

INTRODUCTION AND BASIC CONCEPTS

1.1 Stars in the Broader Context of Modern Astrophysics

The study of stars continues to be at the very core of modern astrophysics more than 5000 years since its inception. Astronomy may well be the oldest science, and when early bronze age man looked up at the night sky he or she saw mostly stars. The signs of the zodiac—the first attempt at classifying, or at least grouping, stars—have their origin in Mesopotamia, the land between the Tigris and Euphrates rivers (modern Iraq) where the Sumerian, Assyrians and Babylonians civilisations flourished.

Today stars play a key role in many branches of astronomy, as well as being of interest in their own right.

- ‘The epoch of reionisation’, which has been described as the ‘last frontier in observational cosmology’ marks the period—approximately 600 million years after the Big Bang—when the Universe experienced its last ‘phase-change’: it changed from being mostly neutral to being mostly ionised (see Figure 1.1).

This transition was caused by the ‘First Stars’ which are believed to have been very massive and highly luminous *at wavelengths below 912 Å*. Photons with such short ultraviolet wavelengths have energies $E = h\nu > 13.6 \text{ eV}$, higher than the ionisation potential of hydrogen. The First Stars were thus able to ionise large volumes around them; when sufficient numbers of such stars had formed, their ionised surroundings, or H II regions, overlapped and most of the volume of the Universe became transparent and accessible to observations from Earth.

A great deal of effort, both theoretical and observational, is currently being devoted to elucidating the nature of the First Stars.

- When massive stars explode, either as supernovae or gamma-ray bursts, they can outshine a whole galaxy. They can thus be observed at the

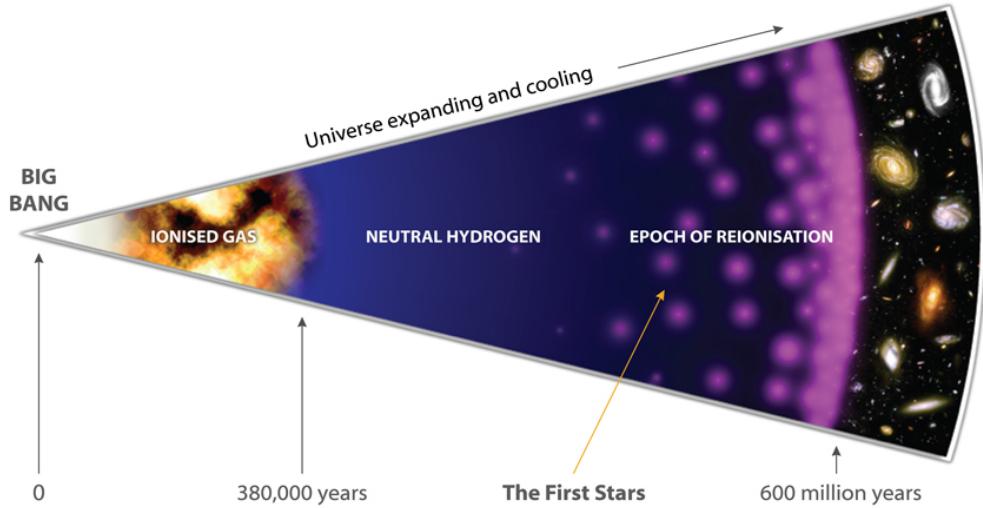


Figure 1.1: Schematic representation of the ‘Epoch of Reionisation’. Time progresses from left to right.

largest distances. At present, the record for the most distant known object in the Universe is held not by a galaxy, nor a quasar, but by a gamma-ray burst (GRB) at redshift $z \sim 9.4$. A gamma-ray burst is thought to mark the end of a *massive and rapidly rotating star*, when its core collapses directly into a black hole and two extremely energetic jets of plasma are emitted from its rotational poles at nearly the speed of light (see Figure 1.2). The massive star that produced GRB 090429B at $z \sim 9.4$ formed when the Universe was only ~ 500 Myr old, or $\sim 4\%$ of its present age. Observations of these very distant GRBs give us a unique window on the early stages in the evolution of the Universe.

- The properties of stars are sufficiently well known that it is possible to reconstruct, using a mixture of theory and observations, the emergent spectrum of a whole population of stars, i.e. a whole galaxy. This ‘spectral synthesis’ technique is now used widely to interpret photometric and spectroscopic observations of galaxies over most of the history of our Universe.
- A relatively recent development in astronomy is the realisation that many (most) stars have their own planetary systems. At the time of writing, there are 1137 planetary systems and 1822 planets (467 multiple planet systems) known (<http://exoplanet.eu/catalog.php>). Nearly all of these have been found by very precise observations of

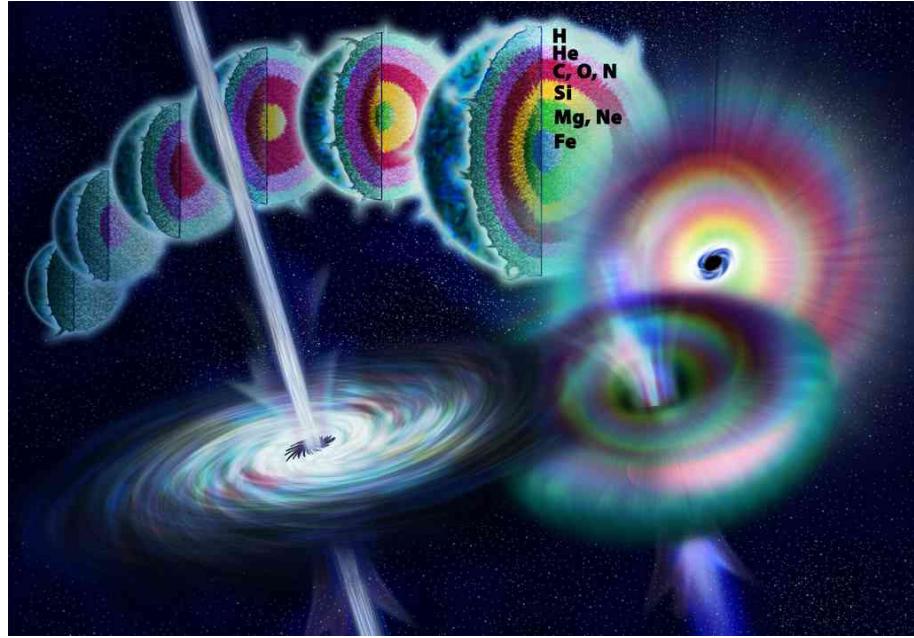


Figure 1.2: Artist’s impression of the late stages in the evolution of a massive star, leading to a gamma-ray burst.

their *stars*, by either radial velocity, photometry (transiting planets), or microlensing measurements.

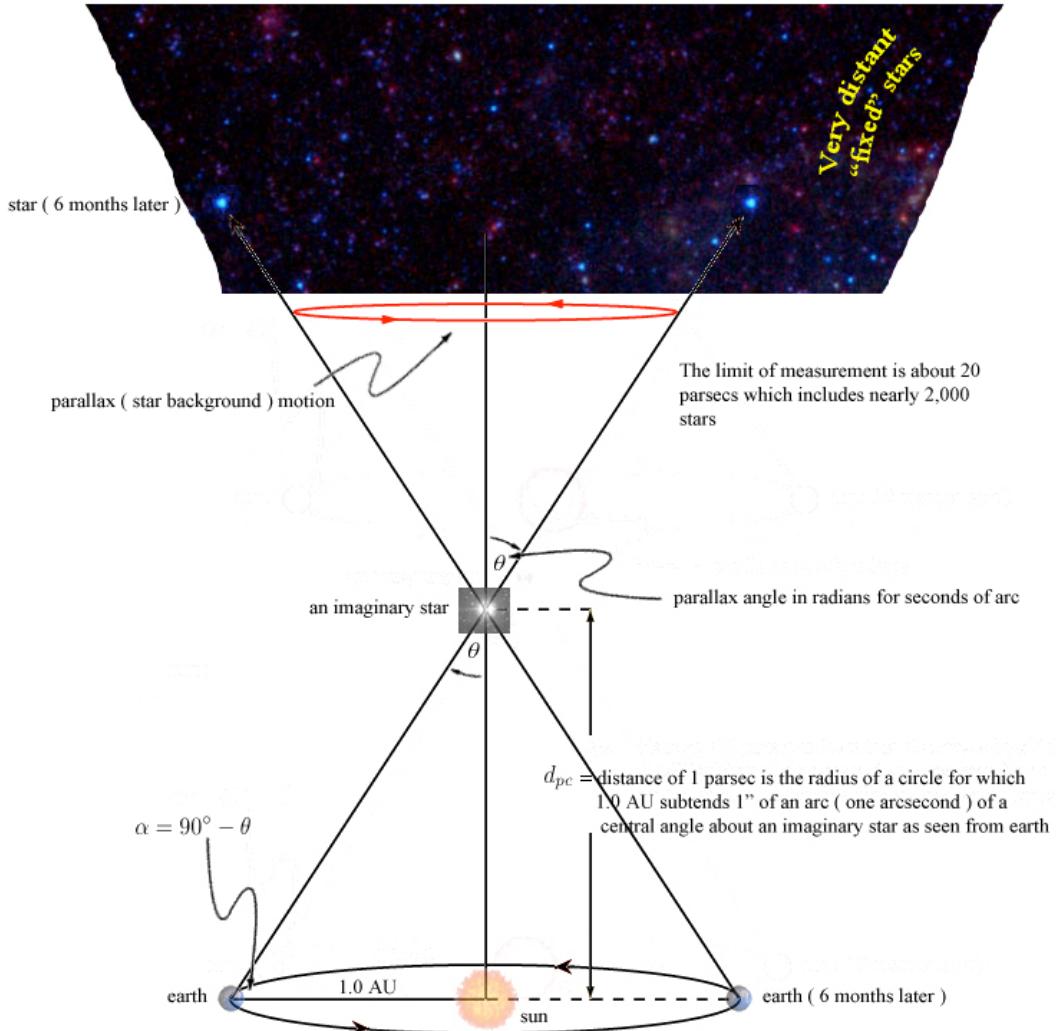
1.2 Useful Astronomical Units

1.2.1 Distances

- Radius of Earth’s orbit around the Sun $\equiv 1$ Astronomical Unit (AU) $= 1.5 \times 10^{13}$ cm
- Distance at which 1 AU subtends an angle of 1 second of arc (see Figure 1.3):

$$\frac{1.5 \times 10^{13} \times 360 \times 60 \times 60}{2\pi} = 3.1 \times 10^{18} \text{ cm} \equiv \boxed{1 \text{ parsec}} \quad (1.1)$$

- pc, kpc, Mpc are the most commonly used distance units in astronomy, e.g.:
 - 1.3 pc = distance from the Sun to the nearest star (Proxima Centauri)
 - 8.5 kpc = distance from the Sun to the Galactic Centre



Where:

θ = angle of parallax in radians for seconds of arc

$d_{pc} = 1/\theta$, distance to an imaginary star in parsecs and is the radius of a circle for which 1.0 AU subtends 1.0'' (one second) of arc of a central angle about an imaginary star as seen from earth

note: the word parsec stands for "Parallax of one arcsecond"

Figure 1.3: Parallax. When the angle θ subtended by a nearby star is 1 arcsec, the star is at a distance of 1 parsec.

- 1 Mpc = size of the Local group of galaxies
- 20 Mpc = distance to the Virgo cluster of galaxies
- 4300 Mpc = Hubble radius (radius of the observable universe)
 $= c/H_0$, where H_0 is the value of the Hubble constant today = $70 \text{ km s}^{-1} \text{ Mpc}^{-1}$



Figure 1.4: The nearby galaxy NGC 6744 observed with the Wide Field Imager on the MPG/ESO 2.2-metre telescope at La Silla. This large spiral galaxy is similar to the Milky Way and is at a distance of about 10 Mpc in the southern constellation of Pavo (The Peacock).

1.2.2 Masses

- 1 solar mass ($1 M_\odot$) = 2×10^{33} g
- Stellar mass of the Milky Way $\simeq 1 \times 10^{11} M_\odot$

1.2.3 Luminosity and Magnitudes

Astronomers measure the brightness of celestial objects on a log scale which goes backwards!

$$F_1/F_2 = 10^{0.4(m_2 - m_1)} \quad (1.2)$$

where F are the fluxes and m the *apparent magnitudes*. We define flux as the energy that passes per unit time through a unit area (so, for example, the energy per unit time, or the power, collected by a telescope of area T is just $F \times T$). The luminosity is the total power emitted by a source,

$L = 4\pi d^2 F$, where d is the distance, assuming that the source radiates isotropically. This assumption is generally a good approximation for stars, but breaks down dramatically for sources of beamed radiation, such as γ -ray bursts.

Magnitudes and luminosities are usually measured over a range of wavelengths (see Lecture 2). Note that:

- $\Delta m = +0.75 \rightarrow (F_1/F_2) = 2$
- $\Delta m = +1 \rightarrow (F_1/F_2) = 2.5$
- $\Delta m = +2.5 \rightarrow (F_1/F_2) = 10$
- $\Delta m = +5 \rightarrow (F_1/F_2) = 100$

The *absolute magnitude* M of an object is the magnitude that object would have if placed at a distance of 10 pc, from which we define a *distance modulus* (in magnitudes)

$$m - M = 2.5 \log(d/10)^2 = 5 \log(d/10) = 5 \log d - 5 \quad (1.3)$$

where d is the distance in parsec.

Distances, and therefore absolute magnitudes, are notoriously difficult to measure in astronomy.

1.3 What is a Star?

A useful definition of a star is an object in which nuclear reactions are (or have been) sufficient to balance surface radiation losses.

1.3.1 The birth of stars

We can understand better this definition in the context of our broad picture of stellar formation and evolution. The starting point is when a cloud of cold, molecular gas is pushed out of equilibrium by some physical process (e.g. a passing shock wave). As the cloud collapses under its own gravity,

gravitational potential energy is turned into kinetic energy of its particles (i.e. heat). As a result the temperature and density of the cloud increase as it collapses. Once the temperature of the cloud core reaches sufficiently high values, in excess of several million degrees, nuclear fusion can begin, converting hydrogen into helium. Thermonuclear reactions now provide a non-gravitational source of energy, whose heating can supply the pressure required to support the gas against further collapse: a star is born!

1.3.2 The death of stars

Nuclear fusion is exothermic because the rest mass of a He nucleus is less than the sum of the rest masses of two neutron and two protons. This mass defect, corresponding to an energy $E = mc^2 = 26 \text{ MeV}$ which is the binding energy of the He nucleus, is released during nuclear fusion as kinetic energy into the surrounding gas. Thermonuclear reactions can proceed in the interior of stars fusing lighter elements into heavier ones

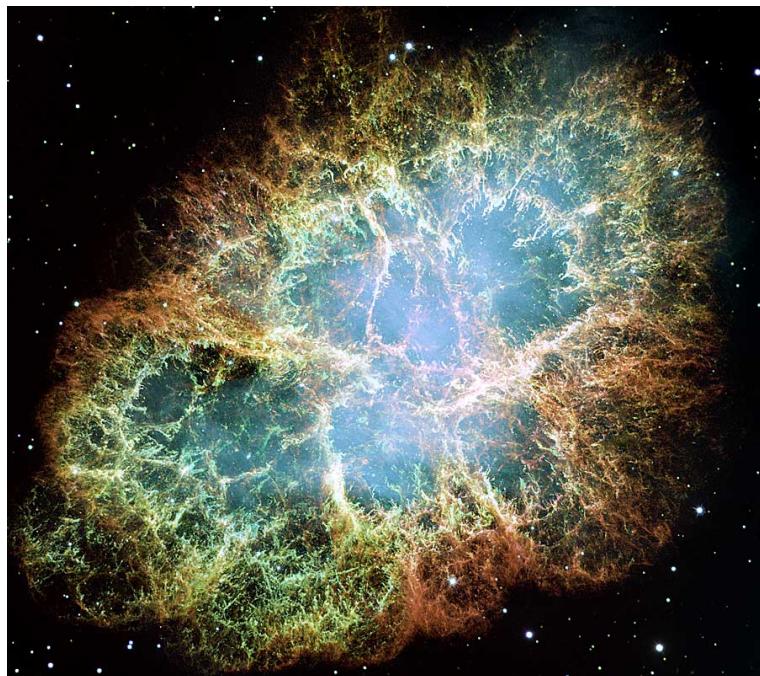


Figure 1.5: *Hubble Space Telescope* image of the Crab Nebula, the remnant of a massive star that exploded as a Type II supernova in 1054 A.D. The nebula is $\sim 2 \text{ kpc}$ from the Sun, in the constellation of Taurus. The ejected gas is still moving outwards with speeds of $\sim 1500 \text{ km s}^{-1}$. At the centre of the nebula lies the Crab Pulsar, a rapidly spinning neutron star.

up to and including Fe. However, elements heavier than Fe have a *lower* binding energy per nucleon than Fe—that is, energy has to be supplied, rather than extracted from the system in order to produce elements heavier than Fe by nuclear fusion.

Thus, when a star has developed an iron core, it can no longer supply the pressure to balance the inward pull of gravity. Most stars end their lives by ejecting their outer layers, either in supernova explosions (see Figure 1.5), or in less energetic mass loss events such as planetary nebulae (see Figure 1.6).



Figure 1.6: *Hubble Space Telescope* image of the Helix Nebula, the prototypical and closest planetary nebula, at a distance of only ~ 200 pc. The nebula consists of the outer layers of a solar-mass star, expelled towards the end of its life and made to glow by the ultraviolet light emitted by the hot core left behind (visible at the centre of the nebula). The expansion velocity is $\sim 20 \text{ km s}^{-1}$. The core, with a mass of $\sim 0.6 M_{\odot}$, will eventually become a white dwarf. Our own Sun may well turn into a planetary nebula in about five billion years.

1.4 Galactic Chemical Evolution

In either case, some fraction of the elements synthesised in the stellar interior during the life of the star and in the supernova explosion which marked its death is returned to the interstellar medium where it will be available for subsequent cycles of star formation. To our knowledge, all the elements of the Periodic Table—including those so central to life as we know it—were manufactured in the interiors of stars. The only exception are the light elements created by Big Bang nucleosynthesis in the first three minutes or so of the existence of our Universe: H, He, Li and some of their isotopes. For convenience, astronomers refer to all elements heavier than Li as ‘metals’, to indicate elements created by stellar, rather than primordial, nucleosynthesis.

There is thus a cycle of Galactic chemical evolution and indeed of ‘Cosmic’ chemical evolution when we consider the Universe as a whole. Presumably the gas in our Galaxy initially had a very small fraction of ‘metals’ which increased over the last 13 billion years or so with the progress of star formation. When we look at a recent region of star formation, such as the Orion nebula, we find the following composition, by mass:

Hydrogen = X = 0.7392

Helium = Y = 0.2486

Everything Else = Z = 0.0122.

Thus, the cumulative effect of all the previous star formation activity in the Milky Way Galaxy—where 80-90% of the baryonic mass is in stars and 10-20% is in gas—has been to enrich the interstellar medium to just over 1% of the mass in metals.

1.4.1 Stellar populations

When we consider the stars in our Galaxy, we find distinct populations, with different kinematics, ages, and metallicities (see Figure 1.7):

- **The disk.** It contains most of the stellar mass of the Milky Way: $M_{\text{disk}} = 4 \times 10^{10} M_{\odot}$. Rotationally supported: at the location of the Sun, $v_{\text{rot}} = 220 \text{ km s}^{-1}$. Metal-rich: $Z \simeq 0.2\text{--}1.5 Z_{\odot}$. Young stars are

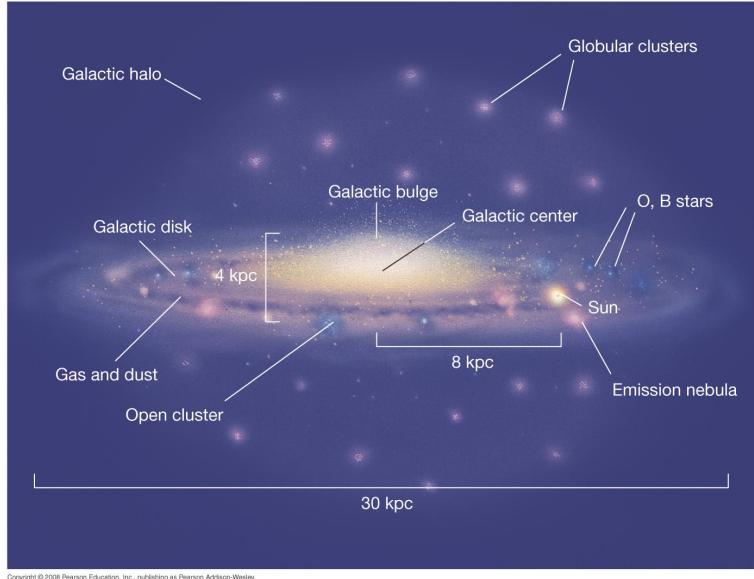


Figure 1.7: Schematic representation of the stellar populations of our Galaxy.

found here; current star-formation rate: $\text{SFR} \simeq 1 M_\odot \text{ yr}^{-1}$.

