\mathrm{pp}}$ and the hydrogen abundance X, we can compute the energy generation rate for hydrogen burning by the combination of pp chains as

$$\epsilon_{\mathrm{pp}} = \psi \, q_{\mathrm{H}} X^2 \frac{\rho}{m_{\mathrm{u}}} \langle \sigma \upsilon \rangle_{\mathrm{pp}}, \tag{6.47}$$

where $q_{\mathrm{H}} = Q_{\mathrm{H}}/4m_{\mathrm{u}}$ is the total energy release per gram of hydrogen burning and ψ is a factor between 1 (for the pp1 chain) and 2 (for the pp2 and pp3 chains), depending on the relative frequency of the chains. Both ψ and q_{H} therefore depend on the temperature, because the three chains have different neutrino losses. The overall temperature dependence of ϵ_{pp} is dominated by the T-dependence of $\langle \sigma \upsilon \rangle_{\mathrm{pp}}$ and is shown in Fig. 6.5. The pp chain is the least temperature-sensitive of all nuclear burning cycles with a power-law exponent ν (eq. 6.30) varying between about 6 at $T_6 \approx 5$ and 3.5 at $T_6 \approx 20$.

The CNO cycle

If some C, N, and O is already present in the gas out of which a star forms, and if the temperature is sufficiently high, hydrogen fusion can take place via the so-called *CNO cycle*. This is a cyclical sequence of reactions that typically starts with a proton capture by a ^{12}C nucleus, as follows:

$$\begin{aligned}
{}^{12}\mathrm{C} + {}^{1}\mathrm{H} &\rightarrow {}^{13}\mathrm{N} + \gamma \\
{}^{13}\mathrm{N} &\rightarrow {}^{13}\mathrm{C} + \mathrm{e}^+ + \nu \\
{}^{13}\mathrm{C} + {}^{1}\mathrm{H} &\rightarrow {}^{14}\mathrm{N} + \gamma \\
{}^{14}\mathrm{N} + {}^{1}\mathrm{H} &\rightarrow {}^{15}\mathrm{O} + \gamma \\
{}^{15}\mathrm{O} &\rightarrow {}^{15}\mathrm{N} + \mathrm{e}^+ + \nu \\
{}^{15}\mathrm{N} + {}^{1}\mathrm{H} &\rightarrow {}^{12}\mathrm{C} + {}^{4}\mathrm{He} \\
&\rightarrow {}^{16}\mathrm{O} + \gamma \\
{}^{16}\mathrm{O} + {}^{1}\mathrm{H} &\rightarrow {}^{17}\mathrm{F} + \gamma \\
{}^{17}\mathrm{F} &\rightarrow {}^{17}\mathrm{O} + \mathrm{e}^+ + \nu \\
{}^{17}\mathrm{O} + {}^{1}\mathrm{H} &\rightarrow {}^{14}\mathrm{N} + {}^{4}\mathrm{He}
\end{aligned} \tag{6.48}$$

[1] For example, if we denote by r_{pD} the rate of $^2\mathrm{H} + {}^1\mathrm{H}$, one has $r_{\mathrm{pD}} = r_{\mathrm{pp}}$, etc. Note that describing the p-p reaction as 'slow' and the $^2\mathrm{H} + {}^1\mathrm{H}$ as 'fast' refers to the difference in cross-section factors $\langle \sigma \upsilon \rangle$ and not to the number of reactions per second r given by eq. (6.8).

The ^{12}C nucleus is reproduced after the first six reactions, and thus only acts as a catalyst for the net hydrogen burning reaction (6.44). This set of six reactions forms the main cycle, also called the CN cycle. The ^{15}N + ^{1}H reaction has a small probability (somewhat less than 10^{-3}) to produce ^{16}O instead of ^{12}C + ^{4}He. This opens up a branching into the second cycle indicated in (6.48). The last three reactions have the effect of transforming ^{16}O, which is initially very abundant, into ^{14}N and thus bringing it into the main CN cycle. The relative proportions of C, N and O nuclei in the cycles change according to the different speeds of the reactions involved, but the total number of CNO-nuclei is always conserved. The three β-decay reactions have neutrino energies between 0.71 and 1.00 MeV and decay times between 10^2 and 10^3 sec. Unless very rapid changes are considered, these β-decays are so fast that one can ignore their detailed rates and the small resulting abundances of ^{13}N, ^{15}O and ^{17}F.

At high enough temperatures, $T \gtrsim 1.5 \times 10^7$ K, all reactions in the cycle come into a steady state or 'equilibrium' where the rate of production of each nucleus equals its rate of consumption. In this situation, as was the case with the p-p chain, the speed of the whole CNO cycle is controlled by the slowest reaction (the one with the smallest cross-section) which is ^{14}N(p, γ)^{15}O. This reaction acts like a bottleneck that congests the nuclei in their flow through the cycle, and ^{14}N thus becomes by far the most abundant of all the CNO nuclei. Looking at this in a bit more detail, the speed of the different reactions in the cycle can be expressed in terms of the nuclear lifetimes $\tau_{\rm p}$ against proton captures, as defined in eq. (6.39). In equilibrium one has d$n(^{12}\mathrm{C})/\mathrm{d}t$ = d$n(^{13}\mathrm{C})/\mathrm{d}t$, etc., so that

$$\left[\frac{n(^{12}\mathrm{C})}{n(^{13}\mathrm{C})}\right]_{\rm eq} = \frac{\langle\sigma v\rangle_{13}}{\langle\sigma v\rangle_{12}} = \frac{\tau_{\rm p}(^{12}\mathrm{C})}{\tau_{\rm p}(^{13}\mathrm{C})}\,, \qquad \text{etc.} \tag{6.49}$$

For the reactions in the CN cycle one typically has

$$\tau_{\rm p}(^{15}\mathrm{N}) \ll \tau_{\rm p}(^{13}\mathrm{C}) < \tau_{\rm p}(^{12}\mathrm{C}) \ll \tau_{\rm p}(^{14}\mathrm{N}) \ll \tau_{\rm nuc}\,.$$

Thus nearly all initially present CNO nuclei are transformed into ^{14}N by the CNO cycle. Therefore, apart from ^{4}He, the second-most important product of the CNO-cycle is ^{14}N – especially because the gas out of which stars form is typically more abundant in carbon and oxygen than in nitrogen.

The energy generation rate of the CNO cycle in equilibrium can be written as

$$\epsilon_{\rm CNO} = q_{\rm H}\, X\, X_{14}\, \frac{\rho}{m_{\rm u}}\, \langle\sigma v\rangle_{\rm pN}, \tag{6.50}$$

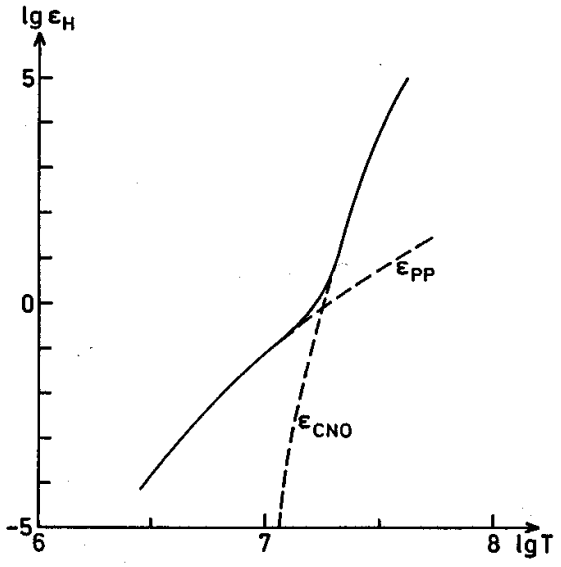


Figure 6.5. Total energy generation rate $\epsilon_{\rm H}$ (in erg g^{-1} s^{-1}) for hydrogen burning as a function of temperature, for ρ = 1 g/cm^3 and abundances X = 1 and $X_{\rm CNO}$ = 0.01. The dashed curves show the contributions of the pp chain and the CNO cycle. Figure from KIPPENHAHN.

where $\langle\sigma v\rangle_{\rm pN}$ is the cross-section factor of the ^{14}N(p, γ)^{15}O reaction which controls the rate of the whole cycle. X_{14} is the ^{14}N mass fraction in the energy-generating zone of the star, which is close to the total abundance $X_{\rm CNO}$ of CNO nuclei once equilibrium is reached in the full CNO cycle. The energy release per unit mass $q_{\rm H} = Q_{\rm H}/4m_{\rm u}$ takes into account the neutrino losses, which for the CNO cycle in equilibrium amounts to $Q_{\rm H} = 24.97$ MeV. The temperature sensitivity of the CNO cycle is much higher than for the pp chain, with ν varying between 23 and 13 for T_7 ranging from 1.0 to 5.0. This is illustrated in Fig. 6.5 where the temperature dependence of $\epsilon_{\rm CNO}$ is compared to that of $\epsilon_{\rm pp}$. For the purpose of very simple approximations one can take

$$\epsilon_{\rm pp} \propto X^2 \rho\, T^4 \qquad \text{and} \qquad \epsilon_{\rm CNO} \propto X X_{14} \rho\, T^{18}. \tag{6.51}$$

The strong difference in temperature sensitivity has the consequence that the pp chain dominates at low temperatures, $T_7 \lesssim 1.5$, while the CNO cycle is dominant at higher temperatures.

6.4.2 Helium burning

Helium burning consists of the fusion of ^{4}He into a mixture of ^{12}C and ^{16}O, which takes place at temperatures $T \gtrsim 10^8$ K. Such high temperatures are needed because (1) the Coulomb barrier for He fusion is higher than that of the H-burning reactions considered above, and (2) fusion of ^{4}He is hindered by the fact that no stable nucleus exists with mass number $A = 8$. Therefore helium burning must occur in two steps:

$$\begin{aligned} &^4\mathrm{He} + {}^4\mathrm{He} \leftrightarrow {}^8\mathrm{Be} \\ &^8\mathrm{Be} + {}^4\mathrm{He} \rightarrow {}^{12}\mathrm{C}^* \rightarrow {}^{12}\mathrm{C} + \gamma \end{aligned} \tag{6.52}$$

The ^{8}Be nucleus temporarily formed in the first reaction has a ground state that is 92 keV higher in energy than that of two separate ^{4}He nuclei. It therefore decays back into two α particles after a few time 10^{-16} s. While extremely short, this time is long enough to build up a very small equilibrium concentration of ^{8}Be, which increases with temperature and reaches about 10^{-9} at $T \approx 10^8$ K. Then the second reaction ^{8}Be(α, γ)^{12}C starts to occur at a significant rate, because of a resonance at just the Gamow peak energy. The result is an excited compound nucleus ^{12}C* which subsequently decays to the ground state of ^{12}C with emission of a γ photon. The corresponding energy level in the ^{12}C nucleus was predicted by Fred Hoyle in 1954, because he could not otherwise explain the existence of large amounts of carbon in the Universe. This excited state of ^{12}C was subsequently found in laboratory experiments.

The net effect of the two reactions (6.52) is called the *triple-α* reaction,

$$3\,{}^4\mathrm{He} \rightarrow {}^{12}\mathrm{C} + \gamma, \tag{6.53}$$

which has $Q = 7.275$ MeV. The energy release per unit mass is $q_{3\alpha} = Q/m(^{12}\mathrm{C}) = 5.9 \times 10^{17}$ erg/g, which is about 1/10 smaller than for H-burning. Since the two reactions need to occur almost simultaneously, the 3α reaction behaves as if it were a three-particle reaction and its rate is proportional to n_α^3. The energy-generation rate can be written as

$$\epsilon_{3\alpha} = q_{3\alpha}\, X_4^3 \rho^2\, \lambda_{3\alpha}, \tag{6.54}$$

where the temperature dependence is described by the factor $\lambda_{3\alpha}$, which depends on the combined cross-sections of the two reactions (6.52). $X_4 \approx Y$ is the mass fraction of ^{4}He. The temperature sensitivity of the 3α rate is extremely high, with $\nu \approx 40$ at $T_8 \approx 1.0$.

When a sufficient amount of ^{12}C has been created by the 3α reaction, it can capture a further α particle to form ^{16}O,

$$^{12}\mathrm{C} + {}^4\mathrm{He} \rightarrow {}^{16}\mathrm{O} + \gamma, \tag{6.55}$$

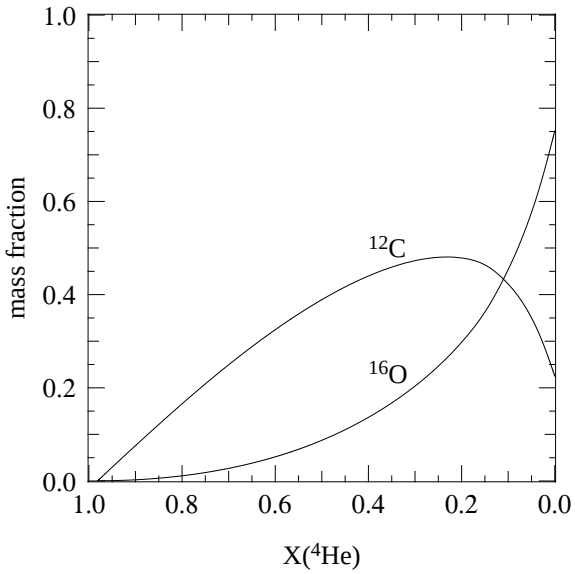


Figure 6.6. Dependence of the mass fractions of ^{12}C and ^{16}O on ^{4}He during He-burning, for typical conditions in intermediate-mass stars.

which has $Q = 7.162\,\mathrm{MeV}$, or $q_{\alpha C} = 4.32 \times 10^{17}$ erg per gram of produced ^{16}O. In principle further α captures on ^{16}O are possible, forming ^{20}Ne, but during normal helium burning conditions these are very rare. The $^{12}\mathrm{C}(\alpha,\gamma)^{16}\mathrm{O}$ reaction is strongly affected by resonances and its rate is quite uncertain. This is important because this reaction competes with the 3α reaction for available ^{4}He nuclei, as illustrated by Fig. 6.6. The final $^{12}\mathrm{C}/^{16}\mathrm{O}$ ratio reached at the end of He-burning is therefore also uncertain.

6.4.3 Carbon burning and beyond

In the mixture of mainly ^{12}C and ^{16}O that is left after helium burning, further fusion reactions can occur if the temperature rises sufficiently. In order of increasing temperature, the nuclear burning cycles that may follow are the following.

Carbon burning When the temperature exceeds $T_8 \gtrsim 5$ the large Coulomb barrier for $^{12}\mathrm{C} + {}^{12}\mathrm{C}$ fusion can be overcome. This is a complicated reaction, in which first an excited compound ^{24}Mg nucleus is formed which can then decay via many different channels. The most important channels are the following:

$$\begin{aligned} {}^{12}\mathrm{C} + {}^{12}\mathrm{C} \rightarrow {}^{24}\mathrm{Mg}^* &\rightarrow {}^{20}\mathrm{Ne} + \alpha \qquad Q = 4.616\,\mathrm{MeV} \quad (\sim 50\%) \\ &\rightarrow {}^{23}\mathrm{Na} + \mathrm{p} \qquad Q = 2.238\,\mathrm{MeV} \quad (\sim 50\%) \end{aligned} \tag{6.56}$$

The protons and α particles released find themselves at extremely high temperatures compared to those needed for hydrogen and helium burning, and will almost immediately react with other nuclei in the mixture, from ^{12}C to ^{24}Mg. Examples are $^{23}\mathrm{Na}(\mathrm{p},\alpha)^{20}\mathrm{Ne}$, $^{20}\mathrm{Ne}(\alpha,\gamma)^{24}\mathrm{Mg}$ and chains such as $^{12}\mathrm{C}(\mathrm{p},\gamma)^{13}\mathrm{N}(\mathrm{e}^+\nu)^{13}\mathrm{C}(\alpha,\mathrm{n})^{16}\mathrm{O}$, where the neutron will immediately react further. The overall energy release is obtained from the combination of all these reactions and is roughly $Q \approx 13\,\mathrm{MeV}$ per $^{12}\mathrm{C}+{}^{12}\mathrm{C}$ reaction. The main products after exhaustion of all carbon are ^{16}O, ^{20}Ne and ^{24}Mg (together 95% by mass fraction). These most abundant nuclei have equal numbers of protons and neutron, but some of the side reactions produce neutron-rich isotopes like ^{21,22}Ne, ^{23}Na and ^{25,26}Mg, so that after C burning the overall composition has a 'neutron excess' (n/p > 1, or $\mu_e > 2$).

Neon burning The next nuclear burning cycle might be expected to be oxygen fusion, but already at somewhat lower temperature ($T_9 \approx 1.5$) a process called 'neon burning' is initiated by the photo-disintegration of ^{20}Ne. At this temperature a sufficient number of photons have energies in the MeV range which is sufficient to break up the relatively fragile ^{20}Ne nucleus into ^{16}O and ^{4}He. This is immediately followed by the capture of the α particle by another ^{20}Ne nucleus, thus:

$$\begin{aligned} &^{20}\mathrm{Ne} + \gamma \leftrightarrow {}^{16}\mathrm{O} + \alpha \qquad Q = -4.73\,\mathrm{MeV} \\ &^{20}\mathrm{Ne} + \alpha \rightarrow {}^{24}\mathrm{Mg} + \gamma \qquad Q = 9.31\,\mathrm{MeV} \end{aligned} \tag{6.57}$$

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

Oxygen burning At $T_9 \approx 2.0$ fusion of ^{16}O nuclei sets in, which is in many ways analogous to the carbon fusion reaction described above. Also in this case there are several reaction channels, the most important ones being:

$$\begin{aligned} {}^{16}\mathrm{O} + {}^{16}\mathrm{O} \rightarrow {}^{32}\mathrm{S}^* &\rightarrow {}^{28}\mathrm{Si} + \alpha \qquad Q = 9.59\,\mathrm{MeV} \quad (\sim 60\%) \\ &\rightarrow {}^{31}\mathrm{P} + \mathrm{p} \qquad Q = 7.68\,\mathrm{MeV} \quad (\sim 40\%) \end{aligned} \tag{6.58}$$

Similar to carbon burning, the p and α particles are immediately captured by other nuclei, giving rise to a multitude of secondary reactions that eventually lead to a composition mostly consisting of ^{28}Si and ^{32}S (together 90% by mass fraction). The net energy release per ^{16}O + ^{16}O reaction is $Q \approx 16\,\mathrm{MeV}$. Since some of the side reactions involve β^+-decays and electron captures, the neutron excess of the final mixture is further increased.

Silicon burning The lightest and most abundant nucleus in the ashes of oxygen burning is ^{28}Si, but the Coulomb barrier for ^{28}Si + ^{28}Si fusion is prohibitively high. Instead silicon burning proceeds by a series of photo-disintegration (γ, α) and α-capture (α, γ) reactions when $T_9 \gtrsim 3$. Part of the silicon 'melts' into lighter nuclei, while another part captures the released ^{4}He to make heavier nuclei:

$$\begin{aligned} &{}^{28}\mathrm{Si}\,(\gamma,\alpha)\,{}^{24}\mathrm{Mg}\,(\gamma,\alpha)\,{}^{20}\mathrm{Ne}\,(\gamma,\alpha)\,{}^{16}\mathrm{O}\,(\gamma,\alpha)\,{}^{12}\mathrm{C}\,(\gamma,\alpha)\,2\alpha \\ &{}^{28}\mathrm{Si}\,(\alpha,\gamma)\,{}^{32}\mathrm{S}\,(\alpha,\gamma)\,{}^{36}\mathrm{Ar}\,(\alpha,\gamma)\,{}^{40}\mathrm{Ca}\,(\alpha,\gamma)\,{}^{44}\mathrm{Ti}\,(\alpha,\gamma)\,\ldots\,{}^{56}\mathrm{Ni} \end{aligned} \tag{6.59}$$

Most of these reactions are in equilibrium with each other, e.g. $^{28}\mathrm{Si} + \gamma \leftrightarrow {}^{24}\mathrm{Mg} + \alpha$, and the abundances of the nuclei can be described by nuclear equivalents of the Saha equation for ionization equilibrium. For $T > 4 \times 10^9\,\mathrm{K}$ a state close to *nuclear statistical equilibrium (NSE)* can be reached, where the most abundant nuclei are those with the lowest binding energy, constrained by the total number of neutrons and protons present. The final composition is then mostly ^{56}Fe because n/p > 1 (due to β-decays and e^--captures during previous burning cycles).

6.5 Neutrino emission

Neutrinos play a special role because their cross-section for interaction with normal matter is extremely small. The neutrinos that are released as a by-product of nuclear reactions have typical energies in the MeV range, and at such energies the interaction cross-section is $\sigma_\nu \sim 10^{-44}\,\mathrm{cm}^2$. The corresponding mean free path in matter at density $\rho = n\mu m_\mathrm{u}$ is $\ell_\nu = 1/(n\sigma_\nu) = \mu m_\mathrm{u}/(\rho\sigma_\nu) \sim 2 \times 10^{20}\,\mathrm{cm}/\rho$, for $\mu \approx 1$. Even at densities as high as $10^6\,\mathrm{g/cm}^3$, this gives $\ell_\nu \sim 3000\,R_\odot$. Therefore

any neutrino produced in the interior of a normal star leaves the star without interaction, carrying away its energy. The energy of neutrinos therefore has to be treated separately from other forms of energy, which are transported by a diffusive process due to a temperature gradient.

As mentioned before, the energy loss by neutrinos that are produced in nuclear reactions are conventionally taken into account by subtracting the neutrino energy from the total energy release of a reaction. In other words, the $\epsilon_{\rm nuc}$ term in the energy balance equation (5.4) is reduced and no separate ϵ_ν term is needed for these neutrinos.

However, also in the absence of nuclear reactions, *spontaneous neutrino emission* can occur at high densities and temperatures as a result of weak interaction processes. Owing to the fundamental coupling of the electromagnetic and weak interactions, for each electronic process that emits a photon, there is a very small but finite probability of emitting a neutrino-antineutrino pair instead of a photon. The theory of weak interactions predicts this probability to be

$$\frac{P(\nu\bar{\nu})}{P(\gamma)} \approx 3 \times 10^{-18} \left(\frac{E_\nu}{m_{\rm e} c^2} \right)^4, \tag{6.60}$$

where E_ν is the neutrino energy. These $\nu\bar{\nu}$ emissions represent a direct loss of energy from the stellar interior (a positive ϵ_ν in eq. 5.4) and thus give rise to *cooling* of the stellar matter.

The following processes of this type are important in stellar interiors (see Maeder Sec. 9.5 or Kippenhahn Sec. 18.6 for more details):

Photo-neutrinos In the process of electron scattering, discussed in Sec. 5.3.1, a photon is scattered by a free electron. There is a tiny probability (6.60) that the outgoing photon is replaced by a neutrino-antineutrino pair: $\gamma + {\rm e}^- \rightarrow {\rm e}^- + \nu + \bar{\nu}$. The average neutrino energy is $E_\nu \sim kT$, and therefore the probability of producing a $\nu\bar{\nu}$ pair instead of a photon is proportional to T^4. The rate of neutrino emission is also proportional to the number density of photons, $n_\gamma \propto T^3$, so that ϵ_ν is a very strong function of temperature, roughly $\epsilon_\nu \propto T^8$. The process of photo-neutrino emission results in significant cooling of stellar matter at $T \gtrsim 2 \times 10^8$ K.

Pair annihilation neutrinos At temperatures, $T \gtrsim 10^9$ K, energetic photons can undergo pair creation (Sec. 3.6.2), quickly followed by annihilation of the electron-positron pair. This normally yields two photons and these processes reach an equilibrium ($\gamma + \gamma \leftrightarrow {\rm e}^+ + {\rm e}^-$). Once in every $\sim 10^{19}$ cases, however, the annihilation produces a neutrino-antineutrino pair: ${\rm e}^+ + {\rm e}^- \rightarrow \nu + \bar{\nu}$, which results in a small one-way leakage out of the equilibrium exchange. This represents an important energy loss in a very hot, but not too dense plasma (ϵ_ν increases even more strongly with T than for photo-neutrinos, but is inversely proportional to ρ).

Plasma-neutrinos In a dense plasma, an electromagnetic wave can generate collective oscillations of the electrons. The energy of these waves is quantized and a quantum of this oscillation energy is called a 'plasmon'. The plasmon usually decays into photons, but again there is a finite probability (6.60) of $\nu\bar{\nu}$ emission. This process of neutrino energy loss dominates at high density, when the electron gas is degenerate.

Bremsstrahlung neutrinos Bremsstrahlung is the emission of a photon by an electron that is slowed down in the Coulomb field of an atomic nucleus (the inverse of free-free absorption, Sec. 5.3.1). The small probability of $\nu\bar{\nu}$ emission instead of a photon gives rise to significant cooling at low temperature and very high density. Unlike the processes discussed above, Bremsstrahlung depends on the presence of nuclei and therefore is more efficient for heavy elements (the neutrino emission rate is $\propto Z^2/A$).

The Urca process This process is different from the ones discussed above in that it involves nuclear transformations. Certain nuclei (Z, A) can capture an electron and subsequently undergo a β-decay back to the original nucleus,

$$\begin{aligned}(Z, A) + \mathrm{e}^- &\rightarrow (Z-1, A) + \nu \\ (Z-1, A) &\rightarrow (Z, A) + \mathrm{e}^- + \bar{\nu}.\end{aligned} \tag{6.61}$$

The net result is that the original particles are restored and two neutrinos are emitted. Only certain nuclei are suitable for this process: the nucleus $(Z-1, A)$ must be β-unstable and have a slightly higher rest energy than (Z, A), and the captured electron must be energetic enough to make the first reaction possible. These conditions are quite restrictive and the Urca process is inconsequential under most conditions found in stars, but it can play a role in very late stages of evolution at very high densities.

Suggestions for further reading

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

Exercises

6.1 Conceptual questions: Gamow peak

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

In the lecture (see eq. 6.22) you saw that the reaction rate is proportional to

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

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

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

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

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

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

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

6.2 Hydrogen burning

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

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

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

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

6.3 Relative abundances for CN equilibrium

Estimate the relative abundances of the nuclei CN-equilibrium if their lifetimes against proton capture at $T = 2 \times 10^7$ K are: $\tau_p(^{15}\mathrm{N}) = 30$ yr, $\tau_p(^{13}\mathrm{C}) = 1600$ yr, $\tau_p(^{12}\mathrm{C}) = 6600$ yr and $\tau_p(^{14}\mathrm{N}) = 6 \times 10^5$ yr.

6.4 Helium burning

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

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

6.5 Comparing radiative and convective cores

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

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

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

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

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

Chapter 7

Stellar models and stellar stability

In the previous chapters we have reviewed the most important physical processes taking place in stellar interiors, and we derived the differential equations that determine the structure and evolution of a star. By putting these ingredients together we can construct models of spherically symmetric stars. Because the complete set of equations is highly non-linear and time-dependent, their full solution requires a complicated numerical procedure. This is what is done in detailed stellar evolution codes, the results of which will be described in later chapters. We will not go into any detail about the numerical methods commonly used in such codes – for those interested, some of these details may be found in Chapter 24.2 of Maeder or Chapter 11 of Kippenhanhn.

The main purpose of this chapter is to briefly analyse the differential equations of stellar evolution and their boundary conditions, and to see how the full set of equations can be simplified in some cases to allow simple or approximate solutions – so-called *simple stellar models*. We also address the question of the stability of stars – whether the solutions to the equations yield a stable structure or not.

7.1 The differential equations of stellar evolution

Let us collect and summarize the differential equations for stellar structure and evolution that we have derived in the previous chapters, regarding m as the spatial variable, i.e. eqs. (2.6), (2.11), (5.4), (5.17) and (6.41):

$$\frac{\partial r}{\partial m} = \frac{1}{4\pi r^2 \rho} \tag{7.1}$$

$$\frac{\partial P}{\partial m} = -\frac{Gm}{4\pi r^4} - \frac{1}{4\pi r^2}\frac{\partial^2 r}{\partial t} \tag{7.2}$$

$$\frac{\partial l}{\partial m} = \epsilon_{\rm nuc} - \epsilon_\nu - T\frac{\partial s}{\partial t} \tag{7.3}$$

$$\frac{\partial T}{\partial m} = -\frac{Gm}{4\pi r^4}\frac{T}{P}\nabla \qquad \text{with} \qquad \nabla = \begin{cases} \nabla_{\rm rad} = \dfrac{3\kappa}{16\pi acG}\dfrac{lP}{mT^4} & \text{if } \nabla_{\rm rad} \le \nabla_{\rm ad} \\ \nabla_{\rm ad} + \Delta\nabla & \text{if } \nabla_{\rm rad} > \nabla_{\rm ad} \end{cases} \tag{7.4}$$

$$\frac{\partial X_i}{\partial t} = \frac{A_i m_u}{\rho}\left(-\sum_j (1+\delta_{ij})\, r_{ij} + \sum_{k,l} r_{kl,i} \right) \quad [+ \text{ mixing terms}] \qquad i = 1 \ldots N \tag{7.5}$$

Note that eq. (7.2) is written in its general form, without pre-supposing hydrostatic equilibrium. In eq. (7.3) we have replaced $\partial u/\partial t - (P/\rho^2)\partial\rho/\partial t$ by $T\partial s/\partial t$, according to the combined first and second laws of thermodynamics. Eq. (7.4) is generalized to include both the cases of radiative and convective energy transport. The term $\Delta\nabla$ is the superadiabaticity of the temperature gradient that must follow from a theory of convection (in practice, the mixing length theory); for the interior one can take $\Delta\nabla = 0$ except in the outermost layers of a star. Finally, eq. (7.5) has been modified to add 'mixing terms' that describe the redistribution (homogenization) of composition in convective regions. There are N such equations, one for each nucleus (isotope) indicated by subscript i.

The set of equations above comprise $4 + N$ partial differential equations that should be solved simultaneously. Let us count the number of unknown variables. Making use of the physics discussed in previous chapters, the functions P, s, κ, $\nabla_{\rm ad}$, $\Delta\nabla$, $\epsilon_{\rm nuc}$, ϵ_ν and the reaction rates r_{ij} can all be expressed as functions of ρ, T and composition X_i. We are therefore left with $4+N$ unknown variables (r, ρ, T, l and the X_i) so that we have a solvable system of equations.

The variables r, ρ, T, l and X_i appearing in the equations are all functions of two *independent* variables, m and t. We must therefore find a solution to the above set of equations on the interval $0 \leq m \leq M$ for $t > t_0$, assuming the evolution starts at time t_0. Note that M generally also depends on t in the presence of mass loss. A solution therefore also requires specification of *boundary conditions* (at $m = 0$ and $m = M$) and of *initial conditions*, for example $X_i(m, t_0)$.

7.1.1 Timescales and initial conditions

Let us further analyse the equations. Three kinds of time derivatives appear:

- $\partial^2 r/\partial t^2$ in eq. (7.2), which describes hydrodynamical changes to the stellar structure. These occur on the dynamical timescale $\tau_{\rm dyn}$ which as we have seen is very short. Thus we can normally assume hydrostatic equilibrium and $\partial^2 r/\partial t^2 = 0$, in which case eq. (7.2) reduces to the ordinary differential equation (2.13). Note that HE was explicitly assumed in eq. (7.4).

- $T\partial s/\partial t$ in eq. (7.3), which is often written as an additional energy generation term (eq. 5.5):

$$\epsilon_{\rm gr} = -T\frac{\partial s}{\partial t} = -\frac{\partial u}{\partial t} + \frac{P}{\rho^2}\frac{\partial \rho}{\partial t}$$

  It describes changes to the thermal structure of the star, which can result from contraction ($\epsilon_{\rm gr} > 0$) or expansion ($\epsilon_{\rm gr} < 0$) of the layers under consideration. Such changes occur on the thermal timescale $\tau_{\rm KH}$. If a star evolves on a much longer timescale than $\tau_{\rm KH}$ then $\epsilon_{\rm gr} \approx 0$ and the star is in thermal equilibrium. Then also eq. (7.3) reduces to an ordinary differential equation, eq. (5.7).

- $\partial X_i/\partial t$ in eqs. (7.5), describing changes in the composition. For the most abundant elements – the ones that affect the stellar structure – such changes normally occur on the longest, nuclear timescale $\tau_{\rm nuc}$.

Because normally $\tau_{\rm nuc} \gg \tau_{\rm KH} \gg \tau_{\rm dyn}$, composition changes are usually very slow compared to the other time derivatives. In that case eqs. (7.5) decouple from the other four equations (7.1–7.4), which can be seen to describe the *stellar structure* for a given composition $X_i(m)$.

For a star in both HE and TE (also called 'complete equilibrium'), the stellar structure equations (7.1–7.4) become a set of ordinary differential equations, independent of time. In that case it is sufficient to specify the initial composition profiles $X_i(m, t_0)$ as initial conditions. This is the case for so-called *zero-age main sequence* stars: the structure at the start of the main sequence depends only

on the initial composition, and is independent of the uncertain details of the star formation process, a very fortunate circumstance!

If a star starts out in HE, but not in TE, then the time derivative represented by $\epsilon_{\rm gr}$ remains in the set of structure equations. One would then also have to specify the specific entropy profile $s(m, t_0)$ as an initial condition. This is the case if one considers pre-main sequence stars. Fortunately, as we shall see later, in this case there is also a simplifying circumstance: pre-main sequence stars start out as fully convective gas spheres. This means that their temperature and pressure stratification is nearly adiabatic, so that s can be taken as constant throughout the star. It then suffices to specify the initial entropy.

7.2 Boundary conditions

The boundary conditions for the differential equations of stellar evolution constitute an important part of the overall problem. Not all boundary conditions can be specified at one end of the interval $[0, M]$: some boundary conditions are set in the centre and others at the surface. This means that direct forward integration of the equations is not possible, and the influence of the boundary conditions on the solutions is not easy to foresee.

7.2.1 Central boundary conditions

At the centre ($m = 0$), both the density and the energy generation rate must remain finite. Therefore, both r and l must vanish in the centre:

$$m = 0 \quad : \qquad r = 0 \quad \text{and} \quad l = 0. \tag{7.6}$$

However, nothing is known a priori about the central values of P and T. Therefore the remaining two boundary conditions must be specified at the surface rather than the centre.

It is possible to get some idea of the behaviour of the variables close to the centre by means of a Taylor expansion. Even though P_c and T_c are unknown, one can do this also for P and T, writing for example

$$P = P_c + m \left[\frac{\mathrm{d}P}{\mathrm{d}m}\right]_c + \tfrac{1}{2} m^2 \left[\frac{\mathrm{d}^2 P}{\mathrm{d}m^2}\right]_c + \ldots$$

and making use of the stellar structure equations for $\mathrm{d}P/\mathrm{d}m$, etc, see Exercise 7.5.

7.2.2 Surface boundary conditions

At the surface ($m = M$, or $r = R$), the boundary conditions are generally much more complicated than at the centre. One may treat the surface boundary conditions at different levels of sophistication.

- The simplest option is to take $T = 0$ and $P = 0$ at the surface (the 'zero' boundary conditions). However, in reality T and P never become zero because the star is surrounded by an interstellar medium with very low, but finite density and temperature.

- A better option is to identify the surface with the *photosphere*, which is where the bulk of the radiation escapes and which corresponds with the visible surface of the star. The photospheric boundary conditions approximate the photosphere with a single surface at optical depth $\tau = \frac{2}{3}$. We can write

$$\tau_{\rm ph} = \int_R^\infty \kappa\rho \, \mathrm{d}r \approx \kappa_{\rm ph} \int_R^\infty \rho \, \mathrm{d}r, \tag{7.7}$$

where $\kappa_{\rm ph}$ is an average value of opacity over the atmosphere (all layers above the photosphere). If the atmosphere is geometrically thin we also have

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{GM}{R^2}\rho \quad \Rightarrow \quad P(R) \approx \frac{GM}{R^2}\int_R^\infty \rho\,\mathrm{d}r. \tag{7.8}$$

Since $\tau_{\rm ph} = \frac{2}{3}$ and $T(R) \approx T_{\rm eff}$ we can combine the above equations to write the photospheric boundary conditions as:

$$m = M\ (r = R) \quad : \qquad P = \frac{2}{3}\frac{GM}{\kappa_{\rm ph}R^2} \qquad \text{and} \qquad L = 4\pi R^2\sigma T^4. \tag{7.9}$$

- The problem with the photospheric boundary conditions above is that the radiative diffusion approximation on which it is based breaks down when $\tau \lesssim$ a few. The best solution is therefore to fit a *detailed stellar atmosphere model* to an interior shell (at $\tau > \frac{2}{3}$) where the radiative diffusion approximation is still valid. This is a more complicated and time-consuming approach, and in many (but not all) practical situations the photospheric conditions are sufficient.

7.2.3 Effect of surface boundary conditions on stellar structure

It is instructive to look at the effect of the surface boundary conditions on the solution for the structure of the outer envelope of a star. Assuming complete (dynamical and thermal) equilibrium, the envelope contains only a small fraction of the mass and no energy sources. In that case $l = L$ and $m \approx M$. It is then better to take P, rather than m, as the independent variable describing depth within the envelope. We can write the equation for radiative energy transport as

$$\frac{\mathrm{d}T}{\mathrm{d}P} = \frac{T}{P}\nabla_{\rm rad} = \frac{3}{16\pi acG}\frac{\kappa l}{mT^3} \approx \text{const}\cdot\frac{L}{M}\frac{\kappa}{T^3} \tag{7.10}$$

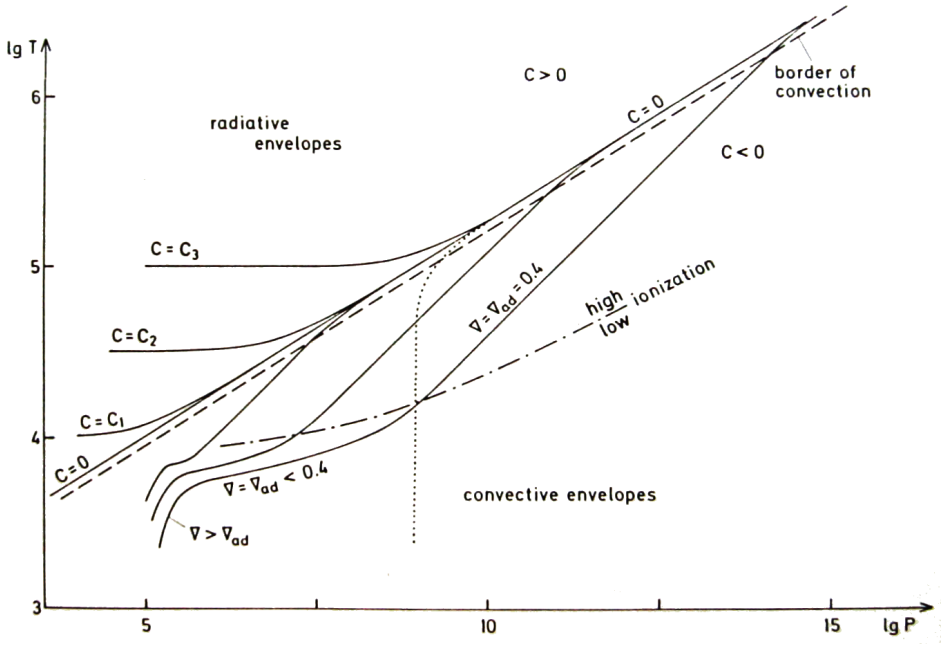


Figure 7.1. Schematic diagram of $\log T$ versus $\log P$ illustrating the different types of envelope structure solutions, as discussed in the text. Figure from Kippenhahn & Weigert.

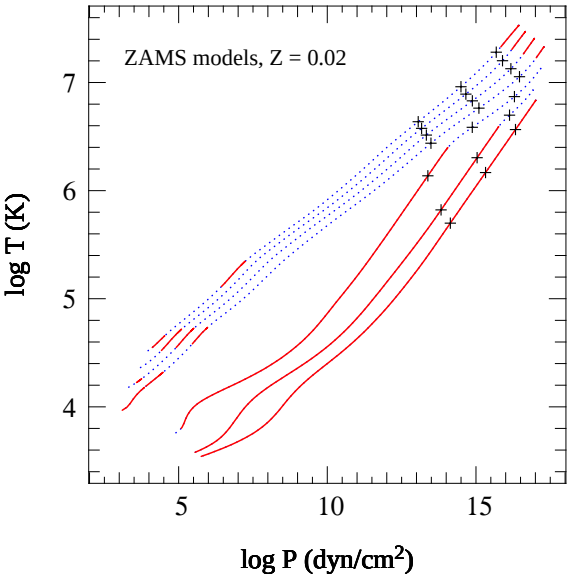


Figure 7.2. Structure of detailed stellar models on the zero-age main sequence in the $\log P$, $\log T$ diagram. Each curve is for a different mass, from top to bottom: 16 $M_\odot$, 8 $M_\odot$, 4 $M_\odot$, 2 $M_\odot$, 1 $M_\odot$, 0.5 $M_\odot$ and 0.25 $M_\odot$. The + symbols on each curve indicate, for increasing P, the part of the envelope containing $0.01M$, $0.1M$ and $0.5M$ (so most of the T and P variation in the envelope occurs in the outer 1 % of the mass). The dotted (blue) parts of each curve indicate radiative regions of the star, the solid (red) parts indicate convective regions.

Stars with $M \leq 1\,M_\odot$ have low $T_{\rm eff}$ and therefore convective envelopes. The depth of the convective envelope increases strongly with decreasing surface temperature (and thus with decreasing mass); the 0.25 $M_\odot$ star is completely convective. On the other hand, more massive stars have higher $T_{\rm eff}$ and mostly radiative envelopes, except for small convective layers near the surface caused by partial ionization.

We also approximate the opacity by a simple law, $\kappa = \kappa_0 P^a T^b$ (this can represent e.g. Kramers opacity, $a = 1$, $b = -4.5$; or electron scattering, $a = b = 0$). We can then integrate $\mathrm{d}T/\mathrm{d}P$ to give

$$T^{4-b} = B\,(P^{1+a} + C) \tag{7.11}$$

where $B \propto L/M$ = constant and C is an integration constant, determined by the boundary conditions. For the Kramers opacity, which is a reasonable approximation for stellar envelopes of moderate temperatures, we find $T^{8.5} = B\,(P^2 + C)$.

The different possible solutions are characterized by the value of C, and the various possibilities are illustrated in Fig. 7.1. At large enough pressure, $P \gg \sqrt{C}$, all solutions approach $T \propto P^{2/8.5} \approx P^{0.235}$. This corresponds to an actual temperature gradient, $\nabla = \mathrm{d}\log T/\mathrm{d}\log P = 0.235 < \nabla_{\rm ad} \approx 0.4$, which is consistent with the assumed radiative transport. There is a fundamental difference between solutions with $C \geq 0$ and $C < 0$, however.

Radiative envelopes correspond to solutions with $C \geq 0$. In the special case $C = 0$ the slope of the solution ∇ remains equal to 0.235 when $P, T \to 0$. This corresponds to the 'zero' boundary conditions discussed above. For $C > 0$ the solutions lie above the $C = 0$ line, and the slope decreases ($\nabla < 0.235$) as the surface is approached. This corresponds to more realistic, e.g. photospheric, boundary conditions with large enough $T_{\rm eff}$, demonstrating that *stars with relatively hot photospheres have radiative envelopes*. In practice this is the case when $T_{\rm eff} \gtrsim 9000$ K. Fig. 7.1 demonstrates that such envelope solutions quickly approach the 'radiative zero' structure, $C = 0$. This means that the envelope structure is insensitive to the assumed surface boundary conditions, and in practice the photospheric BC's are sufficient.

Convective envelopes correspond to solutions with $C < 0$. In this case the solutions lie below the $C = 0$ line, and their slope increases as $\log P$ decreases. This is shown by the dotted line in Fig. 7.1. However, the assumption of radiative transport breaks down when $\nabla > \nabla_{\rm ad} \approx 0.4$ and convection sets in. Therefore, *stars with cool photospheres have convective envelopes*, in practice when $T_{\rm eff} \lesssim 9000$ K. The actual temperature stratification in the envelope is close to adiabatic, $\nabla = \nabla_{\rm ad}$, until the surface is approached. Before this happens, however, partial ionization decreases $\nabla_{\rm ad}$ below 0.4 and gives rise to a much shallower slope. Since the different

solution lie close together near the surface, but further apart in the interior, the structure of a convective envelope *is* sensitive to the surface boundary conditions. This means that the structure also depends on the uncertain details of near-surface convection (see Sec. 5.5). A small change or uncertainty in $T_{\rm eff}$ can have a large effect on the depth of the convective envelope! For small enough $T_{\rm eff}$ the whole star can become convective (leading to the Hayashi line in the H-R diagram, see Sect. 9.1.1).

The approximate description given here is borne out by detailed stellar structure calculations, as demonstrated in Fig. 7.2. Also note that if we assume electron scattering instead of Kramers opacity, the description remains qualitatively the same (the radiative zero solution then has $\nabla = 0.25$ instead of 0.235).

7.3 Equilibrium stellar models

For a star in both hydrostatic and thermal equilibrium, the four partial differential equations for stellar structure (eqs. 7.1–7.4) reduce to ordinary, time-independent differential equations. We can further simplify the situation somewhat, by ignoring possible neutrino losses (ϵ_ν) which are only important in very late stages of evolution, and ignoring the superadiabaticity of the temperature gradient in surface convection zones. We then arrive at the following set of structure equations that determine the stellar structure for a given composition profile $X_i(m)$:

$$\frac{\mathrm{d}r}{\mathrm{d}m} = \frac{1}{4\pi r^2 \rho} \tag{7.12}$$

$$\frac{\mathrm{d}P}{\mathrm{d}m} = -\frac{Gm}{4\pi r^4} \tag{7.13}$$

$$\frac{\mathrm{d}l}{\mathrm{d}m} = \epsilon_{\rm nuc} \tag{7.14}$$

$$\frac{\mathrm{d}T}{\mathrm{d}m} = -\frac{Gm}{4\pi r^4}\frac{T}{P}\nabla \qquad \text{with} \qquad \nabla = \begin{cases} \nabla_{\rm rad} = \dfrac{3\kappa}{16\pi acG}\dfrac{lP}{mT^4} & \text{if } \nabla_{\rm rad} \le \nabla_{\rm ad} \\ \nabla_{\rm ad} & \text{if } \nabla_{\rm rad} > \nabla_{\rm ad} \end{cases} \tag{7.15}$$

We note that the first two equations (7.12 and 7.13) describe the *mechanical structure* of the star, and the last two equations (7.14 and 7.15) describe the *thermal and energetic* structure. They are coupled to each other through the fact that, for a general equation of state, P is a function of both ρ and T.

Although simpler than the full set of evolution equations, this set still has no simple, analytic solutions. The reasons are that, first of all, the equations are very non-linear: e.g. $\epsilon_{\rm nuc} \propto \rho T^\nu$ with $\nu \gg 1$, and κ is a complicated function of ρ and T. Secondly, the four differential equations are coupled and have to be solved simultaneously. Finally, the equations have boundary conditions at both ends, and thus require iteration to obtain a solution.

It is, however, possible to make additional simplifying assumptions so that under certain circumstances an analytic solution or a much simpler numerical solution is possible. We have already discussed one example of such a simplifying approach in Chapter 4, namely the case of *polytropic models* in which the pressure and density are related by an equation of the form

$$P = K\rho^\gamma.$$

Since in this case P does not depend on T, the mechanical structure of a stellar model can be computed in a simple way, independent of its thermal and energetic properties, by solving eqs. (7.12) and (7.13).

Another approach is to consider simple scaling relations between stellar models with different masses and radii, but all having the same (or a very similar) relative density distributions. If a detailed numerical solution can be computed for one particular star, these so-called *homology relations* can be used to find an approximate model for another star.

7.4 Homology relations

Solving the stellar structure equations almost always requires heavy numerical calculations, such as are applied in detailed stellar evolution codes. However, there is often a kind of similarity between the numerical solutions for different stars. These can be approximated by simple analytical scaling relations known as *homology relations*. In past chapters we have already applied simple scaling relations based on rough estimates of quantities appearing in the stellar structure equations. In this section we will put these relations on a firmer mathematical footing.

The requirements for the validity of homology are very restrictive, and hardly ever apply to realistic stellar models. However, homology relations can offer a rough but sometimes very helpful basis for interpreting the detailed numerical solutions. This applies to models for stars on the *main sequence* and to so-called *homologous contraction*.

Definition Compare two stellar models, with masses M_1 and M_2 and radii R_1 and R_2. All interior quantities in star 1 are denoted by subscript '1' (e.g. the mass coordinate m_1), etc. Now consider so-called *homologous mass shells* which have the same relative mass coordinate, $x \equiv m/M$, i.e.

$$x = \frac{m_1}{M_1} = \frac{m_2}{M_2} \tag{7.16}$$

The two stellar models are said to be *homologous* if homologous mass shells within them are located at the same relative radii r/R, i.e.

$$\frac{r_1(x)}{R_1} = \frac{r_2(x)}{R_2} \qquad \text{or} \qquad \frac{r_1(x)}{r_2(x)} = \frac{R_1}{R_2} \tag{7.17}$$

for all x.

Comparing two homologous stars, the ratio of radii r_1/r_2 for homologous mass shells is constant. In other words, two homologous stars have *the same relative mass distribution*, and therefore (as we shall prove shortly) the same relative density distribution.

All models have to obey the stellar structure equations, so that the transition for one homologous model to another has consequences for all other variables. We start by analysing the first two structure equations.

- The first stellar structure equation (7.12) can be written for star 1 as

$$\frac{\mathrm{d}r_1}{\mathrm{d}x} = \frac{M_1}{4\pi {r_1}^2 \rho_1} \tag{7.18}$$

If the stars are homologous, then from eq. (7.17) we can substitute $r_1 = r_2\,(R_1/R_2)$ and obtain

$$\frac{\mathrm{d}r_2}{\mathrm{d}x} = \frac{M_2}{4\pi {r_2}^2 \rho_2} \cdot \left[\frac{\rho_2}{\rho_1}\,\frac{M_1}{M_2}\left(\frac{R_2}{R_1}\right)^3\right]. \tag{7.19}$$

We recognize the structure equation for the radius of star 2 (i.e. eq. 7.18 with subscript '1' replaced by '2'), multiplied by the factor in square brackets on the right-hand side. This equation must hold generally, which is only the case if the factor in square brackets is equal to one for all values of x, that is if

$$\frac{\rho_2(x)}{\rho_1(x)} = \frac{M_2}{M_1}\left(\frac{R_2}{R_1}\right)^{-3}. \tag{7.20}$$

This must hold at any homologous mass shell, and therefore also at the centre of each star. The factor MR^{-3} is proportional to the average density $\bar{\rho}$, so that the density at any homologous shell scales with the central density, or with the average density:

$$\boxed{\rho(x) \propto \rho_c \propto \bar{\rho}} \tag{7.21}$$

Note that, therefore, any two polytropic models with the same index n are homologous to each other.

- We can apply a similar analysis to the second structure equation (7.13) for hydrostatic equilibrium. For star 1 we have

$$\frac{\mathrm{d}P_1}{\mathrm{d}x} = -\frac{G{M_1}^2 x}{4\pi {r_1}^4} \tag{7.22}$$

so that after substituting $r_1 = r_2\,(R_1/R_2)$ we obtain

$$\frac{\mathrm{d}P_1}{\mathrm{d}x} = -\frac{G{M_2}^2 x}{4\pi {r_2}^4}\cdot\left[\left(\frac{M_1}{M_2}\right)^2\left(\frac{R_2}{R_1}\right)^4\right] = \frac{\mathrm{d}P_2}{\mathrm{d}x}\cdot\left[\left(\frac{M_1}{M_2}\right)^2\left(\frac{R_2}{R_1}\right)^4\right], \tag{7.23}$$

where the second equality follows because star 2 must also obey the hydrostatic equilibrium equation. Hence we have $\mathrm{d}P_1/\mathrm{d}x = C\,\mathrm{d}P_2/\mathrm{d}x$, with C equal to the (constant) factor in square brackets. Integrating we obtain $P_1(x) = C\,P_2(x) + B$, where the integration constant $B = 0$ because at the surface, $x \to 1$, for both stars $P \to 0$. Thus we obtain

$$\frac{P_2(x)}{P_1(x)} = \left(\frac{M_2}{M_1}\right)^2\left(\frac{R_2}{R_1}\right)^{-4}, \tag{7.24}$$

at any homologous mass shell. Again this must include the centre, so that for all x:

$$\boxed{P(x) \propto P_c \propto \frac{M^2}{R^4}}. \tag{7.25}$$

The pressure required for hydrostatic equilibrium therefore scales with M^2/R^4 at any homologous shell. Note that we found the same scaling of the *central* pressure with M and R from our rough estimate in Sect. 2.2, and for polytropic models of the same index n.

We can combine eqs. (7.20) and (7.24) to show that two homologous stars must obey the following relation between pressure and density at homologous points,

$$\frac{P_2(x)}{P_1(x)} = \left(\frac{M_2}{M_1}\right)^{2/3}\left(\frac{\rho_2(x)}{\rho_1(x)}\right)^{4/3}, \tag{7.26}$$

or

$$P(x) \propto M^{2/3}\rho(x)^{4/3}. \tag{7.27}$$

7.4.1 Homology for radiative stars composed of ideal gas

In order to obtain simple homology relations from the other structure equations, we must make additional assumptions. We start by analysing eq. (7.15).

- First, let us assume the *ideal gas* equation of state,

$$P = \frac{\mathcal{R}}{\mu}\rho T.$$

Let us further assume that in each star the *composition is homogeneous*, so that μ is a constant for both stars, though not necessarily the same. We can then combine eqs. (7.20) and (7.24) to obtain a relation between the temperatures at homologous mass shells,

$$\frac{T_2(x)}{T_1(x)} = \frac{\mu_2}{\mu_1}\frac{M_2}{M_1}\left(\frac{R_2}{R_1}\right)^{-1} \qquad \text{or} \qquad \boxed{T(x) \propto T_c \propto \mu\frac{M}{R}} \tag{7.28}$$

- Second, we will assume the stars are in *radiative equilibrium*. We can then write eq. (7.15) as

$$\frac{\mathrm{d}(T^4)}{\mathrm{d}x} = -\frac{3M}{16\pi^2 ac}\frac{\kappa l}{r^4} \tag{7.29}$$

This contains two as yet unknown functions of x on the right-hand side, κ and l. We must therefore make additional assumptions about the opacity, which we can very roughly approximate by a power law,

$$\kappa = \kappa_0 \rho^a T^b. \tag{7.30}$$

For a Kramers opacity law, we would have $a = 1$ and $b = -3.5$. However, for simplicity let us assume a *constant opacity* throughout each star (but like μ, not necessarily the same for both stars). Then a similar reasoning as was held above for the pressure, allows us to transform eq. (7.29) into an expression for the ratio of luminosities at homologous points,

$$\left(\frac{T_2(x)}{T_1(x)}\right)^4 = \frac{l_2(x)}{l_1(x)}\frac{M_2}{M_1}\frac{\kappa_2}{\kappa_1}\left(\frac{R_2}{R_1}\right)^{-4} \qquad \Rightarrow \qquad \frac{l_2(x)}{l_1(x)} = \left(\frac{\mu_2}{\mu_1}\right)^4\left(\frac{M_2}{M_1}\right)^3\left(\frac{\kappa_2}{\kappa_1}\right)^{-1} \tag{7.31}$$

making use of eq. (7.28) to obtain the second expression. This relation also holds for the surface layer, i.e. for the total stellar luminosity L. Hence $l(x) \propto L$ and

$$\boxed{L \propto \frac{1}{\kappa}\mu^4 M^3} \tag{7.32}$$

This relation represents a *mass-luminosity relation* for a radiative, homogeneous star with constant opacity and ideal-gas pressure.

Note that we obtained a mass-luminosity relation (7.32) without making any assumption about the mode of energy generation (and indeed, without even having to assume thermal equilibrium, because we have not yet made use of eq. 7.14). We can thus expect a mass-luminosity relation to hold not only on the main-sequence, but for any star in radiative equilibrium. What this relation tells us is that the luminosity depends mainly on how efficiently energy can be transported by radiation: a higher

opacity gives rise to a smaller luminosity, because the nontransparent layers work like a blanket wrapped around the star. In practice, for a star in thermal equilibrium (e.g. on the main sequence) the power generated by nuclear reactions $L_{\rm nuc}$ adapts itself to the surface luminosity L, and thereby also the central temperature needed to make the nuclear reactions proceed at the rate dictated by L.

Note, however, that the simple mass-luminosity relation (7.32) depends on the assumption of constant opacity. If we assume a Kramers opacity law, the mass-luminosity also depends (weakly) on the radius. It is left as an exercise to show that, in this case

$$L \propto \frac{\mu^{7.5} M^{5.5}}{R^{0.5}}. \tag{7.33}$$

This means that if the opacity is not a constant, there is a weak dependence of the luminosity on the mode of energy generation, through the radius dependence (see Sec. 7.4.2).

7.4.2 Main sequence homology

For stars that are in thermal equilibrium we can make use of the last structure equation (7.14) to derive further homology relations for the radius as a function of mass. We then have to assume a specific form for the energy generation rate, say

$$\epsilon_{\rm nuc} = \epsilon_0 \rho T^\nu \tag{7.34}$$

so that eq. (7.14) can be written as

$$\frac{{\rm d}l}{{\rm d}x} = \epsilon_0 M \rho T^\nu \tag{7.35}$$

By making use of the other homology relations, including the mass-luminosity relation eq. (7.32), we obtain for a homogeneous, radiative star with constant opacity and consisting of an ideal gas:

$$\boxed{R \propto \mu^{(\nu-4)/(\nu+3)}\, M^{(\nu-1)/(\nu+3)}} \tag{7.36}$$

The slope of this *mass-radius relation* therefore depends on ν, that is, on the mode of nuclear energy generation. For main-sequence stars, in which hydrogen fusion provides the energy source, there are two possibilities, see Table 7.1.

We can also obtain relations between the central temperature and central density and the mass of a star in thermal equilibrium, by combining the homology relations for the radius (7.36) with those for density and temperature (7.20 and 7.28):

$$\rho_c \propto \mu^{3(4-\nu)/(\nu+3)}\, M^{2(3-\nu)/(\nu+3)} \tag{7.37}$$

$$T_c \propto \mu^{7/(\nu+3)}\, M^{4/(\nu+3)} \tag{7.38}$$

Again, the result depends on the mode of energy generation through the value of ν. For main-sequence stars the possibilities are tabulated in Table 7.1.

The mass-luminosity and mass-radius relations (7.32) and (7.36) can be compared to the observed relations for main-sequence stars that were presented in Chapter 1, and to the results of detailed stellar structure calculations. This comparison is deferred to Chapter 9, where the main sequence is discussed in more detail.

Table 7.1. Homology relations for the radius, central temperature and central density of main-sequence stars

pp-chain	$\nu \approx 4$	$R \propto M^{0.43}$	$T_c \propto \mu M^{0.57}$	$\rho_c \propto M^{-0.3}$
CNO cycle	$\nu \approx 18$	$R \propto \mu^{2/3} M^{0.81}$	$T_c \propto \mu^{1/3} M^{0.19}$	$\rho_c \propto \mu^{-2} M^{-1.4}$

7.4.3 Homologous contraction

We have seen in Chapter 2 that, as a consequence of the virial theorem, a star without internal energy sources must contract under the influence of its own self-gravity. Suppose that this contraction takes place homologously. According to eq. (7.17) each mass shell inside the star then maintains the same relative radius r/R. Writing $\dot{r} = \partial r/\partial t$, etc., this means that

$$\frac{\dot{r}(m)}{r(m)} = \frac{\dot{R}}{R}.$$

Since in this case we compare homologous models with the same mass M, we can replace x by the mass coordinate m. For the change in density we obtain from eq. (7.20) that

$$\frac{\dot{\rho}(m)}{\rho(m)} = -3\frac{\dot{R}}{R}, \tag{7.39}$$

and if the contraction occurs *quasi-statically*, i.e. slow enough to maintain HE, then the change in pressure follows from eq. (7.24),

$$\frac{\dot{P}(m)}{P(m)} = -4\frac{\dot{R}}{R} = \frac{4}{3}\frac{\dot{\rho}(m)}{\rho(m)}. \tag{7.40}$$

To obtain the change in temperature for a homologously contracting star, we have to consider the equation of state. Writing the equation of state in its general, differential form eq. (3.48) we can eliminate $\dot{P}/P$ to get

$$\frac{\dot{T}}{T} = \frac{1}{\chi_T}\left(\frac{4}{3} - \chi_\rho\right)\frac{\dot{\rho}}{\rho} = \frac{1}{\chi_T}\left(3\chi_\rho - 4\right)\frac{\dot{R}}{R}. \tag{7.41}$$

Hence the temperature increases as a result of contraction as long as $\chi_\rho < \frac{4}{3}$. For an ideal gas, with $\chi_\rho = 1$, the temperature indeed increases upon contraction, in accordance with our (qualitative) conclusion from the virial theorem. Quantitatively,

$$\frac{\dot{T}}{T} = \frac{1}{3}\frac{\dot{\rho}}{\rho}.$$

However, for a degenerate electron gas with $\chi_\rho = \frac{5}{3}$ eq. (7.41) shows that the temperature decreases, in other words a degenerate gas sphere will *cool* upon contraction. The full consequences of this important result will be explored in Chapter 8.

7.5 Stellar stability

We have so far considered stars in both hydrostatic and thermal equilibrium. But an important question that remains to be answered is whether these equilibria are *stable*. From the fact that stars can preserve their properties for very long periods of time, we can guess that this is indeed the case. But in order to answer the question of stability, and find out under what circumstances stars may become unstable, we must test what happens when the equilibrium situation is perturbed: will the perturbation be quenched (stable situation) or will it grow (unstable situation). Since there are two kinds of equilibria, we have to consider two kinds of stability:

- dynamical stability: what happens when hydrostatic equilibrium is perturbed?
- thermal (secular) stability: what happens when the thermal equilibrium situation is perturbed?

7.5.1 Dynamical stability of stars

The question of dynamical stability relates to the response of a certain part of a star to a perturbation of the balance of forces that act on it: in other words, a perturbation of hydrostatic equilibrium. We already treated the case of dynamical stability to *local* perturbations in Sec. 5.5.1, and saw that in this case instability gives rise to *convection*. In this section we look at the global stability of concentric layers within a star to radial perturbations, i.e. compression or expansion. A rigorous treatment of this problem is very complicated, so we will only look at a very simplified example to illustrate the principles.

Suppose a star in hydrostatic equilibrium is compressed on a short timescale, $\tau \ll \tau_{\rm KH}$, so that the compression can be considered as adiabatic. Furthermore suppose that the compression occurs *homologously*, such that its radius decreases from R to R'. Then the density at any layer in the star becomes

$$\rho \rightarrow \rho' = \rho \left(\frac{R'}{R} \right)^{-3}$$

and the new pressure after compression becomes P', given by the adiabatic relation

$$\frac{P'}{P} = \left(\frac{\rho'}{\rho} \right)^{\gamma_{\rm ad}} = \left(\frac{R'}{R} \right)^{-3\gamma_{\rm ad}} .$$

The pressure required for HE after homologous contraction is

$$\left(\frac{P'}{P} \right)_{\rm HE} = \left(\frac{\rho'}{\rho} \right)^{4/3} = \left(\frac{R'}{R} \right)^{-4}$$

Therefore, if $\gamma_{\rm ad} > \frac{4}{3}$ then $P' > P'_{\rm HE}$ and the excess pressure leads to re-expansion (on the dynamical timescale $\tau_{\rm dyn}$) so that HE is restored. If, however, $\gamma_{\rm ad} < \frac{4}{3}$ then $P' < P'_{\rm HE}$ and the increase of pressure is not sufficient to restore HE. The compression will therefore reinforce itself, and the situation is unstable on the dynamical timescale. We have thus obtained a criterion for *dynamical stability*:

$$\boxed{\gamma_{\rm ad} > \tfrac{4}{3}} \tag{7.42}$$

It can be shown rigorously that a star that has $\gamma_{\rm ad} > \frac{4}{3}$ everywhere is dynamically stable, and if $\gamma_{\rm ad} = \frac{4}{3}$ it is neutrally stable. However, the situation when $\gamma_{\rm ad} < \frac{4}{3}$ in some part of the star requires further investigation. It turns out that global dynamical instability is obtained when the integral

$$\int \left(\gamma_{\rm ad} - \frac{4}{3} \right) \frac{P}{\rho} {\rm d}m \tag{7.43}$$

over the whole star is negative. Therefore if $\gamma_{\rm ad} < \frac{4}{3}$ in a sufficiently large core, where P/ρ is high, the star becomes unstable. However if $\gamma_{\rm ad} < \frac{4}{3}$ in the outer layers where P/ρ is small, the star as a whole need not become unstable.

Cases of dynamical instability

Stars dominated by an ideal gas or by non-relativistic degenerate electrons have $\gamma_{\rm ad} = \frac{5}{3}$ and are therefore dynamically stable. However, we have seen that for relativistic particles $\gamma_{\rm ad} \rightarrow \frac{4}{3}$ and stars dominated by such particles tend towards a neutrally stable state. A small disturbance of such a star could either lead to a collapse or an explosion. This is the case if *radiation pressure* dominates (at high T and low ρ), or the pressure of relativistically degenerate electrons (at very high ρ).

A process that can lead to $\gamma_{\rm ad} < \frac{4}{3}$ is *partial ionization* (e.g. H $\leftrightarrow$ $H^+ + e^-$), as we have seen in Sec. 3.5. Since this normally occurs in the very outer layers, where P/ρ is small, it does not lead to overall dynamical instability of the star. However, partial ionization is connected to driving oscillations in some kinds of star.

At very high temperatures two other processes can occur that have a similar effect to ionization. These are *pair creation* ($\gamma + \gamma \leftrightarrow e^+ + e^-$, see Sect. 3.6.2) and *photo-disintegration* of nuclei (e.g. γ + Fe $\leftrightarrow \alpha$). These processes, that may occur in massive stars in late stages of evolution, also lead to $\gamma_{\rm ad} < \frac{4}{3}$ but now in the core of the star. These processes can lead to a stellar explosion or collapse (see Chapter 13).

7.5.2 Secular stability of stars

The question of thermal or *secular* stability, i.e. the stability of thermal equilibrium, is intimately linked to the virial theorem. In the case of an ideal gas the virial theorem (Sect. 2.3) tells us that the total energy of a star is

$$E_{\rm tot} = -E_{\rm int} = \tfrac{1}{2} E_{\rm gr}, \tag{7.44}$$

which is negative: the star is bound. The rate of change of the total energy is given by the difference between the rate of nuclear energy generation in the deep interior and the rate of energy loss in the form of radiation from the surface:

$$\dot{E}_{\rm tot} = L_{\rm nuc} - L \tag{7.45}$$

In a state of thermal equilibrium, $L = L_{\rm nuc}$ and $E_{\rm tot}$ remains constant. Consider now a small perturbation of this situation, for instance $L_{\rm nuc} > L$ because of a small temperature fluctuation. This leads to an increase of the total energy, $\delta E_{\rm tot} > 0$, and since the total energy is negative, its absolute value becomes smaller. The virial theorem, eq. (7.44), then tells us that (1) $\delta E_{\rm gr} > 0$, in other words the star will expand ($\delta\rho < 0$), and (2) $\delta E_{\rm int} < 0$, meaning that the overall temperature will decrease ($\delta T < 0$). Since the nuclear energy generation rate $\epsilon_{\rm nuc} \propto \rho T^\nu$ depends on positive powers of ρ and especially T, the total nuclear energy generation will decrease: $\delta L_{\rm nuc} < 0$. Eq. (7.45) shows that the perturbation to $E_{\rm tot}$ will be quenched and the state of thermal equilibrium will be restored.

The secular stability of nuclear burning thus depends on the *negative heat capacity* of stars composed of ideal gas: the property that an increase of the total energy content leads to a decrease of the temperature. This property provides a *thermostat* that keeps the temperature nearly constant and keeps stars in a stable state of thermal equilibrium for such long time scales.

We can generalise this to the case of stars with appreciable radiation pressure. For a mixture of ideal gas and radiation we can write, with the help of eqs. (3.11) and (3.12),

$$\frac{P}{\rho} = \frac{P_{\rm gas}}{\rho} + \frac{P_{\rm rad}}{\rho} = \tfrac{2}{3} u_{\rm gas} + \tfrac{1}{3} u_{\rm rad}. \tag{7.46}$$

Applying the virial theorem in its general form, eq. (2.24), this yields

$$2\, E_{\rm int,gas} + E_{\rm int,rad} = -E_{\rm gr} \tag{7.47}$$

and the total energy becomes

$$E_{\rm tot} = -E_{\rm int,gas} = \tfrac{1}{2}(E_{\rm gr} + E_{\rm int,rad}). \tag{7.48}$$

The radiation pressure thus has the effect of reducing the effective gravitational potential energy. If $\beta = P_{\rm gas}/P$ is constant throughout the star, then eq. (7.48) becomes

$$E_{\rm tot} = \tfrac{1}{2} \beta\, E_{\rm gr} \tag{7.49}$$

This is negative as long as $\beta > 0$. The analysis of thermal stability is analogous to the case of an ideal gas treated above, and we see that stars in which radiation pressure is important, but not dominant, are still secularly stable. However, if $\beta \to 0$ then the thermostatic effect no longer works.

Thermal instability of degenerate gases

In the case of a degenerate electron gas, the pressure and the internal energy are independent of the temperature (Sec. 3.3.5). The mechanical structure of an electron-degenerate star – or the degenerate core of an evolved star – is therefore independent of the thermal-energetic structure (Sec. 7.3). If the same perturbation $L_{\rm nuc} > L$ discussed above is applied to a degenerate gas, the resulting energy input will have no effect on the electron pressure and on the stellar structure. Therefore there will be no expansion and cooling. Instead, there will be a temperature *increase* because the ionized atomic nuclei still behave as an ideal gas, and the energy input will increase their thermal motions. Thus the effect of the perturbation will be $\delta T > 0$, while $\delta\rho \approx 0$.