- **The bulge.** $M_{\text{bulge}} = 2 \times 10^{10} M_\odot$. The disk is not uniformly thick. Near the centre, there is a thicker, roughly spherical region that has different properties from the rest of the disk. It contains a mix of stars, some recently formed and some old, as well as a supermassive black hole ($M_{\text{bh}} \simeq 4.1 \times 10^6 M_\odot$).
 - **The halo.** $M_{\text{halo}} = 1 \times 10^9 M_\odot$. Kinematics are dispersion dominated. Contains some of the oldest objects in the Universe, the Globular Clusters, and some of the most metal-poor stars known, some with less than 10^{-5} the amount of metals (Fe) present in the interstellar gas today. (Partly) made up of the remnants of previous merger events.
- The properties of these different stellar populations give us clues to the processes that led to the formation and evolution of our Galaxy.*

BASIC PROPERTIES OF STARS

2.1 Introduction

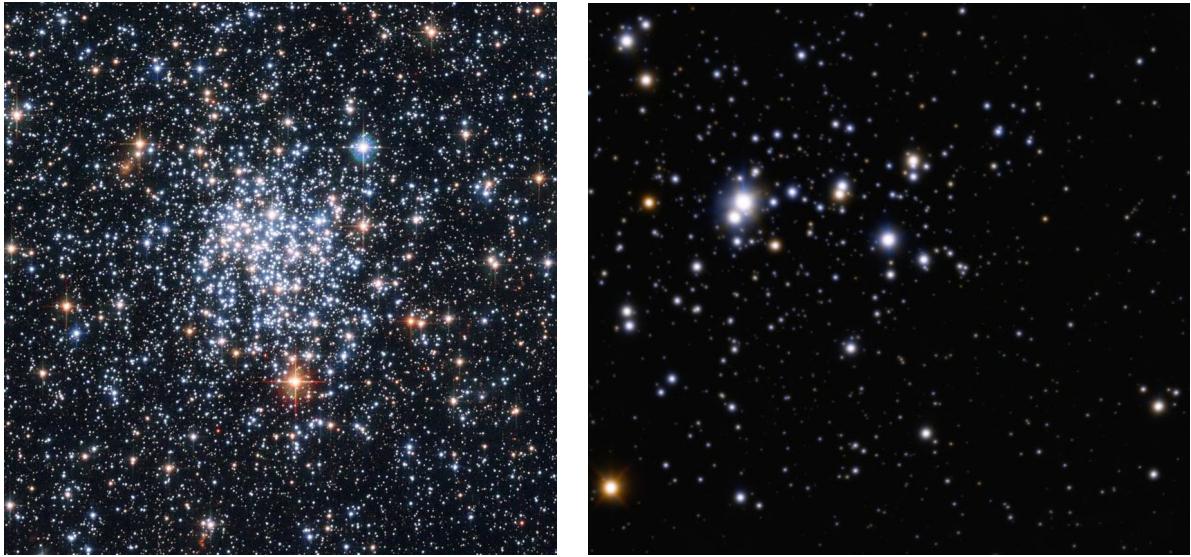


Figure 2.1: Left: *Hubble Space Telescope* image of the star cluster NGC 265 in the Small Magellanic Cloud. Right: ESO *Very Large Telescope* image of the cluster Trumpler 14 in the Carina nebula. This ground-based image has been sharpened with Adaptive Optics techniques.

Figure 2.1 shows two examples of star clusters. Star clusters allow us to appreciate directly some of the physical properties of stars for the simple reasons that, to a first approximation, all the stars of a cluster: (i) are at the same distance from the Sun, and (ii) have the same age. NGC 265 (left panel of Figure 2.1) is approximately 300 million years old and is located in the Small Magellanic Cloud, one of our two companion galaxies at a distance of ~ 60 kpc. Trumpler 14 (right panel of Figure 2.1) is one of the youngest stellar clusters known, with an age of only 1 million years. It is associated with the Carina nebula at a distance of 3.2 kpc.

It is immediately obvious from these images that: (a) stars have a range of colours, and (b) some stars are intrinsically brighter than others. More generally, we can make a list of what we may consider to be the most important physical properties of a star:

1. Mass
2. Temperature
3. Luminosity
4. Gravity
5. Age
6. Chemical Composition

These parameters are all inter-related but, to a first approximation, it is the first one, the mass of a star, that determines its temperature, luminosity, surface gravity and lifetime. The chemical composition is only a second order effect. In this lecture, we'll consider how some of these properties are determined. However, before we can measure the intrinsic characteristics of a star, we need to establish its distance.

2.2 Stellar Distances

We have already encountered the stellar parallax in Lecture 1. Referring to Figure 1.3, with $R = 1 \text{ AU}$, we have:

$$\frac{R}{d} = \tan \theta \simeq \sin \theta \simeq \theta \quad (2.1)$$

for small angles θ , and therefore

$$d = \frac{R}{\theta}. \quad (2.2)$$

By definition, $d = 1 \text{ pc}$ when $\theta = 1 \text{ arcsec}$.

Note that the nearest star, Proxima Centauri, is at a distance $d = 1.3 \text{ pc}$, corresponding to $\theta = 0.764 \text{ arcsec}$, comparable to the size of a stellar image as measured from the ground through the turbulence introduced by the Earth's atmosphere. For this and other complicating reasons, $\theta \gtrsim 0.01 \text{ arcsec}$, $d \lesssim 100 \text{ pc}$ is the limit of parallax measurements from the ground.

The *Hipparcos* satellite, a European Space Agency (ESA) Space Astrometry Mission launched in 1989, successfully observed the celestial sphere for 3.5 years and measured the positions (Right Ascension and Declination) and parallax of over 100 000 stars within $d \lesssim 1000$ pc ($\theta \gtrsim 0.001$ arcsec).

When completed, the current ESA Space Astrometry Mission, *Gaia*, launched in December 2013, will be a major leap forward:

- Positional accuracy: from 0.001 arcsec to 0.00001 arcsec ($\times 10^2$)
- Distance: from 1 kpc to 100 kpc, sufficient to reach stars in neighbouring galaxies!
- Sensitivity: from 10 mag to 20 mag ($\times 10^4$)
- Number of stars: from 10^5 to 10^9 ($\times 10^4$)

Note 1: By combining the parallactic distance with the angular position on the sky (RA and Declination), we have the location of a star in 3-D.

Note 2: Parallactic stellar distances are the first rung of the ‘cosmic distance ladder’. Consider two stars, with observed magnitudes m_1 and m_2 respectively, such that $m_2 > m_1$ (i.e. star 2 is fainter than star 1). Consider the case where star 1 has a parallactic distance, but star 2 is too far away to give a measurable parallax. If we have reasons to believe that the two stars have the same *absolute magnitude*, we can derive a *photometric distance* for star 2 using the inverse square behaviour of stellar fluxes, F :

$$\frac{F_2}{F_1} = 10^{0.4 \times (m_1 - m_2)} = \left(\frac{d_1}{d_2} \right)^2 \quad (2.3)$$

using the relation between stellar fluxes and magnitudes given in Lecture 1, so that

$$d_2 = d_1 \times 10^{-0.4 \times (m_1 - m_2)/2} \quad (2.4)$$

The above equations assume that the dimming of star 2 relative to star 1 is due entirely to its greater distance and neglects other possible sources of dimming, such as interstellar extinction by dust. This is the method used

to derive distances on cosmological scales; it relies on establishing a series of ‘standard candles’, by which astronomers mean astronomical sources of known absolute magnitude, at increasing distances to build a cosmological distance ladder.

2.2.1 Proper Motion

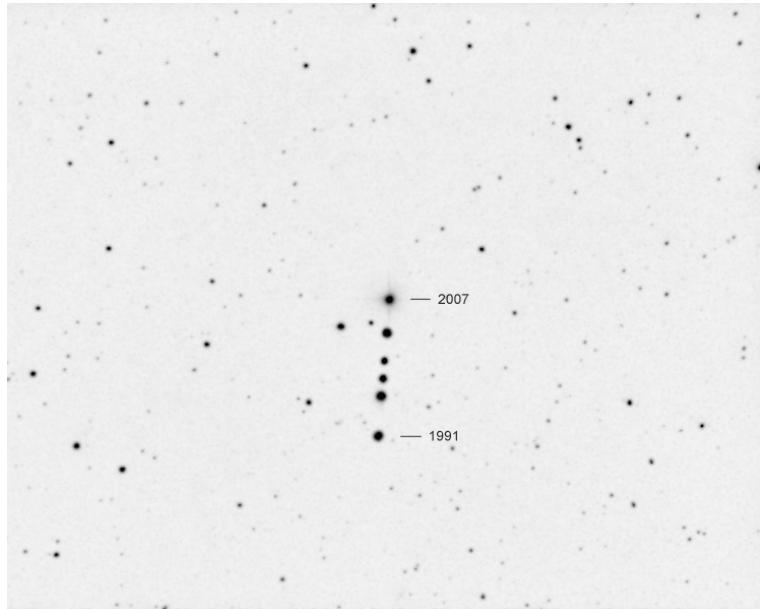


Figure 2.2: Proper motion of Barnard’s star. This low-mass red dwarf is the fourth closest star to the Sun after the three components of the α Cen system, at a distance of 1.83 pc ($\theta = 0.545$ arcsec), and has the highest measured proper motion, $\mu = 10.4$ arcsec yr^{-1} .

If we measure the celestial coordinates (RA and Dec) of a nearby star at the same time each year (or in practice if we correct for the effects of the Earth’s orbit around the Sun), we find that its position in a reference frame based on very distant objects such as quasars is not the same from year to year (see Figure 2.2). This is *proper motion*, reflecting the fact that the positions of stars within the Galaxy are not fixed. For example, the whole Galactic disk rotates with a circular velocity $v_{\text{rot}} = 220 \text{ km s}^{-1}$ at the Sun’s position. Superposed on this regular rotation pattern is a random velocity of individual stars with dispersion $\sigma_{\text{disk}} \simeq 20 \text{ km s}^{-1}$ in the direction perpendicular to the Galactic plane. Stars in the Milky Way halo have much higher random motions: $\sigma_{\text{halo}} \simeq 100 \text{ km s}^{-1}$; when one of these stars intersects the Galactic plane near the Sun’s location its proper motion can substantial.

Note 1: Proper motions are measured in arcsec yr⁻¹; values for all but the nearest stars are $\mu \ll 1$ arcsec yr⁻¹.

Note 2: What we measure as proper motion is the component of the star's motion *perpendicular* to the line of sight from the Sun to the star. Proper motions are generally quoted separately in RA and Dec.

Note 3: If we know the distance to a star, then we can deduce its transverse velocity (in km s⁻¹) from its proper motion.

2.2.2 Doppler Shift and Space Motion

By now you should be familiar with the concept of redshift, the shift to longer wavelengths (perceived as red by the human eye) of light waves as a result of the relative motions (apart) of emitter and receiver. A few points of note:

- In astronomy, in order to measure redshifts, we need to record the spectra of astronomical sources and measure the wavelength(s) of well-defined spectral feature(s), such as emission or absorption lines. Colours are not sufficient because stars and galaxies can appear red because they are cool, or because their light is reddened by interstellar dust.
- When we are dealing with nearby objects, we measure a Doppler (kinematic) redshift or blueshift:

$$z = \frac{\lambda_{\text{obs}} - \lambda_0}{\lambda_0} \quad (2.5)$$

where λ_{obs} is the observed wavelength of a given spectral line, and λ_0 is the rest-frame wavelength of the same atomic transition, as measured in the laboratory.

- With the Doppler redshift/blueshift is associated a *radial velocity*

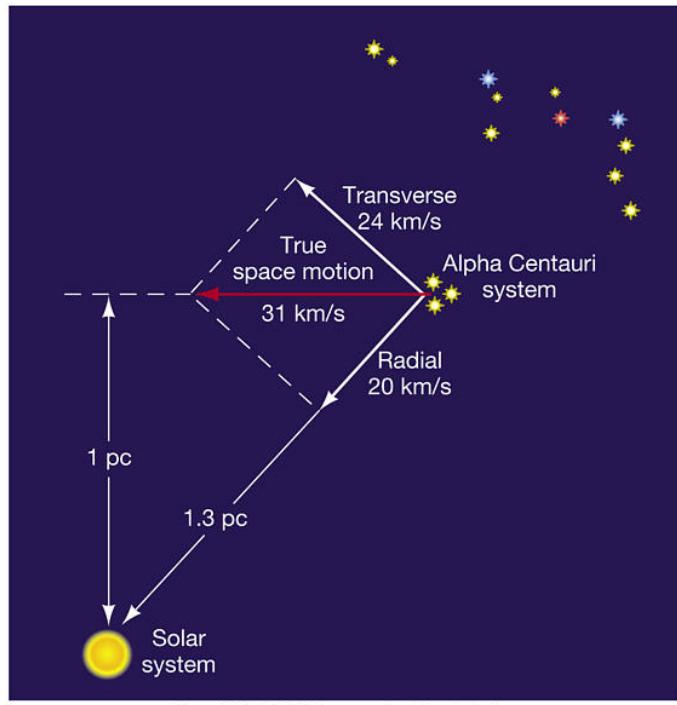
$$v = c \cdot z \quad (2.6)$$

where v is positive for objects moving away from us and negative for objects approaching us.

- Eq. 2.6 is the limiting case (for $v \ll c$) of the special relativity result

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}} \quad (2.7)$$

For stars with measured parallaxes and proper motions, the combination of transverse and radial velocity measurements gives the full space velocity of the stars in 3D (see Figure 2.3).



Copyright © 2005 Pearson Prentice Hall, Inc.

Figure 2.3: By combining proper motion with radial velocity measurements, it is possible to deduce the true space motion of a celestial object.

The velocities and positions of such stars within the Galaxy are fully specified by the six parameters:

l = Galactic longitude (degrees),

b = Galactic latitude (degrees),

d = distance (kpc) from the Galactic Centre,

U = radial velocity relative to the Galactic centre (+ve towards the Galactic centre) ,

V = velocity around the axis of Galactic rotation (+ve in the direction of Galactic rotation), and

W = velocity parallel to the axis of Galactic rotation (+ve in the direction of North Galactic pole).

All six vectors are only available for a relatively small number of stars, because while it is relatively easy to measure celestial coordinates and radial velocities with the necessary accuracy, distances and proper motions are available only for nearby stars. Great improvements expected from the *Gaia* mission which has been designed to measure all six parameters (plus other properties) for one billion stars (still only 1% of the Galactic stellar population!).

2.3 Magnitudes and Luminosities

We have already introduced these concepts in Lecture 1. Magnitudes and luminosities are normally measured through a given filter; Figures 2.4 and 2.5 show the transmission curves of some of the most commonly used optical and near-infrared filters. Note that the near-IR filters are designed to approximately match the transmission windows of the atmosphere at these wavelengths.

It can be appreciated from the Figures that this is hardly a desirable state of affairs. Ideally, we would like the filter transmission curves to be rectangular: 100% transmission over the desired range of wavelengths and 0% everywhere else. Instead, there are ripples and broad tails. There are also pronounced differences between different filter sets, with the potential for error when converting from one set of magnitudes to another.

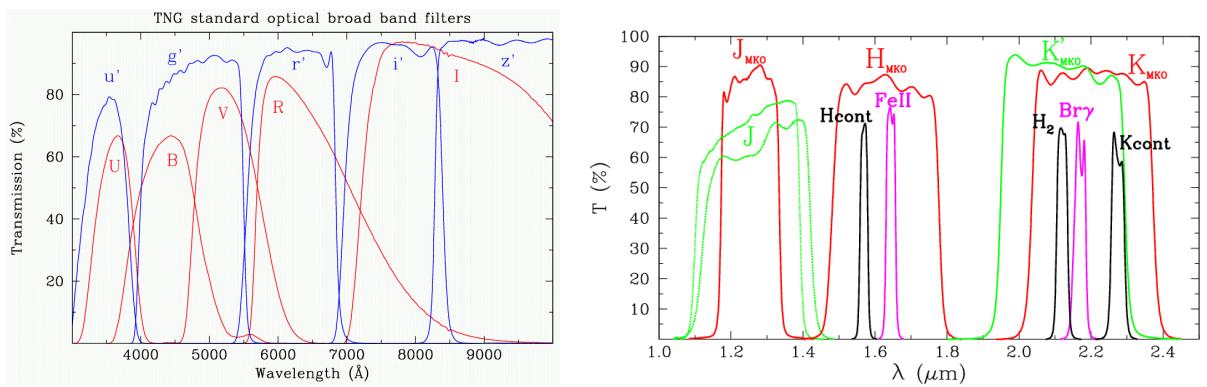


Figure 2.4: Left: Transmission curves of the most commonly used optical broad-band filters. The red curves are for filters in the Johnson-Cousins system, while the blue curves are for the Gunn filters used by the Sloan Digital Sky Survey (SDSS). Right: Transmission curves for commonly used near-IR filters. The narrow-band filters isolate spectral features of particular interest.

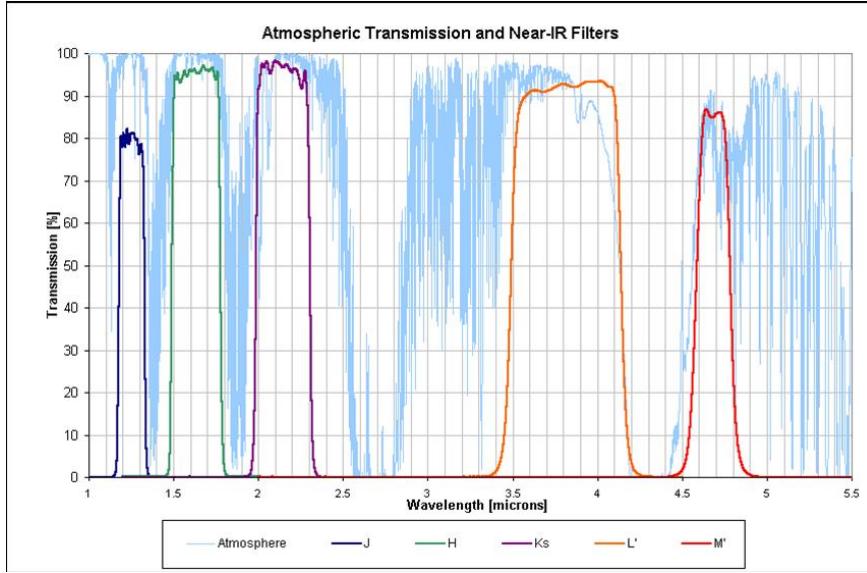


Figure 2.5: Near-IR atmospheric transmission curves and filters.

To circumvent some of these problems, the monochromatic AB magnitude system was defined as:

$$AB = -2.5 \log_{10} f_\nu - 48.60 \quad (2.8)$$

where f_ν is the flux density measured in $\text{ergs s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$. AB magnitudes are used mostly in extragalactic astronomy.

The star α Lyrae provides the zero point of broad-band magnitude systems, since by definition it has $m = 0.00$ in all bands. Again, this is not ideal when one considers that the flux emitted by α Lyrae per unit wavelength (or frequency) interval is far from constant with wavelength (see Figure 2.6).

Some illustrative magnitudes:

Table 2.1 Apparent and absolute magnitudes of selected astronomical sources

Object	Apparent mag m_V	Absolute mag M_V
Sun	-26.7	+4.8
α Canis Majoris (Sirius, brightest star)	-1.4	+1.42
α Lyrae (Vega)	0.0	+0.58
ζ Orionis (in Orion's belt)	+2.0	-5.3
Faintest star visible from Cambridge street	+3.5	
Faintest star visible from dark site, dark adapted	+6.0	
R136 (brightest stellar cluster in LMC)	+9.5	-8.9
Type II Supernova in nearby galaxy	$\sim +14$	-15.3
QSO 3C273 (first known quasar)	+12.9	-26.5
Faintest galaxy in the Hubble Ultra-Deep Field	$\sim +30$	

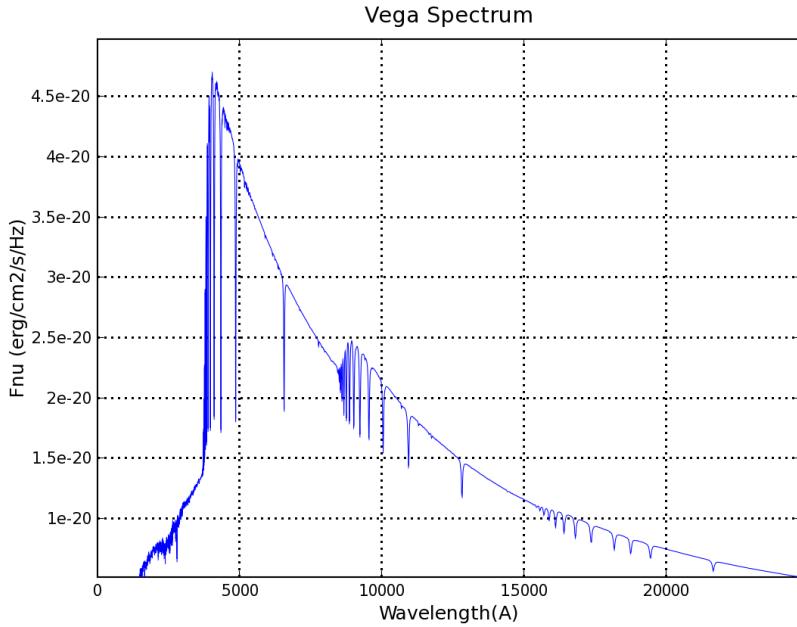


Figure 2.6: The spectrum of α Lyrae, from the near-UV to the near-IR.

Magnitudes can be measured very precisely (to ~ 0.1 millimags from space with the *Kepler* mission), to the extent that it is now possible to infer the existence of Earth-size exoplanets from the shadowing of the stellar disk caused by the planet transiting in front of a solar-type star.

2.4 Effective Temperatures of Stars

The temperature in the interior of stars can reach several 10^8 K. However, of interest here is the *effective temperature*, T_{eff} , of the visible outer layers which radiate the light we see and record with our telescopes. The term *effective temperature* refers to the temperature of a blackbody that most closely approximates the emergent spectrum of a star (see Figure 2.7 left). The continuum spectrum of a blackbody is described by the Planck function (see Figure 2.7 right):

$$B_{\lambda}(T) [\text{erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1} \text{ sr}^{-1}] = \frac{2hc^2/\lambda^5}{e^{hc/\lambda kT} - 1} \quad (2.9)$$

$$B_{\nu}(T) [\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1} \text{ sr}^{-1}] = \frac{2h\nu^3/c^2}{e^{h\nu/kT} - 1} \quad (2.10)$$

Note two important properties of the spectrum described by eq. 2.9:

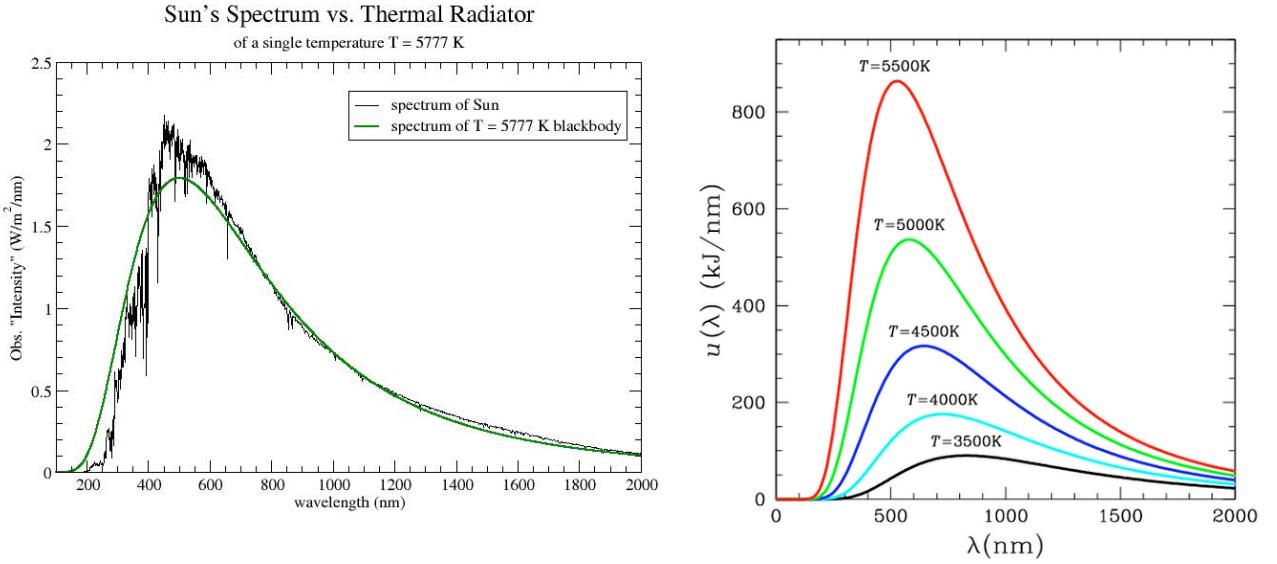


Figure 2.7: Left: The light output from the Sun at visible wavelengths can be approximated by a blackbody spectrum with $T = 5777 \text{ K}$. Right: Blackbody curves for different temperatures.

(1) By considering $d B_\lambda / d \lambda = 0$ to determine the maximum of the function B_λ , we find that

$$\lambda_{\max} T = 0.290 \text{ cm K} \quad (2.11)$$

which is known as Wien's displacement law. For a star like the Sun, with $T_{\text{eff}} = 5770 \text{ K}$, $\lambda_{\max} = 5020 \text{ Å}$.

(2) Figure 2.7 (right) also shows that as its temperature increases, a blackbody emits more energy per second at *all* wavelengths. Integrating eq. 2.9 we find:

$$\int B_\lambda(T) d\lambda = \frac{\sigma}{\pi} T^4 \quad (2.12)$$

where σ is the Stefan-Boltzmann constant ($\sigma = 5.67 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ K}^{-4}$).

To obtain the total luminosity of a blackbody, we multiply by its surface area and integrate over the solid angle:

$$L = 4\pi R^2 \times \int B_\lambda(T) d\lambda \cdot \int d\Omega \quad (2.13)$$

$$L = 4\pi R^2 \times \int B_\lambda(T) d\lambda \cdot \int_0^{2\pi} d\phi \int_0^{\pi/2} \cos \theta \sin \theta d\theta \quad (2.14)$$

$$L = 4\pi R^2 \frac{\sigma}{\pi} T^4 \pi = 4\pi R^2 \sigma T^4 \quad (2.15)$$

Eqs. 2.11 and 2.15 are very powerful. Eq. 2.11 tells us that, if we measure the magnitude of a star through two filters (see Figure 2.4), we can immediately deduce its T_{eff} since, for example:

$$m_B = -2.5 \log F_B + \text{const}$$

$$m_V = -2.5 \log F_V + \text{const}$$

so that:

$$B - V = -2.5 \log(F_B/F_V)$$

where F is the flux through the appropriate filter. Comparison of the measured $B - V$ colour with those of blackbodies of different temperatures then yields a photometric estimate of T_{eff} .

We now have an interpretation for the different colours of stars (e.g. Figure 2.1): *they reflect the fact that stars have different temperatures.* Here are some examples:

Table 2.2 $B - V$ colours and effective temperatures of some nearby stars

Object	$B - V^{\dagger}$ (mag)	T_{eff} (K)
HD 14434	-0.33	47 000
ζ Oph	-0.31	34 000
τ Sco	-0.30	30 000
α Lyr	0.00	9790
51 Aql	+0.30	7300
Sun	+0.63	5777
31 Ori	+1.50	4050
19 Ari	+1.56	3690
α Ori	+1.71	3370
Wolf 359	+2.03	2800

[†] Corrected for interstellar reddening

Astronomers use the term ‘colour’ or ‘colour index’ to refer to the difference in magnitude between any two filters (see Figure 2.4). Conventionally, colours are specified in the sense (shorter wavelength)–(longer wavelength). Thus we have $U - B$, $B - V$, $V - R$, $V - K$, and so on.

Returning to eq. 2.15, we can see that a star’s luminosity depends on both its temperature and size. With a knowledge of a star’s T_{eff} (from its colour)

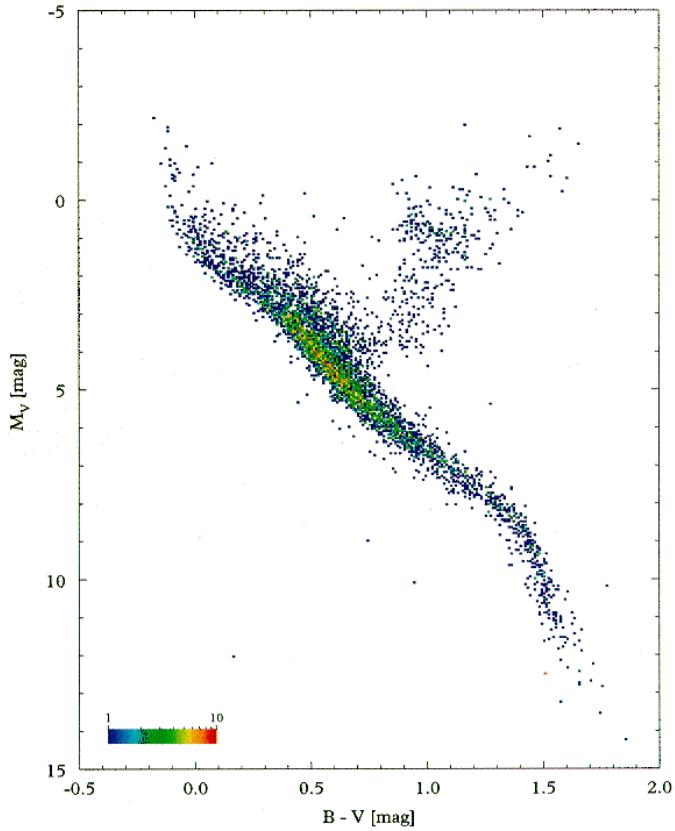


Figure 2.8: H-R diagram for nearby stars with measured parallaxes from the Hipparcos mission.

and luminosity (from its measured magnitude and distance), we can obtain an estimate of its radius.¹

For stars at known distances (or for stars all at the same distance within a cluster), we can construct a diagram plotting their luminosity as a function of colour, as in Figure 2.8. This is undoubtedly the most important diagram in stellar astronomy and we shall explore it in detail during the course of these lectures. It is referred to as the Hertzsprung–Russell diagram (HRD) from the names of the astronomers who first constructed it in the 1910s, or more generally as a colour-magnitude diagram.

We can immediately make the following observations:

- Most of the stars are found in relatively narrow strip on the M_V – $(B - V)$ plane. This strip is the *Main Sequence*.
- Stars of the same $(B - V)$ colour, or equivalently of the same T_{eff} , can

¹The radii of the nearest stars have been measured directly via optical interferometry; these measurements generally agree with the values deduced assuming that the stars radiate as blackbodies.

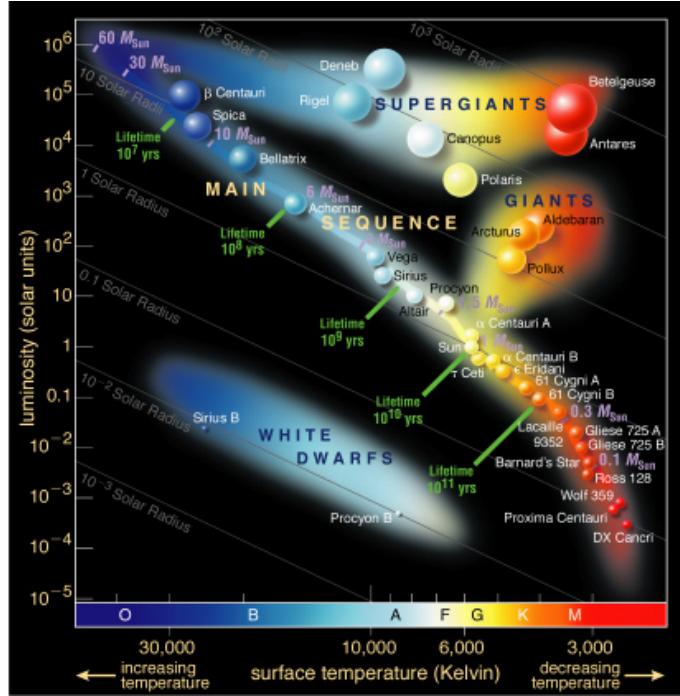


Figure 2.9: Locations of some bright stars on the H-R diagram.

have widely different luminosities. For example, stars with $(B - V) \simeq 1.0$ in Figure 2.8 have values of M_V which span 10 magnitudes, or a factor of 10 000! From eq. 2.15 we understand this to be a consequence of their different sizes; evidently, the radii of such stars span a range of two orders of magnitude. For this reason, stars above the main sequence are termed *giants* and *supergiants*, while stars below the main sequence are *sub-dwarfs*.

- Although the stars collected in Figure 2.8 do not constitute a volume-limited sample (that is, Hipparcos did not measure all the stars within a given distance from the Sun), it is still evident that only a few stars are located away from the Main Sequence. The most straightforward interpretation of this observation is that stars must spend most of their life on the Main Sequence. Thus, the H-R diagram charts stellar evolution: stars move onto the Main Sequence when they are born and evolve off the Main Sequence during the late stages in their evolution.

BASIC PROPERTIES OF STARS. II: SPECTRAL CLASSIFICATION AND THE H-R DIAGRAM

3.1 Introduction

Although the broad spectral distribution of stars approximates that of a blackbody, on closer scrutiny there are deviations (see Figure 2.7). Thus a star colour is not a very precise measure of its temperature. Furthermore, the presence of interstellar dust grains between a star and our telescopes on Earth can extinguish and redden the light (dust grain absorb and scatter photons of shorter wavelengths more readily than those of longer wavelengths). This selective extinction, unless corrected for, can lead us to underestimate the effective temperature of a star.

For all of these reasons, stellar effective temperatures are more accurately determined by measuring discrete spectral features—generally absorption lines—in medium and high resolution spectra of the stars of interest. This is achieved by recording the entire spectrum, normally at optical and infrared wavelengths, of a star with a spectrograph¹ mounted at the focus of a telescope. The move from imaging through just two filters to spectroscopy is time consuming, especially if we are interested in recording the spectra of many stars—within a cluster for example, but the accuracy of the results makes it worthwhile. Indeed, most of what we know about the physical properties of the Universe is derived from the study of spectra of astronomical objects.

We shall return in later lectures to the physical properties of stellar spectra; for the moment we shall limit ourselves to the most important points. The first point to understand is how spectral features are formed, as illustrated in Figure 3.1.

Hot dense objects such as the interiors of stars radiate according to eqs. 2.9 and 2.10. If we were able to view the light from such sources directly,

¹An instrument which includes light-dispersing elements such as prisms or ruled gratings.

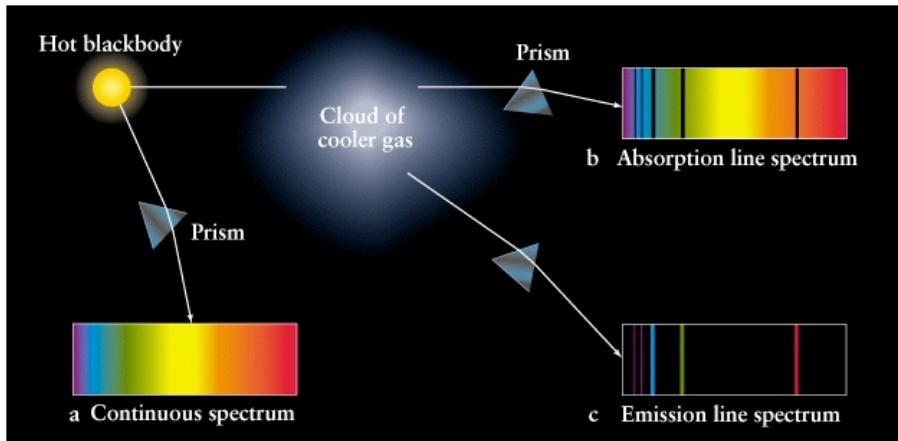


Figure 3.1: Continuum, absorption, and emission spectra of astronomical sources.

without any intervening matter, then the resultant spectrum would appear as a continuum.

Most stars are surrounded by outer layers of gas that are less dense than the core. The photons emitted from the core cover all frequencies (and energies). Photons of specific frequency can be absorbed by electrons in the diffuse outer layer of gas, causing the electron to change energy levels. Eventually the electron will de-excite and jump down to a lower energy level, emitting a new photon of specific frequency. However, the direction of this re-emission is random, so the chance of the re-emitted photon travelling along the same path as the original incident photon is very small. The net effect is that the intensity of light at the wavelength of that photon will be less in the direction of an observer and the resultant spectrum will show dark *absorption* lines. Stellar spectra typically look like this. A digital version of the spectrum, which is just a plot of relative intensity as a function of wavelength, will show reduced intensities—or absorption lines—at specific wavelengths (see, for example, Figure 2.6).

A third possibility occurs if an observer is not looking directly at a hot black body source but instead at a diffuse cloud of gas. If this cloud can be excited by a nearby source of energy, such as hot, young stars or an active galactic nucleus, then the electrons in atoms of the gas cloud can get excited. When they de-excite they emit photons of specific frequency and wavelength. As these photons can re emitted in any direction an external observer will detect light at these wavelengths. The spectrum formed is an *emission* line spectrum. A typical example is the spectrum produced

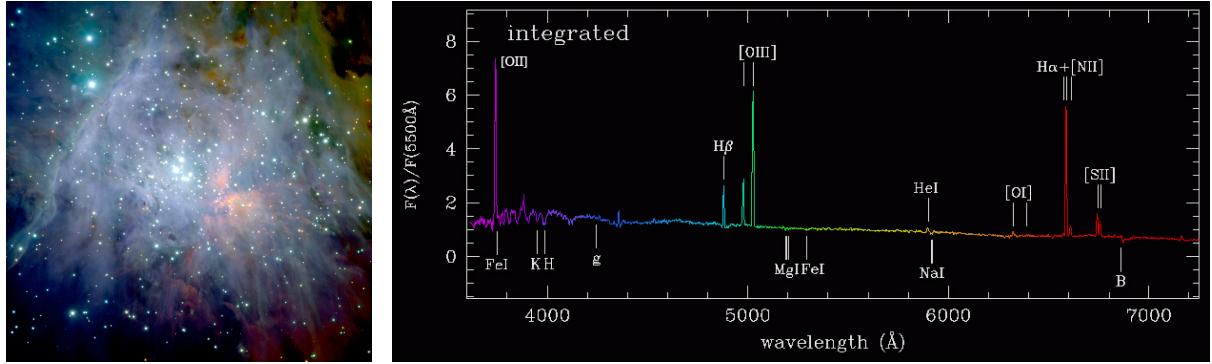


Figure 3.2: Left: Near-IR image of the central part of the Orion nebula, obtained with the ESO VLT. The nebula is diffuse gas ionised by hot stars: the famous Trapezium stars can be seen near the centre together with the associated cluster of about 1,000 stars, approximately one million years old. Right: Typical emission line spectrum of an H II region such as the Orion Nebula. The most important spectral features are labelled.

by the Orion Nebula, our nearest ionised hydrogen, or H II, region (see Figure 3.2).

3.2 Collisional excitation and ionisation

Thermal excitation of different atomic levels and different ionisation stages of the same element becomes significant when the thermal energy kT is comparable to, respectively, the energy difference between different atomic levels and the ionisation potential (binding energy of the electron) of a given ion. For example, in H, with an ionisation potential $\xi_i = 13.6 \text{ eV}$ (see Figure 3.3), this occurs at $T \simeq 10^5 \text{ K}$. At lower temperatures, ionisation or excitation will be incomplete, and the gas will contain a mixture of states.

Referring to Figure 3.3, the number of atoms in an atomic level n with energy E_n is given by the Boltzmann's equation:

$$N_n = A e^{-E_n/kT} g_n, \quad (3.1)$$

where k is Boltzmann's constant ($k = 8.62 \times 10^{-5} \text{ eV deg}^{-1}$), A is a constant of proportionality and g_n is the statistical weight of atomic level n denoting the number of particles which can be in atomic state n . If J_n is the angular momentum of a state, its electronic statistical weight is $g_n = 2J_n + 1$. The relative populations of two levels of the same atom are therefore:

$$\frac{N_n}{N_m} = \frac{g_n}{g_m} e^{-(E_n - E_m)/kT}, \quad (3.2)$$

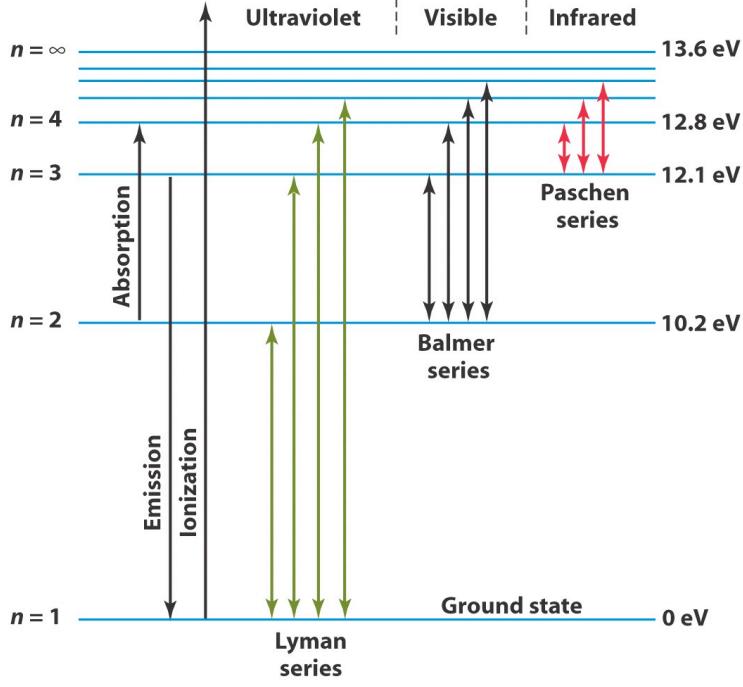


Figure 3.3: Energy levels of the neutral hydrogen atom.

and the total number of atoms in all levels is:

$$N = \sum_{n=1}^{\infty} N_n = A \sum_{n=1}^{\infty} g_n e^{-E_n/kT} = A Z(T), \quad (3.3)$$

where $Z(T) = \sum_n g_n e^{-E_n/kT}$ is the *partition function*.

Eqs. 3.1 and 3.2 have immediate applications to stellar spectra. Consider for example the H α absorption line which occurs when an electron absorbs a photon with energy $E = h\nu = E_3 - E_2 \simeq 12.09 - 10.20 = 1.89$ eV, corresponding to a wavelength $\lambda = 6563\text{ \AA}$ (Figure 3.3). This transition will be strong in the spectrum of a star only if there is an appreciable number of H atoms in the $n = 2$ level capable of absorbing an H α photon. In other words, we must not have $kT \ll E_2$. With $E_2 = 10.20$ eV, we have $T \simeq 1.2 \times 10^5$ K. This temperature is much higher than $T_{\text{eff}} = 5770$ K of the Sun, which explains why the H α absorption line is not very strong in the Sun. Its strength in stars increases in stars hotter than the Sun, and peaks in stars with $T_{\text{eff}} \simeq 10\,000$ K (such as α Lyr; see Table 2.2).