Because of the strong sensitivity of the nuclear energy generation rate to T, the perturbation will now lead to an increase of $L_{\rm nuc}$, and thermal equilibrium will not be restored. Instead, the temperature will continue to rise as a result of the increased nuclear energy release, which in turn leads to further enhancement of the energy generation. This instability is called a *thermonuclear runaway*, and it occurs whenever nuclear reactions ignite in a degenerate gas. In some cases it can lead to the explosion of the star, although a catastrophic outcome can often be avoided when the gas eventually becomes sufficiently hot to behave as an ideal gas, for which the stabilizing thermostat operates. This can be seen from eq. (7.41), valid in the case of homologous expansion, which we can write as

$$\frac{\delta T}{T} = \frac{1}{\chi_T}\left(\frac{4}{3} - \chi_\rho\right)\frac{\delta\rho}{\rho}. \tag{7.50}$$

As soon as the gas is heated enough that it is no longer completely degenerate, $\chi_T > 0$ and some expansion will occur ($\delta\rho < 0$), while χ_ρ decreases below $\frac{5}{3}$. From eq. (7.50) we see that when χ_ρ drops below the critical value of $\frac{4}{3}$, δT changes sign and becomes negative upon further expansion.

We shall encounter several examples of thermonuclear runaways in future chapters. The most common occurrence is the ignition of helium fusion in stars with masses below about $2\,M_\odot$ – this phenomenon is called the *helium flash*. Thermonuclear runaways also occur when hydrogen gas accumulates on the surface of a white dwarf, giving rise to so-called *nova outbursts*.

The thin shell instability

In evolved stars, nuclear burning can take place in a shell around an inert core. If such a burning shell is sufficiently thin the burning may become thermally unstable, even under ideal-gas conditions. We can make this plausible by considering a shell with mass Δm inside a star with radius R, located between a fixed inner boundary at r_0 and outer boundary at r, so that its thickness is $d = r - r_0 \ll R$. If the shell is in thermal equilibrium, the rate of nuclear energy generation equals the net rate of heat flowing out of the shell (eq. 7.14). A perturbation by which the energy generation rate exceeds the rate of heat flow leads to expansion of the shell, pushing the layers above it outward ($\delta r > 0$). This leads to a decreased pressure, which in hydrostatic equilibrium is given by eq. (7.40),

$$\frac{\delta P}{P} = -4\frac{\delta r}{r}. \tag{7.51}$$

The mass of the shell is $\Delta m = 4\pi {r_0}^2 \rho\, d$, and therefore the density varies with the thickness of the shell as

$$\frac{\delta\rho}{\rho} = -\frac{\delta d}{d} = -\frac{\delta r}{r}\,\frac{r}{d}. \tag{7.52}$$

Eliminating $\delta r/r$ from the above equations yields a relation between the changes in pressure and density,

$$\frac{\delta P}{P} = 4\frac{d}{r}\frac{\delta\rho}{\rho}. \tag{7.53}$$

Combining with the equation of state in its general, differential form eq. (3.48) we can eliminate $\delta P/P$ to obtain the resulting change in temperature,

$$\frac{\delta T}{T} = \frac{1}{\chi_T}\left(4\frac{d}{r} - \chi_\rho\right)\frac{\delta\rho}{\rho}. \tag{7.54}$$

The shell is thermally stable as long as expansion results in a drop in temperature, i.e. when

$$4\frac{d}{r} > \chi_\rho \tag{7.55}$$

since $\chi_T > 0$. Thus, for a sufficiently thin shell a thermal instability will develop. (In the case of an ideal gas, the condition 7.55 gives $d/r > 0.25$, but this is only very approximate.) If the shell is very thin, the expansion does not lead to a sufficient decrease in pressure to yield a temperature drop, even in the case of an ideal gas. This may lead to a runaway situation, analogous to the case of a degenerate gas. The thermal instability of thin burning shells is important during late stages of evolution of stars up to about 8 $M_\odot$, during the *asymptotic giant branch*.

Suggestions for further reading

The contents of this chapter are also (partly) covered by Chapter 24 of Maeder, where the question of stability is considered in Section 3.5. A more complete coverage of the material is given in Chapters 9, 10, 19, 20 and 25 of Kippenhahn.

Exercises

7.1 General understanding of the stellar evolution equations

The differential equations (7.1–7.5) describe, for a certain location in the star at mass coordinate m, the behaviour of and relations between radius coordinate r, the pressure P, the temperature T, the luminosity l and the mass fractions X_i of the various elements i.

(a) Which of these equations describe the mechanical structure, which describe the thermal-energetic structure and which describe the composition?

(b) What does ∇ represent? Which two cases do we distinguish?

(c) How does the set of equations simplify when we assume hydrostatic equilibrium (HE)? If we assume HE, which equation introduces a time dependence? Which physical effect does this time dependence represent?

(d) What do the terms $\epsilon_{\rm nuc}$ and $T\partial s/\partial t$ represent?

(e) How does the set of equations simplify if we also assume thermal equilibrium (TE)? Which equation introduces a time dependence in TE?

(f) Equation (7.5) describes the changes in the composition. In principle we need one equation for every possible isotope. In most stellar evolution codes, the nuclear network is simplified. This reduces the number of differential equations and therefore increases speed of stellar evolution codes. The `STARS` code behind *Window to the Stars* only takes into account seven isotopes. Which do you think are most important?

7.2 Dynamical Stability

(a) Show that for a star in hydrostatic equilibrium ($\mathrm{d}P/\mathrm{d}m = -Gm/(4\pi r^4)$) the pressure scales with density as $P \propto \rho^{4/3}$.

(b) If $\gamma_{\mathrm{ad}} < 4/3$ a star becomes dynamically unstable. Explain why.

(c) In what type of stars $\gamma_{\mathrm{ad}} \approx 4/3$?

(d) What is the effect of partial ionization (for example $\mathrm{H} \leftrightarrows \mathrm{H}^+ + \mathrm{e}^-$) on γ_{ad}? So what is the effect of ionization on the stability of a star?

(e) *Pair creation* and *photo-disintegration* of Fe have a similar effect on γ_{ad}. In what type of stars (and in what phase of their evolution) do these processes play a role?

7.3 Mass radius relation for degenerate stars

(a) Derive how the radius scales with mass for stars composed of a *non-relativistic completely degenerate* electron gas. Assume that the central density $\rho_c = a\overline{\rho}$ and that the central pressure $P_c = bGM^2/R^4$, where $\overline{\rho}$ is the mean density, and a and b depend only on the density distribution inside the star.

(b) Do the same for an *extremely relativistic degenerate* electron gas.

(c) The electrons in a not too massive white dwarf behave like a completely degenerate non-relativistic gas. Many of these white dwarfs are found in binary systems. Describe qualitatively what happens if the white dwarf accretes material from the companion star.

7.4 Main-sequence homology relations

We speak of two *homologous stars* when they have the same density distribution. To some extent main sequence stars can be considered as stars with a similar density distribution.

(a) You already derived some scaling relations for main sequence stars from observations in the first set of exercises: the mass-luminosity relation and the mass-radius relation. Over which mass range were these simple relations valid?

(b) During the practicum you plotted the density distribution of main sequence stars of different masses. For which mass ranges did you find that that the stars had approximately the same density distribution.

(c) Compare the L-M relation derived form observational data with the L-M relation derived from homology, eq. (7.32). What could cause the difference? (Which assumptions may not be valid?)

(d) Show that, if we replace the assumption of a constant opacity with a Kramers opacity law, the mass-luminosity-radius relation becomes eq. (7.33),

$$L \propto \frac{\mu^{7.5} M^{5.5}}{R^{0.5}}.$$

(e) Substitute a suitable mass-radius relation and compare the result of (d) with the observational data in Fig. 1.3. For which stars is the Kramers-based L-M relation the best approximation? Can you explain why? What happens at lower and higher masses, respectively?

7.5 Central behaviour of the stellar structure equations

(a) Rewrite the four structure equations in terms of $\mathrm{d}/\mathrm{d}r$.

(b) Find how the following quantities behave in the neighbourhood of the stellar center:
- the mass $m(r)$,
- the luminosity $l(r)$,
- the pressure $P(r)$,
- the temperature $T(r)$.

Chapter 8

Schematic stellar evolution – consequences of the virial theorem

8.1 Evolution of the stellar centre

We will consider the schematic evolution of a star, as seen from its centre. The centre is the point with the highest pressure and density, and (usually) the highest temperature, where nuclear burning proceeds fastest. Therefore, the centre is the most evolved part of the star, and it sets the pace of evolution, with the outer layers lagging behind.

The stellar centre is characterized by the central density ρ_c, pressure P_c and temperature T_c and the composition (usually expressed in terms of μ and/or μ_e). These quantities are related by the equation of state (EOS). We can thus represent the evolution of a star by an evolutionary track in the (P_c, ρ_c) diagram or the (T_c, ρ_c) diagram.

8.1.1 Hydrostatic equilibrium and the P_c-ρ_c relation

Consider a star in hydrostatic equilibrium (HE), for which we can estimate how the central pressure scales with mass and radius from the homology relations (Sec. 7.4). For a star that expands or contracts homologously, we can apply eq. (7.26) to the central pressure and central density to yield

$$\boxed{P_c = C \cdot GM^{2/3}{\rho_c}^{4/3}} \tag{8.1}$$

where C is a constant. This is a fundamental relation for stars in HE: *in a star of mass M that expands or contracts homologously, the central pressure varies as central density to the power* $\frac{4}{3}$. The value of the constant C depends on the density distribution in the star. Note that we found the same relation for polytropic stellar models in Chapter 4, eq. (4.18), where $C = C_n$ depends on the polytopic index. However, the dependence on n, and hence on the density distibution, is only very weak. For polytropic models with index $n = 1.5 - 3$, a range that encompasses most actual stars, C varies between 0.48 and 0.36. Hence relation (8.1) is reasonably accurate, even if the contraction is not exactly homologous. In other words: for a star of a certain mass, the central pressure is almost uniquely determined by the central density.

Note that we have obtained this relation without considering the EOS. Therefore (8.1) defines a *universal relation for stars in HE* that is independent of the equation of state. It expresses the fact that a star that contracts quasi-statically must achieve a higher internal pressure to remain in hydrostatic equilibrium. Eq. (8.1) therefore defines an *evolution track* of a slowly contracting (or expanding) star in the P_c-ρ_c plane.

8.1.2 The equation of state and evolution in the P_c-ρ_c plane

By considering the EOS we can also derive the evolution of the central temperature. This is obviously crucial for the evolutionary fate of a star because e.g. nuclear burning requires T_c to reach certain (high) values. We start by considering lines of constant T, *isotherms*, in the (P, ρ) plane.

We have encountered various regimes for the EOS in Chapter 3:

- Radiation dominated: $P = \frac{1}{3}aT^4$. Hence an isotherm in this region is also a line of constant P.
- (Classical) ideal gas: $P = \dfrac{\mathcal{R}}{\mu}\rho T$. Hence an isotherm has $P \propto \rho$.
- Non-relativistic electron degeneracy: $P = K_{\rm NR}(\rho/\mu_e)^{5/3}$ (eq. 3.35). This is independent of temperature. More accurately: the complete degeneracy implied by this relation is only achieved when $T \to 0$, so this is in fact the isotherm for $T = 0$ (and not too high densities).
- Extremely relativistic electron degeneracy: $P = K_{\rm ER}(\rho/\mu_e)^{4/3}$ (eq. 3.37). This is the isotherm for $T = 0$ and very high ρ.

Figure 8.1 shows various isotherms schematically in the $\log\rho - \log P$ plane. Where radiation pressure dominates (low ρ) the isotherms are horizontal and where ideal-gas pressure dominates they have a slope = 1. The isotherm for $T = 0$, corresponding to complete electron degeneracy, has slope of $\frac{5}{3}$ at relatively low density and a shallower slope of $\frac{4}{3}$ at high density. The region to the right and below the $T = 0$ line is forbidden by the Pauli exclusion principle, since electrons are fermions.

The dashed lines are schematic evolution tracks for stars of different masses M_1 and M_2. According to eq. (8.1) they have a slope of $\frac{4}{3}$ and the track for a larger mass lies at a higher pressure than that for a smaller mass.

Several important conclusions can be drawn from this diagram:

- As long as the gas is ideal, contraction (increasing ρ_c) leads to a higher T_c, because the slope of the evolution track is steeper than that of ideal-gas isotherms: the evolution track crosses isotherms of higher and higher temperature.

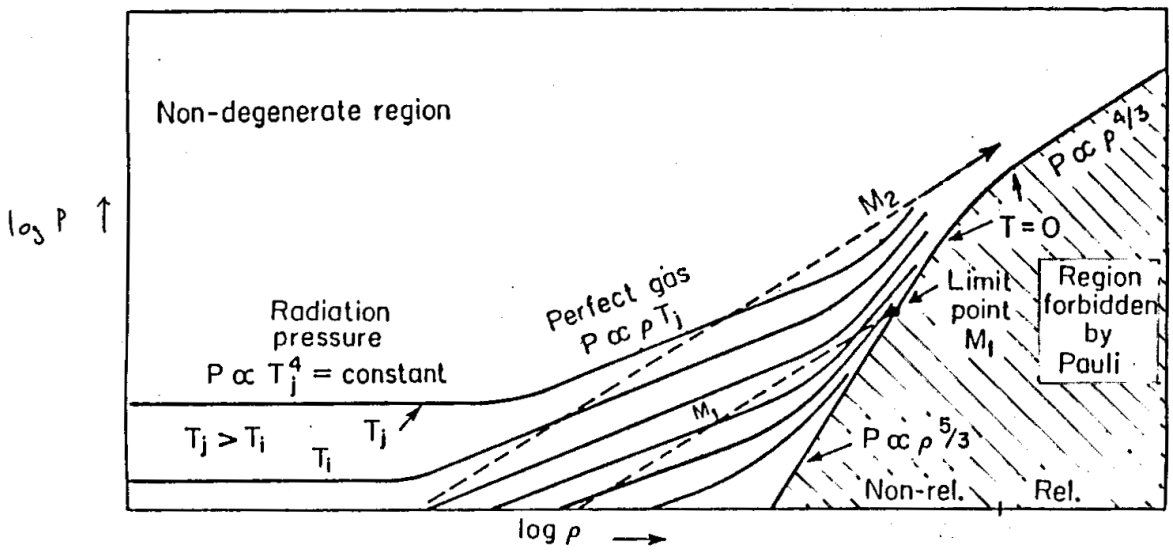


Figure 8.1. Schematic evolution in the $\log\rho - \log P$ plane. Solid lines are isotherms in the equation of state; the dashed lines indicate two evolution tracks of different mass, which have a slope of $\frac{4}{3}$. See the text for an explanation.

Note that this is consistent with what we have already concluded from the virial theorem for an ideal gas: $E_{\rm in} = -\frac{1}{2}E_{\rm gr}$. Contraction (decreasing $E_{\rm gr}$, i.e. increasing $-E_{\rm gr}$) leads to increasing internal (thermal) energy of the gas, i.e. to heating of the stellar gas!

- Tracks for masses lower than some critical value $M_{\rm crit}$, e.g. the track labeled M_1, eventually run into the line for complete electron degeneracy because this has a steeper slope. Hence for stars with $M < M_{\rm crit}$ there exist a maximum achievable central density and pressure, $\rho_{\rm c,max}$ and $P_{\rm c,max}$, which define the endpoint of their evolution. This endpoint is a completely degenerate state, i.e. a white dwarf, where the pressure needed to balance gravity comes from electrons filling the lowest possible quantum states.

  Because complete degeneracy corresponds to $T = 0$, it follows that the evolution track must intersect each isotherm twice. In other words, stars with $M < M_{\rm crit}$ also reach a maximum temperature $T_{\rm c,max}$, at the point where degeneracy starts to dominate the pressure, after which further contraction leads to decreasing T_c. $\rho_{\rm c,max}$, $P_{\rm c,max}$ and $T_{\rm c,max}$ all depend on M and increase with mass.

- Tracks for masses larger than $M_{\rm crit}$, e.g. the one labeled M_2, miss the completely degenerate region of the EOS, because at high ρ this has the same slope as the evolution track. This change in slope is owing to the electrons becoming relativistic and as their velocity cannot exceed c, they exert less pressure than if there were no limit to their velocity. Hence, electron degeneracy is not sufficient to counteract gravity, and a star with $M > M_{\rm crit}$ must keep on contracting and getting hotter indefinitely – up to the point where the assumptions we have made break down, e.g. when ρ becomes so high that the protons inside the nuclei capture free electrons and a neutron gas is formed, which can also become degenerate.

Hence the evolution of stars with $M > M_{\rm crit}$ is qualitatively different from that of stars with $M < M_{\rm crit}$. This critical mass is none other than the *Chandrasekhar mass* that we have already encountered in Chapter 4 (eq. 4.22)

$$M_{\rm Ch} = \frac{5.836}{{\mu_{\rm e}}^2}\, M_\odot. \tag{8.2}$$

It is the unique mass of a completely degenerate and extremely relativistic gas sphere. A star with $M \geq M_{\rm Ch}$ must collapse under its own gravity, but the electrons become extremely relativistic – and, if M is equal to or not much larger than $M_{\rm Ch}$, also degenerate – in the process.

8.1.3 Evolution in the T_c-ρ_c plane

We now consider how the stellar centre evolves in the T_c, ρ_c diagram. First we divide the T, ρ plane into regions where different processes dominate the EOS, see Sec. 3.3.7 and Fig. 3.4, reproduced in Fig. 8.2a.

For a slowly contracting star in hydrostatic equilibrium equation (8.1) implies that, as long as the gas behaves like a classical ideal gas:

$$\frac{\mathcal{R}}{\mu} T_c\, \rho_c = C\, G M^{2/3} {\rho_c}^{4/3} \qquad \rightarrow \qquad \boxed{T_c = \frac{C\,G}{\mathcal{R}}\, \mu\, M^{2/3} {\rho_c}^{1/3}} \tag{8.3}$$

(Compare to Sec. 7.4.3.) This defines an evolution track in the $\log T$, $\log \rho$ plane with slope $\frac{1}{3}$. Stars with different mass evolve along tracks that lie parallel to each other, those with larger M lying at higher T_c and lower ρ_c. Tracks for larger mass therefore lie closer to the region where radiation

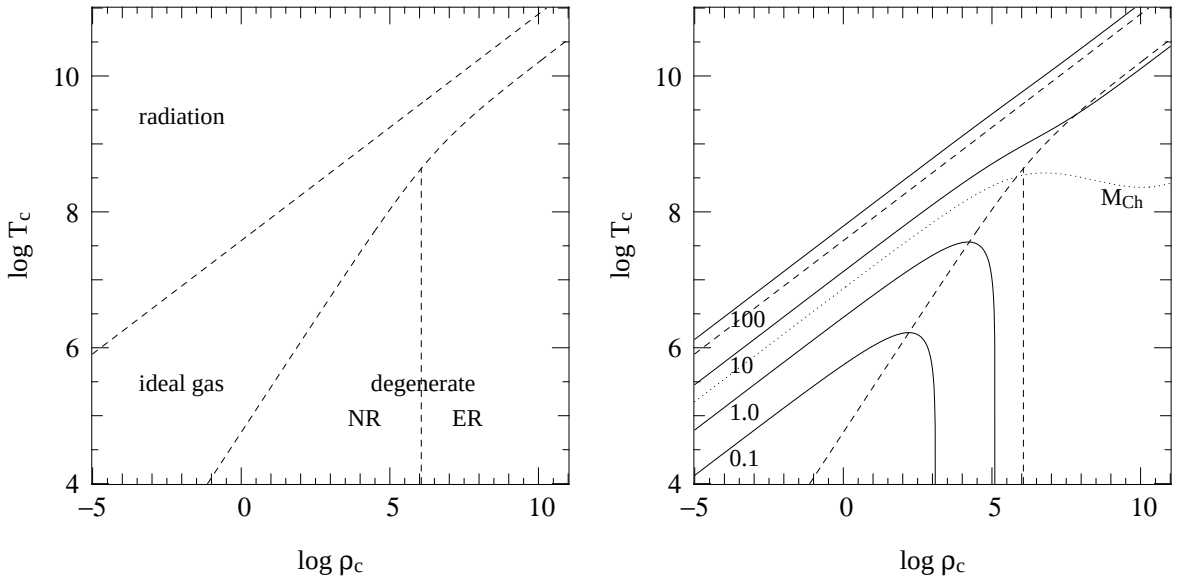


Figure 8.2. The equation of state in the $\log T_c - \log \rho_c$ plane (left panel), with approximate boundaries between regions where radiation pressure, ideal gas pressure, non-relativistic electron degeneracy and extremely relativistic electron degeneracy dominate, for a composition of $X = 0.7$ and $Z = 0.02$. In the right panel, schematic evolution tracks for contracting stars of $0.1 - 100\, M_\odot$ have been added.

pressure is important: *the larger the mass of a star, the more important is the radiation pressure*. Furthermore, the relative importance of radiation pressure does not change as a star contracts, because the track runs parallel to the boundary between ideal gas and radiation pressure.[1]

As the density increases, stars with $M < M_{\rm Ch}$ approach the region where non-relativistic electron degeneracy dominates, because the boundary between ideal gas and NR degeneracy has a steeper slope than the evolution track. Inside this region, equating relation (8.1) to the NR degenerate pressure gives:

$$K_{\rm NR}\frac{{\rho_c}^{1/3}}{{\mu_e}^{5/3}} = C\,G M^{2/3} \qquad \rightarrow \qquad \boxed{\rho_c = \left(\frac{C\,G}{K_{\rm NR}}\right)^3 {\mu_e}^5 M^2} \tag{8.4}$$

When degeneracy dominates the track becomes independent of T_c, and the star moves down along a track of constant ρ_c. This is the $\rho_{\rm c,max}$ we found from the P_c, ρ_c diagram. The larger the mass, the higher this density. (When the electrons become relativistic at $\rho_c \gtrsim 10^6$ g/cm^3, the pressure increases less steeply with density so that the central density for a degenerate star of mass M is in fact larger than given by eq. 8.4).

Equations (8.3) and (8.4) imply that, for a star with $M < M_{\rm Ch}$ that contracts quasi-statically, T_c increases as ${\rho_c}^{1/3}$ until the electrons become degenerate. Then a maximum temperature is reached, and subsequently the star cools at a constant density when degenerate electrons provide the pressure. The schematic evolution tracks for 0.1 and $1.0\, M_\odot$ given in Fig. 8.2 show this behaviour. This can be compared to eq. (7.41) for homologous contraction (Sec. 7.4.3), which indicates that the slope of an evolution track in the $\log T - \log \rho$ plane is equal to $(\frac{4}{3} - \chi_\rho)/\chi_T$. This equals $\frac{1}{3}$ for an ideal gas, but changes sign and becomes negative once $\chi_\rho > \frac{4}{3}$. When degeneracy is almost complete, $\chi_\rho = \frac{5}{3}$ and $\chi_T \rightarrow 0$ such that the slope approaches infinity. The maximum temperature is reached when the

[1] It is easy to show for yourself that the evolution track for a star in which radiation pressure dominates would have the same slope of $\frac{1}{3}$ in the $\log T$, $\log \rho$ plane. However, such stars are very close to dynamical instability.

ideal gas pressure and degenerate electron pressure are about equal, each contributing about half of the total pressure. Combining eqs. (8.3) and (8.4) then implies that the maximum central temperature reached increases with stellar mass as (see Exercise 8.2)

$$T_{\rm c,max} = \frac{C^2 G^2}{4\mathcal{R} K_{\rm NR}} \mu \mu_{\rm e}^{5/3} M^{4/3}. \tag{8.5}$$

For $M > M_{\rm Ch}$, the tracks of contracting stars miss the degenerate region of the T-ρ plane, because at high density the boundary between ideal gas and degeneracy has the same slope as an evolution track. The pressure remains dominated by an ideal gas, and T_c keeps increasing like ${\rho_c}^{1/3}$ to very high values ($> 10^{10}$ K). This behaviour is shown by the schematic tracks for 10 and 100 $M_\odot$.

8.2 Nuclear burning regions and limits to stellar masses

We found that stars with $M < M_{\rm Ch}$ reach a maximum temperature, the value of which increases with mass. This means that only gas spheres above a certain mass limit will reach temperatures sufficiently high for nuclear burning. The nuclear energy generation rate is a sensitive function of the temperature, which can be written as

$$\epsilon_{\rm nuc} = \epsilon_0 \rho^\lambda T^\nu \tag{8.6}$$

where for most nuclear reactions (those involving two nuclei) $\lambda = 1$, while ν depends mainly on the masses and charges of the nuclei involved and usually $\nu \gg 1$. For H-burning by the pp-chain, $\nu \approx 4$ and for the CNO-cycle which dominates at somewhat higher temperature, $\nu \approx 18$. For He-burning by the 3α reaction, $\nu \sim 40$ (and $\lambda = 2$ because three particles are involved). For C-burning and O-burning reactions ν is even larger. As discussed in Chapter 6, the consequences of this strong temperature sensitivity are that

- each nuclear reaction takes place at a particular, nearly constant temperature, and
- nuclear burning cycles of subsequent heavier elements are well separated in temperature

As a star contracts and heats up, nuclear burning becomes important when the energy generated, $L_{\rm nuc} = \int \epsilon_{\rm nuc}\, {\rm d}m$, becomes comparable to the energy radiated away from the surface, L. From this moment on, the star can compensate its surface energy loss by nuclear energy generation: it comes into *thermal equilibrium*. The first nuclear fuel to be ignited is hydrogen, at $T_c \sim 10^7$ K. From the homology relation (7.38) we expect that the central temperature at which hydrogen fusion stabilizes should depend on the mass approximately as

$$T_c = T_{c,\odot} \, (M/M_\odot)^{4/(\nu+3)} \tag{8.7}$$

where $T_{c,\odot} \approx 1.5 \times 10^7$ K, for a composition like that of the Sun ($\mu = \mu_\odot$). We can estimate the minimum mass required for hydrogen burning by comparing this temperature to the maximum central temperature reached by a gas sphere of mass M, eq. (8.5). By doing this (and taking $C = 0.48$ for an $n = 1.5$ polytrope) we find a minimum mass for hydrogen burning of $0.15\, M_\odot$.

Detailed calculations reveal that the minimum mass for the ignition of hydrogen in protostars is about 2 times smaller than this simple estimate, $M_{\rm min} = 0.08\, M_\odot$. Less massive objects become partially degenerate before the required temperature is reached and continue to contract and cool without ever burning hydrogen. Such objects are not stars according to our definition (Chapter 1) but are known as *brown dwarfs*.

We have seen earlier that the contribution of radiation pressure increases with mass, and becomes dominant for $M \gtrsim 100\, M_\odot$. A gas dominated by radiation pressure has an adiabatic index $\gamma_{\rm ad} = \frac{4}{3}$,

which means that hydrostatic equilibrium in such stars becomes marginally unstable (see Sec. 7.5.1). Therefore stars much more massive than 100 $M_\odot$ should be very unstable, and indeed none are known to exist (while those with $M > 50\,M_\odot$ indeed show signs of being close to instability, e.g. they lose mass very readily).

Hence stars are limited to a rather narrow mass range of $\sim 0.1\,M_\odot$ to $\sim 100\,M_\odot$. The lower limit is set by the minimum temperature required for nuclear burning, and the upper limit by the requirement of dynamical stability.

8.2.1 Overall picture of stellar evolution and nuclear burning cycles

As a consequence of the virial theorem, a self-gravitating sphere composed of ideal gas in HE must contract and heat up as it radiates energy from the surface. The energy loss occurs at a rate

$$L = -\dot{E}_{\rm tot} = \dot{E}_{\rm in} = -\tfrac{1}{2}\dot{E}_{\rm gr} \approx \frac{E_{\rm gr}}{\tau_{\rm KH}} \tag{8.8}$$

This is the case for protostars that have formed out of an interstellar gas cloud. Their evolution, i.e. overall contraction, takes place on a thermal timescale $\tau_{\rm KH}$. As the protostar contracts and heats up and its central temperature approaches 10^7 K, the nuclear energy generation rate (which is at first negligible) increases rapidly in the centre, until the burning rate matches the energy loss from the surface:

$$L = -\dot{E}_{\rm nuc} \approx \frac{E_{\rm nuc}}{\tau_{\rm nuc}} \tag{8.9}$$

At this point, contraction stops and T_c and ρ_c remain approximately constant, at the values needed for hydrogen burning. The stellar centre occupies the same place in the T_c-ρ_c diagram for about a nuclear timescale $\tau_{\rm nuc}$. Remember that for a star of a certain mass, L is essentially determined by the opacity, i.e. by how efficiently the energy can be transported outwards.

When H is exhausted in the core – which how consists of He and has a mass typically ~10% of the total mass M – this helium core resumes its contraction. Meanwhile the layers around it expand. This constitutes a large deviation from homology and relation (8.1) no longer applies to the whole star. However the core itself still contracts more or less homologously, while the weight of the envelope decreases as a result of its expansion. Therefore relation (8.1) remains approximately valid for the *core* of the star, i.e. if we replace M by the core mass $M_{\rm c}$. The core continues to contract and heat up at a pace set by its own thermal timescale,

$$L_{\rm core} \approx \dot{E}_{\rm in,core} \approx -\tfrac{1}{2}\dot{E}_{\rm gr,core} \approx \frac{E_{\rm gr,core}}{\tau_{\rm KH,core}} \tag{8.10}$$

as long as the gas conditions remain ideal. It is now the He core mass, rather than the total mass of the star, that determines the further evolution.

Arguments similar to those used for deriving the minimum mass for H-burning lead to the existence a minimum (core) mass for He-ignition, This is schematically depicted in Fig. 8.3, which suggests that this minimum mass is larger than 1 $M_\odot$. However, the schematic tracks in Fig. 8.3 have been calculated for a fixed composition $X = 0.7$, $Z = 0.02$, which is clearly no longer the valid since the core is composed of helium. You may verify that a He-rich composition increases the maximum central temperature reached for a certain mass (eq. 8.5). Detailed calculations put the minimum mass for He-ignition at $\approx 0.3\,M_\odot$. Stars with a core mass larger than this value ignite He in the centre when $T_c \approx 10^8$ K, which stops further contraction while the energy radiated away can be supplied

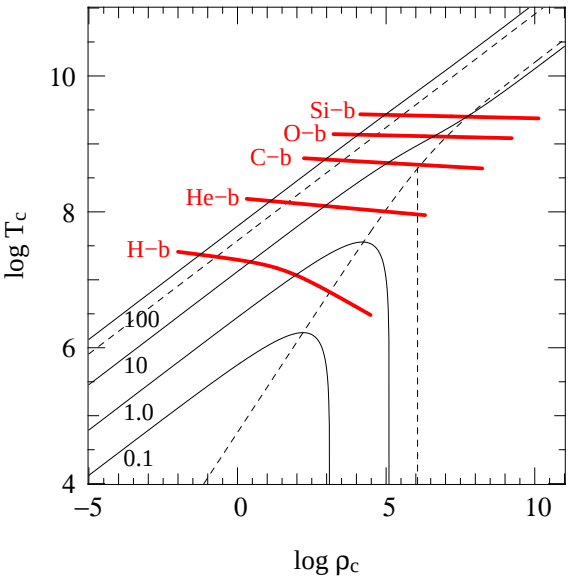


Figure 8.3. The same schematic evolution tracks as in Fig. 8.2, together with the approximate regions in the $\log T_c - \log \rho_c$ plane where nuclear burning stages occur.

by He-burning reactions. This can go on for a length of time equal to the nuclear timescale of He burning, which is about 0.1 times that of H burning. In stars with a He core mass $< 0.3\,M_\odot$ the core becomes degenerate before reaching $T_c = 10^8$ K, and in the absence of a surrounding envelope it would cool to become a white dwarf composed of helium, as suggested by Fig. 8.3. (In practice, however, H-burning in a shell around the core keeps the core hot and when M_c has grown to $\approx 0.5\,M_\odot$ He ignites in a degenerate flash.)

After the exhaustion of He in the core, the core again resumes its contraction on a thermal timescale, until the next fuel can be ignited. Following a similar line of reasoning the minimum (core) mass for C-burning, which requires $T \approx 5 \times 10^8$ K, is $\approx 1.1\,M_\odot$. Less massive cores are destined to never ignite carbon but to become degenerate and cool as CO white dwarfs. The minimum core mass required for the next stage, Ne-burning, turns out to be $\approx M_{\rm Ch}$. Stars that develop cores with $M_c > M_{\rm Ch}$ therefore also undergo all subsequent nuclear burning stages (Ne-, O- and Si-burning) because they never become degenerate and continue to contract and heat after each burning phase. Eventually they develop a core consisting of Fe, from which no further nuclear energy can be squeezed. The Fe core must collapse in a cataclysmic event (a supernova or a gamma-ray burst) and become a neutron star or black hole.

The alternation of gravitational contraction and nuclear burning stages is summarized in Table 8.1, together with the corresponding minimum masses and characteristic temperatures and energies. The schematic picture presented in Fig. 8.3 of the evolution of stars of different masses in the T–ρ diagram can be compared to Fig. 8.4, which shows the results of detailed calculations for various masses.

To summarize, we have obtained the following picture. Nuclear burning cycles can be seen as long-lived but temporary interruptions of the inexorable contraction of a star (or at least its core) under the influence of gravity. This contraction is dictated by the virial theorem, and a result of the fact that stars are hot and lose energy by radiation. If the core mass is less than the Chandrasekhar mass, then the contraction can eventually be stopped (after one or more nuclear cycles) when electron degeneracy supplies the pressure needed to withstand gravity. However if the core mass exceeds the Chandrasekhar mass, then degeneracy pressure is not enough and contraction, interrupted by nuclear burning cycles, must continue at least until nuclear densities are reached.