In stars with $T_{\text{eff}} > 10\,000$ K, the strength of H α decreases again, because of the increasing fraction of H that is ionised, due to collisional ionisation. The relative proportions of ions in two successive stages of ionisation is

given by the Saha equation:

$$\frac{n_e \cdot N_{i+1}}{N_i} = 2 \frac{Z_{i+1}}{Z_i} \frac{(2\pi m_e kT)^{3/2}}{h^3} e^{-\chi_i/kT} \quad (3.4)$$

which can be derived from consideration of the statistical properties of electrons in the continuum of energy levels. Here, Z_i is the partition function of the i th ionisation stage, χ_i is the ionisation potential of the i th ionisation stage (the energy required to ionise ion i to ion $i + 1$), and the other symbols have their usual meaning. Another form of the Saha equation is obtained by multiplying both sides by kT and using the definition of electron pressure $P_e = n_e kT$ to get:

$$\frac{N_{i+1}}{N_i} = 2 \frac{Z_{i+1}}{Z_i} \frac{(2\pi m_e)^{3/2}}{h^3} \frac{(kT)^{5/3}}{P_e} e^{-\chi_i/kT}. \quad (3.5)$$

Trivially, the Saha equation tells us that atoms with low ionisation potentials will be relatively more ionised at any given temperature than atoms of higher ionisation potentials. For the common chemical elements, He and Ne are the hardest to ionise ($\chi_0 > 20$ eV), H, C, N, O are the next (20 eV $> \chi_0 > 10$ eV), and the easiest are Li, Na, Mg, Al, Ca with $\chi_0 \sim 5$ eV. Because χ_0 appears in the exponential in the Saha equation, this difference has a profound effect in the relative abundances of different ions. For example, $N(\text{Na}^+) \propto e^{-5.16 \text{ eV}/kT}$, while $N(\text{H}^+) \propto e^{-13.6 \text{ eV}/kT}$. Thus, in the Sun ($T = 5770$ K), $N(\text{Na}^+)/N(\text{Na}) \sim 10^7 N(\text{H}^+)/N(\text{H})$.

This difference more than makes up the low ratio $N(\text{Na})/N(\text{H}) \sim 10^{-6}$ in the Sun. Thus, in the solar atmosphere, the free electrons present are predominantly from atoms of low ionisation potential, such as Na, Ca, Al, and so on. These free electrons, from the metals, have an impact on the degree of ionisation of H through the n_e term in the Saha equation. The corollary of this statement is that variations in metal abundances from one star to another may have a major effect on the physical state of the gas, especially its ionisation equilibrium.

Returning to our main point, it is now easy to see that by considering the relative strengths of absorption lines of elements in different ionisation and excitation states, it is possible to deduce the effective temperature of the stellar atmosphere. Hotter stars will have absorption lines from highly ionised stages, while the coolest stars can even maintain significant densities of molecules in their atmospheres—see Figure 3.4.

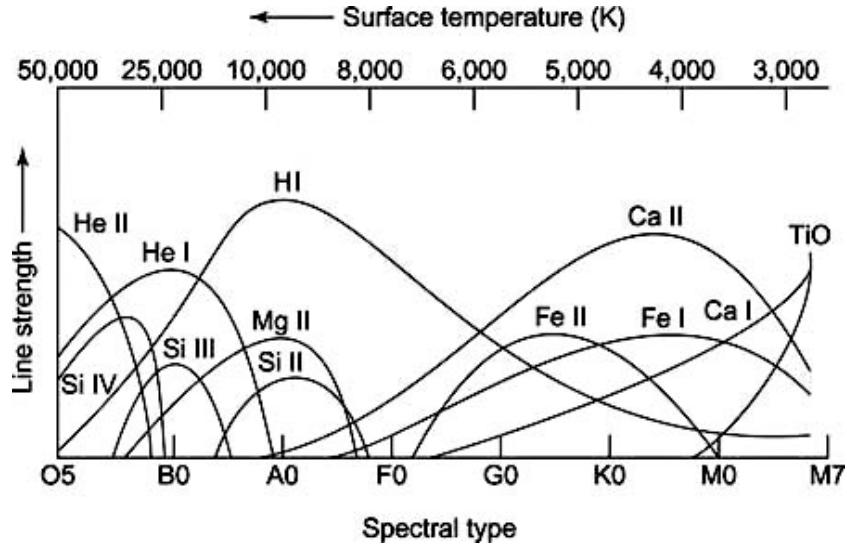


Figure 3.4: Dominant ions in stars of different effective temperatures.

3.3 Stellar Spectral Types

The classification of stellar spectra originates from a time (late 1800s–early 1900s) when their physical meaning was not understood. The Harvard classification scheme is a one-dimensional scheme which divides stars into a sequence of seven classes, O, B, A, F, G, K, M, based on their effective temperature, as follows:

Class	T_{eff} (K)
O	$\geq 33\,000$
B	10 000 – 33 000
A	7500 – 10 000
F	6000 – 7500
G	5200 – 6000
K	3700 – 5200
M	≤ 3700

Typical spectra for stars of different classes are shown in Figure 3.5. Each class is subdivided into 10 sub-classes, with T_{eff} decreasing from sub-class 0 to 9. Thus we have spectral classes O9 and B0, the former being just hotter than the latter. Division between subclasses can be finer, e.g. class O9.5. O and B stars are sometimes referred to as ‘early-type’, while K and M are ‘late-type’.

The Sun has spectral class G2. More recently the classification has been

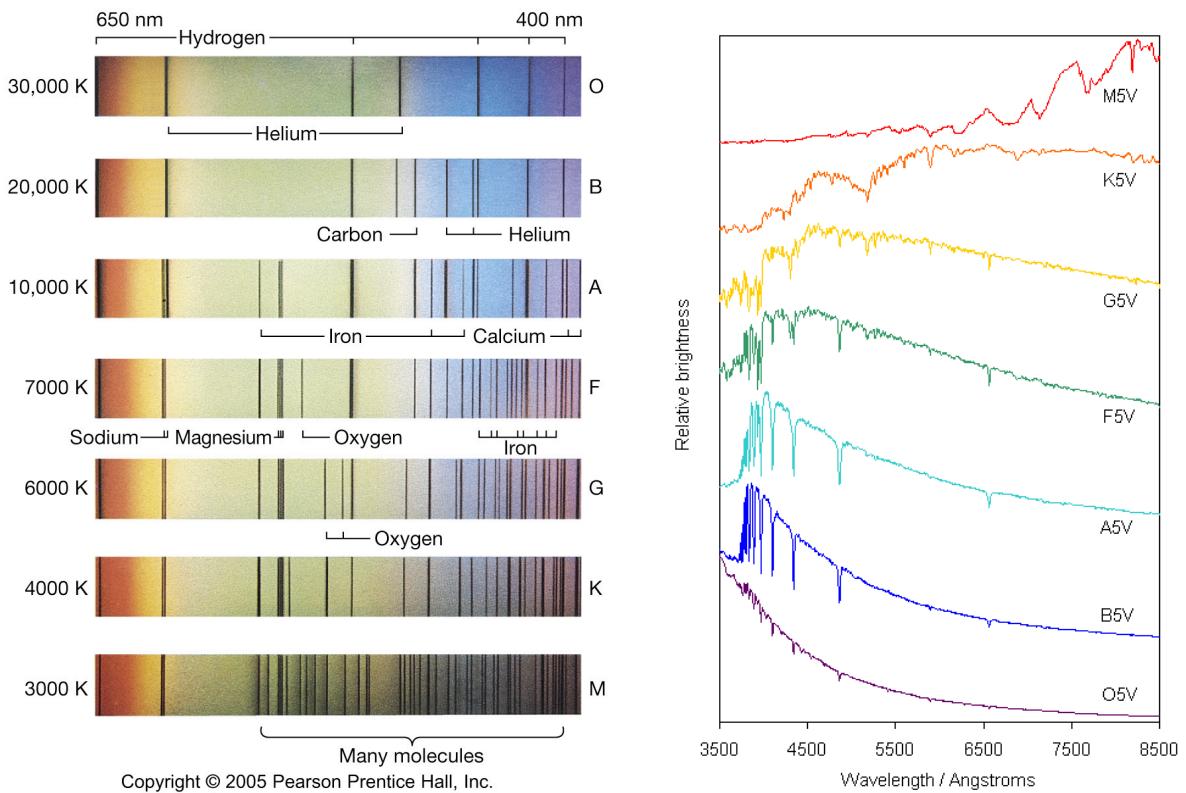


Figure 3.5: Typical spectra of different stellar spectral classes .

extended to cooler stars, so cool that they are brightest at infrared wavelengths, with classes L and T (mostly brown dwarfs which do not achieve sufficiently high temperatures in their cores to ignite thermonuclear fusion).

3.4 Luminosity Classes

As we shall study in more detail later in the course, the spectra of stars give a great deal more information on their physical properties than just a measure of T_{eff} . For our present purposes all we have to appreciate is that the *profile*, as well as the overall strength, of an absorption line gives information about conditions in the gas that produced it. In particular, the width of an absorption line depends on the density of the stellar atmosphere where it was formed. If we consider stars of the same T_{eff} but different radii, the surface density and gravity are lower in stars of larger radii. As a consequence, more luminous stars (at a given T_{eff} , recall eq. 2.15) have narrower absorption lines. Thus, the absorption line widths provide a measure of the ‘Luminosity Class’.

Luminosity classes are indicated by roman numerals, from I to V in order

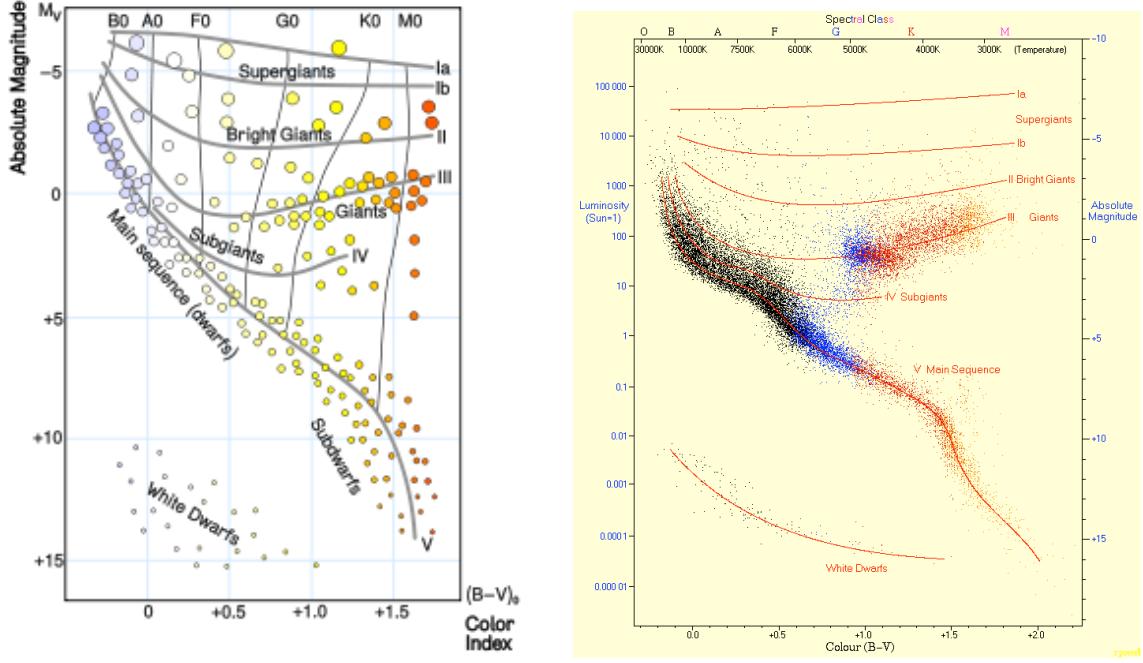


Figure 3.6: Left: Schematic representation of the MK stellar classification scheme. Right: Luminosity classes in the Hipparcos H-R diagram for nearby stars.

of *decreasing* luminosity. Class I is further sub-divided into classes Ia and Ib—these are the ‘*supergiants*’. Stars of luminosity class III are referred to as ‘*giants*’. Stars on the main sequence are given luminosity class V and, confusingly, are sometimes referred to as ‘*dwarfs*’. The luminosity class, together with the spectral class, results in the two-dimensional M-K stellar classification scheme (see Figure 3.6). Table 3.1 now includes the M-K spectral type for the same nearby stars considered in Table 2.2.

Table 3.1 Spectral types, $B - V$ colours and effective temperatures of some nearby stars

Object	Spectral Type	$B - V^\dagger$ (mag)	T_{eff} (K)
HD 14434	O5 V	-0.33	47 000
ζ Oph	O9 V	-0.31	34 000
τ Sco	B0 V	-0.30	30 000
α Lyr	A0 V	0.00	9790
51 Aql	F0 V	+0.30	7300
Sun	G2 V	+0.63	5777
31 Ori	K5 III	+1.50	4050
19 Ari	M0 III	+1.56	3690
α Ori	M2 I	+1.71	3370
Wolf 359	M6 V	+2.03	2800

[†] Corrected for interstellar reddening

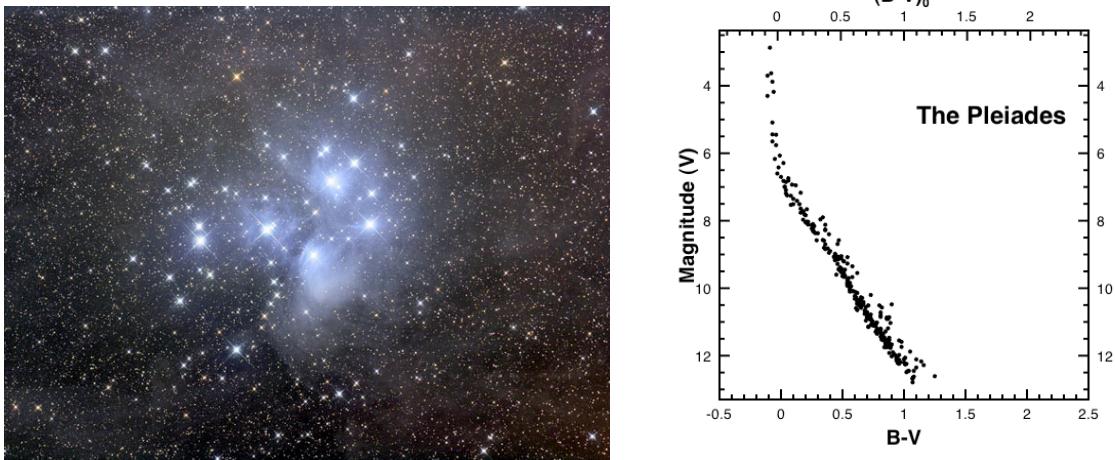


Figure 3.7: The Pleiades: a young star cluster and its colour-magnitude diagram.

3.5 Cluster HR Diagrams

Constructing colour-magnitude diagrams for stellar clusters where all the stars are coeval and have the same chemical composition (since they were presumably born out of the same interstellar cloud) allows us to consider evolutionary effects on the HR diagram.

The Pleiades (Figure 3.7) are a famous stellar cluster visible from northern skies in the winter. At a distance of only 120 pc, it is one of the closest open star cluster. It consists of over 1000 stars, of which the brightest are mid- to late-B type stars. Its colour-magnitude diagram (right panel of Figure 3.7) shows a well populated main sequence up to stars with $(B - V) \simeq 0.0$, corresponding to $T_{\text{eff}} \simeq 10\,000$ K. The earliest main sequence star in the Pleiades is of spectral type B8V (18 Tau). The absence of hotter stars in the colour-magnitude diagram indicates an age of $\sim 10^8$ years; hotter and more massive stars evolve off the main sequence in less than 10^8 years.

Contrast Figure 3.7 with Figure 3.8, an old globular cluster in the halo of the Milky Way. At a distance of 10 kpc from the Sun, M3 contains approximately half a million stars. Its age is estimated to be $\sim 8 \times 10^9$ years from the fact that only relatively cool stars, with $(B - V) \simeq +0.4$ (late F) are found on the main sequence.

The right panel of Figure 3.8 illustrates the nomenclature applied to different regions of the HR diagrams, in accordance with our understanding of the stages they represent in the evolution of a star. Anticipating the material to be considered in future lectures, I briefly describe them here.

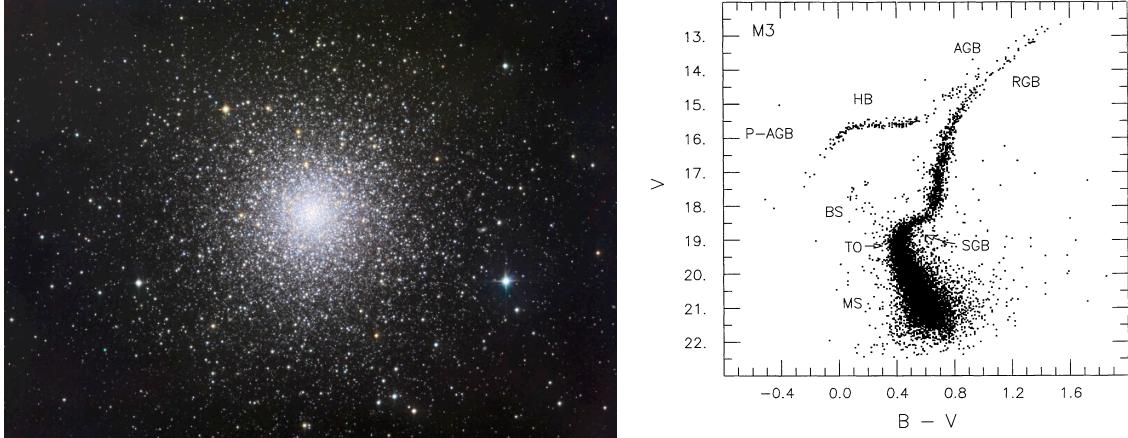


Figure 3.8: M3: an old halo globular cluster and its colour-magnitude diagram.

In the case of a halo globular cluster, with little foreground contamination, the density of points in the HR diagram approximately reflects the length of time a star spends in different evolutionary stages. Stars spend most of their lives on the main sequence (MS), burning hydrogen in their cores. Then turn-off point (TO) marks the end of the MS phase. The lifetime of the most massive star still on the MS then defines the age of the cluster. The TO point is the start of the subgiant branch (SBG) with shell H burning; the star increases in brightness and evolves to the red giant branch (RGB). Here the star loses mass and eventually becomes a horizontal branch (HB) star. The final evolution of the star may lead up to the asymptotic giant branch (AGB) and into the very hot post-asymptotic giant branch (P-AGB). Blue stragglers (BS) are thought to be stars that have been severely affected by close encounters in the dense cluster region or by mass transfer in binary systems.

Thus, a star of a given mass travels through a well specified path, or track, in the colour-magnitude diagram during its lifetime, from its birth to its death, be it a supernova or a cooling white dwarf.

On the other hand, an isochrone is a line on the CMD which defines all stars of a given age. Fitting theoretical isochrones, computed according to stellar evolution theory, to accurate photometry of globular cluster stars gives us estimates of the cluster ages (see Figure 3.9). For a long time, there was a conflict between the ages of the oldest Galactic globular clusters so deduced and estimates of the age of the Universe based on the favoured values of the Hubble constant and of the matter density $\Omega_{M,0}$. This conflict has now been largely resolved: the age of the Universe in today's consensus

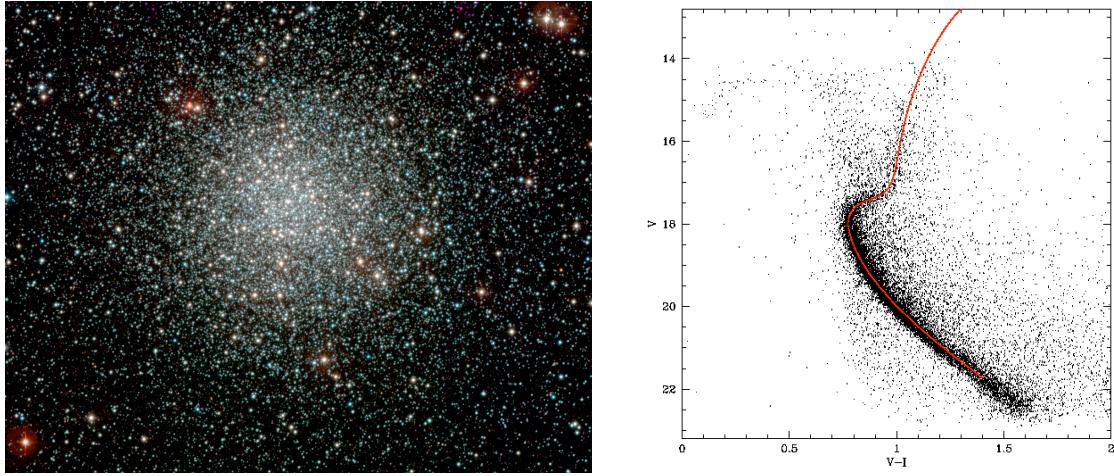


Figure 3.9: NGC 3201, a low Galactic latitude globular cluster at a distance of 5.2 kpc from the Sun. This globular cluster has a relatively low concentration of stars near its center. By fitting its reddening-corrected colour-magnitude diagram with a theoretical isochrone appropriate to its low metallicity ($[{\rm Fe/H}] = -1.54$), a best fitting age $t = 12$ Gyr is deduced (von Braun & Mateo 2001, AJ, 121, 1522).

cosmology, $t = 13.7$ Gyr, is ‘comfortably’ greater than the age $t = 12.4$ Gyr of the oldest globular clusters of the Milky Way.

Isochrone fitting to colour-magnitude diagrams is a powerful way to unravel the past history of star formation in a galaxy, as it can reveal distinct episodes of star formation, particularly if they were separated by quiescent periods. Figure 3.10 shows an example for a rich star cluster in our neighbouring galaxy, the Large Magellanic Cloud.

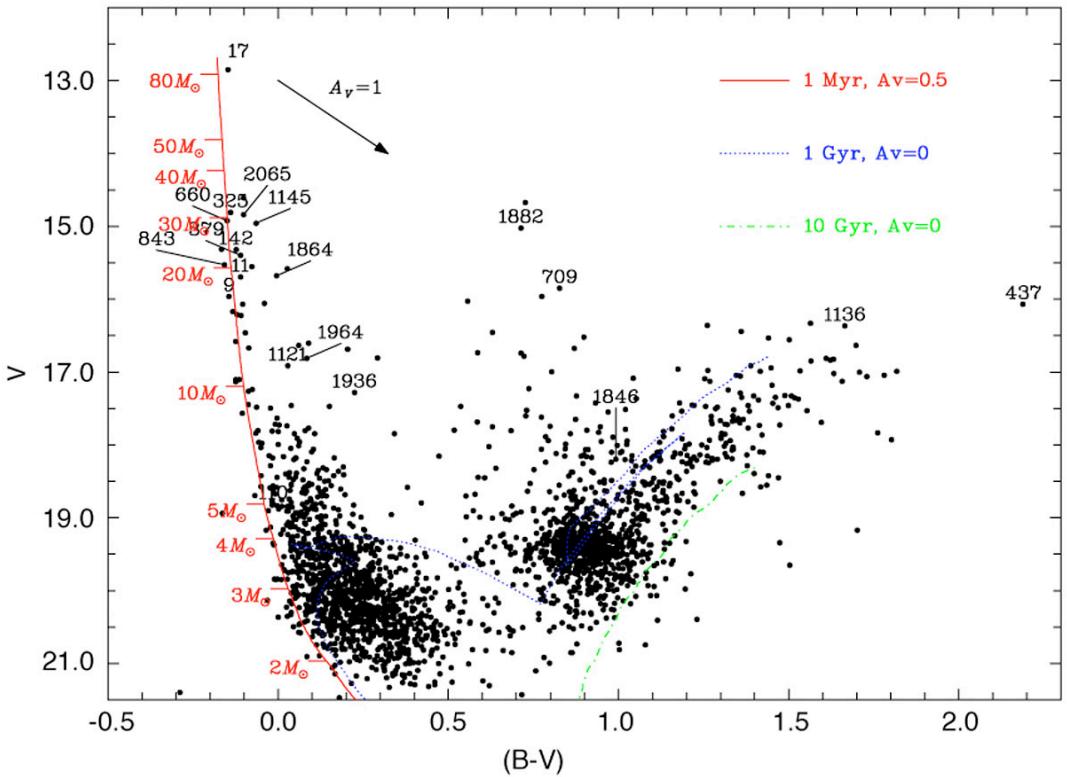


Figure 3.10: Colour-magnitude diagram of 2341 stars observed in the cluster N214C of the Large Magellanic Cloud. Three theoretical isochrones are shown, representing the positions of stars with ages of 1×10^6 years (red curve), 1×10^9 years (dotted blue), and 1×10^{10} years (dashed-dotted green), computed for the LMC metallicity and distance. It is clear from this diagram that N214C is composed of two populations: a very young one, containing very massive stars, and an older one. Star numbered 17 is of spectral type O2 V, one of only a handful of stars known with such an early spectral type. (Figure reproduced from Meynadier et al. 2005, A&A, 436, 117).

BASIC PROPERTIES OF STARS. III: MASS

4.1 Introduction

It was stated earlier that the most important property of a star is its mass. Stellar masses can be determined directly by studying their gravitational interaction with other objects. As at least half of all (nearby) stars are thought to be in multiple systems, there are many opportunities to monitor the motions of binary stars to deduce their masses. Considering for simplicity only binary stars (as opposed to triple and quadruple systems), we distinguish three main classes: visual binaries, eclipsing binaries and spectroscopic binaries. We now consider them in turn.

4.2 Visual Binaries

Visual binaries tend to be systems that are relatively close to us so that the individual stars can be resolved. They are systems in which the component stars are also physically widely separated, tens to a few hundred AUs. The stars in such systems are gravitationally bound to each other but otherwise do not ‘interact’ as do other close binaries where one star may draw material off the surface of the other. The brightest component in the system has the suffix ”A”, the next ”B” and so on. Systems with three or four components have been identified. Less than 1,000 visual binary systems have been detected. Two out of the three brightest stars in the sky, α CMa and α Cen, are binaries.

Referring to Figure 4.1, with the two stars orbiting about the common centre of mass, we have straightforwardly:

$$\frac{m_1}{m_2} = \frac{r_2}{r_1} = \frac{a_2}{a_1} \quad (4.1)$$

where m is the stellar mass, r is the star’s distance from the centre of mass and a is the semi-major axis of the elliptical orbit. By monitoring over the years (provided the period is not too long compared to human

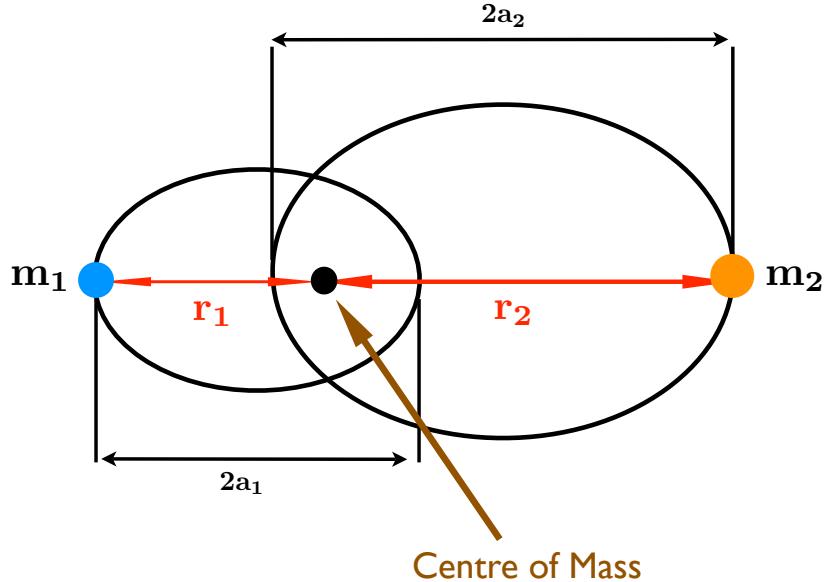


Figure 4.1: Schematic of a binary star system viewed face-on. In this example, $m_1 = 2m_2$.

timescales!) the relative positions on the sky of the two stars, it is possible to determine the orientation of the orbits and the system's centre of mass (see Figure 4.2). The distances r_1 , r_2 from the common centre of mass subtend angles $\theta_1 = r_1/d$ and $\theta_2 = r_2/d$ at the star's distance d . It is therefore possible to deduce the mass ratio immediately from observations of the orbits:

$$\frac{m_1}{m_2} = \frac{\theta_2}{\theta_1} \quad (4.2)$$

If we know d , for example from parallax measurements, we can deduce the

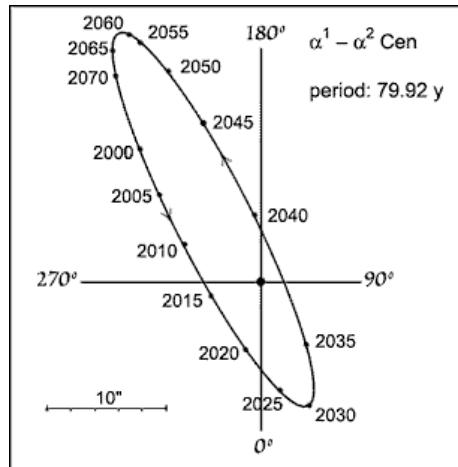


Figure 4.2: Sky projection of the orbit of α Cen B relative to α Cen A. The predicted positions of B relative to A for the current orbit are shown by year.

individual masses using Kepler's third law:

$$P^2 = \frac{4\pi^2}{G(m_1 + m_2)} a^3 \quad (4.3)$$

where P is the period (the same for both orbits) and $a = a_1 + a_2$ is the semimajor axis of the orbit of the reduced mass μ ,

$$\mu = \frac{m_1 \cdot m_2}{m_1 + m_2}.$$

Recall that, in general, a two-body problem may be treated as an equivalent one-body problem with the reduced mass μ moving about a fixed mass $M = m_1 + m_2$ at a distance $r = |\mathbf{r}_2 - \mathbf{r}_1|$.

In order to deduce m_1 and m_2 from observations of θ_1 , θ_2 and P it is necessary to correct for: (i) the parallax of the whole system, (ii) the proper motion of the centre of mass, and (iii) the inclination of the plane of the orbit relative to the plane of the sky. (i) is easy: just observe a binary system for more than one year cycle. (ii) is also relatively simple, since the centre of mass must move at constant velocity. (iii) is trickier.

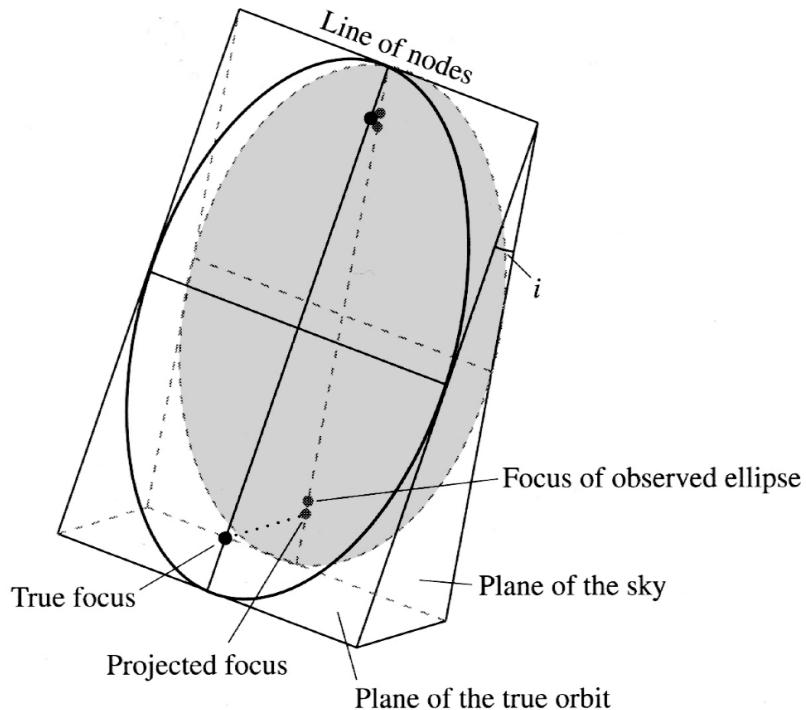


Figure 4.3: The projection of an elliptical orbit inclined by the angle i to the plane of the sky is also an elliptical orbit. However, the real foci of the ellipse do not project to the foci of the observed ellipses. (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

Consider the special case where the orbital plane is inclined at angle i to the plane of the sky (that is, it is inclined by an angle $90^\circ - i$ to the line of sight) and the two planes intersect along a line parallel to the minor axis of the stellar orbit, forming a *line of nodes*, as in Figure 4.3. What we observe in this case are angles $\theta'_1 = \theta_1 \cos i$ and $\theta'_2 = \theta_2 \cos i$. The unknown inclination doesn't affect the estimate of the mass ratio, since the $\cos i$ factors cancel out in eq. 4.2. However, they can make a significant difference in the estimate of a in eq. 4.3, which now becomes (solving for the sum of the masses):

$$m_1 + m_2 = \frac{4\pi^2}{G} \frac{(\theta d)^3}{P^2} = \frac{4\pi^2}{G} \left(\frac{d}{\cos i} \right)^3 \frac{\theta'^3}{P^2} \quad (4.4)$$

where θ is in radians and $\theta' = \theta'_1 + \theta'_2$.

Thus, in order to evaluate the sum of the masses properly, we need to know the angle of inclination i . This can be deduced by careful observation of the centre of mass which, as shown in Figure 4.3, will not coincide with the focus of the projected ellipse. The geometry of the true ellipse may be determined by comparing the observed stellar positions with mathematical projections of various ellipses onto the plane of the sky. The real situation is of course more complicated because in general the orbital plane may be inclined about both the minor and major axes.

In cases where the distance to a visual binary is not known, it may still be possible to deduce a_1 and a_2 and solve for m_1 and m_2 using radial velocity measurements, which give the projections of the velocity vectors along the line of sight.

Several hundred visual pairs are known, but in most cases it has not yet been determined whether they are bound binary systems or chance superpositions. Many visual binaries have long orbital periods of several centuries or millennia and therefore have orbits which are uncertain or poorly known. For this reason, they only sample rather sparsely the HR diagram, with a strong bias towards the more common (and therefore more likely to occur in the solar vicinity) low mass stars. Fortunately, other types of binary stars help us expand the range of reliable stellar mass determinations.

4.3 Spectroscopic Binaries

When two stars in a binary system are too far away to be resolved even with the largest telescopes on Earth, the binarity of the system can still be inferred from consideration of the spectrum, which will be the superposition of two sets of spectral features (which may be different if the stars are of different spectral types). In *double-lined spectroscopic binaries*, the absorption lines in the composite spectrum will be seen to move in wavelength, as each star moves in its orbit towards us and away from us (see Figure 4.4). The maximum blueshift and redshift we measure within an orbit are lower limits to the true velocities because of the unknown inclination i of the orbital plane to the line of sight: $v_1 r_{\max} = v_1 \sin i$ and $v_2 r_{\max} = v_2 \sin i$.

Many spectroscopic binaries have nearly circular orbits because the timescales of tidal interactions which tend to circularise the orbits are short compared to the stellar lifetimes. When the eccentricities are small ($\epsilon \ll 1$), the orbital speed is essentially constant: $v = 2\pi a/P$ and where P is the period and the semi-major axis a is now the radius. Substituting into eq. 4.1, we

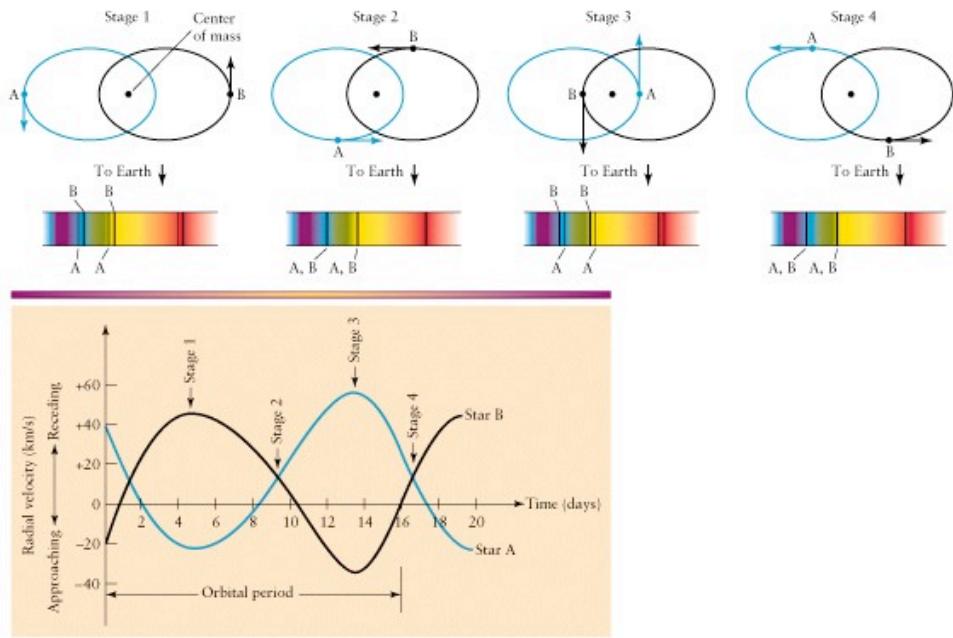


Figure 4.4: Schematic diagram of a double-lined spectroscopic binary, showing the orbits and the resultant composite spectrum produced at different orbital phases. Note that the centre of mass of the system has a radial velocity $v_r \simeq +15 \text{ km s}^{-1}$.

now have:

$$\frac{m_1}{m_2} = \frac{v_2}{v_1} \quad (4.5)$$

or, in terms of the observables:

$$\frac{m_1}{m_2} = \frac{v_{2r}/\sin i}{v_{1r}/\sin i} = \frac{v_{2r}}{v_{1r}}. \quad (4.6)$$

Thus, as in the case of visual binaries, the mass ratio can be deduced independently of the unknown inclination angle i .

However, the sum of the masses does require knowledge of $\sin i$. Replacing a with:

$$a = a_1 + a_2 = \frac{P}{2\pi} (v_1 + v_2), \quad (4.7)$$

substituting into eq. 4.3 and solving for the sum of the masses, we obtain:

$$m_1 + m_2 = \frac{P}{2\pi G} (v_1 + v_2)^3, \quad (4.8)$$

or, in terms of the observables:

$$m_1 + m_2 = \frac{P}{2\pi G} \frac{(v_{1r} + v_{2r})^3}{\sin^3 i}. \quad (4.9)$$

Since the inclination angle is generally unknown, eq. 4.9 is usually solved statistically. That is, we assume that the orbits are randomly inclined relative to our line of sight and use the integral average of $\sin^3 i$ between 0 and 90° , $(\langle \sin^3 i \rangle) = 3\pi/16 \simeq 0.589$, to deduce the average mass of stars in a given luminosity or T_{eff} class. A selection effect correction is usually applied to account for the fact that when the orbits are nearly face-on (i less than a few degrees, $\sin i \ll 0.1$), it is much more difficult to recognise that a star is a spectroscopic binary. Thus the larger value $(\langle \sin^3 i \rangle) \simeq 2/3$ is usually taken to be representative in this statistical approach.

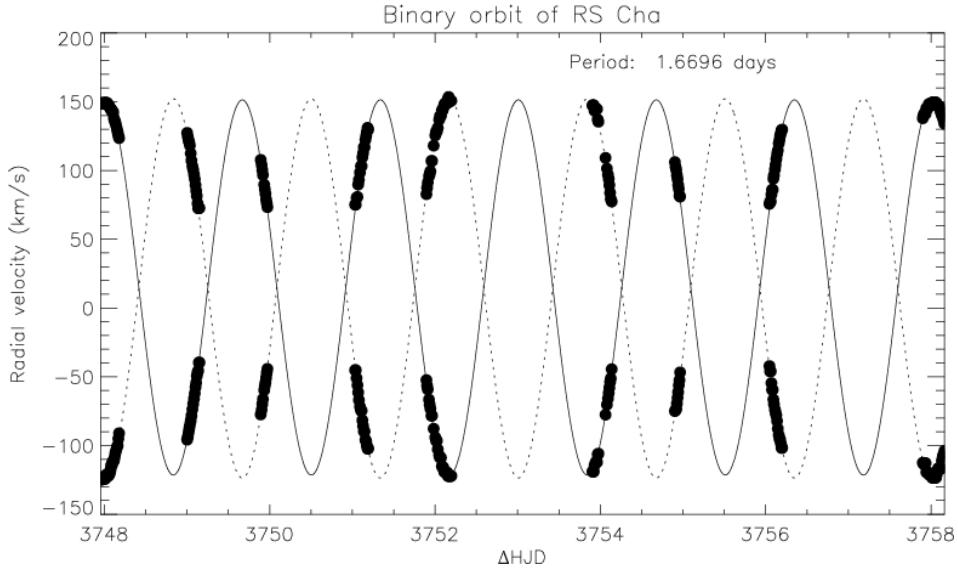


Figure 4.5: RS Cha belongs to an enigmatic group of pre-main sequence (PMS) stars of intermediate mass, between 2 and $8 M_{\odot}$, referred to as Herbig Ae/Be stars from the American astronomer who first identified them in 1960. At a parallactic distance of 93 ± 2 pc, RS Cha is a bright spectroscopic eclipsing binary star, with both components being Herbig Ae PMS stars of similar mass ($\sim 1.9 M_{\odot}$). Their age is 6^{+2}_{-1} Myr; other physical parameters are collected in Table 4.1. This Figure, reproduced from Böhm et al. (2009), shows the radial velocities of the two stars measured with near-continuous observations during 14 nights at the 1 m Mt John (New Zealand) telescope using a high resolution spectrograph. Note the sinusoidal velocity curves, indicative of nearly circular orbits. The orbital plane lies close to our line of sight ($i = 83.4$); thus, the measured values of v_{1r} and v_{2r} are very close to the real velocities v_1 and v_2 ($\sin i = 0.99$).

Table 4.1. Parameters of RS Cha. References: [1] Alecian et al., 2005, [2] Ribas et al., 2000, [3] Clausen & Nordstrom, 1980.

Parameter	Primary	Secondary	References
M/M_{\odot}	1.89 ± 0.01	1.87 ± 0.01	[1]
R/R_{\odot}	2.15 ± 0.06	1.87 ± 0.01	[1]
T_{eff} [K]	7638 ± 76	7228 ± 72	[2]
$\log(L/L_{\odot})$	1.15 ± 0.09	1.13 ± 0.09	$L = 4\pi R^2 \sigma T_{\text{eff}}^4$
$\log(g)$ [cm s^{-2}]	4.05 ± 0.06	3.96 ± 0.06	$g = MG/R^2$
$v \sin i$ [km s^{-1}]	64 ± 6	70 ± 6	[1]
P_{orb} [d]		1.67	[1]
i [deg]		83.4 ± 0.3	[3]
$[\text{Fe}/\text{H}]$		0.17 ± 0.01	[1]

Of much interest in astronomy are *single-lined spectroscopic binaries*. These are cases where only the spectrum of one of the pair is observed, but the periodic variations in its radial velocity indicate the presence of an unseen companion. This could be the case if: (a) the second star is very much fainter than the first—Sirius A and B are a good example; (b) the companion is a dark object, such as a neutron star or a black hole—such systems provide some of the most compelling evidence for the existence of stellar-mass black holes; and (c) if the secondary is a planet. In this case, the radial velocity amplitudes are only m s^{-1} , rather than km s^{-1} .