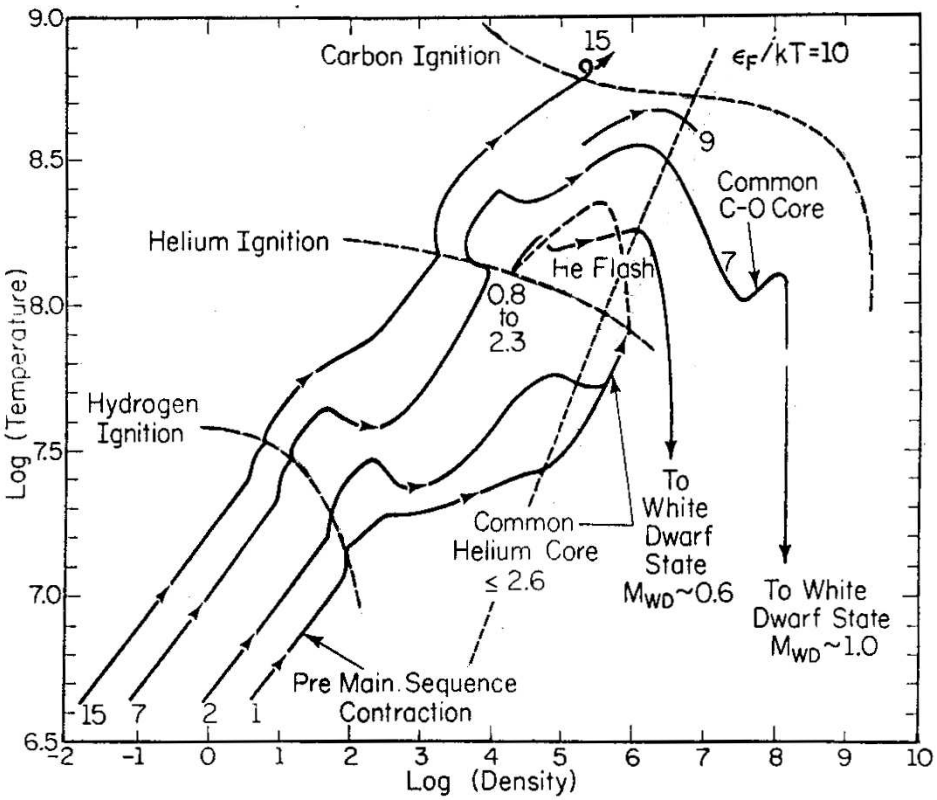


Figure 8.4. Detailed evolution tracks in the $\log\rho_c$ - $\log T_c$ plane for masses between 1 and 15 $M_\odot$. The initial slope of each track (labelled pre-main sequence contraction) is equal to $\frac{1}{3}$ as expected from our simple analysis. When the H-ignition line is reached wiggles appear in the tracks, because the contraction is then no longer strictly homologous. A stronger deviation from homologous contraction occurs at the end of H-burning, because only the core contracts while the outer layers expand. Accordingly, the tracks shift to higher density appropriate for their smaller (core) mass. These deviations from homology occur at each nuclear burning stage. Consistent with our expectations, the most massive star (15 $M_\odot$) reaches C-ignition and keeps evolving to higher T and ρ. The core of the 7 $M_\odot$ star crosses the electron degeneracy border (indicated by $\epsilon_F/kT = 10$) before the C-ignition temperature is reached and becomes a C-O white dwarf. The lowest-mass tracks (1 and 2 $M_\odot$) cross the degeneracy border before He-ignition because their cores are less massive than 0.3 $M_\odot$. Based on our simple analysis we would expect them to cool and become He white dwarfs; however, their degenerate He cores keep getting more massive and hotter due to H-shell burning. They finally do ignite helium in an unstable manner, the so-called He flash.

Suggestions for further reading

The schematic picture of stellar evolution presented above is very nicely explained in Chapter 7 of Prialnik, which was one of the sources of inspiration for this chapter. The contents are only briefly covered by Maeder in Sec. 3.4, and are somewhat scattered throughout Kippenhahn & Weigert, see sections 28.1, 33.1, 33.4 and 34.1.

Table 8.1. Characteristics of subsequent gravitational contraction and nuclear burning stages. Column (3) gives the total gravitational energy emitted per nucleon since the beginning, and column (5) the total nuclear energy emitted per nucleon since the beginning. Column (6) gives the minimum mass required to ignite a certain burning stage (column 4). The last two columns give the fraction of energy emitted as photons and neutrinos, respectively.

phase	T (10^6 K)	total $E_{\rm gr}$/n	main reactions	total $E_{\rm nuc}$/n	$M_{\rm min}$	γ (%)	ν (%)
grav.	$0 \rightarrow 10$	~ 1 keV/n				100	
nucl.	$10 \rightarrow 30$		^{1}H $\rightarrow$ ^{4}He	6.7 MeV/n	0.08 $M_\odot$	~95	~5
grav.	$30 \rightarrow 100$	~ 10 keV/n				100	
nucl.	$100 \rightarrow 300$		^{4}He $\rightarrow$ ^{12}C, ^{16}O	$\approx$ 7.4 MeV/n	0.3 $M_\odot$	~100	~0
grav.	$300 \rightarrow 700$	~ 100 keV/n				~50	~50
nucl.	$700 \rightarrow 1000$		^{12}C $\rightarrow$ Mg, Ne	$\approx$ 7.7 MeV/n	1.1 $M_\odot$	~0	~100
grav.	$1000 \rightarrow 1500$	~ 150 keV/n					~100
nucl.	$1500 \rightarrow 2000$		^{16}O $\rightarrow$ S, Si	$\approx$ 8.0 MeV/n	1.4 $M_\odot$		~100
grav.	$2000 \rightarrow 5000$	~ 400 keV/n	Si $\rightarrow \ldots \rightarrow$ Fe	$\approx$ 8.4 MeV/n			~100

Exercises

8.1 Homologous contraction (1)

(a) Explain in your own words what *homologous contraction* means.

(b) A real star does not evolve homologously. Can you give a specific example? [Think of core versus envelope]

(c) Fig. 8.3 shows the central temperature versus the central density for schematic evolution tracks assuming homologous contraction. Explain qualitatively what we can learn form this figure (nuclear burning cycles, difference between a 1 $M_\odot$ and a 10 $M_\odot$ star, ...)

(d) Fig. 8.4 shows the same diagram with evolution tracks from detailed (i.e. more realistic) models. Which aspects were already present in the schematic evolution tracks? When and where do they differ?

8.2 Homologous contraction (2)

In this question you will derive the equations that are plotted in Figure 8.2b.

(a) Use the homology relations for P and ρ to derive eq. (8.1),

$$P_c = CGM^{2/3}\rho_c^{4/3}$$

To see what happens qualitatively to a contracting star of given mass M, the total gas pressure can be approximated roughly by:

$$P \approx P_{\rm id} + P_{\rm deg} = \frac{\mathcal{R}}{\mu}\rho T + K\left(\frac{\rho}{\mu_e}\right)^\gamma \tag{8.11}$$

where γ varies between $\frac{5}{3}$ (non-relativistic) and $\frac{4}{3}$ (extremely relativistic).

(b) Combine this equation, for the case of NR degeneracy, with the central pressure of a contracting star in hydrostatic equilibrium (eq. 8.1, assuming $C \approx 0.5$) in order to find how T_c depends on ρ_c.

(c) Derive an expression for the maximum central temperature reached by a star of mass M.

8.3 Application: minimum core mass for helium burning

Consider a star that consists completely of helium. Compute an estimate for the minimum mass for which such a star can ignite helium, as follows.

- Assume that helium ignites at $T_c = 10^8$ K.
- Assume that the critical mass can be determined by the condition that the ideal gas pressure and the electron degeneracy pressure are equally important in the star at the moment of ignition.
- Use the homology relations for the pressure and the density. Assume that $P_{c,\odot} = 10^{17}\,\mathrm{g\,cm^{-1}\,s^{-2}}$ and $\rho_{c,\odot} = 60\,\mathrm{g\,cm^{-3}}$.

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 ~ 10 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*.

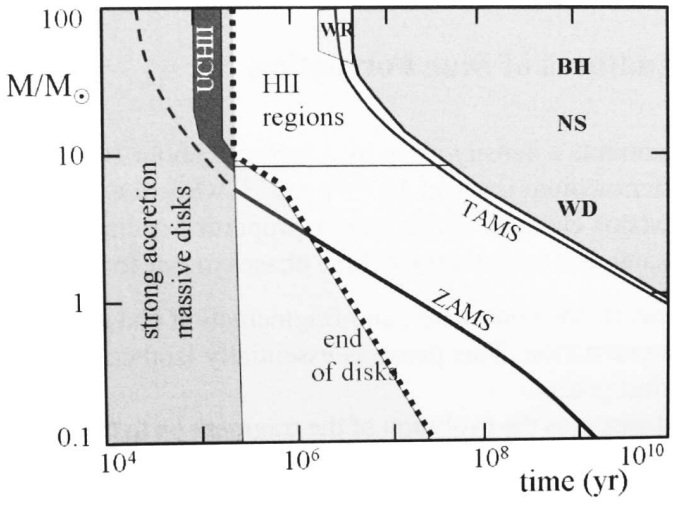


Figure 9.1. Timescales and properties of stars of mass M on the main sequence. Time along the abscissa is in logarithmic units to highlight the early phases, $t = 0$ corresponds to the formation of a hydrostatic core (stage 3 in the text). Initially the star is embedded in a massive accretion disk for $(1 - 2) \times 10^5$ 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.

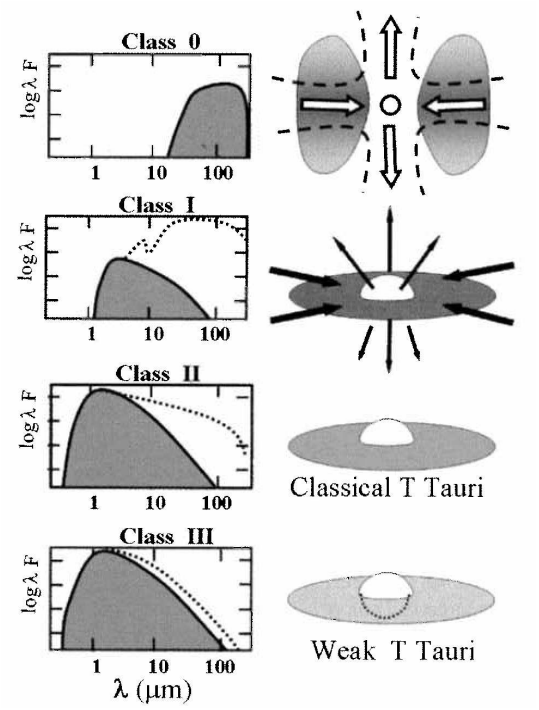


Figure 9.2. Schematic illustration of four stages in the evolution of protostars and their circumstellar disks. On the left, the stellar flux is depicted (shaded area) and the contribution from the disk (dotted line). On the right the corresponding geometry of the object is shown.

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

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

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 $\tau_{\rm acc} = M/\dot{M}$ is much smaller than the thermal timescale $\tau_{\rm KH}$.

Dissociation and ionization The gas initially consists of molecular hydrogen and behaves like an ideal gas, such that $\gamma_{\rm ad} > \frac{4}{3}$ and the protostellar core is dynamically stable. When the core temperature reaches ~ 2000 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 $\gamma_{\rm ad}$ 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 $\tau_{\rm KH}$.

A rough estimate of the radius $R_{\rm p}$ 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 $\chi_{\rm H_2} = 4.48\,{\rm eV}$ per H_2 molecule) and ionization of hydrogen ($\chi_{\rm H} = 13.6\,{\rm eV}$) and helium ($\chi_{\rm He} = 79\,{\rm eV}$). 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, $E_{\rm tot} = \frac{1}{2}E_{\rm gr}$. Taking $E_{\rm gr}$ as given by eq. (2.28), we can write

$$\frac{\alpha}{2}\frac{GM^2}{R_{\rm p}} \approx \frac{M}{m_{\rm u}}\left(\frac{X}{2}\chi_{\rm H_2} + X\chi_{\rm H} + \frac{Y}{4}\chi_{\rm He}\right) \equiv \frac{M}{m_{\rm u}}\chi. \tag{9.3}$$

Taking $X = 0.72$ and $Y = 1 - X$, we have $\chi = 16.9\,{\rm eV}$ 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

$$R_{\rm p} \approx \frac{\alpha}{2}\frac{GMm_{\rm u}}{\chi} \approx 50\,R_\odot\left(\frac{M}{M_\odot}\right). \tag{9.4}$$

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

$$\bar{T} \approx \frac{\alpha}{3}\frac{\mu}{\mathcal{R}}\frac{GM}{R_{\rm p}} = \frac{2}{3}\frac{\mu}{k}\chi \approx 8\times10^4\,{\rm K}, \tag{9.5}$$

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 $T_{\rm eff}$ 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 $\nabla - \nabla_{\rm ad}$ 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 $\nabla_{\rm ad}$. 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 $T_{\rm eff} \approx$ 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 $T_{\rm eff}$) 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 $\nabla = \nabla_{\rm ad}$. For an ideal gas we have a constant $\nabla_{\rm ad} = 0.4$, if we ignore the variation of $\nabla_{\rm ad}$ 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 $\gamma_{\rm ad} = \frac{5}{3}$. 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 $T_{\rm eff}$, we can write the boundary conditions as

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

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

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

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

The equation of state in the photospheric layer is

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

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, $T_{\rm eff}$, L and R. The solution thus always contains one free parameter, that is, the solution is a relation between two quantities, say L and $T_{\rm eff}$. 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 $T_{\rm eff}$ after eliminating all other unknowns. The solution can be written as

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

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

$$A = \frac{\frac{3}{2}a - \frac{1}{2}}{9a + 2b + 3} \qquad \text{and} \qquad B = \frac{a+3}{9a+2b+3}. \tag{9.13}$$

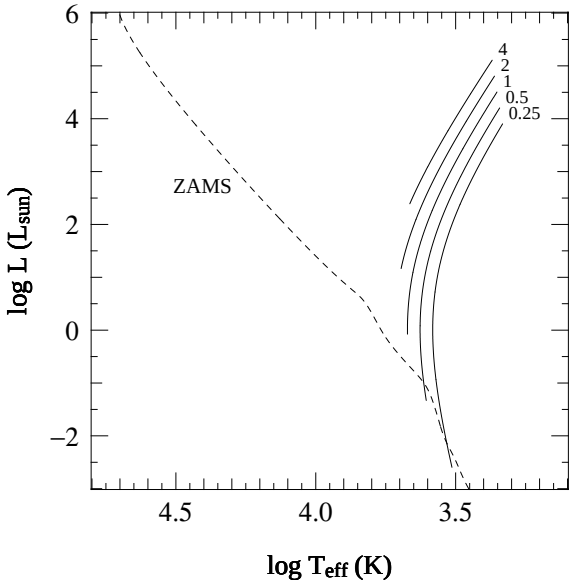


Figure 9.3. The position of the Hayashi lines in the H-R diagram for masses M = 0.25, 0.5, 1.0, 2.0 and 4.0 $M_\odot$ as indicated. The lines are analytic fits to detailed models computed for composition $X = 0.7, Z = 0.02$. The zero-age main sequence (ZAMS) for the same composition is shown as a dashed 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 $\nabla_{\rm ad} < 0.4$) 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 $T_{\rm eff}$ 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 $\nabla = \nabla_{\rm ad}$ throughout, because otherwise they would be *on* the Hayashi line. Defining $\overline{\nabla}$ 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 $T_{\rm eff}$) have either $\overline{\nabla} > \nabla_{\rm ad}$ or $\overline{\nabla} < \nabla_{\rm ad}$. It turns out (after more tedious analysis of the above equations and their dependence on polytropic index n) that models with $\overline{\nabla} < \nabla_{\rm ad}$ lie at higher $T_{\rm eff}$ than the Hayashi line (to its left in the HRD) while models with $\overline{\nabla} > \nabla_{\rm ad}$ lie at lower $T_{\rm eff}$ (to the right in the HRD).

Now consider the significance of $\overline{\nabla} \neq \nabla_{\rm ad}$. If on average $\overline{\nabla} < \nabla_{\rm ad}$ then some part of the star must have $\nabla < \nabla_{\rm ad}$, 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).

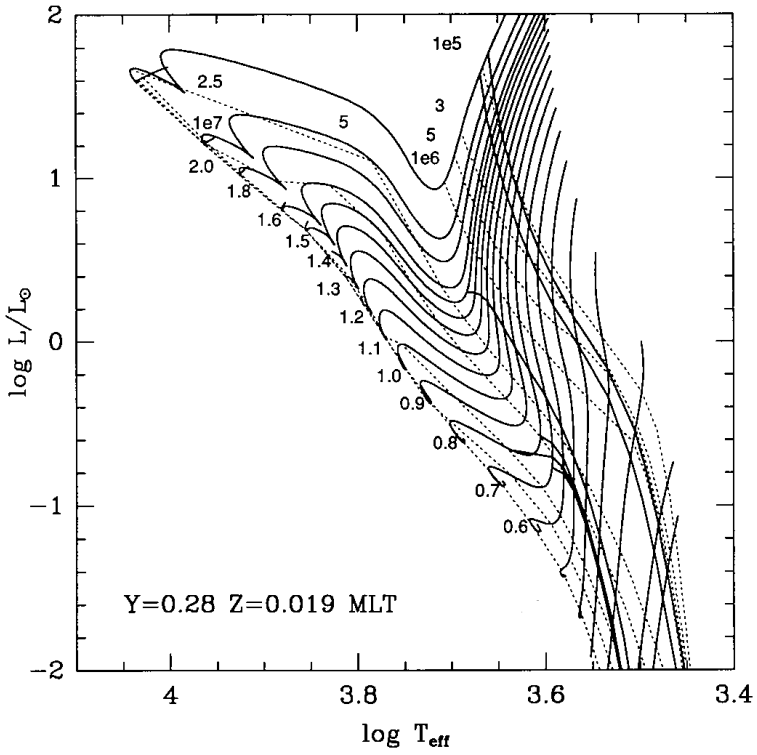


Figure 9.4. Pre-main-sequence evolution tracks for $0.3 - 2.5\,M_\odot$, according to the calculations of D'Antona & Mazzitelli (1994). The dotted lines are isochrones, connecting points on the tracks with the same age (between $t = 10^5$ yrs and 10^7 yrs, as indicated). Also indicated as solid lines 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 $\bar{\nabla} > \nabla_{\rm ad}$ 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 $\nabla - \nabla_{\rm ad}$ 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 $\nabla = \nabla_{\rm ad}$ 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 $T_{\rm eff} < T_{\rm eff,HL}$, 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 $\nabla_{\rm rad}$) decreases, until at some point $\nabla_{\rm rad} < \nabla_{\rm ad}$ 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 $T_{\rm eff}$ (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.

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 ($^2\mathrm{H} + {}^1\mathrm{H} \rightarrow {}^3\mathrm{He} + \gamma$, the second reaction in the pp chain). This reaction destroys all deuterium present in the star when $T \approx 1.0 \times 10^6\,\mathrm{K}$, while the protostar is still on the Hayashi line. The energy produced (5.5 MeV per reaction) is large enough to halt the contraction of the PMS star for a few times 10^5 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 $^{12}\mathrm{C}(\mathrm{p},\gamma)^{13}\mathrm{N}$ 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 $\tau_{\mathrm{PMS}} \approx 10^7 (M/M_\odot)^{-2.5}$ 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),

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

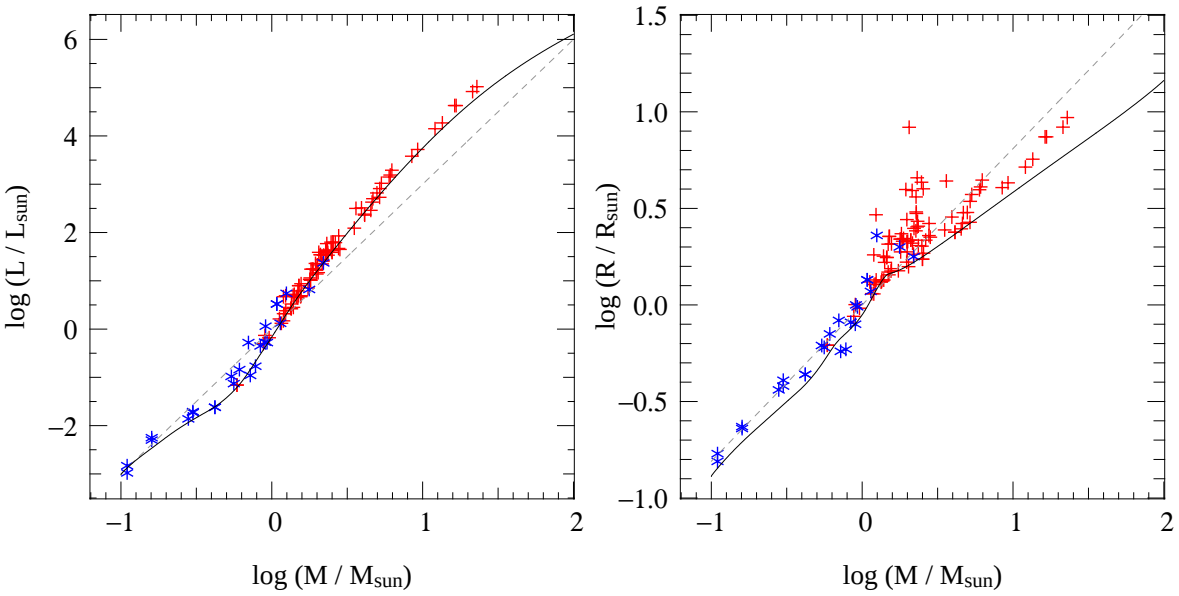


Figure 9.5. ZAMS mass-luminosity (left) and mass-radius (right) relations from detailed structure models with $X = 0.7, Z = 0.02$ (solid lines) and from homology relations scaled to solar values (dashed lines). For the radius homology relation, a value $\nu = 18$ appropriate for the CNO cycle was assumed (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).

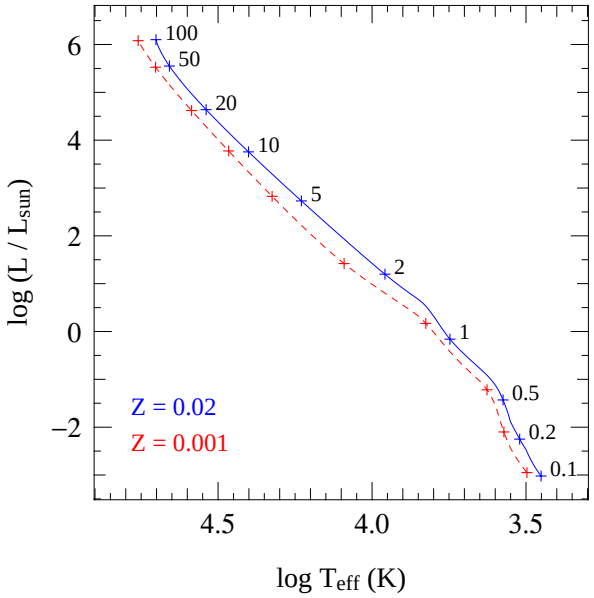


Figure 9.6. The location of the zero-age main sequence in the Hertzsprung-Russell diagram for homogeneous, detailed stellar models with $X = 0.7, Z = 0.02$ (blue solid line) and with $X = 0.757, Z = 0.001$ (red dashed line). Plus symbols indicate models for specific masses (in units of $M_\odot$). ZAMS 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.

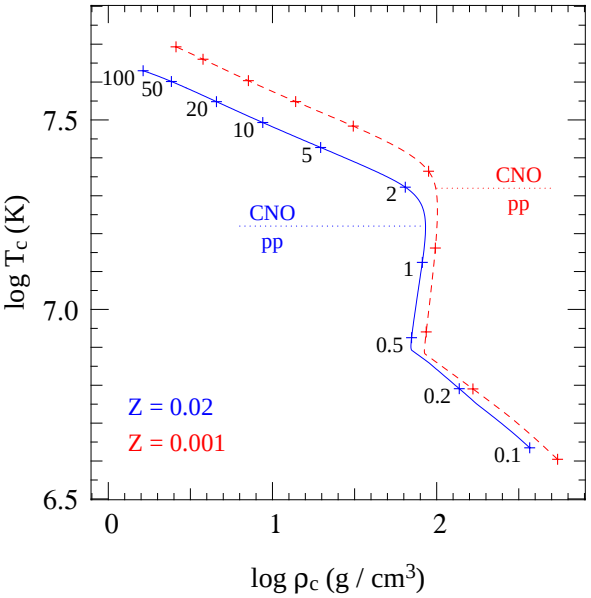


Figure 9.7. Central temperature versus central density for detailed ZAMS models with $X = 0.7, Z = 0.02$ (blue solid line) and with $X = 0.757, Z = 0.001$ (red dashed line). Plus symbols indicate models for specific masses (in units of $M_\odot$). The dotted lines indicate the approximate temperature 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

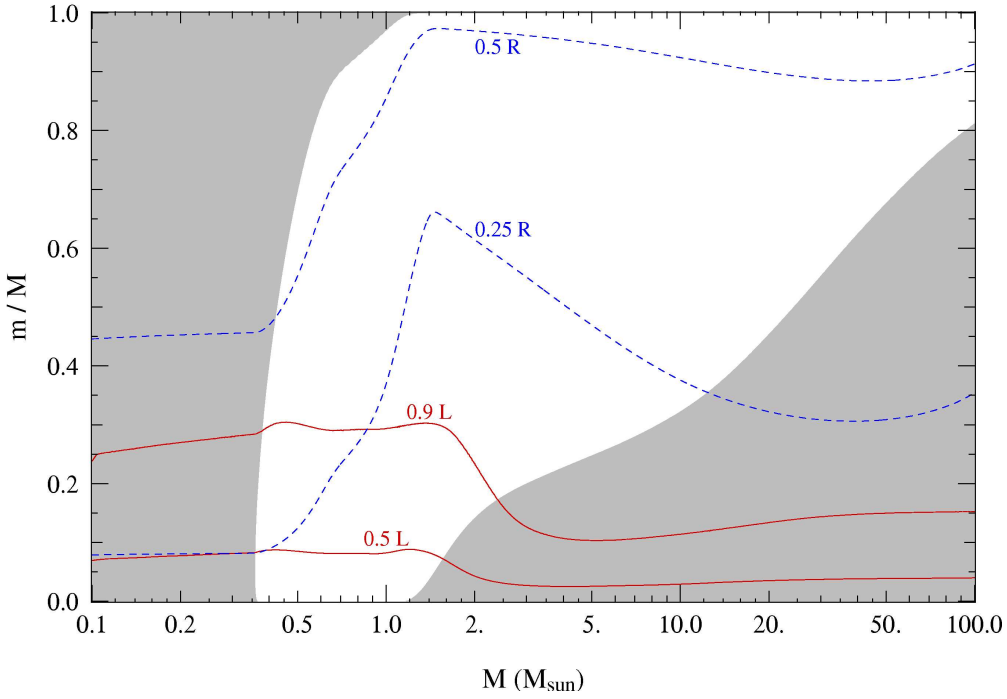


Figure 9.8. Occurrence of convective regions (gray shading) on the ZAMS in terms of fractional mass coordinate m/M as a function of stellar mass, for detailed stellar models with a composition $X = 0.70$, $Z = 0.02$. The solid (red) lines show the mass shells inside which 50% and 90% of the total luminosity 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, $\nabla_{\rm ad}$ 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 $\nabla_{\rm ad} \approx 0.25$.

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 $\nabla_{\rm rad}$ 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 cm^2/g which results in enormous values of $\nabla_{\rm rad}$ (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 $T_{\rm eff}$ 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 $T_{\rm eff}$, and their radius increase is modest. Higher-mass stars, on the other hand, evolve towards lower $T_{\rm eff}$ 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 $\epsilon_{\rm nuc}$ 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 $\epsilon_{\rm pp} \propto T^4$ and $\epsilon_{\rm CNO} \propto T^{18}$, 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

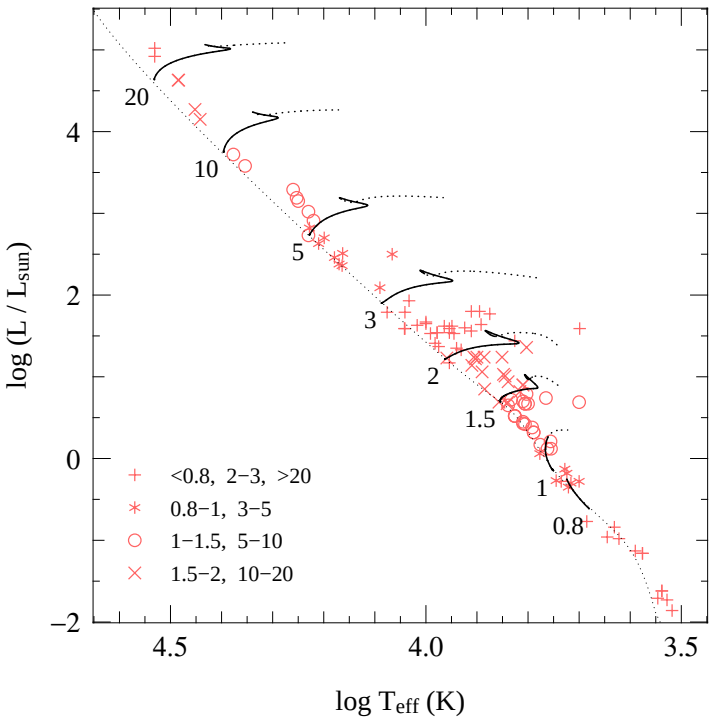


Figure 9.9. Evolution tracks in the H-R diagram during central hydrogen burning for stars of various masses, as labelled (in $M_\odot$), and for a composition $X = 0.7, Z = 0.02$. The dotted portion of each track shows the continuation of the evolution after central hydrogen exhaustion; the evolution 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:

$$P_{\rm env} = \int_{m_c}^{M} \frac{Gm}{4\pi r^4}\,{\rm d}m \tag{9.14}$$

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, $\epsilon_{\rm CNO} \propto \rho\, T^{18}$, 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 $P_{\rm env}$ decreases, i.e. the outer layers must expand to keep the star in HE ($\rho_{\rm env} \downarrow$ 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 $\epsilon_{\rm nuc}$ towards the centre. This gives rise to a large central $\nabla_{\rm rad} \propto l/m$ and hence to *convective cores*, which are mixed homogeneously ($X(m)$ = constant within the convective core mass $M_{\rm cc}$). This increases the amount of fuel available and therefore the lifetime of central hydrogen burning (see Fig. 9.10). In general $M_{\rm cc}$ decreases during the evolution, which is a consequence of the fact that $\nabla_{\rm rad} \propto \kappa$ 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 $M_{\rm cc}$ and $\epsilon_{\rm nuc}$ 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 $T_{\rm eff}$ 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 $\epsilon_{\rm pp} \propto \rho\, T^4$ 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.

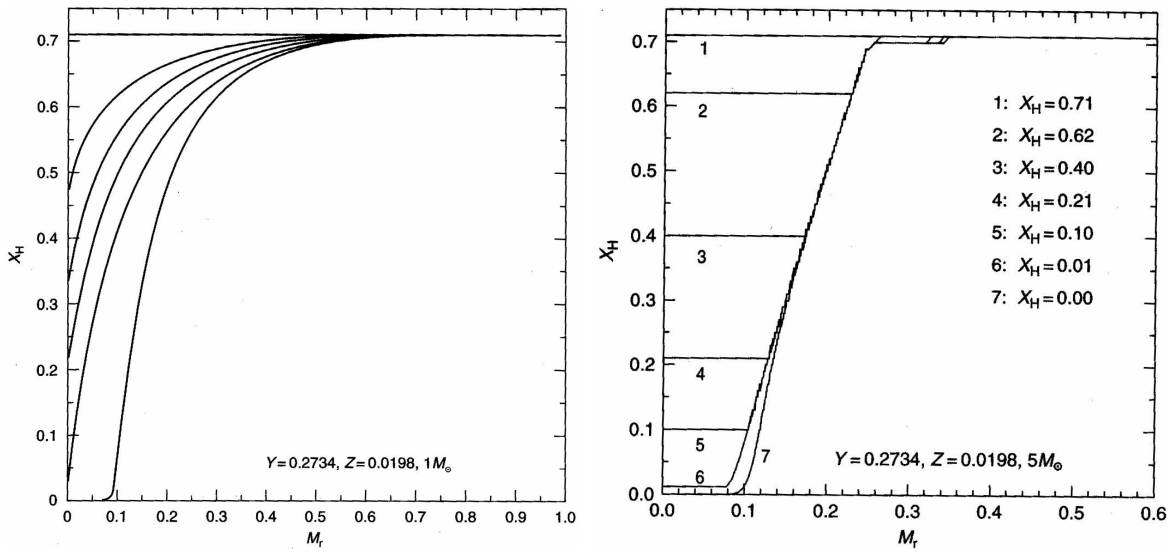


Figure 9.10. Hydrogen abundance profiles at different stages of evolution for a $1\,M_\odot$ star (left panel) and a $5\,M_\odot$ star (right panel) at quasi-solar composition. Figures reproduced from Salaris & Cassisi.

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 $\tau_{\rm MS}$ 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 $\epsilon_{\rm nuc}$ by eq. (6.43),

$$\frac{{\rm d}X}{{\rm d}t} = -\frac{\epsilon_{\rm nuc}}{q_{\rm H}}. \tag{9.15}$$

Here $q_{\rm H} = Q_{\rm H}/4m_{\rm u}$ is the effective energy release per unit mass of the reaction chain ($4\,^1{\rm H} \rightarrow {}^4{\rm He} + 2\,{\rm e}^+ + 2\,\nu$), corrected for the neutrino losses. Hence $q_{\rm H}$ is somewhat different for the pp chain and the CNO cycle. Note that $q_{\rm H}/c^2$ 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,

$$\frac{{\rm d}M_{\rm H}}{{\rm d}t} = -\frac{L}{q_{\rm H}}, \tag{9.16}$$

where $M_{\rm H}$ 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

$$\Delta M_{\rm H} = \frac{1}{q_{\rm H}} \int_0^{\tau_{\rm MS}} L\,{\rm d}t = \frac{\langle L \rangle\, \tau_{\rm MS}}{q_{\rm H}}, \tag{9.17}$$

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 $q_{\rm c}$ inside which all hydrogen is consumed, so that

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

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 $q_{\rm c}$ 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.

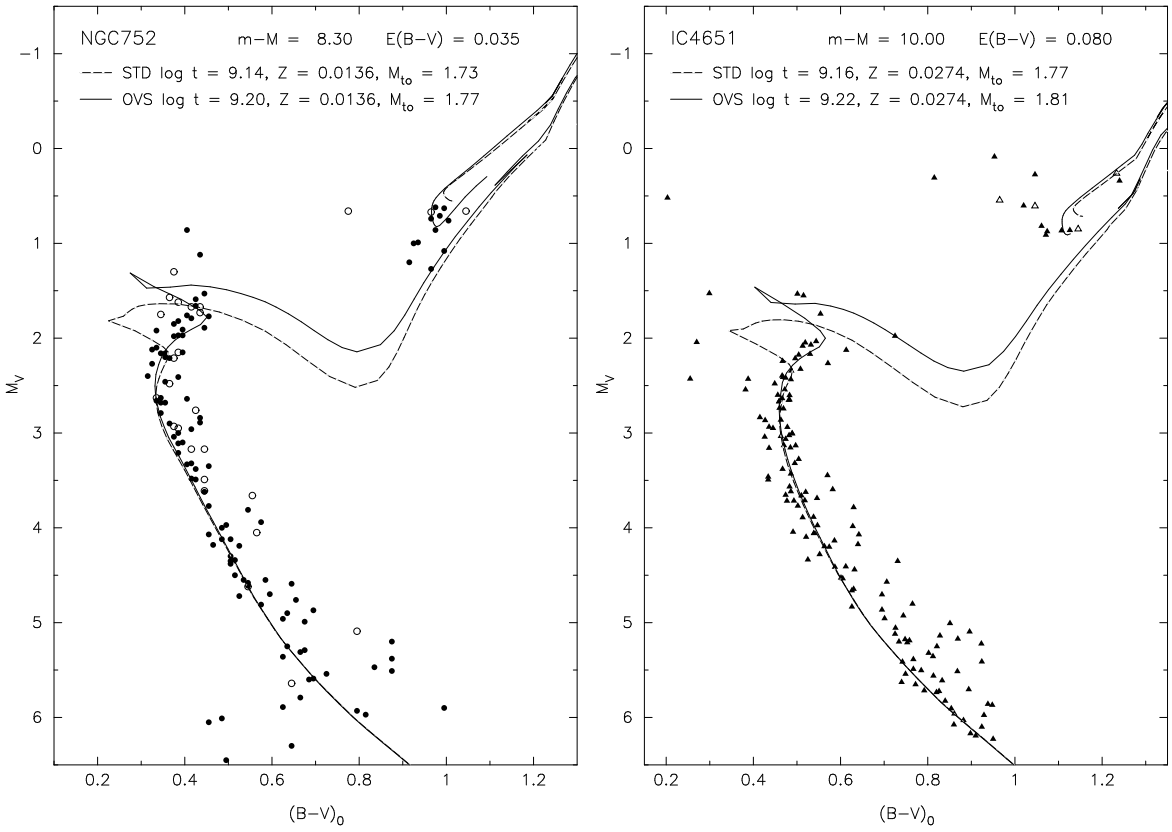


Figure 9.11. Two examples of *isochrone fitting* to the colour-magnitude diagrams of open clusters, NGC 752 and IC 4651. The distribution of stars in the turn-off region is matched to isochrones for standard stellar evolution models (STD) and for models with convective overshooting (OVS). The overshooting 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 $\alpha_{\rm ov}$ 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 $\alpha_{\rm ov}$. Such tests indicate that $\alpha_{\rm ov} \approx 0.25$ is appropriate in the mass range $1.5 - 8\,M_\odot$. For larger masses, however, $\alpha_{\rm ov}$ 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.