In single-lined binaries, where we cannot measure v_{2r} , we can substitute the relation $v_{2r} = v_{1r} m_1 / m_2$ (eq. 4.6) into eq. 4.9 to obtain:

$$m_1 + m_2 = \frac{P}{2\pi G} \frac{v_{1r}^3}{\sin^3 i} \left(1 + \frac{m_1}{m_2}\right)^3 \quad (4.10)$$

which can be rearranged in a form which groups together all the observables on the right-hand side of the equation:

$$\frac{m_2^3}{(m_1 + m_2)^2} \sin^3 i = \frac{P}{2\pi G} v_{1r}^3. \quad (4.11)$$

The left-hand side of this equation is known as the mass function. Even if m_1 is not known, the mass function can still provide interesting lower limits to the mass of the unseen companion, since $m_1 > 0$ and $\sin i \leq 1$, and therefore:

$$\frac{P}{2\pi G} v_{1r}^3 < m_2 \quad (4.12)$$

If the condition $m_2 \ll m_1$ is satisfied, which is the case of the secondary component of the binary system is a planet, then $m_1 + m_2 \approx m_1$. Substituting into 4.11, we now have:

$$m_2^3 \sin^3 i \approx \frac{P}{2\pi G} v_{1r}^3 m_1^2 \quad (4.13)$$

While there is still an inclination uncertainty for any particular system, statistical results can be obtained for large sample of stars with measured oscillations attributable to planet-mass companions.

4.4 Eclipsing Binaries

The ambiguities associated with the unknown orientation can be removed in cases where we see occultations of one of the stars by the other. Provided that the separation between the two stars is much greater than the sum of their radii (a condition which is *not* satisfied in contact binaries), then it must be the case that the inclination of the orbital plane to the sky is close to 90° (see Figure 4.6). Note also that for $i > 75^\circ$, $\sin i > 0.9$, so that the error in the masses deduced with the assumption that $i = 90^\circ$ is less than 10%.

Comparing the light curves for the cases of complete (Figure 4.6) and partial (Figure 4.7) eclipse, it can be appreciated that it is possible to recognise the cases where $i < 90^\circ$.

When the eclipse is total, we can deduce the radii of both stars from accurate timing of the phases of the eclipse. With the assumption that the smaller star is moving perpendicularly to our line of sight during the duration of the eclipse, its radius can be straightforwardly derived from

$$r_s = \frac{v}{2} (t_b - t_a) \quad (4.14)$$

where t_a and t_b are the times of first contact and minimum light respectively (see Figure 4.6) and $v = v_s + v_l$ is the relative velocity of the two stars.

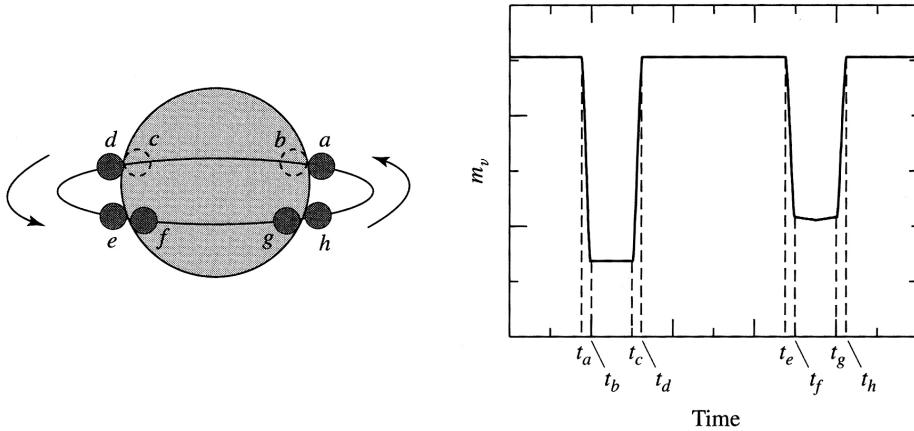


Figure 4.6: Schematic diagram of an eclipsing binary. The smaller star is assumed to be hotter than the larger one. (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

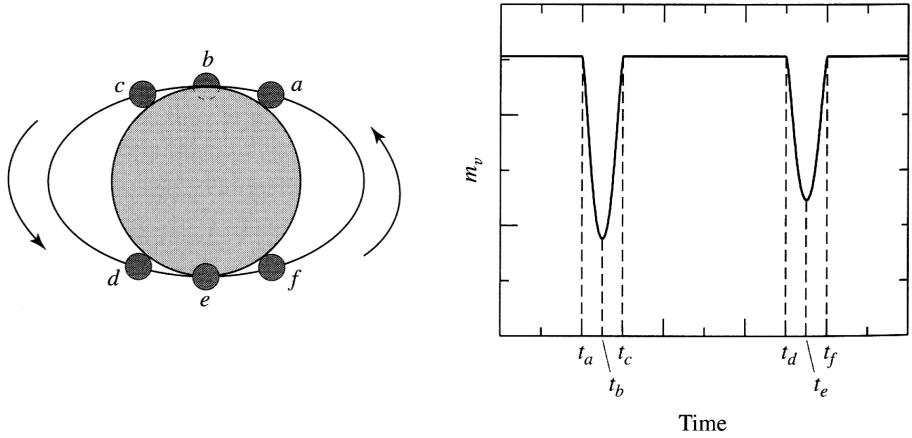


Figure 4.7: Schematic diagram of a partial eclipsing binary. The smaller star is assumed to be hotter than the larger one. (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

Similarly:

$$r_l = \frac{v}{2} (t_c - t_a) = r_s + \frac{v}{2} (t_c - t_b) \quad (4.15)$$

The light curve of eclipsing binaries gives information not only on the radii of the two stars but also on the ratio of their effective temperatures. This follows directly from eq. 2.13, $L = 4\pi R^2 \sigma T^4$; as when an area πR^2 is eclipsed from the system, the drop in flux will be different depending on whether the hotter star of the two is in front or behind the cooler one (see Figure 4.6). Assuming for simplicity a uniform flux across the stellar disk, we have:

$$F_0 = A (\pi R_l^2 F'_l + \pi R_s^2 F'_s) \quad (4.16)$$

where F' is the radiative surface flux, F_0 is the measured flux when there is no eclipse, and A is a proportionality constant to account for the fact that we register only a fraction of the flux emitted (due to distance, intervening absorption and limited efficiency of the instrumentation). The deeper, or primary, minimum in the light curve occurs when the hotter star is eclipsed by the cooler one. In the example shown in Figure 4.6, this is the smaller star. Then, during the primary minimum we have:

$$F_1 = A \pi R_l^2 F'_l, \quad (4.17)$$

while during the secondary minimum:

$$F_2 = A (\pi R_l^2 - \pi R_s^2) F'_l + A \pi R_s^2 F'_s. \quad (4.18)$$

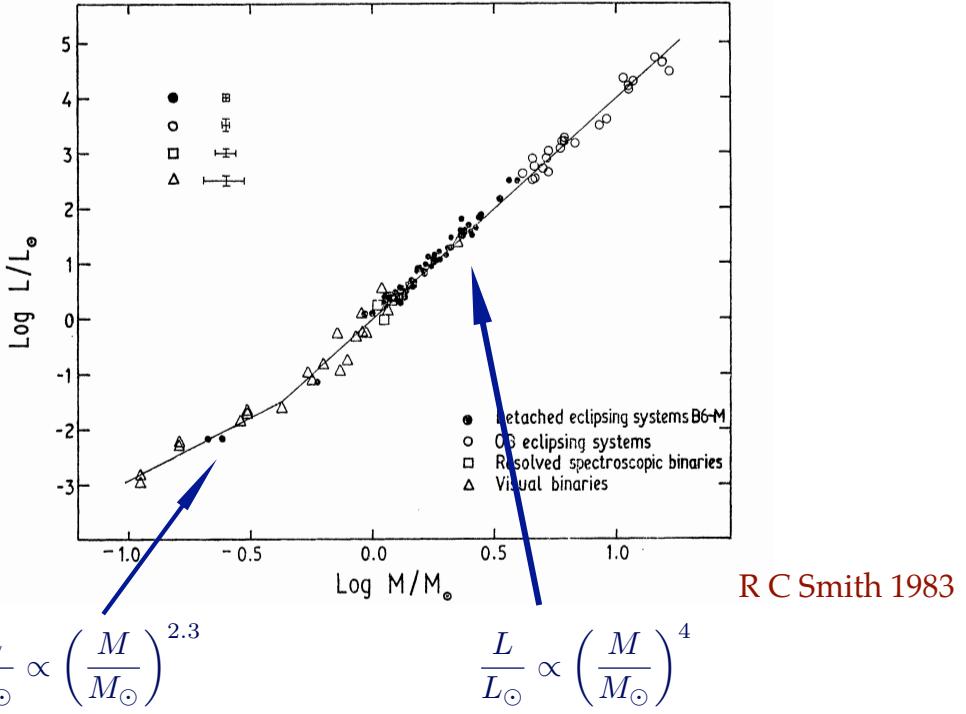


Figure 4.8: The empirical stellar mass-luminosity relation constructed from observations of different types of binary stars (from Smith 1983).

To circumvent uncertainties in the constant A , we concern ourselves with the ratio of the two fluxes:

$$\frac{F_0 - F_1}{F_0 - F_2} = \frac{F'_s}{F'_l} = \left(\frac{T_s}{T_l}\right)^4 \quad (4.19)$$

What eq. 4.19 tells us is that the ratio of the measured fluxes during the primary and secondary eclipses gives a direct measure of the ratio of the effective temperatures of the two stars in the eclipsing binary system.

4.5 The Stellar Mass-Luminosity Relation

When we bring together the best determinations of stellar masses from different types of binary stars, we find a well defined mass-luminosity relation for hydrogen burning dwarfs. Figure 4.8 shows the empirical mass-luminosity relation constructed from data available in the late 1970s-early 1980s. Thirty years later, the number of stars with direct measurements of mass and radius has increased considerably, thanks in part to the advent of long-baseline optical interferometry which can resolve the stellar disks. Figure 4.9, reproduced from the review by Torres et al. 2010 (A&ARv,

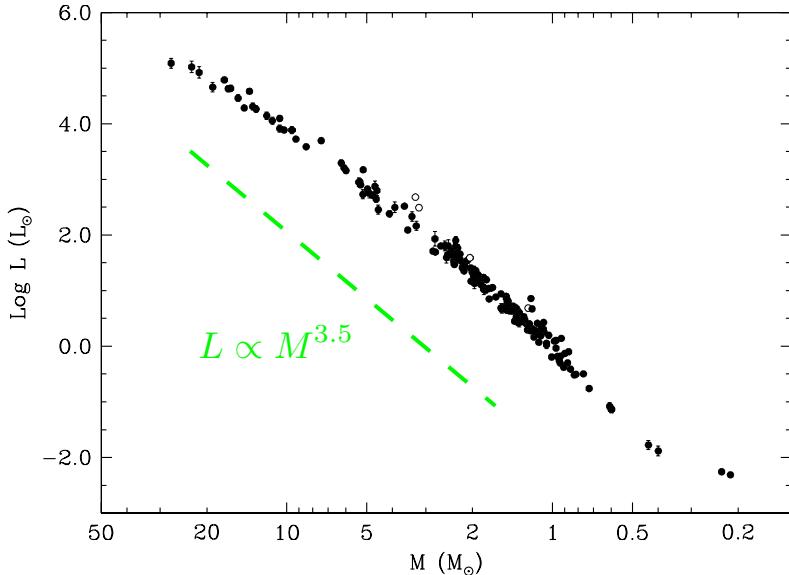


Figure 4.9: The empirical stellar mass-luminosity relation from observations of 190 stars in 95 detached binary systems, all with masses and radii known with an accuracy of 3% or better (data from Torres et al. 2010).

18, 67), is a compilation of measurements for 95 detached binary systems containing 190 stars satisfying the criterion that the mass and radius of both stars be known with an accuracy of 3% or better.

Any theory of stellar structure must be able to reproduce such a relation in order to be deemed valid; we shall return to this point in Lecture 10. Here we limit ourselves to some preliminary considerations. First of all, such a clear-cut $M - L$ relation provides a natural explanation for the existence of a prominent main sequence in the HR diagram. After forming within a collapsing interstellar cloud, stars begin their hydrogen-burning lives on the main sequence, at a location on the $M_V - (B - V)$ plane *determined by their mass*. Stars do *not* evolve along the main sequence, they evolve *off* the main sequence.

A rough approximation to the slope of the mass-luminosity relation over the full range of stellar masses is $L \propto M^{3.5}$. If stars shine through nuclear fusion, we can write:

$$\frac{dM}{dt} = k L$$

where L is the luminosity and k is a constant of proportionality. Integrating, we have:

$$t \propto \frac{M}{L} \propto \frac{M}{M^{3.5}} \propto M^{-2.5}.$$

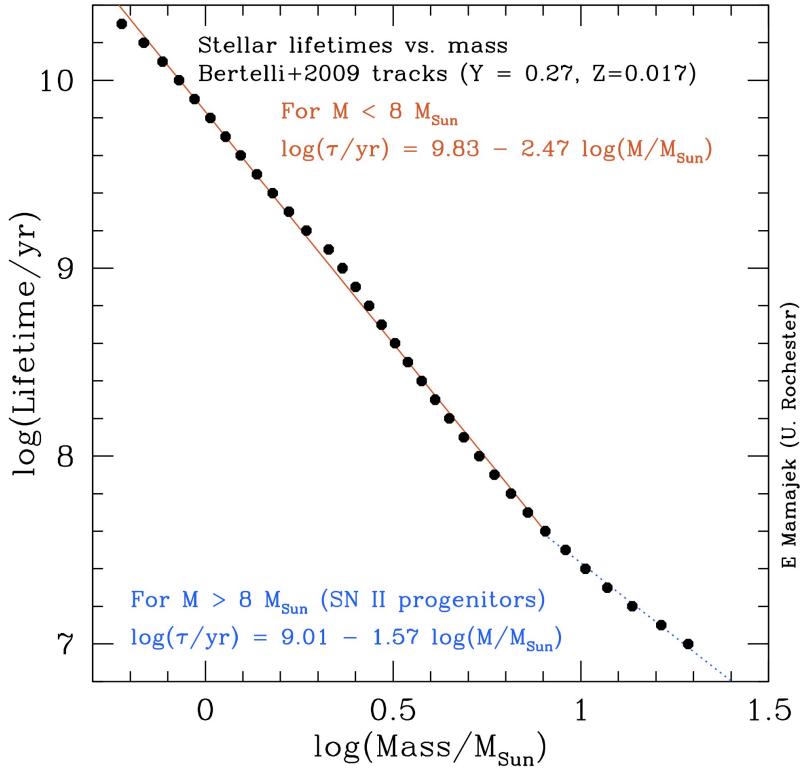


Figure 4.10: Stellar lifetimes as a function of mass from the stellar evolution models of Bertelli et al. (2009). Stars more massive than $8M_{\odot}$ (with lifetimes $t < 4 \times 10^7$ years) are thought to end their lives as Type II supernovae.

In other words, the steep slope of the stellar mass-luminosity relation implies a very strong dependence of the stellar lifetimes on their mass. While a $1M_{\odot}$ star will burn hydrogen for 10^{10} years before evolving off the main sequence, a $20M_{\odot}$ star has sufficient fuel to last for only 10^7 years (see Figure 4.10).

STELLAR ATMOSPHERES. I: STELLAR OPACITY

5.1 Introduction and Definitions

The light we receive from a star originates in its atmosphere, the layers of gas overlying the opaque interior. The photons produced there carry away the gravitational energy released when the star forms from a collapsing gas cloud, the energy released by the thermonuclear reactions taking place in the core of the star during its lifetime and, once thermonuclear reactions stop altogether, the energy released by the cooling of the star's interior, as is the case in white dwarfs. The photons collected by our telescopes carry with them information about the temperature, density, and chemical composition of the atmospheric layers from which they have escaped (see Figure 5.1). In order to decode this information we must understand how light travels through the gas which makes up a star.

It is useful at this point to clarify what we mean by some of the quantities we are going to use. We define the specific intensity (or simply the

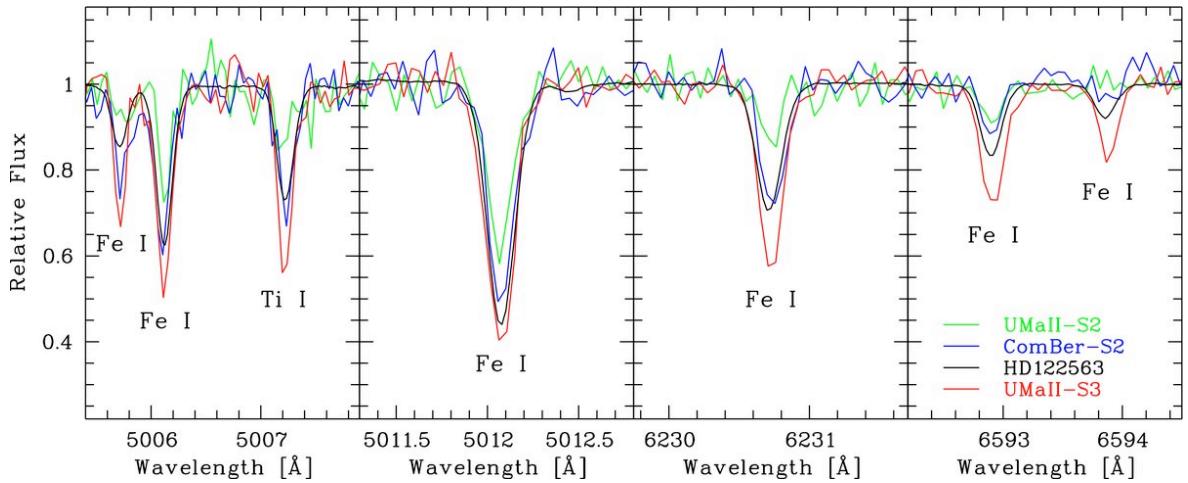


Figure 5.1: Absorption lines in four stars of similar effective temperatures by different Fe abundance, as follows: UMa II-S2 (green), $[\text{Fe}/\text{H}] = -3.2$; ComBer-S2 (blue), $[\text{Fe}/\text{H}] = -2.9$; HD 122563 (black), $[\text{Fe}/\text{H}] = -2.8$; and UMa II-S3 (red), $[\text{Fe}/\text{H}] = -2.3$. The iron abundance is measured on a log scale relative to its value in the Sun; thus a star with $[\text{Fe}/\text{H}] = -3$ has only 1/1000 of the solar Fe/H ratio (by number), one with $[\text{Fe}/\text{H}] = -2$ has 1/100 the solar ratio and so on.

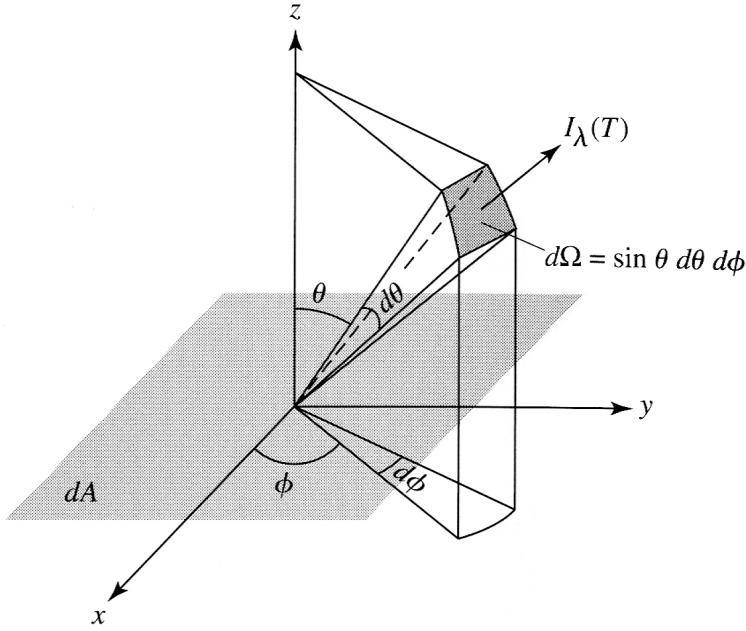


Figure 5.2: Nomenclature for the definition of specific intensity I_λ . (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

intensity) I_λ ,

$$E_\lambda d\lambda = I_\lambda d\lambda dt dA \cos \theta \sin \theta d\theta d\phi \quad (5.1)$$

as the amount of electromagnetic radiation energy with wavelength between λ and $\lambda + d\lambda$ that passes in time dt through the star surface area dA into a solid angle $d\Omega = \sin \theta d\theta d\phi$ (see Figure 5.2). Thus the units of intensity are (in the cgs systems used by astronomers) $\text{erg s}^{-1} \text{cm}^{-2} \text{\AA}^{-1} \text{sr}^{-1}$.

The mean intensity J_λ is obtained by integrating the intensity over all directions and dividing by the solid angle enclosed by a sphere, which is $4\pi \text{ sr}$. Thus:

$$J_\lambda = \frac{1}{4\pi} \int I_\lambda d\Omega = \frac{1}{4\pi} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} I_\lambda \sin \theta d\theta d\phi \quad (5.2)$$

For an isotropic radiation field, $J_\lambda = I_\lambda$. A blackbody radiates isotropically by definition; thus $J_\lambda = B_\lambda$.

The specific energy density u_λ measures how much energy is contained within the radiation field, that is the energy per unit volume having a wavelength between λ and $\lambda + d\lambda$:

$$u_\lambda d\lambda = \frac{1}{c} \int I_\lambda d\lambda d\Omega = \frac{4\pi}{c} J_\lambda d\lambda \quad (5.3)$$

Thus, for a blackbody radiation field,

$$u_\lambda d\lambda = \frac{4\pi}{c} B_\lambda d\lambda = \frac{8\pi hc/\lambda^5}{e^{hc/\lambda kT} - 1} d\lambda \quad (5.4)$$

and in frequency, rather than wavelength, units:

$$u_\nu d\nu = \frac{4\pi}{c} B_\nu d\nu = \frac{8\pi h\nu^3/c^3}{e^{h\nu/kT} - 1} d\nu. \quad (5.5)$$

The total energy density is found by integrating over all wavelengths (or frequencies):

$$u = \int_0^\infty u_\lambda d\lambda. \quad (5.6)$$

For a blackbody:

$$u = \frac{4\pi}{c} \int_0^\infty B_\lambda(T) d\lambda = \frac{4\sigma}{c} T^4 = aT^4 \quad (5.7)$$

where $a \equiv 4\sigma/c$ is the radiation constant.

5.2 Local Thermodynamic Equilibrium

We can define the temperature of a star in many different ways:

1. *The effective temperature*, defined in terms of the luminosity of the star (and its radius) according to eq. 2.13;
2. *The excitation temperature*, defined by the relative populations of different excited levels of an atom or ion according to Boltzmann equation (eq. 3.2);
3. *The ionisation temperature*, defined by the relative populations of different ionisation stages of an atom according to Saha equation (eq. 3.4);
4. *The kinetic temperature*, defined by the Maxwell-Boltzmann distribution:

$$n_v dv = n \left(\frac{m}{2\pi kT} \right)^{3/2} e^{-mv^2/2kT} 4\pi v^2 dv \quad (5.8)$$

where n_v is the number of particles per unit volume with speeds between v and $v + dv$, n is the volume density (of all particles), m is the mass and the other symbols have their usual meanings;

5. *The colour temperature*, being the temperature of the blackbody whose spectral energy distribution resembles most closely that of the star.

Except for the effective temperature, which is an important global descriptor of the star, all of the other temperatures apply to any location within the stellar interior and vary according to the physical conditions of the gas. Of course, all of them become the same in the ideal situation of the particles and radiation being in equilibrium. In such a steady-state condition, there is no net flow of energy in nor out of a volume element, nor any transfer of energy between matter and radiation. Every process, such as the absorption of a photon, occurs at the same rate as its inverse process, such as the emission of a photon. Such an idealised condition is referred to as thermodynamic equilibrium. A blackbody is by definition in thermodynamic equilibrium.

Clearly this ideal condition does not apply to stars. There is a net outward flow of energy through a star. The temperature varies from millions of degrees in the core to thousands of degrees in the atmosphere. As the gas particles collide with one another and interact with the radiation field by absorbing and emitting photons, the description of the processes of excitation and ionisation becomes quite complex. Despite all of this complications, we can still apply the the approximation of Local Thermodynamic Equilibrium (LTE), *provided the typical distance travelled by particles and photons between collisions—their mean free path—is small compared to the scale over which the temperature changes significantly*. This situation applies to most of the stellar interior, where density and temperature are high, so that the mean distance between collisions is small. You can think of this situation as the particles and photons being confined to a limited volume of nearly constant temperature.

5.3 Opacity

As a beam of light of intensity I_λ travels through a gas, some of the photons will be removed through scattering off and absorption by ions, atoms and molecules in the gas. We can write:

$$dI_\lambda = -\kappa_\lambda \rho I_\lambda ds, \quad (5.9)$$

where ds is the distance travelled, ρ is the density of the gas and κ_λ is the absorption coefficient, or opacity. In terms of the photon mean free path:

$$\mu = \frac{1}{\kappa_\lambda \rho} = \frac{1}{\sigma_\lambda n}$$

where σ is the cross-section for interaction and n is the number density of particles. With these definitions, κ has units $\text{cm}^2 \text{ g}^{-1}$, ρ has units g cm^{-3} , σ is in cm^2 , and n is in cm^{-3} . Thus, both products on the denominators of the above equation can be thought of as the fraction of photons scattered off the beam per cm travelled.

We define the optical depth as:

$$\tau_\lambda = \int_0^s \kappa_\lambda \rho ds \quad (5.10)$$

In the case of a photon travelling from the stellar interior to the surface, $s = 0$ at the starting point, and $\tau_\lambda = 0$ at the surface of the star. Thus, we can think of the optical depth as the number of mean free paths for the

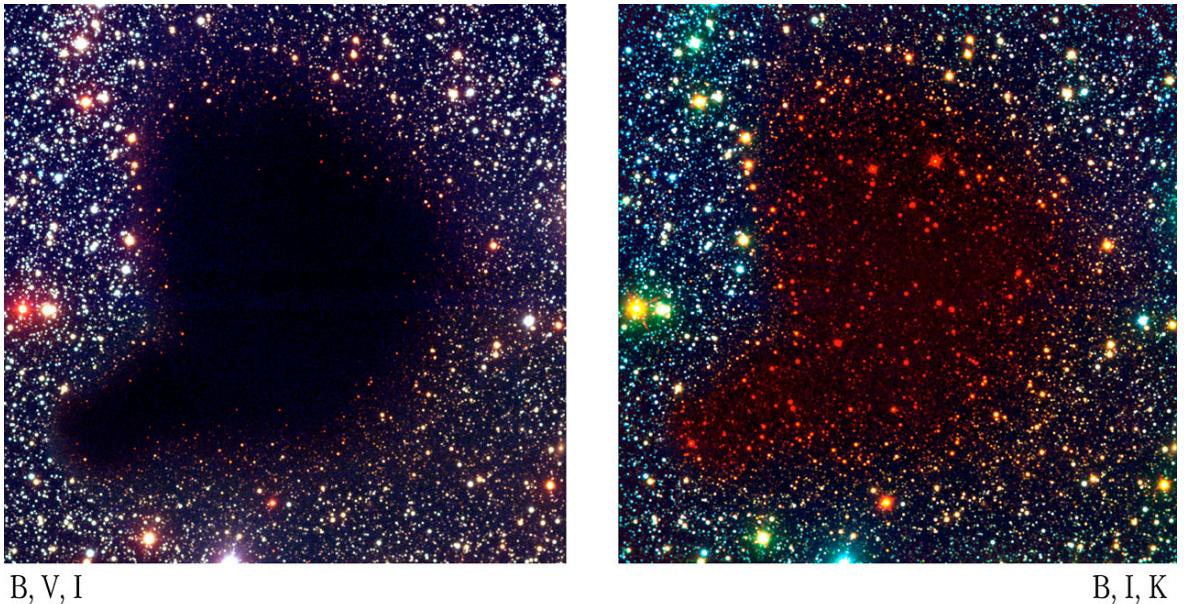


Figure 5.3: Colour composites of ESO images of the Dark Cloud Barnard 68 taken through different optical and near-IR filters. Dust in the cloud hides stars located behind it; because dust scatters infrared light less efficiently than optical light, the opacity of the cloud is reduced in the colour composite that includes the near-IR K -band filter. The cloud is located at a distance of 160 pc and is seen against the background of a rich star field in the Milky Way. Curiously, there are no foreground stars. Barnard 68 seems to be a molecular cloud in the earliest phase of collapse to form new stars; for this reason it is the subject of many studies at a variety of wavelengths.

photon, from a given location in the star's interior to the surface. Note that in all the above equations we have explicitly indicated that κ_λ , σ_λ and τ_λ are all functions of wavelength. Figure 5.3 is a vivid demonstration of the wavelength dependence of the opacity, in this case of interstellar dust.

Integrating eq. 5.9 and with the definition of eq. 5.10, we have:

$$I_\lambda = I_{\lambda,0} e^{-\tau_\lambda} \quad (5.11)$$

where $I_{\lambda,0}$ is the intensity at wavelength λ that would be measured in the absence of absorption/scattering. Gas with $\tau_\lambda \gg 1$ is said to be optically thick; conversely, if $\tau_\lambda \ll 1$ the gas is optically thin. From the above, it can be appreciated that we typically do not see deeper into a stellar atmosphere (at a given wavelength) than unit optical depth, i.e. $\tau_\lambda \approx 1$.

5.4 Sources of Opacity

When dealing with stellar atmospheres, the opacity can be due to one of four main physical processes (or a combination of them):

1. Bound-bound transitions. These are the familiar transitions between different energy levels which cause absorption (or emission) lines at discrete wavelengths. Thus $\kappa_{\lambda,bb}$ is small or zero at all wavelengths except those which correspond to the energy difference between two atomic levels.

It is of interest to consider what happens to the photon after it is absorbed. If the electron spontaneously returns to the same energy level from which the upward transition took place, a photon with the same energy is re-emitted but in a random direction. So, the net effect is one of scattering. On the other hand, if the electron returns to its original state via two or more transitions to intermediate energy levels, then the net effect is to degrade the average energy of the photons in the radiation field.

2. Bound-free absorption. This is the familiar process of photoionisation which will occur for all photon energies $h\nu > \chi_n$, where χ_n is the ionisation potential of a given atomic energy level. The difference $\Delta E = h\nu - \chi_n$ is the kinetic energy of the free electron. Thus, $\kappa_{\lambda,bf}$

is one source of continuum opacity (as opposed to opacity at discrete wavelengths as is the case for $\kappa_{\lambda,bb}$).

3. Free-free absorption. This is the inverse process of free-free emission (*bremssstrahlung*) in which a free electron is decelerated by the electric potential of an ion and, as a result, radiates. Thus, in free-free absorption, a photon is absorbed by a free electron and an ion, which share the photon's momentum and energy. $\kappa_{\lambda,ff}$ also contributes to the continuum opacity.
4. Electron scattering. This is the scattering of photons by free electrons without change of photon energy (Thompson scattering). The cross-section for Thompson scattering is independent of wavelength, so that κ_{es} is also a source of continuum opacity. Its value is very small, $\sigma_T = 6.7 \times 10^{-25} \text{ cm}^2$, some seven orders of magnitude smaller, for example, than the cross-section for photoionisation of neutral hydrogen from the ground state, $\sigma_0 = 6.3 \times 10^{-18} \text{ cm}^2$. Thus, Thomson scattering is important only when the density of free electrons is very high, that is when the gas is fully ionised (or nearly so), as is the case in the atmospheres of the hottest stars.

5.4.1 H⁻ and Molecules

The temperatures of stars later than spectral type F0 ($T_{\text{eff}} \lesssim 7300 \text{ K}$) are sufficiently low for non-negligible concentrations of H⁻ to be present. With a binding energy of only 0.754 eV ($\sim 1/20$ the binding energy of the electron in neutral hydrogen), corresponding to a wavelength of 16.4 μm , H⁻ is an important source of continuum opacity. Even at longer wavelengths, the H⁻ ion adds to the continuum opacity through free-free absorption. In A and B stars, photoionisation of H and free-free absorption are the main sources of continuum opacity. In O stars, He photoionisation and Thomson scattering also contribute. In the coolest stars, significant concentrations of molecules, as opposed to atoms and ions, can survive, providing additional sources of opacity through bound-bound (many molecular bands), bound-free, and photodissociation (the absorption of a photon breaks the weak molecular bond) transitions (see Figure 5.4).

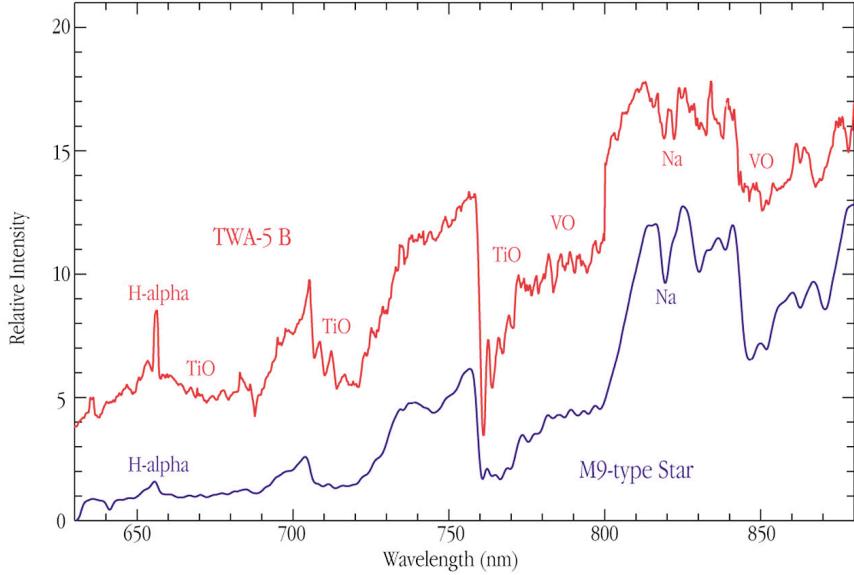


Figure 5.4: At optical and infrared wavelengths the spectra of cool stars are dominated by strong molecular bands.

5.5 Mean Opacities

It can be appreciated that the full calculation of stellar opacities, which depend on the chemical composition, pressure and temperature of the gas, as well as the wavelength of the incident light, is a complex endeavour. The problem can be simplified by using a mean opacity averaged over all wavelengths, so that only the dependence on the gas physical properties remains. The most commonly used is the Rosseland mean opacity, defined as:

$$\frac{1}{\langle \kappa \rangle} = \frac{\int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B_\nu(T)}{\partial T} d\nu}{\int_0^\infty \frac{\partial B_\nu(T)}{\partial T} d\nu}. \quad (5.12)$$

Note that this is an harmonic mean, in which the greatest contribution comes from the lowest values of opacity, weighted by a function that depends on the rate at which the blackbody spectrum varies with temperature.

Returning to the different sources of opacity considered earlier, let us consider their Rosseland means. There is no simple analytical expression that describes the multitudes of absorption lines resulting from bound-bound transitions. However, we can do somewhat better for the bound-free and free-free transitions, whose mean opacities can be approximated

by Kramers opacity law:

$$\langle \kappa_{bf} \rangle = \kappa_{0,bf} \rho T^{-3.5} \quad (5.13)$$

and

$$\langle \kappa_{ff} \rangle = \kappa_{0,ff} \rho T^{-3.5} \quad (5.14)$$

where $\kappa_{0,bf}$ and $\kappa_{0,ff}$ are constants that depend on the composition of the gas.

For electron scattering, which has no wavelength, density, nor temperature dependence, the Rosseland mean takes a particularly simple form:

$$\kappa_{es} = \kappa_{0,es} \frac{1}{\mu_e} \quad (5.15)$$

where $1/\mu_e$ is the number of electrons per nucleon and $\kappa_{0,es}$ is the opacity of a fully ionised, pure hydrogen gas (in which case $1/\mu_e = 1$, clearly). Recalling (lecture 1) that we use the symbols X , Y , and Z to denote the mass fractions of, respectively, H, He and everything else (metals),

$$\frac{1}{\mu_e} = X + \frac{1}{2}Y + \frac{1}{2}(1 - X - Y) \simeq \frac{1}{2}(1 + X)$$

since helium and the most abundant heavier elements are (nearly) fully ionised in the interior of stars, so that the average number of electrons per nucleon is $\approx 1/2$.

The mean opacity due to H^- does not have a simple analytical approximation.

The full calculation of stellar opacities is a major endeavour. For this reason, the *Opacity Project* (OP) was set-up in 1984 as an international collaboration to calculate the extensive atomic data required to estimate stellar envelope opacities and to compute Rosseland mean opacities and other related quantities. It brought together the work of research groups from France, Germany, the United Kingdom, the United States and Venezuela, under the leadership of M. J. Seaton at University College London. The project involved, among other aspects, the computation by *ab initio* methods of accurate atomic properties such as energy levels, oscillator strengths and photoionization cross sections. Similar work conducted at the Lawrence Livermore National Laboratory, the OPAL project, gave results generally in good agreement with those of the OP.

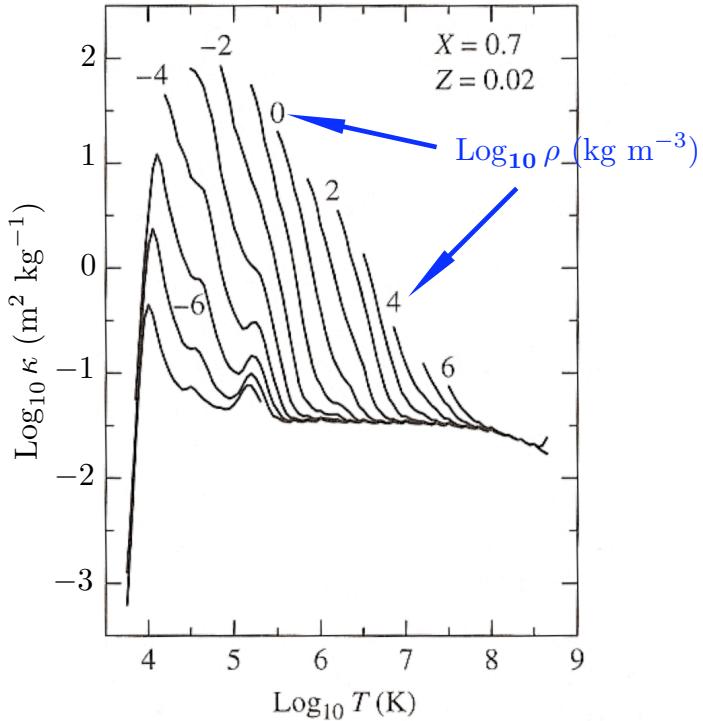


Figure 5.5: Rosseland mean opacities calculated by the OPAL project for ‘standard’ solar abundances (Iglesias & Rogers 1996).

Figure 5.5 shows the OPAL calculations of the Rossland mean opacity as a function of density and temperature for a gas of solar composition. It is worthwhile considering the behaviour of $\langle \kappa \rangle$ even though we cannot derive the curves ourselves. The first thing to notice is that at a fixed temperature the opacity increases with density (as one may expect). Focussing on one of the curves, note that $\langle \kappa \rangle$ at first increases rapidly with increasing T , reflecting the increase in the number of free electrons produced by ionisation of H and He. The decline with increasing T beyond the peak value of $\langle \kappa \rangle$ follows a Kramers law, $\langle \kappa \rangle \propto T^{-3.5}$, and is due mostly to bound-free and free-free absorption. All the curves come together at very high temperatures, when the gas is fully ionised and the main source of opacity is electron scattering (eq. 5.15). The small bumps superimposed on some of the curves occur at the temperatures where He and higher mass elements (mostly Fe) are ionised, releasing additional electrons into the plasma.

Opacities are an essential ingredient of model stellar atmospheres.

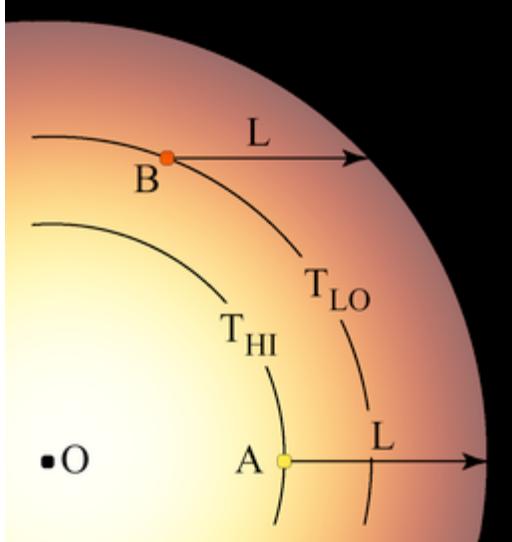


Figure 5.6: A given linear distance L probes to different depths into a stellar atmosphere depending on the angle of the line of sight to the radial coordinate.

5.6 Limb Darkening

It was stated earlier that we typically see no further into a star than unity optical depth. A more careful treatment actually shows that the level within a stellar atmosphere from which most of the photons of wavelength λ escape is at optical depth $\tau_\lambda \simeq 2/3$. Indeed, the condition $\tau_\lambda \simeq 2/3$ defines the stellar photosphere—the layer of a star’s atmosphere from which the light we see originates.

There are two consequences of this realisation. First, the condition applies to all viewing angles; therefore, the distance ds corresponding to the condition $\tau_\lambda = 2/3$ will probe further into the star’s interior at the centre of a stellar disk than at its edges (see Figure 5.6). Second, recalling the definition of optical depth,

$$\tau_\lambda = \int_0^s \kappa_\lambda \rho \, ds \quad (5.10)$$

it is obvious that if the opacity κ_λ increases at some wavelength, then ds must be smaller to satisfy the condition $\tau_\lambda \simeq 2/3$. Thus, we see further into a star in its continuum light than at the wavelengths of discrete absorption lines.