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 $M_{\rm HeF}$).

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 $M_{\rm HeF}$ and $M_{\rm up} \approx 8\,M_\odot$. 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 $M_{\rm up} \approx 8\,M_\odot$ 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 ($\tau_{\rm nuc} \gg \tau_{\rm KH}$) 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, $l(m) = \int_m \epsilon_{\rm nuc}\,{\rm d}m = 0$ and hence ${\rm d}T/{\rm d}r \propto l = 0$. 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, $\mathrm{d}P/\mathrm{d}r = (\mathcal{R}T/\mu)\,\mathrm{d}\rho/\mathrm{d}r$, 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 $q_\mathrm{c} = M_\mathrm{c}/M$ to be

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

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 $q_\mathrm{SC} \approx 0.10$ is appropriate for a helium core with $\mu_\mathrm{c} = 1.3$ and a H-rich envelope. (A simple, qualitative derivation of eq. 10.1 can be found in Maeder Section 25.5.1.)

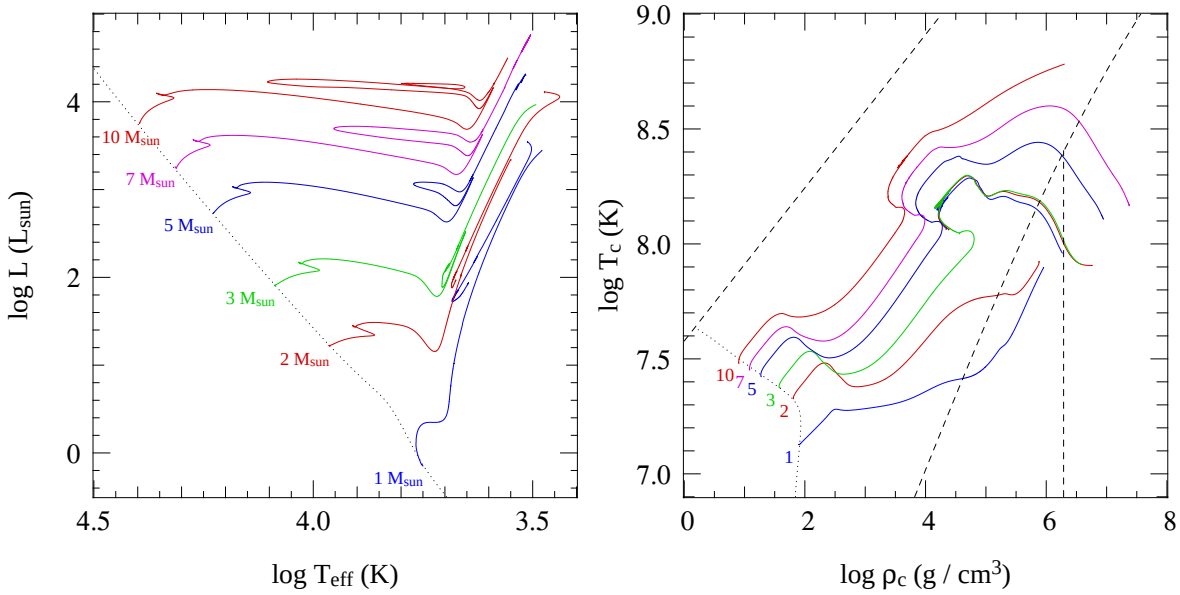


Figure 10.1. Evolution tracks for stars of quasi-solar composition ($X = 0.7$, $Z = 0.02$) and masses of 1, 2, 3, 5, 7 and 10 $M_\odot$ in the H-R diagram (left panel) and in the central temperature versus density plane (right panel). Dotted lines in both diagrams show the ZAMS, while the dashed lines 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, $\rho_{\rm shell}$ must therefore decrease and hence also the pressure in the burning shell must decrease. Therefore the pressure $P_{\rm env}$ 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*!

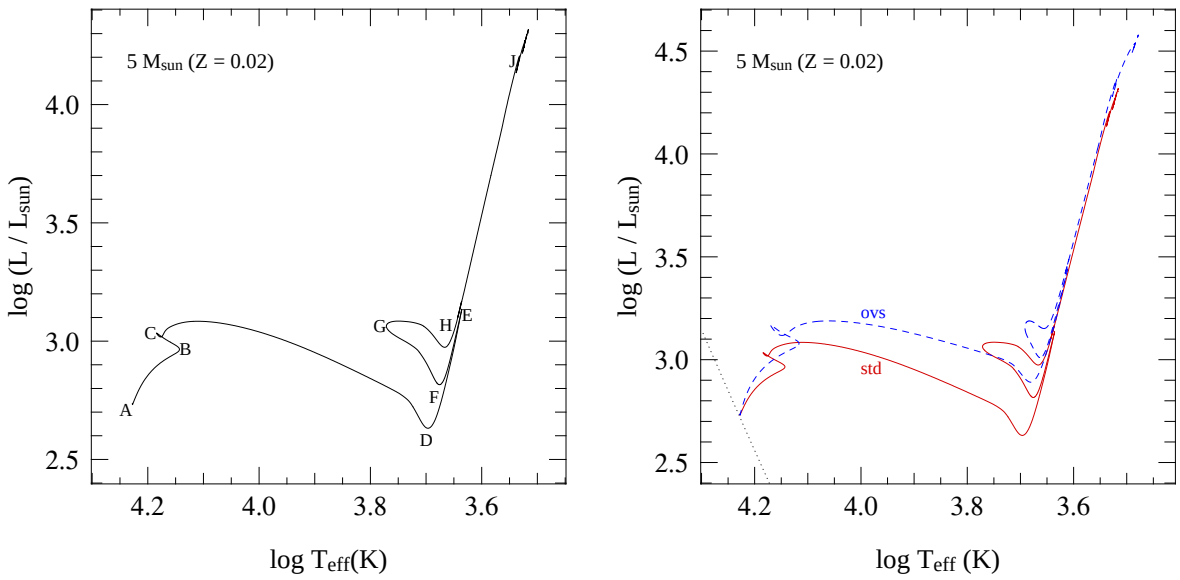


Figure 10.2. Evolution track in the Hertzsprung-Russell diagram of a $5\,M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. See text for details. The evolution track in the left panel was computed without convective overshooting. The right panel shows a comparison between this track and the evolution of the same star computed with moderate overshooting ($\alpha_{\rm ov} = l_{\rm ov}/H_P \approx 0.25$; 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-Chandrasekhar 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

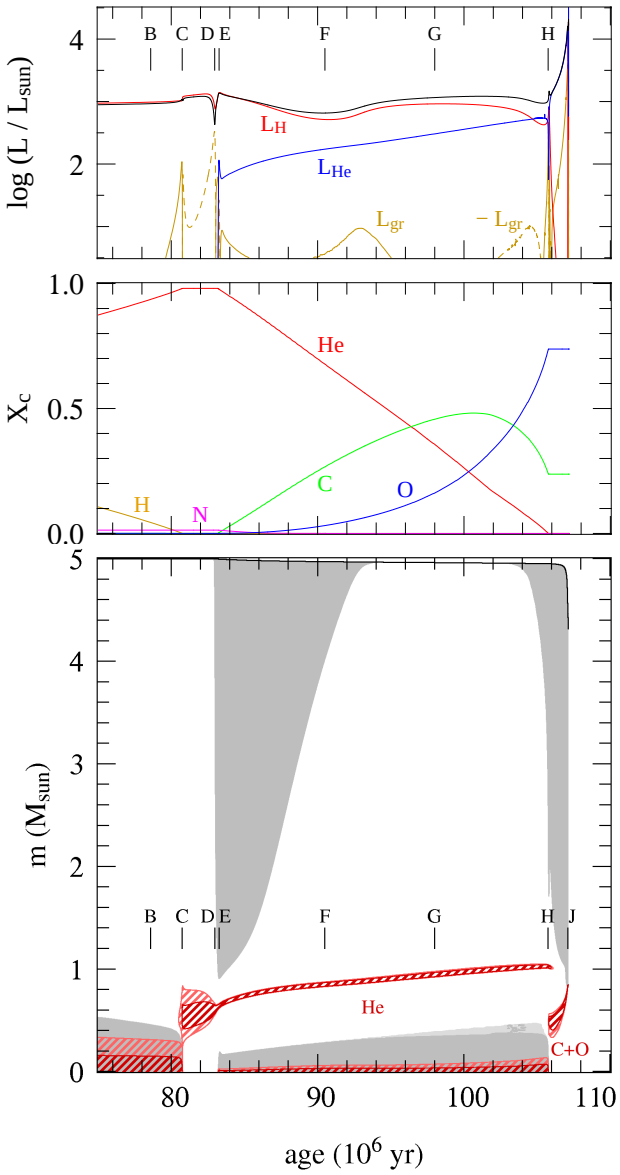


Figure 10.3. Internal evolution of a 5 $M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. The panels show various internal quantities as a function of time, from top to bottom:

(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 $\epsilon_{\rm nuc} > 10\,L/M$ (dark red) and $\epsilon_{\rm nuc} > 2\,L/M$ (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 $\log T_{\rm eff} > 4.05$. The phase of expansion from $\log T_{\rm eff} \approx 4.05$ to point D at $\log T_{\rm eff} \approx 3.7$ 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 $T_{\rm eff} \approx 5000$ 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

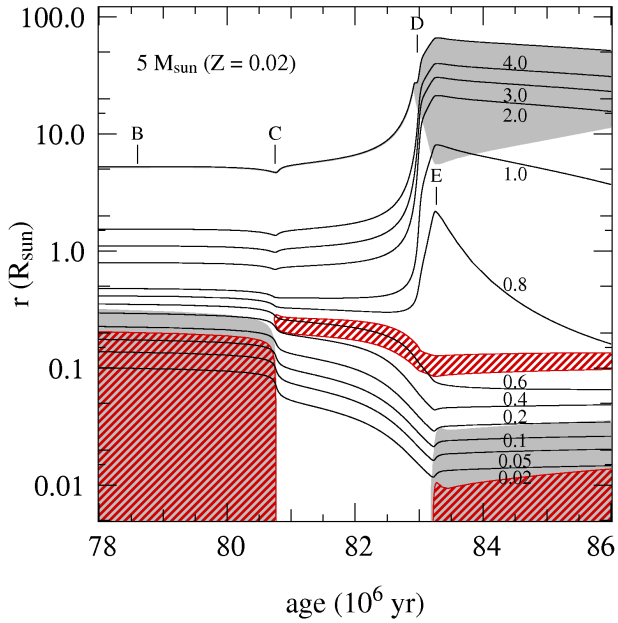


Figure 10.4. Radial variation of various mass shells (solid lines) in the $5\,M_\odot$ ($Z = 0.02$) star of Fig. 10.3, during the early post-main sequence evolution. Each line is labelled with its mass coordinate m in units of $M_\odot$; the top-most curve indicates the total radius R. Gray areas indicate convection and red cross-hatched areas have intense nuclear burning ($\epsilon_{\rm nuc} > 10\,L/M$). 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 $M_{\rm c}/M < 0.1$ 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

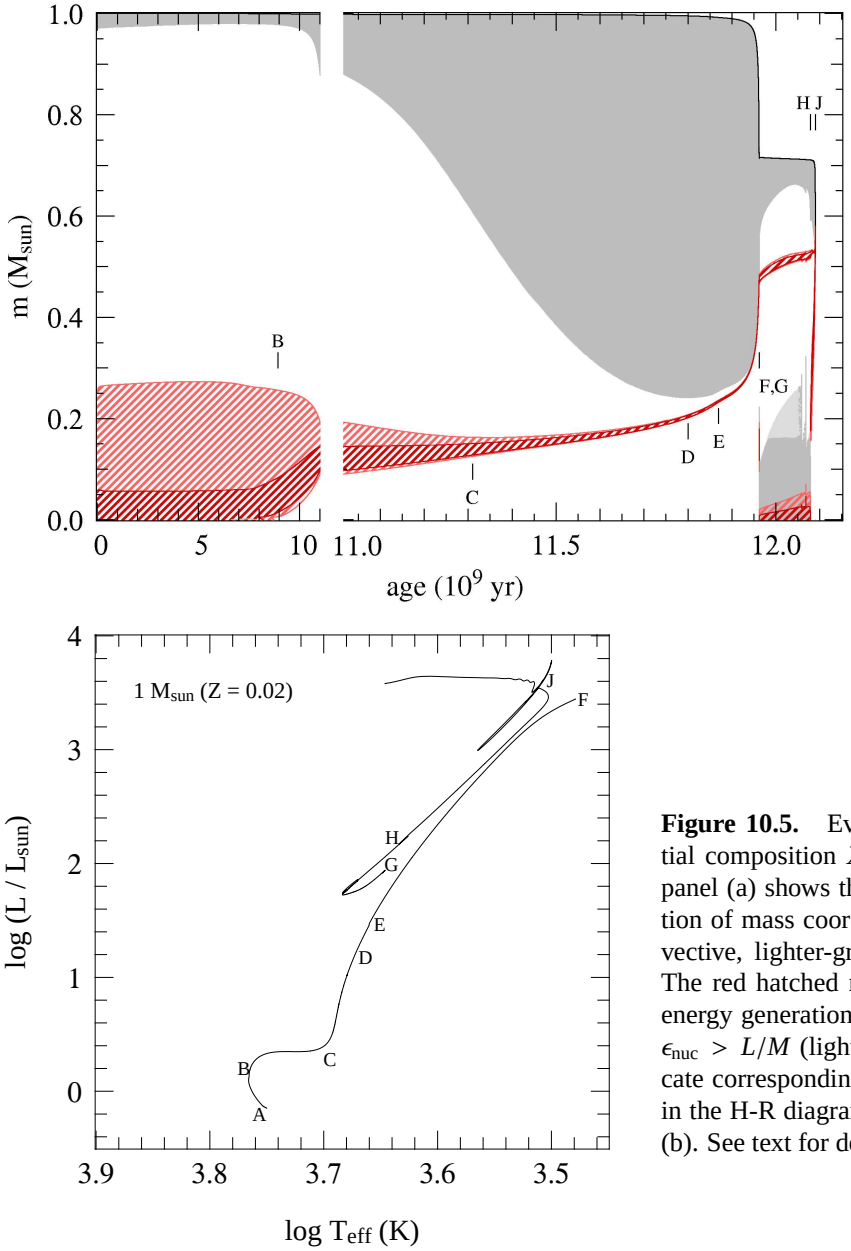


Figure 10.5. Evolution of a $1\,M_\odot$ star of initial composition $X = 0.7$, $Z = 0.02$. The top panel (a) shows the internal structure as a function of mass coordinate m. Gray areas are convective, lighter-gray areas are semi-convective. The red hatched regions show areas of nuclear energy generation: $\epsilon_{\rm nuc} > 5\,L/M$ (dark red) and $\epsilon_{\rm nuc} > L/M$ (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 $T_{\rm eff}$ 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, $M_{\rm c} \lesssim 0.5\,M_\odot$ and can be approximately described by a power law

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

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 $M_{\rm c} \approx 0.1M$, 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 $T_{\rm eff}$ 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 $T_{\rm eff}$. 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 $M_{\rm c}$-L relation and depends slightly on the total mass.

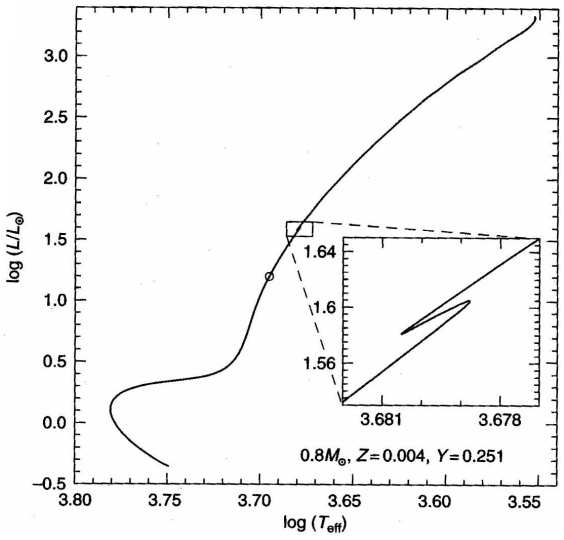


Figure 10.6. Evolution track of a $0.8\,M_\odot$ star of rather low metallicity, $Z = 0.004$. The inset shows the temporary decrease of luminosity when the H-burning shell crosses the hydrogen discontinuity left by the first dredge-up (corresponding to point E in Fig. 10.5). The open circle indicates 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:

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

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 $^{12}\mathrm{C}+\alpha$ 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 $X_\mathrm{He} < 0.2$ 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 $^{12}\mathrm{C}/^{16}\mathrm{O}$ 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 $^{12}\mathrm{C}/^{16}\mathrm{O}$ ratio depends on the uncertain rate of the $^{12}\mathrm{C}(\alpha,\gamma)$ 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 $0.27 \times \tau_\mathrm{MS}$. 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 $L_{\rm He}$ is a steep function of $M_{\rm c}$, 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, $M_{\rm c}$ grows with time during the He-burning phase and $L_{\rm He}$ increases accordingly. Another consequence is that in models computed with convective overshooting $L_{\rm He}$ 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, $\tau_{\rm nuc} \propto M_{\rm c}/L_{\rm He}$) is *shorter* in models with overshooting. A $5\,M_\odot$ star of the same composition computed with overshooting has a main-sequence lifetime $\tau_{\rm MS} = 100\,$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 $X_{\rm He} \approx 0.3$. 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 $X_{\rm He} \approx 0.05$. 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 $T_{\rm eff}$ 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 $M_{\rm c} \approx 0.45\,M_\odot$ (Sect. 10.2.3). The luminosity of low-mass He-burning stars is therefore almost independent

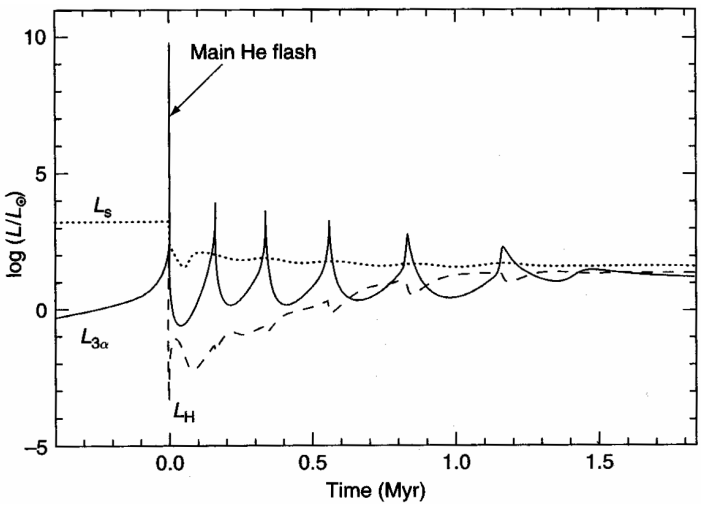


Figure 10.7. The helium flash. Evolution with time of the surface luminosity (L_s), the He-burning luminosity ($L_{3\alpha}$) and the H-burning luminosity ($L_{\rm H}$) during the onset of He burning at the tip of the RGB in a low-mass star. Time $t = 0$ corresponds to the start of the main helium flash. 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 $T_c \approx 10^8\,$K and $\rho_c \approx 10^6\,$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 $T \approx 3 \times 10^8\,$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 $T_c \approx 10^8\,$K and $\rho_c \approx 2 \times 10^4\,$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.

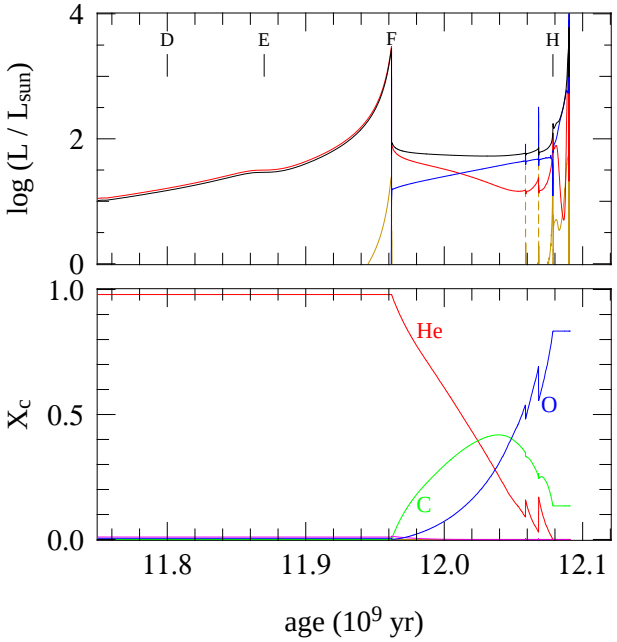


Figure 10.8. Evolution with time of the luminosities and central abundances in a $1\,M_\odot$ star during the late part of the red giant branch and during helium burning. Letters D…H correspond to the same evolution phases as in Fig. 10.5.

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 ≈ 1.5 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

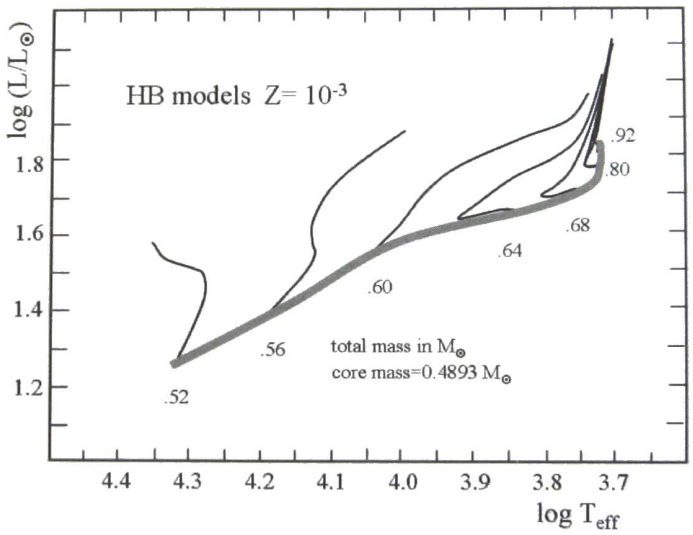


Figure 10.9. Location of the zero-age horizontal branch (think gray line) for a metallicity $Z = 0.001$ typical of globular clusters. These models have the same core mass ($0.489\,M_\odot$) but varying total (i.e. envelope) mass, which determines their position in the H-R diagram. Evolution tracks during 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/`.

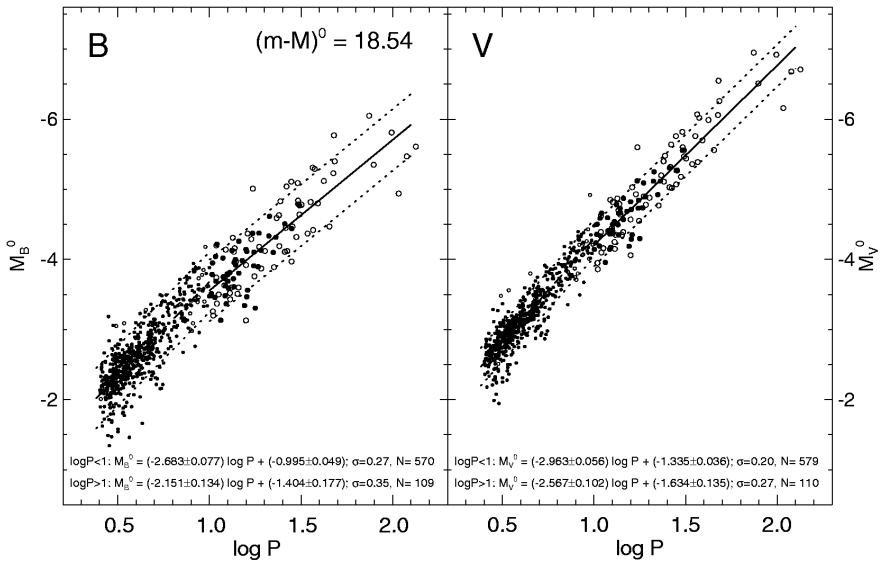


Figure 10.10. The period-luminosity relation for classical Cepheids in the Large Magellanic Cloud. Luminosity is expressed as absolute magnitude in the B band (left) and in the V band. Figure from Sandage et al. (2004, A&A 424, 43).

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 $\gamma_{\rm ad} > \frac{4}{3}$. In this situation a perturbation of pressure equilibrium will be restored, the restoring force being larger the more $\gamma_{\rm ad}$ 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

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

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

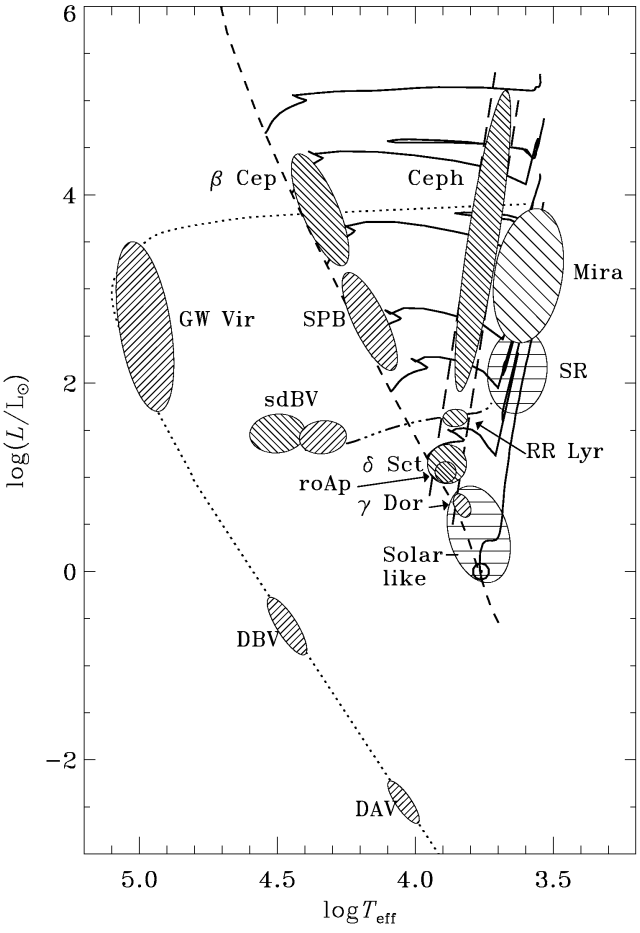


Figure 10.11. Occurrence of various classes of pulsating stars in the H-R diagram, overlaid on stellar evolution tracks (solid lines). Cepheid variables are indicated with 'Ceph', they lie within the pulsational instability strip in the HRD (long-dashed lines). Their equivalents are the RR Lyrae variables 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).

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 $\gamma_{\rm ad}$ constant throughout the star. This yields

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

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, $\oint P\,{\rm d}V > 0$. By the first law of thermodynamics, this work is provided by a net amount of heat being absorbed by the element during the cycle,

$$\oint {\rm d}Q = \oint P\,{\rm d}V > 0.$$

The change in entropy of the mass element is ${\rm d}S = {\rm d}Q/T$. Since entropy is a state variable, $\oint {\rm d}Q/T = 0$ 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

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

or

$$\oint {\rm d}Q \approx \oint {\rm d}Q\,\frac{\delta T}{T_0}. \tag{10.8}$$

Eq. (10.8) means that heat must enter the element (${\rm d}Q > 0$) when the temperature is high ($\delta T > 0$), i.e. when the layer is compressed, and/or heat must leave the layer (${\rm d}Q < 0$) 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 $\oint {\rm d}Q = \oint P\,{\rm d}V < 0$). To determine the overall effect, the contributions $\oint P\,{\rm d}V$ 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 $\epsilon_{\rm nuc}$. 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 $(\mathrm{d}\ln\kappa/\mathrm{d}\ln P)_{\rm ad} > 0$. We can write this as

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

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

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

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 $\nabla_{\rm ad} = 0.4$, we normally have $\kappa_P + \kappa_T\,\nabla_{\rm ad} < 0$, 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 $T_{\rm eff}$ for this to be important.

- In case of a Kramers-like opacity, a small value of $\nabla_{\rm ad}$ can lead to pulsation instability. For $\kappa_P \approx 1$ and $\kappa_T \approx -4.5$, eq. (10.10) implies $\nabla_{\rm ad} \lesssim 0.22$. Such small values of $\nabla_{\rm ad}$ 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 $T \approx 1.5\times10^4$ K where both $\mathrm{H} \leftrightarrow \mathrm{H}^+ + \mathrm{e}^-$ and $\mathrm{He} \leftrightarrow \mathrm{He}^+ + \mathrm{e}^-$ occur, and one at $T \approx 4\times10^4$ K where helium becomes twice ionized ($\mathrm{He}^+ \leftrightarrow \mathrm{He}^{++} + \mathrm{e}^-$). These partial ionization zones can explain the location of the instability strip in the H-R diagram, as follows.

- At large $T_{\rm eff}$ (for $T_{\rm eff} \gtrsim 7500$ 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 $T_{\rm eff}$ 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 $T_{\rm eff}$ (for $T_{\rm eff} \lesssim 5500\,{\rm K}$, 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 $T_{\rm eff}$ 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 $L \propto R^2 T_{\rm eff}^4$, we can write

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

With $\alpha \approx 3$ and $T_{\rm eff} \approx$ 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 $E_{\rm gr} = -\frac{3}{5}GM^2/R$. 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

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

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 $\mathrm{d}M_{\rm c}/\mathrm{d}t$. Use the energy released per gram in hydrogen shell burning.

(b) Derive how the core mass evolves in time, i.e, $M_{\rm c} = M_{\rm c}(t)$.

(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:

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

(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: $\kappa_e = 0.2(1+X)$

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

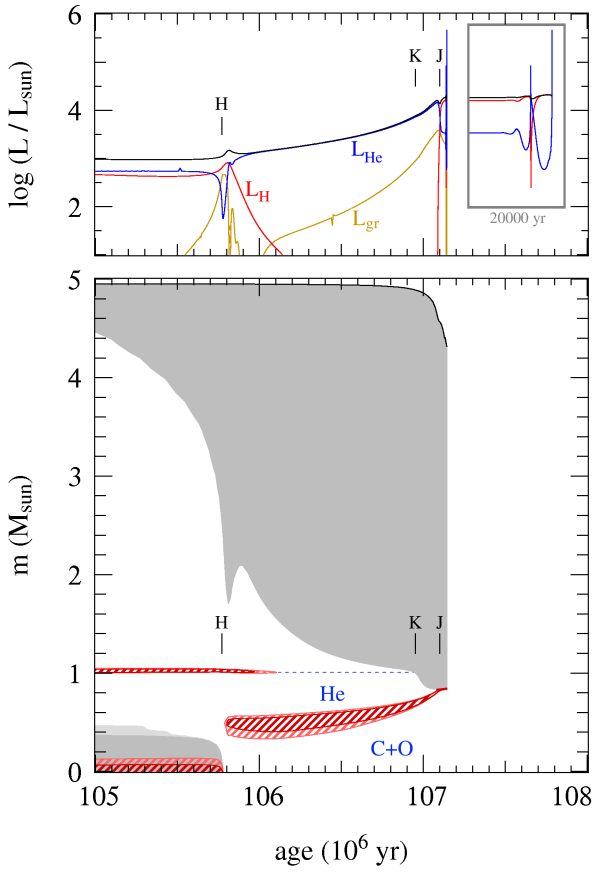


Figure 11.1. Evolution of luminosities (upper panel) and internal structure (lower panel) with time in a 5 $M_\odot$ star (with composition $X = 0.70$, $Z = 0.02$) during the last stages of helium burning and on the AGB. Compare with Fig. 10.3 for the same star. The early AGB starts at point H, when 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, $M_{\rm up}$, 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,

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

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

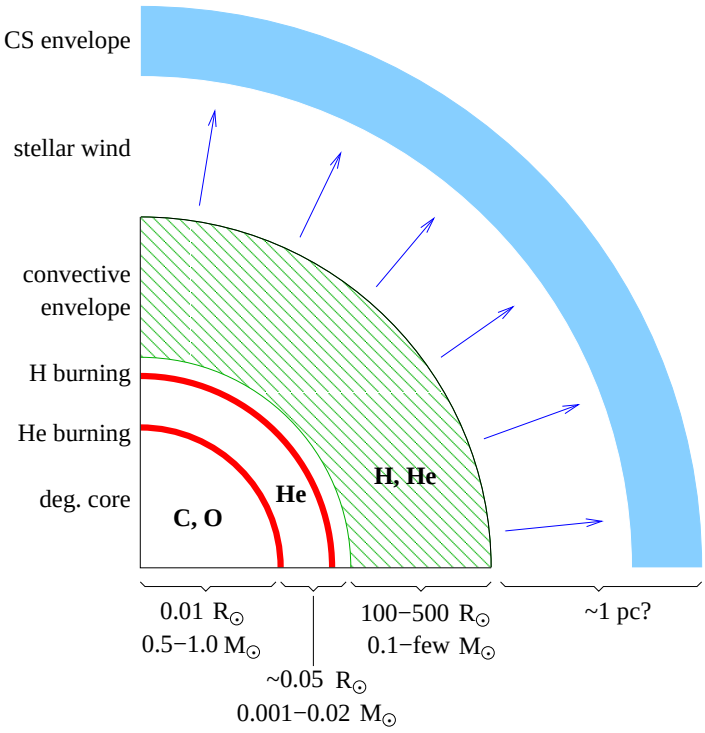


Figure 11.2. Schematic structure of an AGB star during its thermally pulsing phase. The CO core is degenerate and very compact, and is surrounded by two burning shells very close together in mass coordinate. The convective envelope by contrast is very extended and tenuous, having a radius 10^4–10^5 times 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 $L_{\rm He} \approx 10^8\,L_\odot$ are reached during $\sim$ 1 year. The large energy flux drives convection in the whole intershell region (producing an *intershell convection zone*, ICZ).

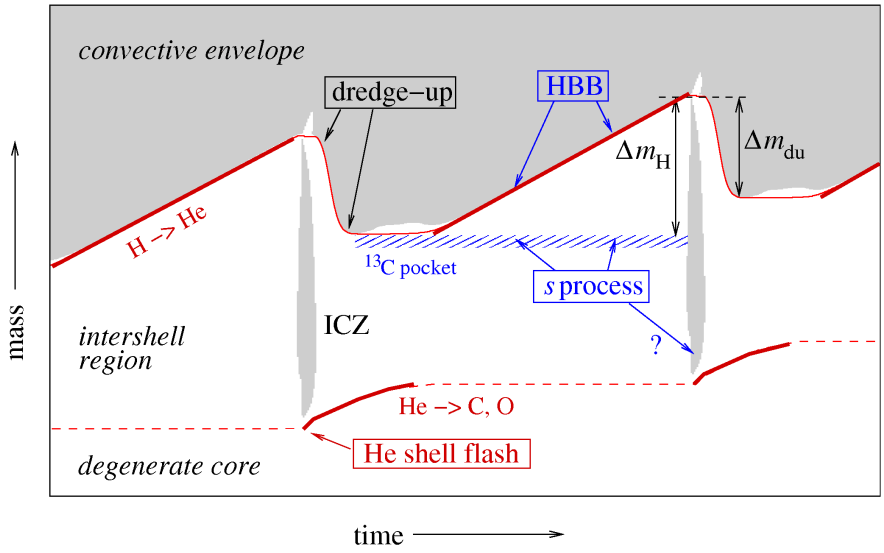


Figure 11.3. Schematic evolution of an AGB star through two thermal-pulse cycles. Convective regions are shown as gray shaded areas, where 'ICZ' stands for the intershell convection zone driven by the He-shell flash. The H-exhausted core mass is shown as a thin red solid line and the He-exhausted core 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 $\sim$ 100 years) are expanded relative to the interpulse phase (10^4 – 10^5 years).

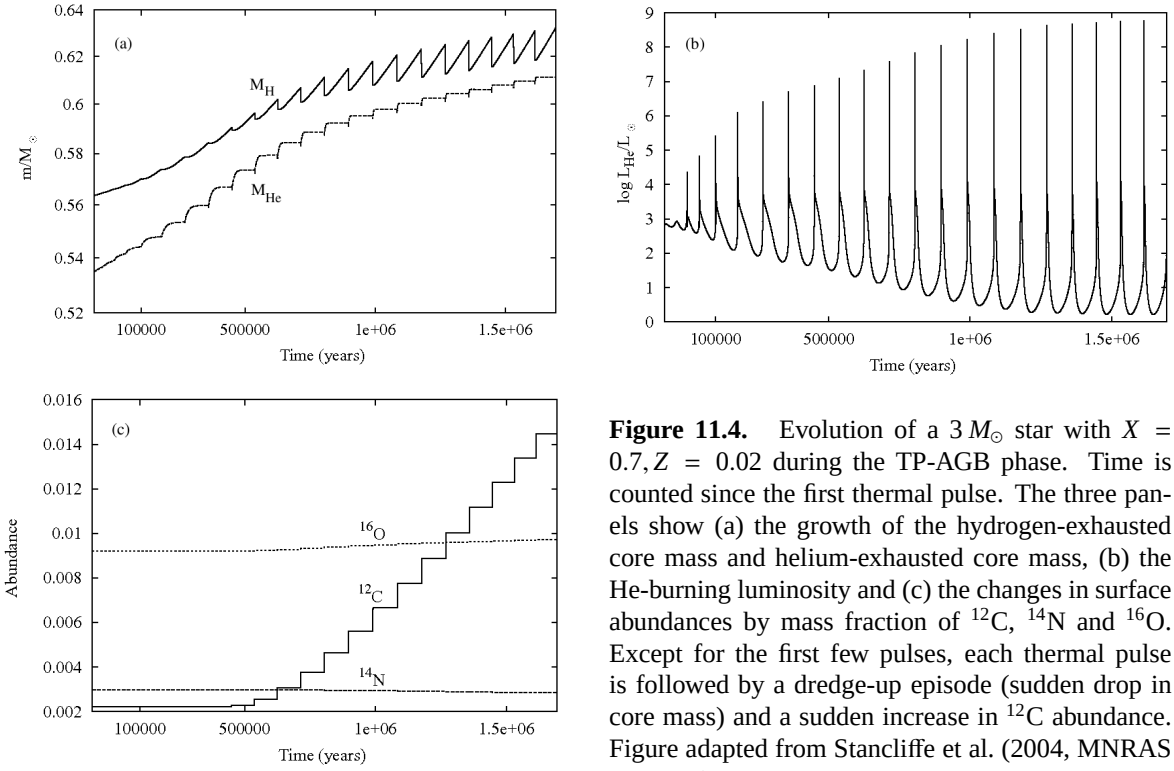


Figure 11.4. Evolution of a $3\,M_\odot$ star with $X = 0.7, Z = 0.02$ during the TP-AGB phase. Time is counted since the first thermal pulse. The three panels show (a) the growth of the hydrogen-exhausted core mass and helium-exhausted core mass, (b) the He-burning luminosity and (c) the changes in 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),

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

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 $n(\mathrm{C})/n(\mathrm{O}) < 1$ (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: $^{13}\mathrm{C}(\alpha,\mathrm{n})^{16}\mathrm{O}$ and $^{22}\mathrm{Ne}(\alpha,\mathrm{n})^{25}\mathrm{Mg}$. 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: $^{14}\mathrm{N}(\alpha,\gamma)^{18}\mathrm{F}(\beta^+)^{18}\mathrm{O}(\alpha,\gamma)^{22}\mathrm{Ne}$.

The main neutron source in low-mass stars (up to $3\,M_\odot$) is probably the $^{13}\mathrm{C}(\alpha,\mathrm{n})^{16}\mathrm{O}$ 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 ($T_{\rm BCE} \gtrsim 3 \times 10^7$ 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*, $M_{\rm Ch} \approx 1.46\,M_\odot$, 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 $M_{\rm Ch}$. 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.

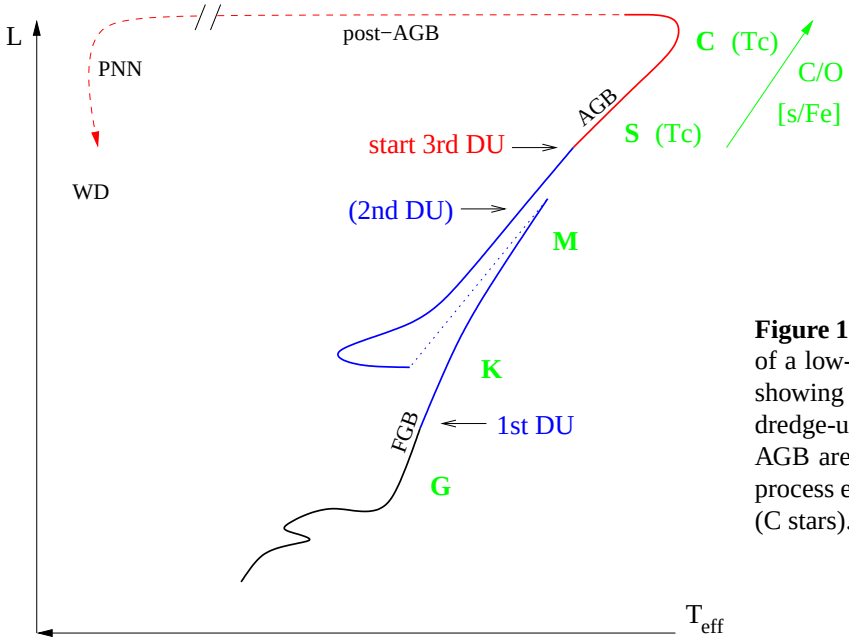
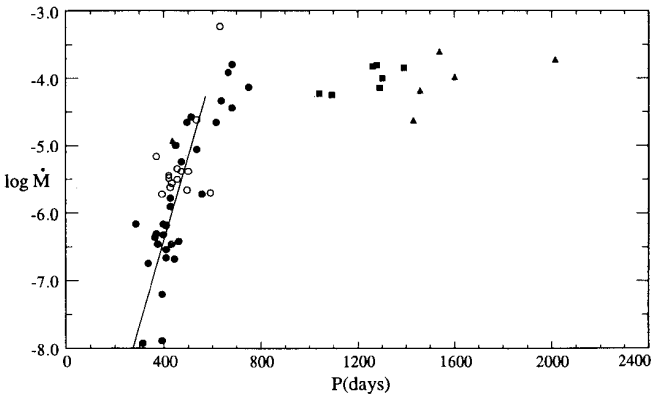


Figure 11.5. Schematic evolution track of a low-mass star in the H-R diagram, showing the occurrence of the various dredge-up episodes. Stars on the upper AGB are observed to be enriched in s-process elements (S stars) and in carbon (C stars).



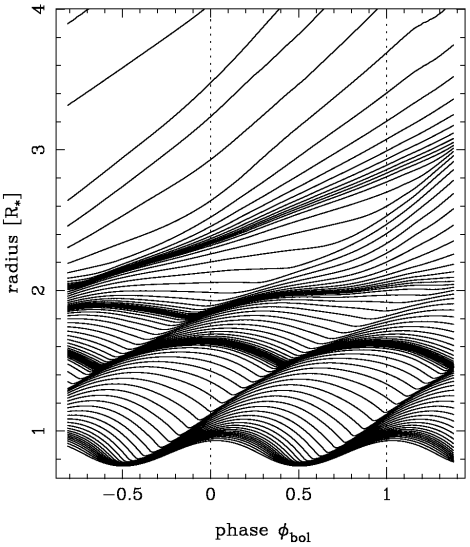


Figure 11.6. Mass loss of AGB stars. *Left:* the observed correlation between the pulsation period P of Mira variables and their mass-loss rate $\dot{M}$ (in $M_\odot$/yr) (from Vassiliadis & Wood 1993, ApJ 413, 641). *Right:* a theoretical model (Nowotny et al. 2005, A&A 437, 273) showing streamlines 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 (~ 1500 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

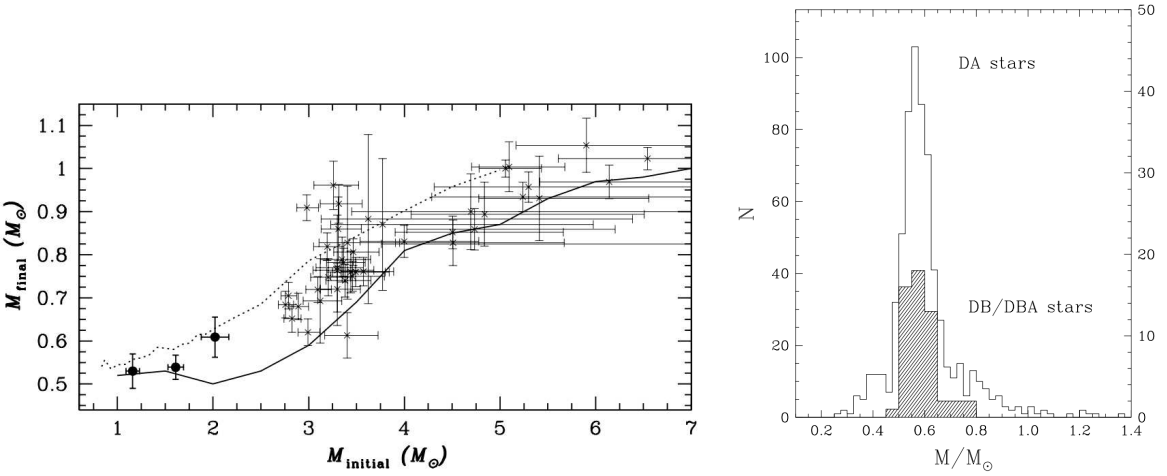


Figure 11.7. *Left:* Relation between between the initial and final mass of low- and intermediate-mass stars, from Kalirai et al. (2008, ApJ 676, 594). The data points represent white dwarfs observed in open clusters, for which the mass has been determined from their spectra. The age of the cluster $t_{\rm cl}$ 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),

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

the radius goes to zero as all electrons become extremely relativistic. White dwarfs more massive than $M_{\rm Ch}$ must collapse as the relativistic degeneracy pressure is insufficient to balance gravity.

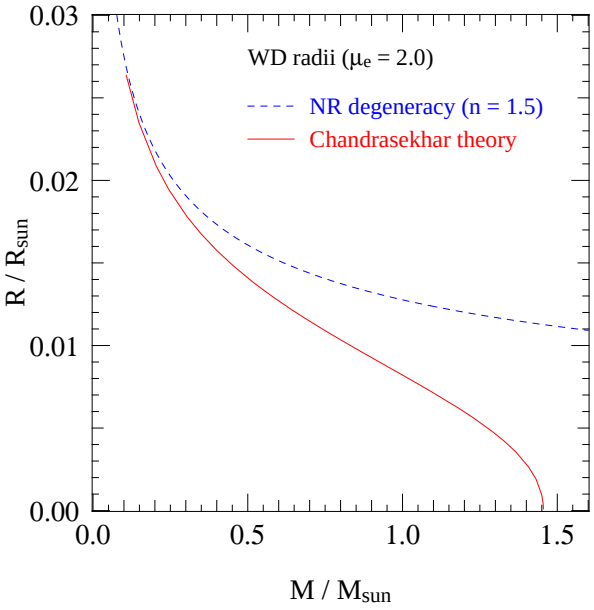


Figure 11.8. Comparison of the radius-mass relation of a completely degenerate star computed using Chandrasekhar's theory for white dwarfs (taking into account the partly relativistic velocities of the electrons in the Fermi sea) and an approximation based on non-relativistic degeneracy.

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

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

A neutron-rich nucleus such as ^{24}Na is normally unstable to β-decay ($^{24}\mathrm{Na} \rightarrow {}^{24}\mathrm{Mg} + \mathrm{e}^- + \bar{\nu}$), 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:

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

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

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

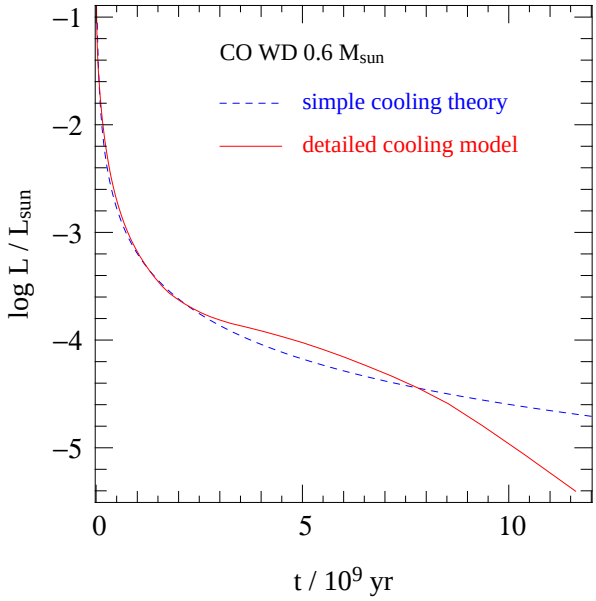


Figure 11.9. Theoretical cooling curves for a CO white dwarf with a typical mass of $0.6\,M_\odot$. The dashed (blue) line shows the evolution of luminosity with time based on the simple cooling theory by Mestel, which yields $L \propto t^{-7/5}$. The solid (red) line is a detailed cooling model for 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, $P_\mathrm{e} = (\mathcal{R}/\mu_\mathrm{e})\rho T$, since the ions are non-degenerate everywhere. At this point, denoted with subscript 'b', we have

$$\frac{\mathcal{R}}{\mu_\mathrm{e}}\rho_\mathrm{b} T_\mathrm{b} = K_\mathrm{NR}\left(\frac{\rho_\mathrm{b}}{\mu_\mathrm{e}}\right)^{5/3}.$$

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

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

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 $T_\mathrm{c}^{7/2} = \alpha\, L/M$. 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 $\mu_\mathrm{e} = 2$ 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,

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

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

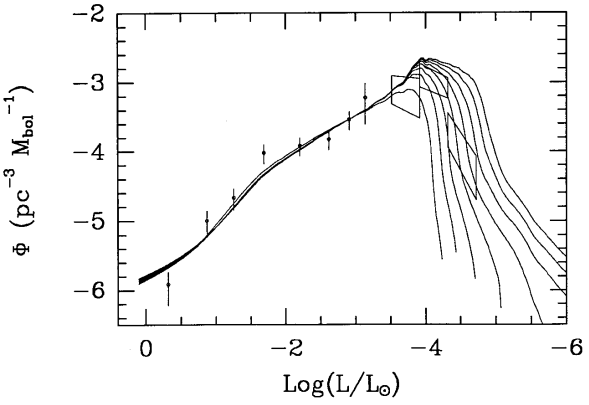


Figure 11.10. Observed and theoretical distributions of white dwarf luminosities in the Galactic disk, from Wood (1992, ApJ 386, 539), based on cooling models similar to the one shown in Fig. 11.9. The curves are for assumed ages of the Galactic disk between 6 and 13 Gyr. The paucity of observed white 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

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

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

$$L = -\frac{{\rm d}E_{\rm in}}{{\rm d}t} = -c_V M \frac{{\rm d}T_c}{{\rm d}t}, \tag{11.9}$$

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

$$T_c^{7/2} = -\alpha c_V \frac{{\rm d}T_c}{{\rm d}t},$$

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 $c_V = \frac{3}{2}\mathcal{R}/\mu_{\rm ion}$, we can write this relation as

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

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 $\mu_{\rm ion}$ 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, $\mu_{\rm ion} \approx 14$.

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, $c_V > \frac{3}{2}\mathcal{R}/\mu_{\rm ion}$) 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 $\dot{M}_c/M_\odot = 1.0\times10^{-11}(L_*/L_\odot)$. 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 $T_{\rm eff}$ 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: $-\dot{M}M = 0.5\,{\rm d}(M^2)/{\rm d}t$).

(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 ($M_{\rm env} \simeq 0$).
- 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., $\eta = 1$.

Chapter 12

Pre-supernova evolution of massive stars

We have seen that low- and intermediate-mass stars (with masses up to $\approx 8\,M_\odot$) develop carbon-oxygen cores that become degenerate after central He burning. As a consequence the maximum core temperature reached in these stars is smaller than the temperature required for carbon fusion. During the latest stages of evolution on the AGB these stars undergo strong mass loss which removes the remaining envelope, so that their final remnants are C-O white dwarfs.

The evolution of *massive stars* is different in two important ways:

- They reach a sufficiently high temperature in their cores ($> 5{\times}10^8$ K) to undergo *non-degenerate carbon ignition* (see Fig. 12.1). This requires a certain minimum mass for the CO core after central He burning, which detailed evolution models put at $M_{\mathrm{CO-core}} > 1.06\,M_\odot$. Only stars with initial masses above a certain limit, often denoted as M_{up} in the literature, reach this critical core mass. The value of M_{up} is somewhat uncertain, mainly due to uncertainties related to mixing (e.g. convective overshooting), but is approximately $8\,M_\odot$.

  Stars with masses above the limit $M_{\mathrm{ec}} \approx 11\,M_\odot$ also ignite and burn fuels heavier than carbon until an Fe core is formed which collapses and causes a supernova explosion. We will explore the evolution of the cores of massive stars through carbon burning, up to the formation of an iron core, in the second part of this chapter.

- For masses $M \gtrsim 15\,M_\odot$, *mass loss by stellar winds* becomes important during all evolution phases, including the main sequence. For masses above $30\,M_\odot$ the mass-loss rates $\dot{M}$ are so large that the timescale for mass loss, $\tau_{\mathrm{ml}} = M/\dot{M}$, becomes smaller than the nuclear timescale τ_{nuc}. Therefore mass loss has a very significant effect on their evolution. The stellar wind mechanisms involved are in many cases not well understood, so that $\dot{M}$ is often quite uncertain. This introduces substantial uncertainties in massive star evolution. The effect of mass loss on massive star evolution is discussed in the first part of this chapter.

12.1 Stellar wind mass loss

Observations in the ultraviolet and infrared part of the spectrum show that luminous stars, with masses above about $15\,M_\odot$, undergo rapid mass outflows (stellar winds) that gradually erode their outer layers. An empirical formula that fits the average observed mass-loss rates of stars of roughly solar metallicity in the upper part of the HR diagram ($L \gtrsim 10^3\,L_\odot$) was derived by De Jager and others in 1988:

$$\log(-\dot{M}) \approx -8.16 + 1.77 \log\left(\frac{L}{L_\odot}\right) - 1.68 \log\left(\frac{T_{\mathrm{eff}}}{\mathrm{K}}\right) \qquad (\text{in } M_\odot/\mathrm{yr}). \tag{12.1}$$

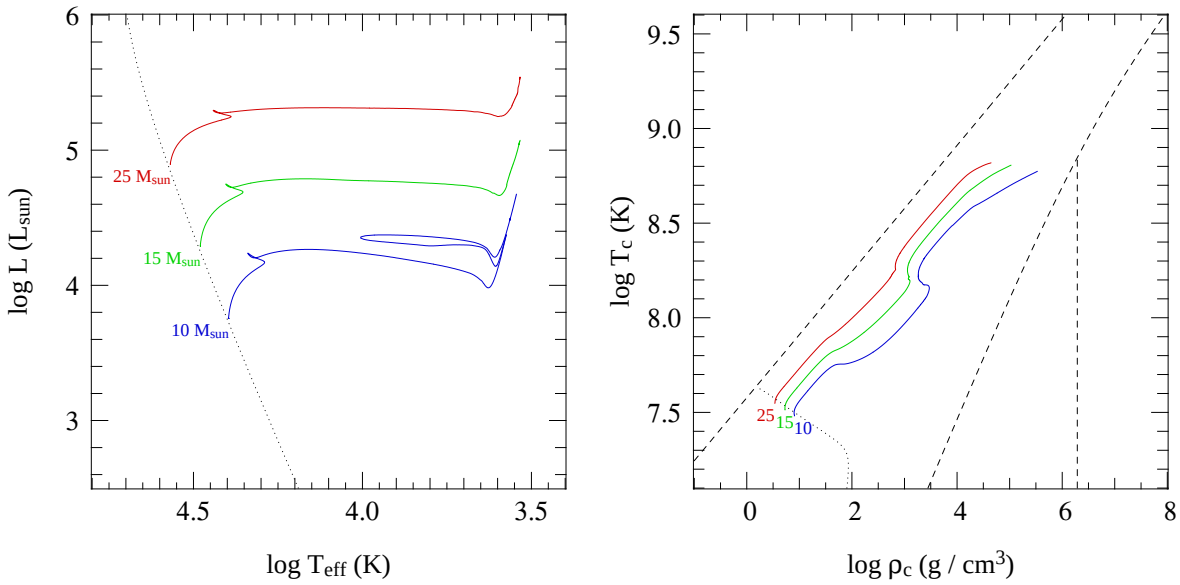


Figure 12.1. Evolution tracks in the HR diagram (left panel) and in the $\log \rho_c$-$\log T_c$ diagram (right panel) for stars with $Z = 0.02$ and $M = 10$, 15 and 25 $M_\odot$, computed with a moderate amount of overshooting. The tracks end when carbon is ignited in the centre, under non-degenerate conditions.

For example, for the 25 $M_\odot$ star depicted in Fig. 12.1 you can check by estimating L and $T_{\rm eff}$ from the graph that this implies a mass loss of 5×10^{-8} $M_\odot$/yr at the ZAMS, increasing up to 5×10^{-7} $M_\odot$/yr at the end of the main sequence. By the end of the evolution track, when the star is a red supergiant, the mass-loss rate implied by the above formula has increased to 5×10^{-5} $M_\odot$/yr.

The observed strong mass loss is probably caused by different mechanisms in different parts of the HR diagram.

Radiation-driven stellar winds

Hot, luminous stars (OB-type main-sequence stars and *blue supergiants*, BSG) undergo a fast *radiation-driven stellar wind*. Radiation pressure at frequencies corresponding to absorption lines in the spectrum, where the interaction between photons and matter is strong, causes an outward acceleration. An upper limit to the mass-loss rate that can be driven by radiation is obtained by assuming that the photons transfer their entire momentum to the outflowing matter:

$$\dot{M} v_\infty < \frac{L}{c} \tag{12.2}$$

where v_∞ is the terminal wind velocity at large distance from the star ('infinity'). A typical value of the terminal velocity is about three times the escape velocity, $v_\infty \approx 3 v_{\rm esc}$ (about 1000–3000 km/s for O-type stars). Comparing the mass-loss rates from eq. (12.1) with the upper limit shows that the empirical rates are indeed smaller, but only by a factor 1/3 to 1/6: apparently momentum is transferred quite efficiently from the photons to the wind. This can be attributed to the acceleration of the wind: the associated Doppler broadening of the spectral lines means a large part of the flux can be used (the outflowing atoms can absorb photons of a different, higher frequency as they get accelerated). This is a positive feedback mechanism that reinforces the wind driving.

The theory for radiation-driven winds is quite well developed, but the theoretical predictions for $\dot{M}$ are uncertain due to inhomogeneities in the wind ('clumping'). The uncertain clumping factor also affects the mass-loss rates inferred from observations, and current estimates are typically a factor

~ 3 lower than the the empirical rate of eq. (12.1). Radiation-driven mass loss is also dependent on metallicity, because it is mostly the lines of the heavier elements that contribute to the line driving. A dependence $\dot{M} \propto Z^{0.7}$ has been inferred both theoretically and from observations.

Red supergiant mass loss

Cool, luminous stars known as *red supergiants* (RSG) undergo a slow but copious stellar wind that is probably driven by the same mechanism as the 'superwind' of AGB stars: a combination of stellar pulsations and radiation pressure on dust particles that form in the cool outer atmosphere. There are no theoretical predictions, so we must rely on observations which imply very high values of $\dot{M}$ up to $10^{-4} M_\odot$/yr (eq. 12.1).

Stars with $M \lesssim 40\,M_\odot$ spend a large fraction of their *core He-burning* phase as red supergiants. During this phase a large part or even the entire envelope can be evaporated by the wind, exposing the helium core of the star as a Wolf-Rayet (WR) star (see Sect. 12.1.2).

12.1.1 The Humphreys-Davidson limit and luminous blue variables

Observations of the most luminous stars in our Galaxy and in the Magellanic Clouds have revealed a clear upper limit to stellar luminosities that depends on the effective temperature (see Fig. 12.2). In particular there are no red supergiants in HR diagram with $\log(L/L_\odot) > 5.8$, which corresponds to the expected RSG luminosity of a star of $40\,M_\odot$. Apparently stars with $M \gtrsim 40\,M_\odot$ do not become red supergiants.

The upper limit in the HRD is known as the *Humphreys-Davidson limit* after its discoverers, Roberta Humphreys and Kris Davidson. At $T_{\rm eff}$ above 10 000 K the maximum luminosity increases gradually to $\log(L/L_\odot) = 6.8$ at 40 000 K (O stars).

The existence of the HD limit is interpreted as a (generalized) Eddington limit. We have seen in Sec. 5.4 that when the luminosity of a star exceeds the classical Eddington limit (eq. 5.38),

$$L_{\rm Edd} = \frac{4\pi c G M}{\kappa_{\rm e}}, \tag{12.3}$$

where $\kappa_{\rm e}$ is the electron-scattering opacity, the outward force due to radiation pressure on the free electrons exceeds the gravitational force (on the nuclei) inwards. The electrostatic coupling between

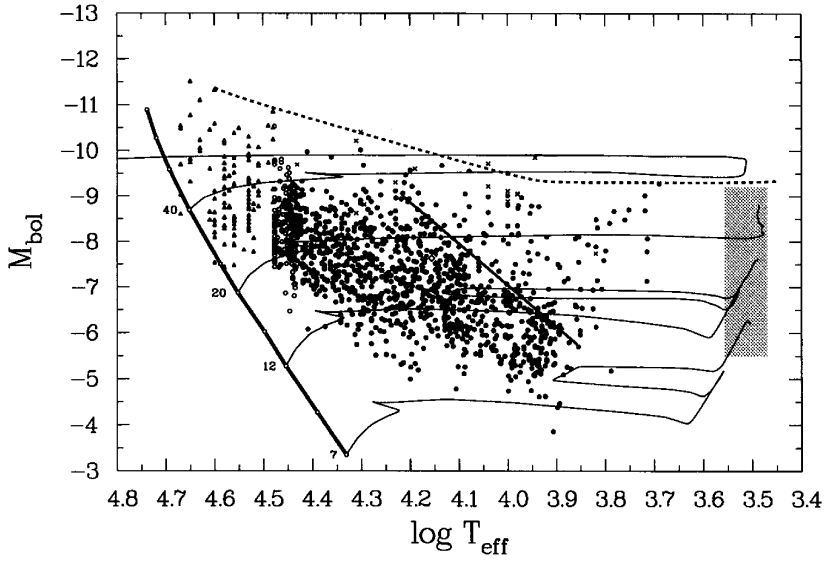


Figure 12.2. The HRD of the brightest supergiants in the LMC. The shaded region contains several hundred red supergiants that are not individually shown. The upper envelope of observed stars traced by the dotted line is known as the Humphreys-Davidson limit (the lower envelope is simply an observational cut-off). Figure adapted from Fitzpatrick & Garmany (1990, ApJ 363, 119).

electrons and ions means that the outer layers are accelerated outwards and the star becomes unstable. However, the *actual* opacity in the atmosphere is larger than the electron-scattering opacity, and decreases with temperature. Therefore the luminosity at which the radiation-pressure limit is reached is lower than the classical Eddington limit, and the decrease of the HD limit with decreasing $T_{\rm eff}$ can be explained at least qualitatively by this effect.

Luminous stars located near the HD limit are indeed observed to be very unstable, undergoing large excursions in the HRD and episodic mass loss with $\dot{M} \gtrsim 10^{-3}\,M_\odot$/yr during outbursts. These stars are known as *luminous blue variables* (LBVs), examples of which in our Galaxy are η Carinae and P Cygni. The remnants of the vigorous mass loss episodes are seen as circumstellar nebulae, which in the extreme case of η Car contains $\sim 10\,M_\odot$ ejected during an outburst in the mid-1800s. The nebula is considerably enriched in nitrogen, showing that the layers processed by CNO-cycling are being ejected. Stars losing mass due to LBV outbursts are destined to become *Wolf-Rayet stars*. The strong LBV mass loss prevents them from ever becoming red supergiants.

12.1.2 Wolf-Rayet stars

Wolf-Rayet (WR) stars are hot, very luminous stars with bright emission lines in their spectra. The emission indicates very strong, optically thick stellar winds, with mass-loss rates of $\dot{M} \sim 10^{-5} - 10^{-4}\,M_\odot$/yr. They are often surrounded by circumstellar nebulae of ejected material. The winds are probably driven by radiation pressure as for O stars, but multiple photon scattering in the optically thick outflow can increase the mass-loss rate to well above the single-scattering limit (eq. 12.2).

The spectra of WR stars reveal increased CNO abundances, indicating that they are the exposed H- or He-burning cores of massive stars. On the basis of the surface abundances they are classified into several subtypes:

WNL stars have hydrogen present on their surfaces (with $X_{\rm H} < 0.4$) and increased He and N abundances, consistent with equilibrium values from the CNO-cycle

WNE stars are similar to WNL stars in terms of their He and N abundances, but they lack hydrogen ($X_{\rm H} = 0$)

WC stars have no hydrogen, little or no N, and increased He, C and O abundances (consistent with partial He-burning)

WO stars are similar to WC stars with strongly increased O abundances (as expected for nearly complete He burning)

This is interpreted as an *evolutionary sequence* of exposure of deeper and deeper layers, as a massive star is peeled off to a larger and larger extent by mass loss (see Sec. 12.2).

12.2 Evolution of massive stars with mass loss in the HR diagram

Fig. 12.3 shows evolution tracks in the HRD for massive stars calculated with mass loss at metallicity $Z = 0.02$. As revealed by this figure, the evolutionary journey of a massive star through the HRD can be rather complicated. Evolution proceeds at nearly constant luminosity, because massive stars do not develop degenerate cores and most of the mass is in radiative equilibrium. However, the evolution track shows several left-right excursions and loops which depend on the mass of the star. The relation between the theoretical evolution tracks and the zoo of observed types of massive star encountered in Sec. 12.1 is described by the following *evolution scenario*, originally proposed by Peter Conti:

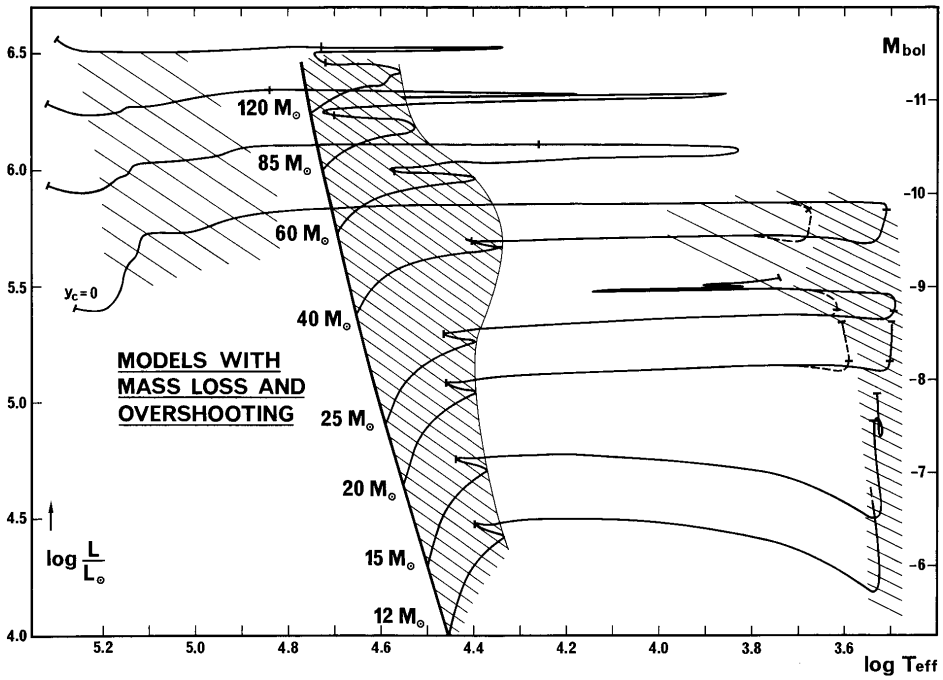


Figure 12.3. Evolution tracks of massive stars ($12 - 120\,M_\odot$) calculated with mass loss and a moderate amount of convective overshooting ($0.25\,H_P$). The shaded regions correspond to long-lived evolution phases on the main sequence, and during core He burning as a RSG (at $\log T_{\rm eff} < 4.0$) or as a WR star (at $\log T_{\rm eff} > 4.8$). Stars with initial mass $M > 40\,M_\odot$ are assumed to lose their entire envelope due to LBV episodes and never become RSGs. Figure from Maeder & Meynet (1987, A&A 182, 243).

$M \lesssim 15\,M_\odot$	MS (OB) → RSG (→ BSG in blue loop? → RSG) → SN II mass loss is relatively unimportant, $\lesssim$ few $M_\odot$ is lost during entire evolution
$15\,M_\odot \lesssim M \lesssim 25\,M_\odot$	MS (O) → BSG → RSG → SN II mass loss is strong during the RSG phase, but not strong enough to remove the whole H-rich envelope
$25\,M_\odot \lesssim M \lesssim 40\,M_\odot$	MS (O) → BSG → RSG → WNL → WNE → WC → SN Ib the H-rich envelope is removed during the RSG stage, turning the star into a WR star
$M \gtrsim 40\,M_\odot$	MS (O) → BSG → LBV → WNL → WNE → WC → SN Ib/c an LBV phase blows off the envelope before the RSG can be reached

The limiting masses given above are only indicative, and approximately apply to massive stars of Population I composition ($Z \sim 0.02$). Since mass-loss rates decrease with decreasing Z, the mass limits are higher for stars of lower metallicity. The relation of the final evolution stage to the supernova types indicated above will be discussed in Chapter 13.

The scenario for the most massive stars is illustrated in Fig. 12.4 for a $60\,M_\odot$ star. After about 3.5 Myr, while the star is still on the main sequence, mass loss exposes layers that formerly belonged to the (large) convective core. Thus CNO-cycling products (nitrogen) are revealed, and the surface He abundance increases at the expense of H. During the very short phase between central H and He burning ($t = 3.7$ Myr), several $M_\odot$ are rapidly lost in an LBV phase. During the first part of core He burning (3.7 – 3.9 Myr) the star appears as a WNL star, and subsequently as a WNE star (3.9 – 4.1 Myr) after mass loss has removed the last H-rich layers outside the H-burning shell. After 4.1 Myr

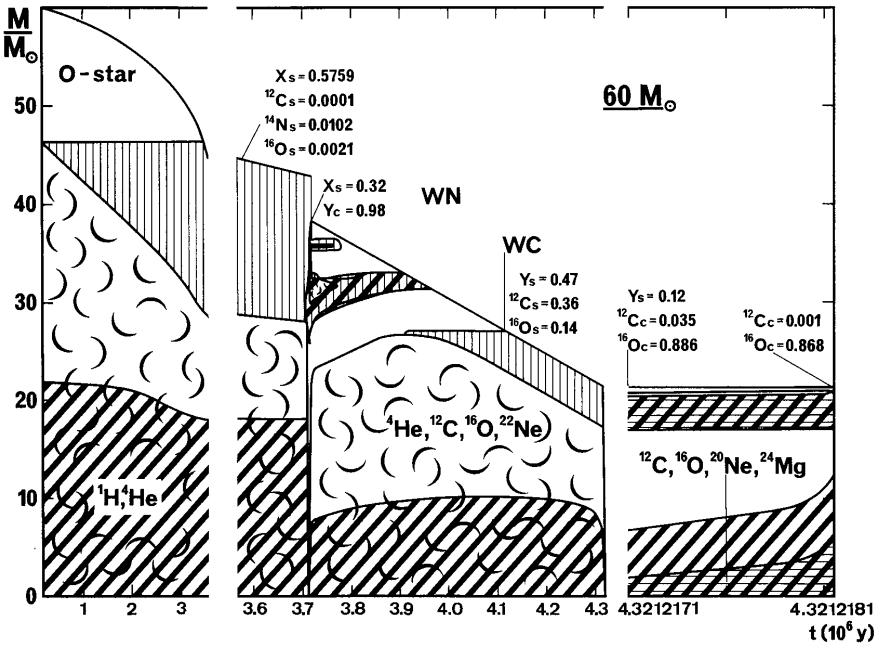


Figure 12.4. Kippenhahn diagram of the evolution of a $60\,M_\odot$ star at $Z = 0.02$ with mass loss. Cross-hatched areas indicate where nuclear burning occurs, and curly symbols indicate convective regions. See text for details. Figure from Maeder & Meynet (1987).

material that was formerly in the He-burning convective core is exposed at the surface: N, which was consumed in He-burning reactions, disappears while the products of He-burning, C and O, appear. The last 0.2 Myr of evolution this star spends as a WC star.

In general, mass-loss rates during all evolution phases increase with stellar mass, resulting in timescales for mass loss that are less that the nuclear timescale for $M \gtrsim 30\,M_\odot$. As a result, there is a *convergence* of the final (pre-supernova) masses to $\sim 5 - 10\,M_\odot$. However, this effect is much diminished for metal-poor stars because the mass-loss rates are generally lower at low metallicity.

12.3 Advanced evolution of massive stars

The evolution of the surface properties described in the previous section corresponds to the hydrogen and helium burning phases of massive stars. Once a carbon-oxygen core has formed after He burning, which is massive enough ($> 1.06\,M_\odot$) to undergo carbon ignition, the subsequent evolution of the *core* is a series of alternating nuclear burning and core contraction cycles in quick succession (see Fig. 12.5). The overall evolution trend is an increase of central temperature and central density, roughly following $T_c \propto {\rho_c}^{1/3}$ as expected from homologous contraction in our schematic evolution picture (Chapter 8). For central temperatures $\gtrsim 5\times10^8$ K, the evolution tracks deviate from this trend, sloping towards somewhat higher ρ_c and lower T_c. This is the result of cooling of the core by strong *neutrino emission* (see Sect 12.3.1).

The main effect of neutrino energy losses, however, is not visible in Fig. 12.5: they speed up the evolution of the core enormously. Less than a few thousand years pass between the onset of carbon burning until the formation of an iron core. During this time the mass of the C-O core remains fixed. Furthermore, the stellar *envelope* hardly has time to respond to the rapid changes in the core, with the consequence that the evolution of the envelope is practically disconnected from that of the core. As a result the position of a massive star in the HR diagram remains almost unchanged during carbon burning and beyond. We can thus concentrate on the evolution of the core of the star from this point onwards.

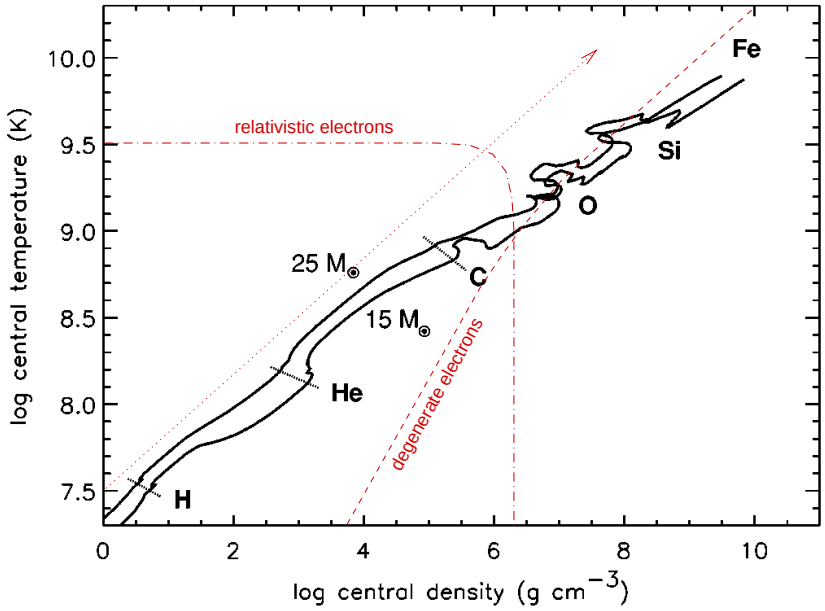


Figure 12.5. Evolution of central temperature and density of 15 $M_\odot$ and 25 $M_\odot$ stars at $Z = 0.02$ through all nuclear burning stages up to iron-core collapse. The dashed line indicated where electrons become degenerate, and the dash-dotted line shows where electrons become relativistic ($\epsilon_e \approx m_e c^2$). The dotted line and arrow indicates the trend $T_c \propto {\rho_c}^{1/3}$ that is expected from homologous contraction. Non-monotonic (non-homologous) behaviour is seen whenever nuclear fuels are ignited and a convective core is formed. Figure adapted from Woosley, Heger & Weaver (2002, Rev. Mod. Ph. 74, 1015).

12.3.1 Evolution with significant neutrino losses

In Sect. 6.5 we discussed several weak interaction processes that result in spontaneous neutrino emission at high temperatures and densities, such as photo-neutrinos, plasma-neutrinos and pair annihilation neutrinos. When the central temperature exceeds $\sim 5 \times 10^8$ K, these neutrino losses are the most important *energy leak* from the stellar centre, taking away energy much more rapidly than photon diffusion or even convection can transport it to the surface. From this point onwards the neutrino luminosity from the core far exceeds the luminosity radiated from surface, $L_\nu \gg L$.

The dependence of the nuclear energy generation rate $\epsilon_{\rm nuc}$ and the neutrino loss rate ϵ_ν on temperature are depicted in Fig. 12.6, for the centre of a typical massive star (i.e. following an evolution track approximating those shown in Fig. 12.5). Both ϵ_ν and $\epsilon_{\rm nuc}$ increase strongly with temperature, but the T-dependence of $\epsilon_{\rm nuc}$ is larger than that of ϵ_ν. During nuclear burning cycles energy production and neutrino cooling are in balance, $\epsilon_{\rm nuc} = \epsilon_\nu$, and this condition (the intersection of the two lines) defines the temperature at which burning takes place.[1]

During each nuclear burning phase, $L_{\rm nuc} = \dot{E}_{\rm nuc} \approx L_\nu$, which thus results in a much shorter nuclear timescale than if neutrino losses were absent: $\tau_{\rm nuc} = E_{\rm nuc}/L_\nu \ll E_{\rm nuc}/L$. Similarly, in between burning cycles the rate of core contraction (on the thermal timescale) speeds up: $\dot{E}_{\rm gr} \approx L_\nu$ so that $\tau_{\rm th} = E_{\rm gr}/L_\nu \ll E_{\rm gr}/L$. Therefore the evolution of the core speeds up enormously, at an accelerating rate as the core continues to contract and heat up. The lifetime of each nuclear burning stage can be estimated from Fig. 12.6 by approximating $\tau_{\rm nuc} \sim q/\epsilon_{\rm nuc}$, where q is the energy gain per unit mass from nuclear burning ($\sim 4.0, 1.1, 5.0$ and 1.9×10^{17} erg/g for C-, Ne-, O- and Si-burning,

[1] Note that because $\epsilon_{\rm nuc}$ is a steeper function of T than ϵ_ν, nuclear burning is stable also in the presence of neutrino losses: a small perturbation $\delta T > 0$ would increase the local heat content ($\epsilon_{\rm nuc} > \epsilon_\nu$), leading to expansion and cooling of the core until thermal equilibrium is re-established.

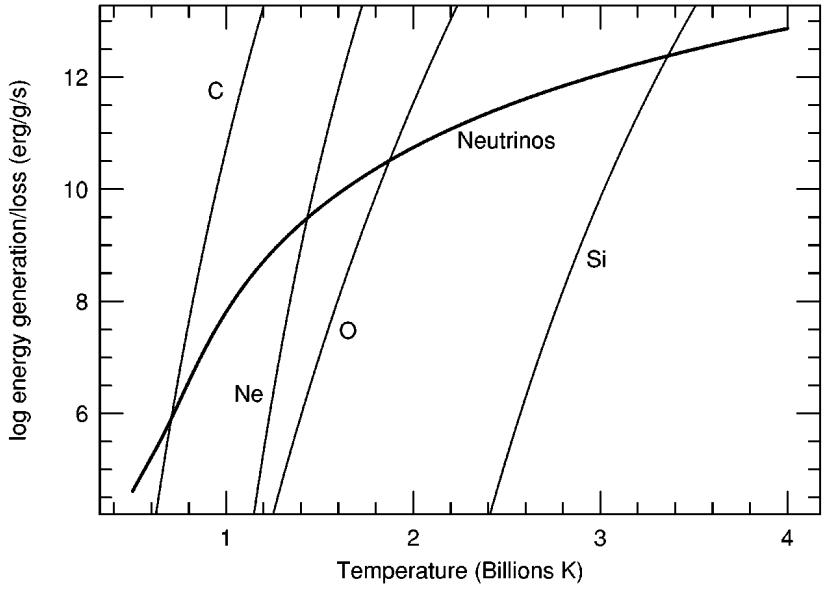


Figure 12.6. Energy generation rate and neutrino loss rate during the advanced evolution of a massive star. The stellar center is assumed to follow a track approximating that shown in Fig. 12.5. The intersections of the nuclear burning lines with the neutrino loss line define the burning temperature of the corresponding fuel. Figure from Woosley, Heger & Weaver (2002).

respectively) and $\epsilon_{\rm nuc}$ is the energy generated per gram and per second at the intersection with ϵ_ν in Fig. 12.6. Thus the lifetime ranges from several 10^3 years for C-burning to about a day for Si-burning!

12.3.2 Nuclear burning cycles: carbon burning and beyond

When the temperature in the contracting C-O core reaches $5 - 8 \times 10^8$ K (depending on the mass of the core), carbon is the first nuclear fuel to be ignited. The reactions involved in carbon burning and further nuclear burning cycles were treated in Sec. 6.4.3. In the following sections we briefly review these and discuss the consequences for the structure and evolution of the star. A typical example of the interior evolution is shown in Fig. 12.7 for a 15 $M_\odot$ star, and the corresponding stellar properties are given in Table 12.1.

Table 12.1. Properties of nuclear burning stages in a 15 $M_\odot$ star (from Woosley et al. 2002).

burning stage	T (10^9 K)	ρ (g/cm^3)	fuel	main products	timescale
hydrogen	0.035	5.8	H	He	1.1×10^7 yr
helium	0.18	1.4×10^3	He	C, O	2.0×10^6 yr
carbon	0.83	2.4×10^5	C	O, Ne	2.0×10^3 yr
neon	1.6	7.2×10^6	Ne	O, Mg	0.7 yr
oxygen	1.9	6.7×10^6	O, Mg	Si, S	2.6 yr
silicon	3.3	4.3×10^7	Si, S	Fe, Ni	18 d

Carbon burning

Carbon burning proceeds via the $^{12}C + {}^{12}C$ reaction, which produces a mixture of products, mainly ^{20}Ne and some ^{24}Mg. Most of the energy produced escapes in the form of neutrinos and only a small fraction is carried away by photons. In stars with masses up to about 20 $M_\odot$ the photon luminosity is large enough to produce a convective core (as shown in Fig. 12.7) of about 0.5 $M_\odot$. In more massive stars carbon burns radiatively, because the initial ^{12}C abundance is smaller and the luminosity not

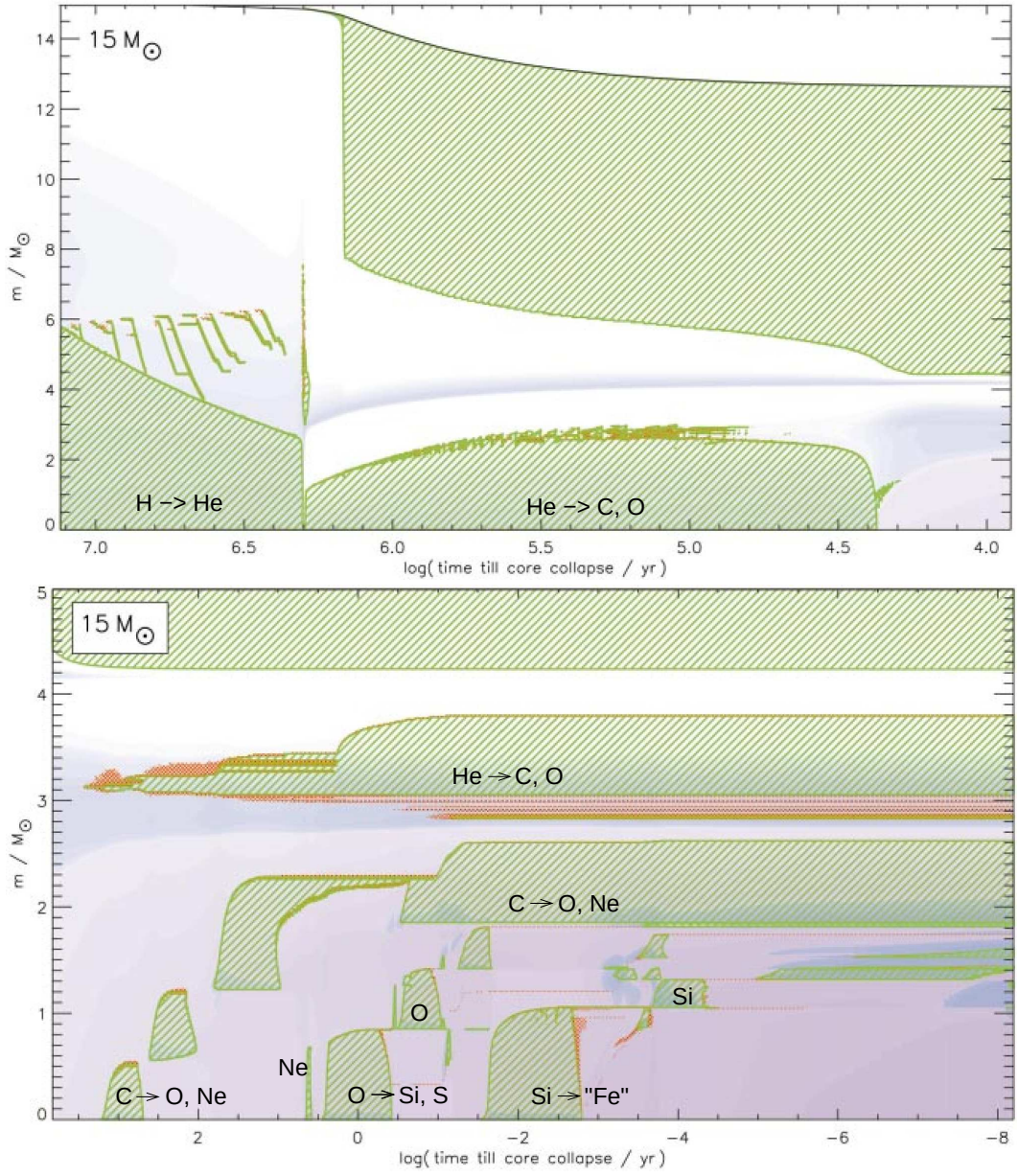


Figure 12.7. Kippenhahn diagram of the evolution of a 15 $M_\odot$ star showing convective regions (cross-hatching) and nuclear burning intensity (blue shading) during central H and He burning (top panel) and during the late stages in the inner 5 $M_\odot$ of the star (bottom panel). A complicated series of convective burning cores and shells appear, due to respectively carbon burning (around $\log t \sim 3$), neon burning (around $\log t \sim 0.6$), oxygen burning (around $\log t \sim 0$) and silicon burning (around $\log t \sim -2$). Figure from Woosley, Heger & Weaver (2002.)

carried away by neutrinos can all be transported by radiation. The duration of the C-burning phase is of the order of 10^3 yrs. It should be noted that these results are sensitive to the uncertain rate of the $^{12}C(\alpha,\gamma)^{16}O$ rate, which determines the ^{12}C abundance left after He-burning – a lower rate will leave more ^{12}C to be burned and this increases both the size of the convective core and the duration of the C-burning phase.

Following carbon exhaustion in the centre, the core – which is now composed mostly of O and Ne – contracts on its neutrino-accelerated Kelvin-Helmholtz timescale and carbon burning continues in a convective shell around this core. Several such convective shell-burning episodes can occur in succession, as shown in Fig. 12.7, their number depending on the mass of the star. The discrete nature of these shell burning events can also produce a discrete (discontinuous) dependence of the final state of the core on the initial stellar mass.

In stars with masses up to about $11\,M_\odot$ (C-O core masses less than $1.38\,M_\odot$) carbon burning proceeds somewhat differently. The C-O core becomes partially degenerate and neutrino losses effectively cool the centre of the star, so that carbon does not ignite in the centre but in an off-centre shell in a mildly degenerate flash (analogous to, but much less violent than the He flash in low-mass stars). After one or more of these mild carbon flashes the burning front moves to the centre and stable carbon burning in a convective core follows. After carbon burning, the O-Ne core becomes degenerate and no further nuclear fuels are ignited. The structure of these stars is then very similar to those of AGB stars with degenerate CO cores, discussed in Ch. 11, and such stars have been named *super-AGB stars*. The fate of such stars is uncertain and depends on whether the O-Ne core can reach the Chandrasekhar limit by shell burning. If this is the case the core eventually collapses, producing a supernova explosion. On the other hand, if mass loss is strong enough to remove the envelope before the Chandrasekhar limit is reached, the final remnant is a O-Ne white dwarf.

Neon and oxygen burning

In stars with masses $\gtrsim 11\,M_\odot$, once the temperature in the contracting O-Ne core reaches $\approx 1.5{\times}10^9$ K neon is 'burned' into oxygen and magnesium by a combination of photo-disintegration and α-capture reactions (Sec. 6.4.3). Neon burning always occurs in a convective core, regardless of stellar mass. By this time increased neutrino losses have accelerated the rate of evolution by a factor $\sim 10^3$ compared to the carbon-burning phase (see Fig. 12.6). The duration of the neon-burning phase is therefore very short, of order 1 year. Neon burning then shifts to a shell, as was the case for carbon burning, but in this case the time left until the next fuel is ignited is so short that no significant shell burning occurs.

When $T_9 \approx 2.0$ oxygen is ignited in the core by means of the $^{16}O + {}^{16}O$ reaction, producing mostly ^{28}Si and ^{32}S with a significant admixture of other isotopes (see below). Oxygen burning also occurs in a convective core, with a typical mass of $\approx 1.0\,M_\odot$ (see Fig. 12.7). The duration is somewhat longer than that of neon burning, of order 1 year, despite the higher neutrino loss rate at this stage. The reasons for this longer duration are the large oxygen mass fraction, ~ 0.7, and the large energy gain per gram compared to Ne burning. Similar to carbon burning, after central oxygen burning a number of convective oxygen-burning shells appears in quick succession. By this point the remaining time until core collapse (< 1 yr) is so short that the overlying helium- and carbon-burning shells remain frozen into the stellar structure.

Apart from ^{28}Si and ^{32}S, oxygen burning produces several neutron-rich nuclei such as ^{30}Si, ^{35}S and ^{37}Cl. Partly these result from α-captures on n-rich isotopes already present during C-burning, and partly from weak interactions (electron captures) such as $^{30}P(e^-,\nu)^{30}Si$. As a result the overall number of neutrons in the remnant Si-S core exceeds the number of protons (n/p > 1) and therefore that of electrons (implying that $\mu_e > 2$).

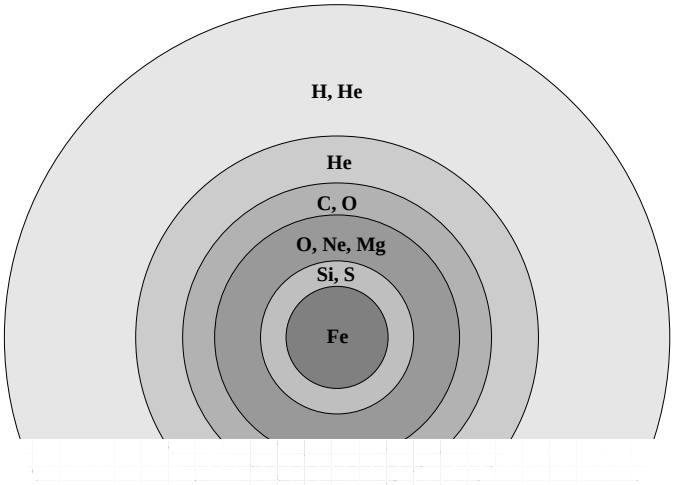


Figure 12.8. Schematic overview of the onion-skin structure of a massive star at the end of its evolution.

Silicon burning

When the central temperature exceeds 3×10^9 K, a process known as silicon burning starts. Rather than a fusion reaction this is a complex combination of photo-disintegration and α-capture reactions. Most of these reactions are in equilibrium with each other, and their abundances can be described by nuclear equivalents of the Saha equation for ionization equilibrium. For $T > 4 \times 10^9$ K a state close to *nuclear statistical equilibrium* (NSE) can be reached, where the most abundant nuclei are those with the lowest binding energy, i.e. isotopes belonging to the *iron group*. The abundances are further constrained by the total number of neutrons and protons present. Due to the neutron excess of the oxygen burning ashes (see above), the final composition is mostly ^{56}Fe and ^{52}Cr.

Silicon burning also occurs in a convective core of $\approx 1\,M_\odot$ and its duration is extremely short, of order 10^{-2} yr. As in previous phases, several convective shell-burning episodes usually follow in quick succession. The precise extent and number of these convective events determines the exact value of the final mass of the *iron core*, which has important consequences for the following core collapse and supernova phase (see Sec. 13.2).

12.3.3 Pre-supernova structure

We have obtained the following general picture of the final stages in the life of a massive star. The C-O core left by helium burning goes through a rapid succession of nuclear burning stages, during which the stellar envelope (and the star's position in the H-R diagram) remains largely unchanged. After the exhaustion of a fuel (e.g. carbon) in the centre, the remaining core contracts and burning continues in a surrounding shell. Neutrino losses speed up the contraction and heating of the core, until the next fuel (e.g. neon) is ignited in the centre. At each subsequent burning stage the outer burning shells have advanced outward, while neutrino cooling has become more efficient, resulting in a smaller burning core (in mass) than during the previous stage. Eventually this leads to an *onion-skin* structure, depicted schematically in Fig. 12.8. The star is composed of different concentric shells, which consist of heavier nuclei as one moves from the from the envelope towards the centre, and which are separated by burning shells. Often the nuclear burning, both in the centre and in shells, causes convective regions to appear that partially mix the various onion-skin layers. This leads to rather complicated abundance profiles at the moment when the inner core has gone through silicon burning and is composed of iron-group elements. An example of this structure is shown in Fig. 12.9 for a $15\,M_\odot$ star.

At this point the mixture of nuclei in the inner core has reached the minimum possible nuclear binding energy, given the ratio of neutrons to protons that is present, i.e. the composition is mainly

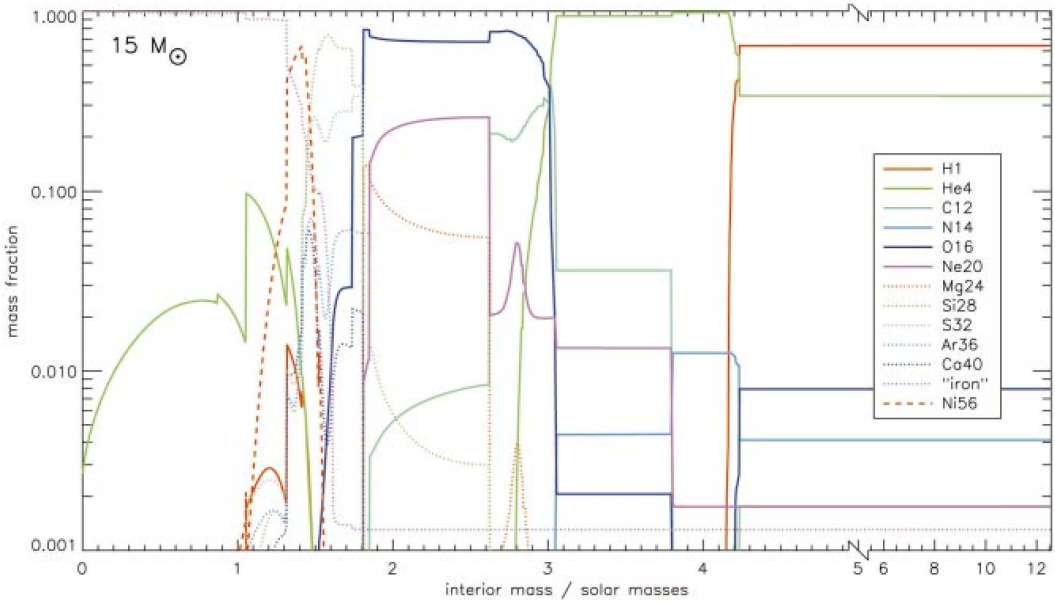


Figure 12.9. Final composition profiles of a 15 $M_\odot$ star (see Fig. 12.7), just before core collapse. "Iron" refers to the sum of neutron-rich nuclei of the iron group, especially ^{56}Fe. Figure from Woosley, Heger & Weaver (2002).

^{56}Fe. From this iron core no further energy can be extracted by nuclear fusion: it has become inert. The iron core rapidly becomes unstable and starts collapsing, giving rise to a supernova explosion. The collapse of the core and its consequences are discussed in the next chapter.

Suggestions for further reading

The evolution of massive stars, including the effects of mass loss and rotation, is treated in detail in Chapters 27 and 28.1-4 of Maeder. A thorough review of the current state of our understanding of the evolution of massive stars, their explosions and nucleosynthesis is given by Woosley, Heger & Weaver (2002, Rev. Mod. Ph., 74, 1015). Several of the figures from this article are reproduced in this chapter.

Exercises

12.1 Mass loss of massive stars during the main sequence

The mass-luminosity relation for massive stars on the main sequence is approximately

$$\log\left(\frac{L}{L_\odot}\right) \approx 0.781 + 2.760 \times \log\left(\frac{M_i}{M_\odot}\right),$$

where M_i is the initial mass. The mass loss rate of massive stars can roughly be approximated by

$$\log \dot{M} \approx -12.76 + 1.3 \times \log\left(\frac{L}{L_\odot}\right).$$

The duration of the main sequence phase $\tau_{\rm MS}$ in years is approximately

$$\log \tau_{\rm MS} \approx 7.719 - 0.655 \times \log\left(\frac{M_i}{M_\odot}\right).$$

(a) Calculate the fraction of mass that is lost by massive stars with $M_i = 25, 40, 60, 85$ and $120\,M_\odot$ during the main sequence phase.

(b) A star with an initial mass of $85\,M_\odot$ on the zero age main sequence has a convective core that contains 83 % of the mass. Calculate the time at which products of nuclear burning will appear at the surface.

(c) *Wolf-Rayet* (WR) stars are massive stars that have lost practically their complete hydrogen rich envelope. They can be classified according to their surface abundances:

WC No hydrogen, high abundances of He, C and O

WNE No Hydrogen, N/He ratio consistent with CNO equilibrium

WNL Some Hydrogen, N/He ratio consistent with CNO equilibrium

Put the sub-classifications in 'chronological order'. What type of WR is the star in question b)?

12.2 Maximum mass loss rate for a radiation driven wind

(a) Assume that all photons transfer their entire *momentum* to the outflowing wind. Show that the maximum mass loss rate that can be driven by radiation is given by

$$\dot{M} < \dot{M}_{\rm max} = \frac{L}{v_\infty c},$$

where v_∞ is the velocity of the wind at a large distance of the star.

(b) Show that with this maximum mass loss rate, the *kinetic energy* of the wind is only a small fraction of the luminosity, i.e.

$$\frac{1}{2}\dot{M}_{\rm max} v_\infty^2 \ll L \quad (v_\infty \approx 3 v_{\rm esc})$$

12.3 Burning stages

(a) Explain why the timescales of the burning stages from C-burning onward are very short compared to the H- and He-burning phases.

(b) Why does neon burning precede oxygen burning (why does it occur at a lower temperature) even though ^{20}Ne is a heavier nucleus than ^{16}O?

(c) The end result of nuclear burning in a massive star is an onion-like structure of the ashes of the various nuclear burning stages. Try to identify these layers, and the nuclear reactions that are responsible for them, in Figure 12.9.

Chapter 13

Stellar explosions and remnants of massive stars

13.1 Supernovae

Supernovae are stellar explosions during which the luminosity of a star reaches $10^9 - 10^{10}\,L_\odot$ at maximum, remaining bright for several months afterward. At least eight supernovae have been observed in our Galaxy over the past 2000 years, by Chinese and in some cases also by Japanese, Korean, Arabian and European astronomers (see Table 13.1). The remnants of these supernovae are in most cases still visible as luminous expanding nebulae, containing the matter that was expelled in the explosion. The supernova that left the remnant known as Cas A has not been reported, its explosion date has been inferred from the expansion rate of the nebula. Recently, however, the *light echo* of this supernova, as well as that of Tycho's supernova of 1572, have been detected from which the supernova type has been determined. No supernova is known to have occurred in our Galaxy in the last 340 years. Most of our observational knowledge comes from extragalactic supernovae, the first of which was discovered in 1885 in the Andromeda galaxy, and which are currently discovered at a rate of several hundred per year thanks to dedicated surveys. A Galactic supernova rate of about 1 every 30 years has been inferred from this.

Table 13.1. Historical supernovae.

year (AD)	V (peak)	SN remnant	SN type	compact object
185	−2	RCW 86	Ia?	–
386		?	?	
393	−3	?	?	
1006	−9	PKS 1459-41	Ia?	–
1054	−6	Crab nebula	II	NS (pulsar)
1181	−1	3C 58	II	NS (pulsar)
1572	−4	'Tycho'	Ia	–
1604	−3	'Kepler'	Ia?	–
~1667	$\gtrsim$+6	Cas A	IIb	NS

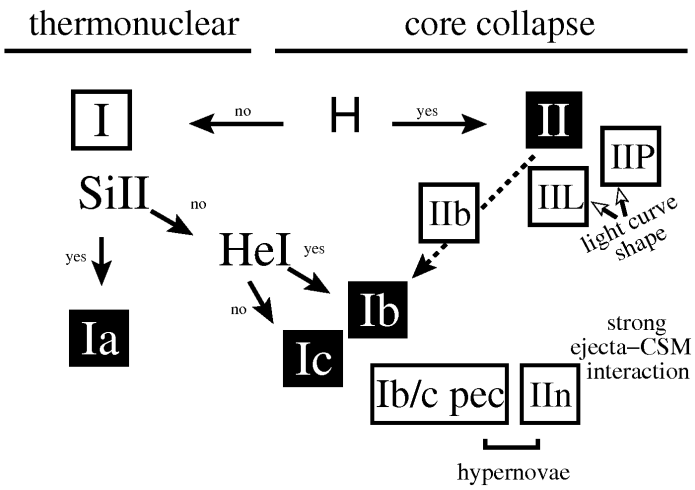


Figure 13.1. Classification of supernovae, based on their spectra and lightcurve shapes. The main supernova types are shown as black squares. Figure from Turatto (2003, LNP 598, 21).

Supernova classification

On the basis of their spectra, supernovae (SNe) have been historically classified into Type I (those that do not show hydrogen lines) and Type II (those that do). A more detailed classification scheme that is currently used, is shown in Fig. 13.1.

Type Ia The main spectral features are the lack of H lines and the presence of strong Si II lines around maximum brightness. After several months, lines of Fe and Co appear in the spectra. Type Ia supernovae occur in galaxies of all types, including elliptical galaxies which only contain old stellar populations, indicating that SNe Ia can have long-lived, low-mass progenitors. They are caused by the *thermonuclear explosion* of a CO white dwarf that reaches the Chandrasekhar limit $M_{\rm Ch}$ by mass accretion in a binary system (see Sec. 13.3). About 25–30% of observed supernovae are of Type Ia. They are (on average) the most luminous of all supernova types and their lightcurves (see Fig. 13.2) form a rather homogeneous group, which makes them of great interest as cosmological probes.

Type II The spectra of Type II supernovae are dominated by H lines, while lines of Ca, O and Mg are also present. SNe II occur in the spiral arms of galaxies where star formation takes place, and therefore correspond to the explosion of massive stars with short lifetimes. With about 50% of all supernovae, these are the most common type of stellar explosion. SNe II form the main class of explosions associated with the *core collapse* of massive stars that have hydrogen-rich envelopes. In several cases, the progenitor stars of Type II supernovae have been detected before the explosion. With the notable exception of SN 1987A (see Sec. 13.2.3), these progenitor stars were *red supergiants* with masses $8\,M_\odot \lesssim M \lesssim 16\,M_\odot$.

Type II supernovae show a variety of lightcurve shapes (Fig. 13.2), on the basis of which they are often sub-classified into **Type II-P** (showing, after an initial rapid rise and decline in brightness, a long 'plateau' phase of almost constant luminosity lasting 2–3 months, before a slow exponential decay) and **Type II-L** (which lack the plateau phase). In addition, one distinguishes **Type IIb**, in which the spectral signatures change from Type II to Type Ib (see below); and **Type IIn**, showing narrow emission lines on top of broad emission lines, which are interpreted as resulting from heavy mass loss prior to the explosion.

Type Ib and Ic Type Ib supernovae have strong He lines in their spectra, which are lacking in Type Ic supernovae. Both types show a lack of hydrogen, and strong lines of O, Ca and Mg are present.

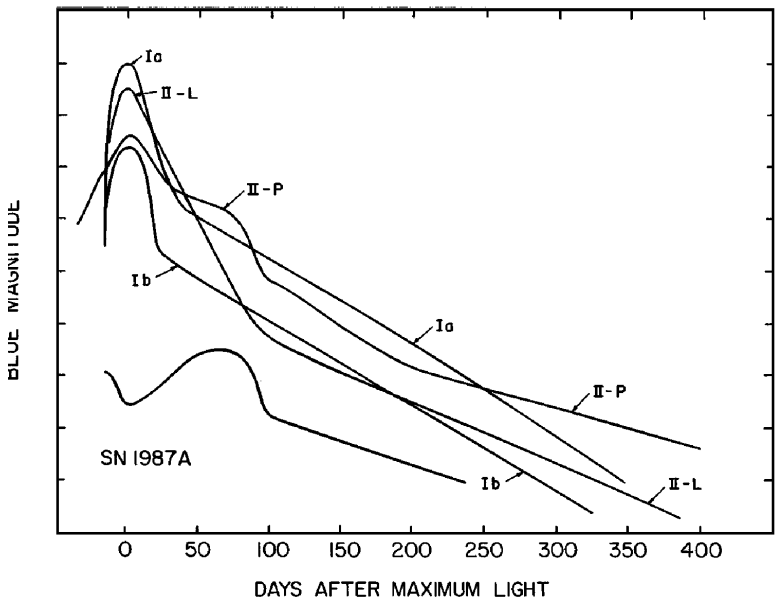


Figure 13.2. Schematic supernova lightcurves. Typical maximum *B*-band magnitudes are −19.5 for SNe Ia, −17.6 for both SNe Ibc and II-L, and −17.0 for SNe II-P. The lightcurves of SNe Ic resemble those of SNe Ib. Figure from Filippenko (1997, ARA&A 35, 309).

Similar to SNe II, they are found in star-forming regions, and their late-time spectra are also similar to Type II. Hence Type Ib/c supernovae are also associated with core collapse of massive stars, namely those that have lost their H-envelopes prior to explosion. Together they constitute about 20% of all supernovae. A subclass of very bright Type Ic supernovae, known as *hypernovae*, may be associated with gamma-ray bursts.

13.2 Core collapse and explosion of massive stars

As indicated in Fig. 13.1, essentially all types of supernova – except Type Ia – appear to be associated with the core collapse of massive stars ($\gtrsim 8\,M_\odot$) at the end of their evolution. The distinction between the different types and subtypes of core-collapse supernovae is related to differences in the structure and composition of the envelopes of the progenitor stars. For example, the progenitors of Type II supernovae are still surrounded by a massive H-rich envelope at the moment of explosion, while SN Ib progenitors have lost their H-rich envelopes and SN Ic progenitors have also lost most of the He layers surrounding the core. This sequence can be the consequence of mass loss from stars of increasing initial mass (see Sec. 12.2), but can also result from interaction with a binary companion.

13.2.1 The collapse of the iron core

Despite these differences in appearance, the *physical mechanism* is similar in all core-collapse supernovae. We have seen in Sec. 12.3.3 that stars with $M \gtrsim 11\,M_\odot$ form an inner core composed of iron-group elements (mostly ^{56}Fe) at the end of their evolution. From this iron core no further energy can be extracted by nuclear fusion: it has become inert.

The iron core is in a peculiar state in several respects. Because of neutrino cooling during the late evolution stages, the core typically has a considerable degree of electron degeneracy, except for the largest stellar masses (see Fig. 12.5). However, the high temperature and density ($\gtrsim 10^9$ g/cm^3) mean that the electrons are always relativistic (their typical energy exceeds $m_e c^2$). In that case contraction cannot be stopped, even if the core is degenerate, and must continue on the very rapid, neutrino-mediated thermal timescale. Furthermore, since the relativistic electron gas dominates the pressure,

the adiabatic exponent $\gamma_{\rm ad}$ is close to $\frac{4}{3}$. The iron core is therefore very close to a state of dynamical instability. Two processes occur at high density and temperature that contribute to accelerating the (already rapid) contraction into a dynamical collapse of the core.

Electron captures At very high density free electrons can be captured and bound into otherwise β-unstable heavy nuclei. This process, also known as inverse β-decay, occurs when the most energetic electrons have energies high enough to overcome the difference in nuclear binding energy (see also Sec. 11.2). As a result, the composition becomes increasingly neutron-rich – a process known as *neutronization*. Furthermore, the electron pressure decreases which can destroy the precarious state of hydrostatic equilibrium and trigger the collapse of the core.

If the core is significantly degenerate, the Chandrasekhar mass plays an important role. For a composition of predominantly ^{56}Fe one would expect $M_{\rm Ch} = 5.83\,\mu_{\rm e}^{-2}\,M_\odot \approx 1.26\,M_\odot$. Electron captures increase the average mass per free electron ($\mu_{\rm e}$) and thus decrease the effective Chandrasekhar mass. This can bring the core mass above this critical mass and facilitate its collapse.

Electron captures can also trigger the collapse of stars with initial masses below $\approx 11\,M_\odot$, which develop degenerate O-Ne cores at the end of their lives. If the mass of this core can grow (through shell burning) to $1.37\,M_\odot$, electrons are captured by ^{24}Mg and ^{20}Ne which initiates the collapse of the core. Stellar explosions caused by this mechanism are called *electron-capture supernovae*.

Photo-disintegration If the temperature in the contracting core reaches values close to 10^{10} K, the energy of the photons becomes large enough to break up the heavy nuclei into lighter ones, in particular ^{56}Fe is disintegrated into α particles and neutrons:

$$^{56}\mathrm{Fe} + \gamma \leftrightarrow 13\,{}^4\mathrm{He} + 4\,\mathrm{n} \tag{13.1}$$

This reaction is in statistical equilibrium and the abundances of the nuclei involved are determined by a Saha-type equation, the balance shifting more towards the right-hand side the higher the temperature. The process is thus similar to the ionization of hydrogen, and results in lowering $\gamma_{\rm ad}$ to below the critical value of $\frac{4}{3}$. The core therefore becomes dynamically unstable. This process dominates in relatively massive iron cores.

The photo-disintegration of ^{56}Fe requires a lot of energy, about 2 MeV per nucleon. This energy is absorbed from the radiation field and thus ultimately from the internal energy of the gas. As a result the pressure decreases quite drastically, triggering an almost free-fall collapse of the core.

The collapse is extremely rapid, taking of the order of 10 msec, because of the short dynamical timescale at the high density ($\sim 10^{10}$ g/cm^3) when the collapse is initiated. During the collapse the temperature and pressure keep rising, but never enough to reverse the collapse until nuclear densities are reached. Further photo-disintegrations can occur due to the increasing photon energies, which was once thought (prior to 1980) to dissociate even the α particles completely into free protons and neutrons ($^4\mathrm{He} + \gamma \rightarrow 2\,\mathrm{p} + 2\,\mathrm{n}$, which would require another 7 MeV per nucleon of internal energy from the gas). It has since become clear that full dissociation of ^{56}Fe into α particles and free nucleons does not occur during the collapse. On the other hand, electron captures onto protons ($\mathrm{p} + \mathrm{e}^- \rightarrow \mathrm{n} + \nu$) inside the heavy nuclei continue the process of neutronization, creating more and more neutron-rich nuclei. These eventually merge, creating what is essentially a gigantic stellar-mass nucleus, when ρ

approaches nuclear densities of the order $10^{14}\,\mathrm{g/cm^3}$. The composition inside the core is predominantly neutrons, which become degenerate and thereby modify the equation of state to suddenly become 'stiff', i.e. the neutron gas becomes almost incompressible. This terminates the collapse at a core radius of $R_\mathrm{c} \approx 20\,\mathrm{km}$.

Energetics of core collapse and supernova explosion

The gravitational energy released during the collapse of the core can be estimated as

$$E_\mathrm{gr} \approx -\frac{GM_\mathrm{c}^2}{R_\mathrm{c,i}} + \frac{GM_\mathrm{c}^2}{R_\mathrm{c,f}} \approx \frac{GM_\mathrm{c}^2}{R_\mathrm{c,f}} \approx 3 \times 10^{53}\,\mathrm{erg}, \tag{13.2}$$

assuming homologous collapse of a core of $M_\mathrm{c} \approx 1.4\,M_\odot$ from an initial radius $R_\mathrm{c,i} \sim 3000\,\mathrm{km}$ to a final radius $R_\mathrm{c,f} \approx 20\,\mathrm{km} \ll R_\mathrm{c,i}$. Let us compare this with the energy necessary to expel the envelope, which has no time to follow the core collapse,

$$E_\mathrm{env} = \int_{M_\mathrm{c}}^{M} \frac{Gm}{r}\,dm \ll \frac{GM^2}{R_\mathrm{c,i}}. \tag{13.3}$$

The upper limit (13.3) is $\approx 10^{53}\,\mathrm{erg}$ for $M = 10\,M_\odot$, but taking into account a realistic mass distribution in the envelope, this estimate comes down to $E_\mathrm{env} \sim 10^{50}\,\mathrm{erg}$. Only a very small fraction of the energy released in the collapse of the core is needed to blow away the envelope. Part of the energy goes into kinetic energy of the ejected envelope and energy radiated away by the supernova. For a typical Type II supernova, the ejected envelope is $\sim 10\,M_\odot$ and observed ejecta velocities are about $10^4\,\mathrm{km/s}$, giving $E_\mathrm{kin} \sim 10^{51}\,\mathrm{erg}$. The supernova has a luminosity $L \approx 2 \times 10^8\,L_\odot$ for up to several months, so that the total energy lost in the form of radiation is $E_\mathrm{ph} \sim 10^{49}\,\mathrm{erg}$. Therefore

$$E_\mathrm{ph} \sim 0.01 E_\mathrm{kin} \sim 10^{-4} E_\mathrm{gr} \qquad \text{and} \qquad E_\mathrm{gr} \gg E_\mathrm{env} + E_\mathrm{kin} + E_\mathrm{ph},$$

such that only a small fraction of the energy released in the collapse is used in the actual explosion. The question is how this fraction of about 1% of the gravitational energy can be converted into kinetic energy of the envelope, which turns out to be a very difficult problem.

13.2.2 The explosion mechanism

When the collapsing core reaches nuclear densities ($\rho_\mathrm{nuc} \approx 3 \times 10^{14}\mathrm{g/cm^3}$) the neutrons become degenerate, resulting in a strong increase in pressure. Furthermore, nuclear forces between the nucleons become important. These effects reverse the collapse. When the inner part of the core is compressed to ~ 1.5 times nuclear density, it bounces back like a spring – an event named *core bounce*.

As the velocity of the inner core material is reversed, it encounters matter from the still free-falling outer part of the core. If the collision were perfectly elastic, the outer core would bounce back to its initial radius even if the inner core were stationary. The outward motion of the inner core thus gives the possibility of a 'super-elastic' core bounce that might conceivably explode the star. The impact of the infalling matter is supersonic and creates a shock wave that steepens as it travels outward into regions of lower density. The kinetic energy stored in the shock wave was once thought to be sufficient to blow off the envelope, giving rise to a so-called *prompt explosion*. However, two problems arise that prevent such a prompt explosion to occur.

First, as the shock wave travels through the infalling matter which still mostly consists of iron-group nuclei, it heats up the nuclei and disintegrates them effectively into protons and neutrons. We can estimate the energy spent in photo-disintegration by noting that the binding energy of an ^{56}Fe nucleus is about 9 MeV/nucleon, so that the disintegration of an iron core of $1.4\,M_\odot$ (1.7×10^{57}

nucleons) requires about 2×10^{52} erg. Note that this amounts to absorbing, during a fraction of a second, practically all of the nuclear energy that was released during the lifetime of the star!

Second, electron captures on the free protons created behind the shock produce energetic neutrinos by means of

$$\mathrm{p} + \mathrm{e}^- \rightarrow \mathrm{n} + \nu.$$

These neutrinos take away the largest fraction, about 90%, of the energy released in the collapse, especially as the shock moves into relatively low-density regions ($< 10^{12}$ g/cm^3) from where they can easily escape. In the case of supernova 1987A these neutrinos have been detected (see Sec. 13.2.3). As a result, the shock wave fizzles out before it reaches the envelope of the star and no prompt explosion occurs.

Effects of neutrinos

The role played by neutrinos during core collapse requires closer examination. The neutrinos produced *before* the collapse set in had typical energies of the order of the thermal energy of the electrons (see Sect. 12.3.1). During the collapse neutrino production by neutronization (electron captures) dominates. The typical energy of these neutrinos is of the order of the Fermi energy of the relativistic electrons,

$$\frac{E_\nu}{m_\mathrm{e}c^2} \approx \frac{E_\mathrm{F}}{m_\mathrm{e}c^2} = \frac{p_\mathrm{F}}{m_\mathrm{e}c} = \frac{h}{m_\mathrm{e}c}\left(\frac{3}{8\pi}\,\frac{\rho}{\mu_\mathrm{e}m_\mathrm{u}}\right)^{1/3} \approx 10^{-2}\left(\frac{\rho}{\mu_\mathrm{e}}\right)^{1/3}, \tag{13.4}$$

using eq. (3.33) and the relation $\rho = \mu_\mathrm{e} m_\mathrm{u} n_\mathrm{e}$, and with ρ in g/cm^3 in the last equality. In the presence of heavy nuclei, the neutrinos interact mainly through so-called coherent scattering with these nuclei, with a typical cross section of the order

$$\sigma_\nu \approx 10^{-45} A^2 \left(\frac{E_\nu}{m_\mathrm{e}c^2}\right)^2 \mathrm{cm}^2, \tag{13.5}$$

which gives together with eq. (13.4),

$$\sigma_\nu \approx 10^{-49} A^2 \left(\frac{\rho}{\mu_\mathrm{e}}\right)^{2/3} \mathrm{cm}^2. \tag{13.6}$$

If $n = \rho/(Am_\mathrm{u})$ is the number density of nuclei, the mean free path ℓ_ν of the neutrinos in the collapsing core can then be estimated as

$$\ell_\nu \approx \frac{1}{n\sigma_\nu} \approx 2 \times 10^{25}\,\frac{1}{\mu_\mathrm{e}A}\left(\frac{\rho}{\mu_\mathrm{e}}\right)^{-5/3} \mathrm{cm}. \tag{13.7}$$

Taking $\mu_\mathrm{e} \approx 2$ and $A \approx 100$, we find with eq. (13.7) that when $\rho/\mu_\mathrm{e} \approx 4 \times 10^9$ g/cm^3, the neutrino mean free path $\ell_\nu \approx 10^7$ cm, which is the typical dimension of the collapsing core. Apparently, neutrinos can no longer escape freely at the high densities prevailing in the collapsing core. The core becomes opaque for neutrinos, which can only diffuse out of the core via many scattering events. Towards the end of the collapse phase, when $\rho > 3 \times 10^{11}$ g/cm^3, the diffusion velocity even becomes smaller than the infall velocity of the gas, so that neutrinos are *trapped* in the core. Analogous to the photosphere of a star, one can define a 'neutrinosphere' in the outer layers of the core where the density is low enough for the neutrinos to escape. Interior to this, there is a 'neutrino trapping surface' below which the neutrinos are trapped.

The real situation is much more complicated because σ_ν depends on the neutrino energy, so that the neutrino transport problem has to be solved in an energy-dependent way. The congestion of neutrinos in the core causes them to become degenerate (since neutrinos are fermions) with a high Fermi energy. Electron capture becomes less probable, because the new neutrinos have to be raised to the top of the Fermi sea. Neutronization therefore effectively stops when $\rho \approx 3 \times 10^{12}$ g/cm^3. Only after some neutrinos have diffused out of the core can further neutronization take place. The process of neutronization therefore takes about 3–10 seconds, while the collapse only takes a few milliseconds.

The deposition of neutrino energy in the core provides an energy source that may revive the shock wave and cause an explosion. Neutrinos diffusing out of the collapsed core (the proto-neutron star) heat the region through which the former shock wave has passed, which stretches from ∼30 km to 100–300 km, and cause it to become convectively unstable. Convection thus provides a way to convert some of the thermal energy from neutrino deposition into kinetic energy. Multi-dimensional hydrodynamical calculations show that the outward force thus created can overcome the ram pressure of the outer layers that are still falling onto the core and launch a successful explosion, but only for rather low initial stellar masses (up to $\sim 11\,M_\odot$).

A recently proposed alternative is that the proto-neutron star becomes unstable to g-mode oscillations, which generate acoustic energy that builds up in the shocked region. These acoustic waves eventually cause an anisotropic explosion, whereby the core still accretes on one side while the explosion occurs in the other direction. The asymmetric explosion that results may help explain the large space velocities of radio pulsars, which indicate that neutron stars receive a 'kick' at birth.

13.2.3 Lightcurves of core-collapse supernovae

The main physical parameters that determine the appearance of a supernova are:

- the total kinetic energy imparted by the explosion into the envelope,
- the structure (density profile and chemical composition) of the pre-supernova star, as well as the possible presence of circumstellar material lost by the star earlier in its evolution,
- energy input by decay of radioactive isotopes ejected in the explosion.

As we have estimated above, the typical kinetic energy of the explosion is of the order of $E_{\rm kin} \approx 10^{51}$ erg.[1] Given the uncertainty in the precise physical mechanism that converts ∼1% of the core-collapse energy into an explosion (Sec. 13.2.2), one usually models these explosions by injecting a specified amount of energy at the bottom of the envelope by means of a 'piston'. Both $E_{\rm kin}$ and the mass boundary between core and envelope (or 'mass cut') are uncertain and are usually treated as free parameters.

The visible supernova explosion starts when the shock wave induced by the piston reaches the stellar surface, giving rise to a short pulse (∼30 minutes) of soft X-ray emission. The luminosity then declines rapidly as the stellar surface expands and cools. The expanding envelope remains optically thick for a sufficient amount of time that most of the explosion energy is converted into kinetic energy of the outflow. When the envelope has expanded enough to become optically thin, only ∼1% of the initial kinetic energy has been converted into radiation, as the total amount of energy radiated away in the supernova is about 10^{49} erg.

When a massive H-rich envelope is present, the recombination of ionized hydrogen provides an additional source of energy once the envelope has become optically thin and cools efficiently. As the

[1] The quantity of 10^{51} erg is sometimes referred to as 'f.o.e.' in the supernova literature, and has recently been defined as a new unit 'bethe' (1 B = 10^{51} erg) after Hans Bethe, a pioneer in supernova studies.

envelope keeps expanding, a recombination wave moves inward in mass coordinate, while staying at roughly the same radius and temperature. This gives rise to the plateau phase in the lightcurve of a Type II-P supernova. This phase ends when the recombination wave dies out as it meets the denser material of the inner envelope which expands at smaller velocity ($< 10^3$ km/s). If the H-rich envelope is not massive enough to sustain such a recombination wave, the plateau phase is absent (Type II-L lightcurves).

In the last phase of the supernova the lightcurve is determined by the radioactive decay of isotopes released in the explosion. The main source of radioactive energy is ^{56}Ni, which undergoes two electron captures to produce stable ^{56}Fe:

$$^{56}\mathrm{Ni} + \mathrm{e}^- \rightarrow {}^{56}\mathrm{Co} + \nu + \gamma \qquad (\tau_{1/2} = 6.1\,\mathrm{d})$$
$$^{56}\mathrm{Co} + \mathrm{e}^- \rightarrow {}^{56}\mathrm{Fe} + \nu + \gamma \qquad (\tau_{1/2} = 77\,\mathrm{d})$$

The exponential decline of the luminosity after 50–100 days corresponds to the decay of ^{56}Co. The amount of ^{56}Ni ejected in the explosion, required to explain the observed lightcurves, is about 0.07 $M_\odot$ for a typical Type II supernova. This puts constraints on the position of the 'mass cut' between the collapsing core and the ejected envelope (the remainder of the ^{56}Ni synthesized is locked up in the collapsed compact object). The lightcurves of Type Ib and Ic supernovae are completely dominated by this radioactive decay, after the initial luminosity peak caused by shock breakout. Other radioactive isotopes (with longer half-lives than ^{56}Co) can also play a role in the lightcurve at later stages.

SN 1987A

This supernova (Type II) in the Large Magellanic Cloud was the nearest supernova observed since Kepler's supernova in 1604. Its progenitor is known from images taken before the supernova: surprisingly it was a *blue* supergiant, with $L \approx 1.1 \times 10^5\,L_\odot$ and $T_{\rm eff} \approx 16\,000$ K, and a probable initial mass of about 18 $M_\odot$. Its relative faintness at peak magnitude is probably related to the compactness of the progenitor star compared to the red supergiant progenitors of SNe II. SN 1987A is the only supernova from which *neutrinos* have been detected, shortly before the visible explosion. During 10 seconds, detectors in Japan and the USA detected 20 neutrinos with energies between 8 and 40 MeV. These energies and the 10 sec time span correspond to the transformation of an Fe core into a hot proto-neutron star during core collapse (see Sec. 13.2.2).

13.2.4 Final masses and remnants

Figure 13.3 shows the possible relation between the initial mass of a star of solar metallicity, the mass just before the moment of explosion and the final mass of the remnant. This figure is based on a particular set of stellar models, and the precise masses are dependent on the assumed mass-loss rates, convective overshooting etc., and also depend on metallicity. The pre-supernova mass is determined by mass loss during the evolution of the star, which becomes important for initial masses $\gtrsim 15\,M_\odot$ (Sec. 12.2). For masses above $\sim$30 $M_\odot$, mass loss is strong enough to remove the H-rich envelope as well part of the material that was inside the He core and even the C-O core, shown as green shading.

The type of stellar remnant left behind depends, first of all, on whether the collapse of the iron core successfully generates a supernova explosion. As discussed above, this is still an area of large uncertainty. Detailed numerical simulations do indicate, however, that a successful explosion is more likely the lower the initial mass of the star, or rather, the lower the mass of its C-O core. Stars with initial masses up to 20 $M_\odot$ probably leave *neutron star* remnants. With increasing mass, the amount of kinetic energy generated by the collapse decreases, while the binding energy of the envelope increases. If only a weak explosion is generated, some of the material ejected may fall back onto the

proto-neutron star. If accretion causes the mass to exceed the maximum possible mass of a neutron star – which is uncertain but probably lies in the range 2–3 $M_\odot$– then the proto-neutron star will collapse and form a *black hole*. The mass limit separating stars that form neutron stars and those that leave black holes is probably in the range 20–25 $M_\odot$, but is sensitive to the details of the explosion mechanism as well as to the maximum neutron-star mass. It is even possible that, due to the non-linear behaviour of mass loss, the relation between initial mass and final remnant is non-monotonic and that stars above a certain mass again leave neutron stars (as suggested in Fig. 13.3). On the other hand, if mass loss is weak and a massive C-O core is left prior to core collapse, a successful supernova shock may not develop at all and the entire star may collapse directly into a black hole.

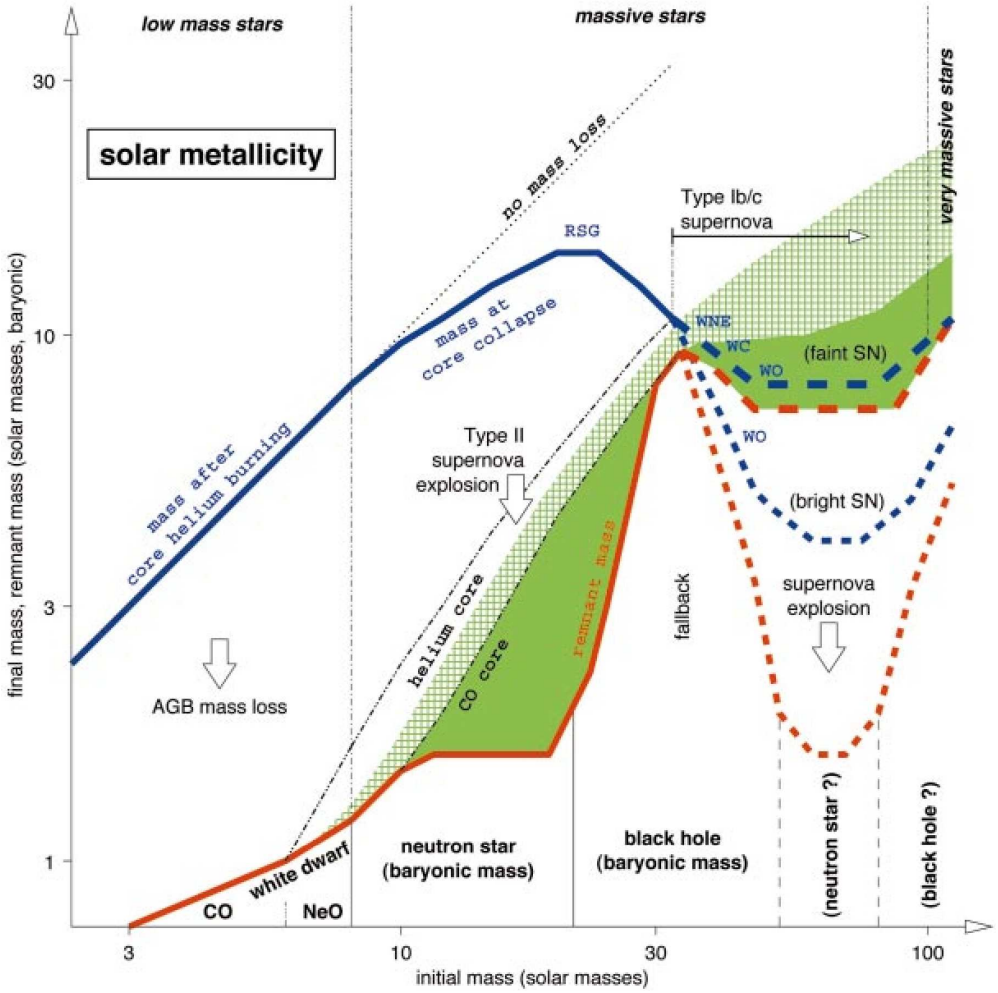


Figure 13.3. Initial-final mass relation for stars of solar composition. The blue line shows the stellar mass after core helium burning, reduced by mass loss during earlier phases. For $M \gtrsim 30\,M_\odot$ the helium core is exposed as a WR star, the dashed line gives two possibilities depending on the uncertain WR mass-loss rates. The red line indicates the mass of the compact stellar remnant, resulting from AGB mass loss in the case of intermediate-mass stars, and ejection of the envelope in a core-collapse supernova for massive stars. The green areas indicate the amount of mass ejected that has been processed by helium burning and more advanced nuclear burning. (Figure from Woosley et al. 2002).

13.3 Type Ia supernovae

Type Ia supernovae are fundamentally different from other SN types, because they are not associated with the core collapse of massive stars. Instead they are caused by the *thermonuclear explosion* of a CO white dwarf that reaches a critical mass for carbon ignition.

Carbon-burning reactions can occur in a low-temperature degenerate gas if the density is sufficiently high, about 2×10^9 g/cm^3 (these are known as *pycno-nuclear reactions*, see Sec. 6.2.3). These densities are reached in the centre when the mass is very close to the Chandrasekhar mass of 1.4 $M_\odot$. Because the gas is strongly degenerate, carbon burning is unstable and leads to a strong increase in temperature at constant density and pressure. This is analogous to what happens during the core He flash in low-mass stars, except that the degeneracy is so strong that it can only be lifted when the temperature has reached about 10^{10} K. The ignition of carbon therefore causes the incineration of all material in the core of the white dwarf to Fe-peak elements (in nuclear statistical equilibrium). An explosive burning flame starts to propagate outwards, behind which material undergoes explosive nuclear burning. The composition of the ashes depends on the maximum temperature reached behind the flame, which decreases as the burning front crosses layers of lower and lower densities (although still degenerate). The composition is mainly ^{56}Ni in the central parts, with progressively lighter elements (Ca, S, Si, etc) in more external layers. The total energy released by nuclear burning is of order 10^{51} erg, which is sufficient to overcome the binding energy of the white dwarf in the explosion. Therefore no stellar remnant is left.

The lightcurve of a Type Ia supernova is powered by the radioactive decay of the ^{56}Ni formed in the explosion. The nickel mass is a substantial fraction of the mass of the white dwarf, $0.5 - 1.0\ M_\odot$, which is the main reason for the greater peak luminosities of SNe Ia compared to most core-collapse supernovae. About 50 days after maximum brightness, an exponential decay of the lightcurve occurs due to radioactive decay of ^{56}Co into ^{56}Fe.

In single stars of intermediate mass, the degenerate CO core cannot grow to the Chandrasekhar limit because mass loss quickly removes the envelope during the AGB phase (Ch. 11). Even if the Chandrasekhar limit were reached, the remaining H-rich envelope would cause a strong hydrogen signature in the supernova spectrum which is not seen in SNe Ia. Therefore it is commonly agreed upon that the CO white dwarfs that cause SN Ia explosions grow by accreting mass in a binary

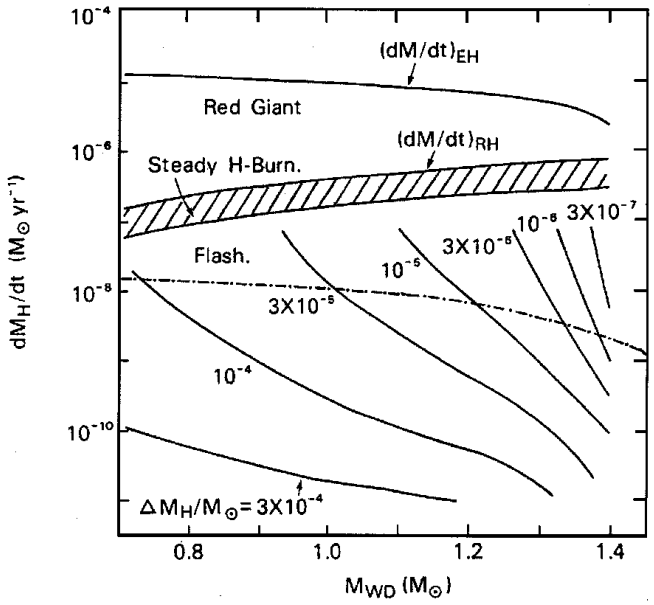


Figure 13.4. Critical mass transfer rates for hydrogen-accreting white dwarfs, as a function of the WD mass. Only for a small range of mass transfer rates (hatched area) can the material quietly burn on the WD surface, and thus lead to a growth of the WD mass towards the Chandrasekhar mass and a SN Ia explosion. (Figure from Kahabka & van den Heuvel 1997).

system. However, the exact mechanism by which this happens is still a matter of debate. Two types of progenitor scenarios are being discussed:

The single degenerate scenario In this scenario the white dwarf accretes H- or He-rich matter from a non-degenerate binary companion star: a main-sequence star, a red giant or a helium star (the stripped helium core of an initially more massive star). The difficulty is that steady burning of H and He, leading to growth of the mass of the white dwarf, is possible only for a narrow range of accretion rates (see Fig. 13.4). If accretion is too fast, a H-rich envelope is formed around the white dwarf (which would have an observable signature if the WD explodes). If accretion is too slow, the accreted matter burns in unstable flashes (*nova* outbursts) that throw off almost as much mass as has been accreted, such that the WD mass hardly grows. At present such models are too restrictive to explain the observed rate of SN Ia in galaxies.

The double degenerate scenario In this case the Chandrasekhar limit is reached by the merging of two CO white dwarfs in a close binary system. Such a close double WD can form as a result of strong mass and angular momentum loss during binary evolution (a process called *common envelope* evolution). Once a close double WD system is formed, angular momentum loss by gravitational waves can bring about the eventual merger of the system. Although at present no convincing evidence exists for a double WD binary with a total mass in excess of $M_{\rm Ch}$, the theoretical merger rate expected from binary evolution models appears sufficient to explain the observed SN Ia rate (note, however, that these models have large uncertainties). The main doubt about this scenario is whether the C-burning initiated by the WD merger leads to the required incineration and explosion of the merged white dwarf, or proceeds quiescently and results in a core collapse.

Suggestions for further reading

See Chapter 28.4-6 of Maeder.

Exercises

13.1 Energy budget of core-collapse supernovae

(a) Neutron stars have a radius of about 10 km. Use this to estimate the energy generated during a core collapse supernova (Hint: assume that before the collapse the core is like a white dwarf with mass $M_{\rm c} = M_{\rm Ch}$, where $M_{\rm Ch}$ is the Chandrasekhar mass, and that it suffers no significant mass loss after the collapse).

(b) The kinetic energy measured in the supernovae ejecta is about $E_{\rm kin} = 10^{51}$ erg. What is a typical velocity of the ejecta, if the original star was one of $10\,M_\odot$?

(c) The supernova shines with a luminosity of $2 \times 10^8\,L_\odot$ for about two months. Estimate the total energy in form of photons.

(d) Which particles carry away most of the energy of the supernova? Assuming an average energy of 5 MeV of those particles, how many of them are created by a supernova?

13.2 Neutrino luminosity by Si burning

Silicon burning forms iron out of silicon. Assume that 5 MeV of energy are liberated by creating one ^{56}Fe nucleus from silicon, and that the final result of this burning is an iron core of about 2 $M_\odot$. Silicon burning only lasts about one day, as most of the liberated energy is converted into neutrinos (of about 5 MeV each).

(a) Compare the corresponding neutrino luminosity to that of Supernova 1987A, which can be well approximated by the calculations in Exercise 13.1.

(b) Now, knowing that this supernova was 50 kpc away, and that about 10 neutrinos were detected during one second: how close does the silicon burning star have to be, such that we can detect the neutrino emission?

13.3 Carbon ignition in a white dwarf

When a white dwarf approaches the Chandrasekhar mass, its central density exceeds 2×10^9 g/cm^3, carbon is ignited under degenerate conditions which will quickly burn the whole white dwarf to iron-group elements (mainly ^{56}Ni). Compare the energy liberated by nuclear fusion to the gravitational binding energy of the white dwarf. What will be the outcome? (Use the masses of ^{12}C, ^{16}O and ^{56}Ni nuclei in Table 6.1, and assume that the white dwarf is composed of equal mass fractions of ^{12}C and ^{16}O.)

Contents