These two effects explain a phenomenon known as ‘limb darkening’, first recognised in the Sun, whereby the light emitted in successive annuli from the centre decreases in intensity and becomes progressively redder, as

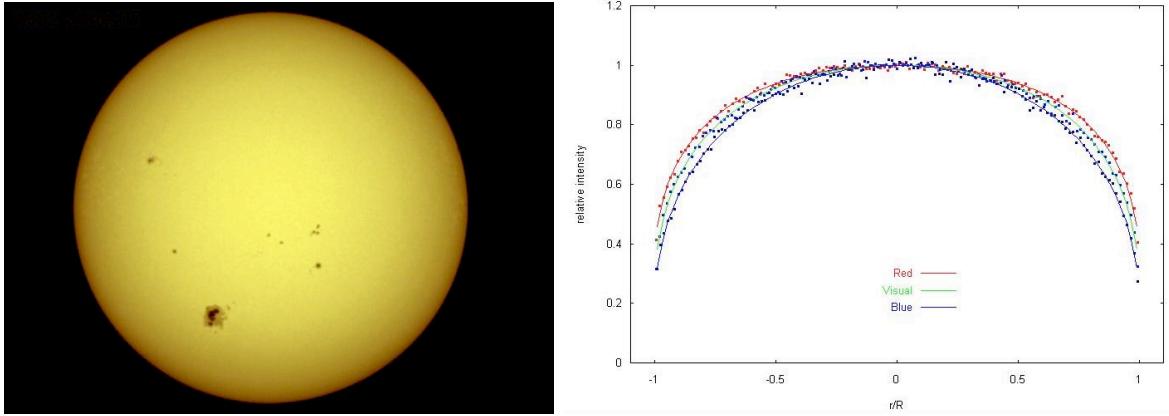


Figure 5.7: Limb darkening in the Sun r/R is the fractional distance from the centre of the solar disk. Note that the effect is more pronounced at shorter wavelengths: near the edge of the disk, the Sun is not only dimmer but also redder.

shown in Figure 5.7. Sightlines near the limb do not penetrate as deeply into the Sun’s atmosphere by the time $\tau_\lambda \simeq 2/3$ is reached; since the Sun’s temperature decreases outwards from the centre, such sightlines see light from cooler regions of the Sun’s atmosphere.

While most easily seen in the Sun because of its proximity to Earth, limb darkening has now been seen in a handful of nearby stars, and its effects have been recognised in the light curves of eclipsing binaries and of microlensing events (when one star passes in front of a second, more distant, star whose brightness can be greatly amplified as light rays are bent towards us by gravitational lensing).

5.7 Summary of Definitions

I_λ : Intensity ($\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ sr}^{-1}$)

B_λ : Mean intensity of blackbody radiation ($\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1} \text{ sr}^{-1}$)

u : energy density (erg cm^{-3})

κ_λ : opacity ($\text{cm}^2 \text{ g}^{-1}$)

τ_λ : optical depth

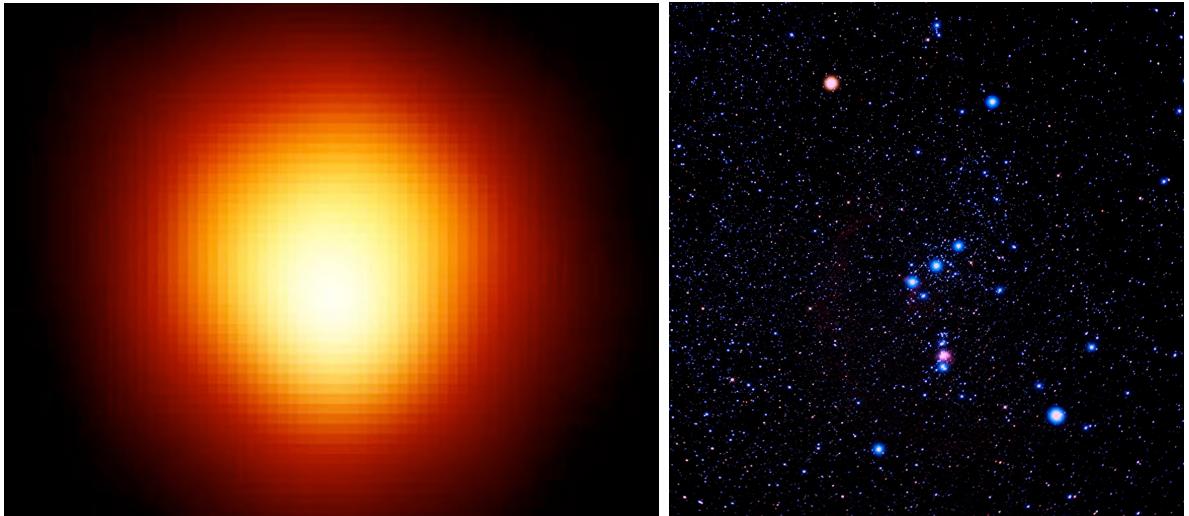


Figure 5.8: This is the first direct image of the surface of a star other than the Sun. The star is Betelgeuse (α Ori), a red supergiant (M2 Iab) at a distance of ~ 150 pc; among the ten brightest stars in the night sky, it is easily recognised in the constellation of Orion (right). The image on the left was obtained with the *Hubble Space Telescope* at near-ultraviolet wavelengths. The diameter of the star at 2500 \AA is 2.2 times larger than at optical wavelengths, as a result of the higher UV opacity. Limb-darkening is evident in the *HST* image. The diameter of α Ori is ~ 1000 times larger than that of the Sun.

STELLAR ATMOSPHERES. II: THE ANALYSIS OF ABSORPTION LINES

6.1 Introduction

Much of our knowledge of the physical conditions of astrophysical systems relies on spectroscopic observations. The absorption lines formed in the atmospheres of stars contain a wealth of information, on the temperature, gravity, rotational velocity and chemical composition of the star. The last one in particular is of great interest as it reflects (for stars on the main sequence at least) the integrated yields of nucleosynthesis by previous generations of stars. For example, if we can decipher the clues encrypted in the chemical composition of different stellar populations of the Milky Way, we can begin to reconstruct the past history of star formation of our Galaxy (see Figure 6.1).

Another example is offered by the most metal-poor star known, with metallicities of only 10^{-4} to 10^{-5} of solar metallicity or, in the shorthand commonly used by astronomers, with $[Fe/H] = -4$ to -5 . Thus, $[Fe/H]_{\text{star}} = \log (Fe/H)_{\text{star}} - \log (Fe/H)_{\odot}$. The relative proportions of different elements of the periodic table in the atmospheres of these stars can be very different from their proportions today, as measured in the Sun or in the gas from which stars have recently formed, such as the Orion nebula. Possibly these most metal-poor stars have been enriched by only one previous generation of stars which themselves presumably formed out of pristine gas. Thus, the chemical composition of the most metal-poor stars known is one of the means at our disposal to shed light on the nature of the ‘First Stars’, sometimes referred to as Population III stars.

A further example is provided by the relative element ratios in the atmospheres of stars that have evolved off the main sequence. In some cases, material deep in the stellar core is ‘dredged up’ to the surface. The composition of this gas can be quite different from that of the interstellar cloud from which the star originally formed, as a result of the nuclear reactions which have powered the star over the course of its lifetime. Measurements

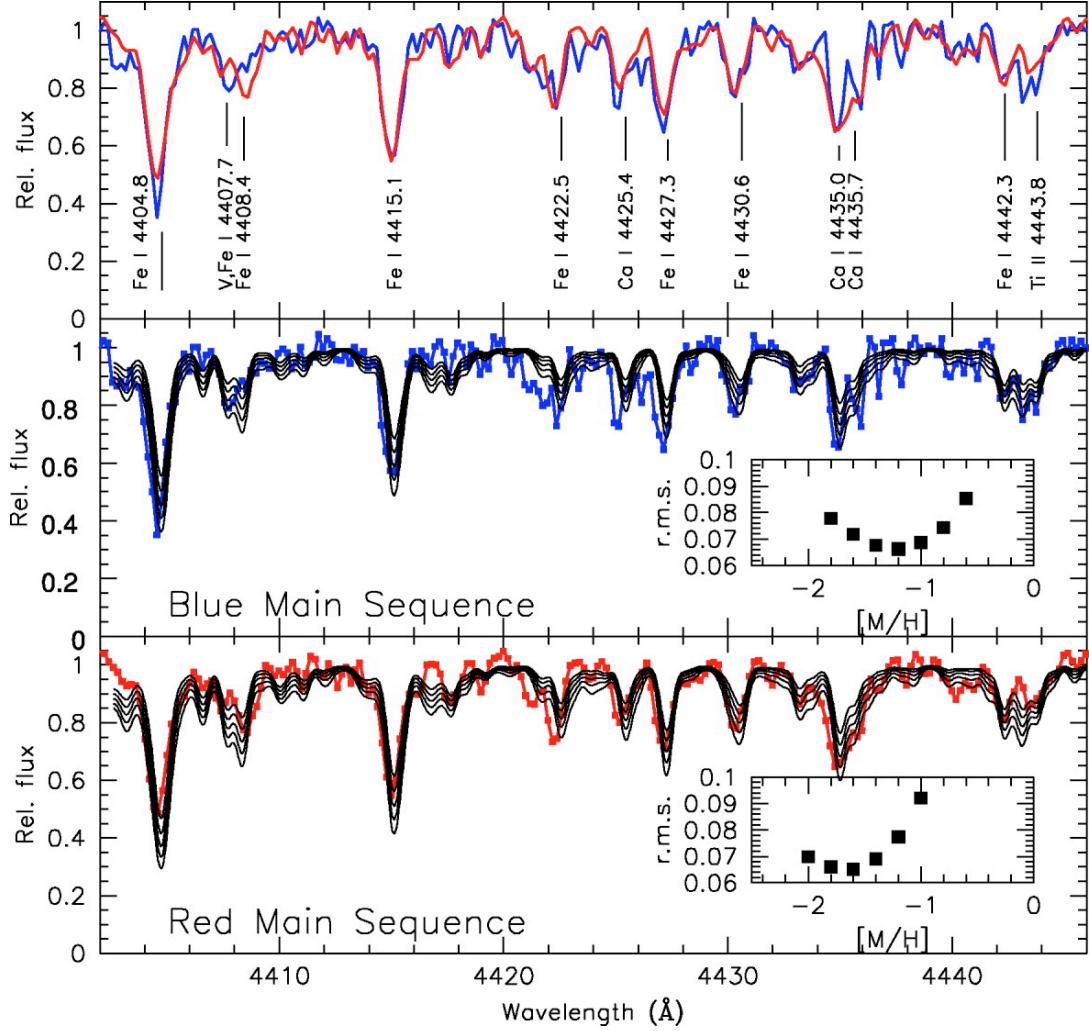


Figure 6.1: The globular cluster ω Centauri (at a distance of 5 kpc in the Galactic halo, with a mass of $10^7 M_\odot$) is unique among Galactic globular clusters in showing two distinct populations of stars, a bluer population and a redder one, with quite distinct main sequences in the colour-magnitude diagram. The origin of these two populations is still unclear. This figure compares a small portion of the average spectrum of 17 blue and 17 red stars, recorded at high resolution with the ESO Very Large Telescope. The top panel shows that there are clear differences in most spectral lines between the two spectra. The middle and bottom panels compare the observed spectra with those produced by model atmospheres with different metallicities $[M/H]$, where M is a ‘metal’ (Fe, Co, Ti) and $[M/H]_{\omega\text{Cen}} = \log(M/H)_{\omega\text{Cen}} - \log(M/H)_\odot$, as indicated. The surprising result is that the blue population is less metal-poor than the red population ($[M/H] = -1.2$ compared to $[M/H] \simeq -1.6$), adding to the puzzle that is ω Cen.

of element ratios in the atmospheres of these evolved stars give us the means to check empirically the validity of our ideas as to which nuclear reactions are dominant in stars of different masses.

In order to extract such information from the analysis of stellar spectra, we need to understand how absorption lines are formed in stellar atmospheres.

6.2 Line Equivalent Width

We have already encountered the optical depth τ_λ in the preceding lecture, where we saw that a light beam emerges from the stellar surface with an intensity:

$$I_\lambda = I_{\lambda,0} e^{-\tau_\lambda} \quad (6.1)$$

where $I_{\lambda,0}$ is the intensity at its starting point within the star's interior. We also saw that we can write the optical depth in terms of the opacity and the distance travelled

$$\tau_\lambda = \int_0^s \kappa_\lambda \rho \, ds = \int_0^s n \sigma_\lambda \, ds. \quad (6.2)$$

In this lecture we shall use the second equality, where n is the volume density of particles (atoms and ions) and σ_λ is the cross-section for the interaction which, in the case of discrete absorption lines, corresponds to bound-bound transitions, that is electronic transitions between different energy levels.

We define the equivalent width W_λ of an absorption line as the width, in wavelength units, of a rectangular strip of spectrum having the same area

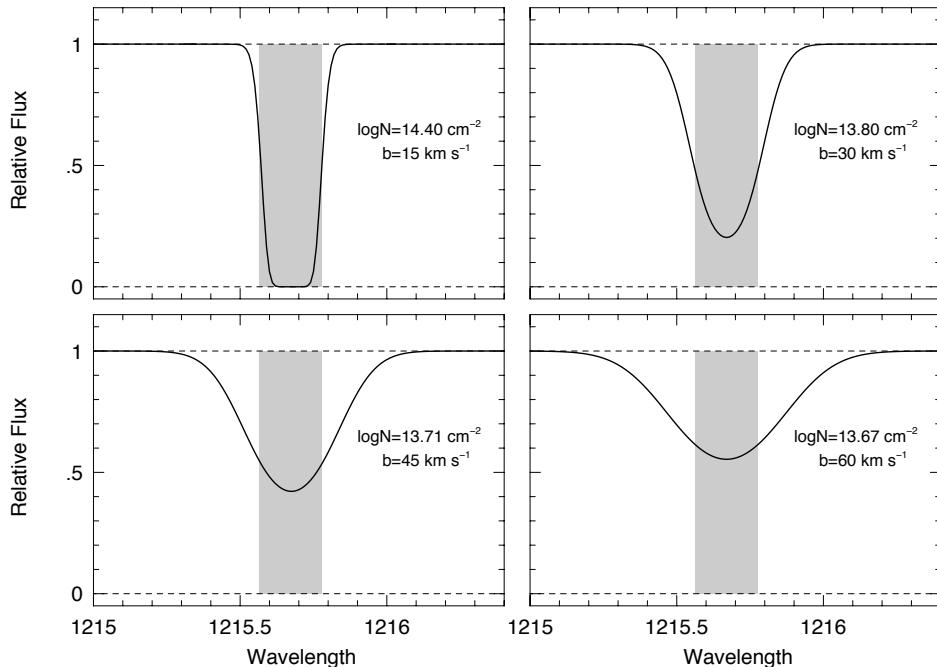


Figure 6.2: Four absorption lines with the same equivalent width (shaded grey area) but different widths, as measured by the value of b (km s^{-1}). N (cm^{-2}) is column density, defined as the number of absorbers in a column of unit cross-sectional area. Figure courtesy of Chris Churchill.

as the absorption line:

$$W_\lambda = \int_{-\infty}^{\infty} \frac{I_{\lambda,0} - I_\lambda}{I_{\lambda,0}} d\lambda = \int_{-\infty}^{\infty} (1 - e^{-\tau_\lambda}) d\lambda \quad (6.3)$$

The equivalent width is a convenient measure of the strength of an absorption line. Its value lies in the fact that historically the resolution of spectrographs was too coarse to resolve the intrinsic profiles of the absorption lines. What all instrument record is the *convolution* of the intrinsic line shape with the instrumental broadening function; if the latter is broader than the intrinsic width of the absorption line recorded, much of the information encoded in the line profile itself is lost. However, the equivalent width is invariant to the convolution, and is thus a conserved quantity (modulo complications introduced by limited signal-to-noise ratio of the spectra). Some stellar parameters of interest can be deduced from measurements of the line equivalent widths, particularly the relative fractions of atoms and ions in different excitation and ionisation stages and the relative abundances of different elements.

6.3 Line Broadening Processes

The cross-section for absorption, σ_λ , in eq. 6.2 can be written as:

$$\sigma_\lambda = \sigma_0 \cdot \Phi_\lambda \quad (6.4)$$

where σ_0 includes the atomic parameters of the transition

$$\sigma_0 = \frac{\lambda^4}{8\pi c} \frac{g_u}{g_l} a_{ul} \quad (6.5)$$

where g_u and g_l are the statistical weights of the upper and lower electronic energy levels, a_{ul} is the transition probability and the other symbols have their usual meaning.

Φ_λ is the broadening function, defined so that if an absorption does take place in the line, $\Phi_\lambda d\lambda$ is the probability that the wavelength of the absorbed photon lies between λ and $\lambda + d\lambda$. Thus:

$$\int_{-\infty}^{+\infty} \Phi_\lambda d\lambda = 1 \quad (6.6)$$

The value of Φ_λ is large near λ_0 (the line centre) and falls off rapidly at longer and shorter wavelengths.

In stellar atmospheres, there are a number of physical processes that broaden the absorption lines. The two main processes are:

1. Natural broadening, due to the uncertainty ΔE in the energy of the upper atomic level k , and
2. Doppler broadening due to the motions of the absorbers.

6.3.1 Natural and Pressure Broadening

Considering natural broadening first, we have (for an atom at rest):

$$\phi_\lambda(v = 0) = \frac{1}{\pi} \frac{\delta_k}{\delta_k^2 + (\lambda - \lambda_0)^2} \quad (6.7)$$

where δ_k is the radiation damping constant in wavelength units:

$$\delta_k = \frac{\lambda^2}{4\pi c} \sum_{E_r < E_k} a_{kr} \quad (6.8)$$

Thus, the radiation damping constant is inversely proportional to the lifetime of the upper level k , which is related to an intrinsic ‘width’ of the energy level according to Heisenberg uncertainty principle:

$$\Delta E \approx \frac{\hbar}{\Delta t} \quad (6.9)$$

A line profile given solely by the natural broadening function in eq. 6.7 is sometimes referred to as a damping (or Lorentz) profile because its inverse square functional form is characteristic of the spectrum of radiation emitted by an electric charge undergoing damped simple harmonic motion.

In the dense atmospheres of stars, the electronic energy levels can be perturbed by collisions with neutral atoms and by close encounters with the electric field of an ion. Both processes are often referred to as pressure broadening. The full calculation of pressure broadening effects is a complex matter. For our purposes, let us consider the case where the time

interval between collisions is small compared to the lifetime of the energy level. Collisional de-excitation of the electron will then shorten the lifetime of the atomic level; referring back to eq. 6.9, Δt is reduced and ΔE is therefore increased.

Pressure broadening also produces a damping profile and the overall effect is an increase in the damping constant $\delta'_k = \delta_k + \delta_p$, where δ_p is the number of collisionally induced transitions per unit time. We can obtain an approximate estimate of δ_p by considering the average time between collisions:

$$\delta_p = \frac{1}{\Delta t} \approx \frac{v}{\mu} = \sqrt{\frac{2kT}{m}} \cdot n\sigma_c, \quad (6.10)$$

where, as we saw earlier, the mean free path is $\mu = 1/n\sigma_c$ (where n is the particle density and σ_c the collisional cross-section), and the most probable speed is $v = \sqrt{2kT/m}$ for a Maxwell-Botzmann distribution.

In cooler stars, where pressure broadening is most important, collisions with neutral hydrogen atoms are the main source of pressure broadening. Note the dependence of δ_p on n . This explains why the absorption lines in giant and supergiant stars (luminosity classes III to Ia in the MK classification scheme) are narrower than in dwarfs (luminosity class V): the extended atmospheres of (super)giants have lower densities than their counterpart main-sequence stars of the same spectral class.

6.3.2 Doppler Broadening

No atoms is at rest. In reality, the absorbing atoms will have a distribution of radial velocities along the line of sight $\Psi(v)$. There can be several contributions to $\Psi(v)$. At the atomic level, the temperature of the gas results in a corresponding Maxwellian distribution of velocities:

$$\Psi(v) = \frac{1}{\sqrt{\pi}b} \exp \left[-\frac{(v - v_0)^2}{b^2} \right] \quad (6.11)$$

where b , sometimes referred to as the Doppler width, is related to the temperature via:

$$b_{\text{th}} = \left(\frac{2kT}{m} \right)^{1/2} \quad (6.12)$$

$b = \sqrt{2}\sigma$ where σ here is the one-dimensional rms velocity projected along the line of sight, while the three-dimensional rms velocity is $v_{\text{rms}} = \sqrt{3} \times (b/\sqrt{2})$.

In addition to motions at the atomic level, there can be bulk motions resulting, for example, from different parts of the stellar disk having different velocities. Large-scale convective motions within the photosphere, such as those responsible for the granulation seen on the surface of the Sun, are an example of such macroturbulence. Such macroturbulence can often also be approximated by a velocity distribution similar to eq. 6.11, with an equivalent b_{turb} . Other types of large-scale motion, such as rotation, cannot; they give characteristically broadened line profiles that have to be treated separately.

Finally, in the analysis of stellar spectra a parameter termed ‘microturbulence’ is also used, generally as a ‘fudge’ to explain the fact that the lines appear broader than expected on the basis of the temperature derived by other means. Physically, this additional turbulence can be understood if, along any sightline from optical depth $\tau_\lambda = 2/3$ to the surface of the star, the photons encounter regions of different velocity, introducing an additional Doppler broadening. The total Doppler broadening is obtained by adding the different b -values in quadrature:

$$b^2 = b_{\text{th}}^2 + b_{\text{turb}}^2 + b_{\text{micro}}^2 \quad (6.13)$$

Including Doppler broadening into the broadening function, we now have:

$$\Phi_\lambda = \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{\delta'_k}{\delta_k'^2 + [\lambda - \lambda_0(1 + \frac{v}{c})]^2} \Psi(v) dv \quad (6.14)$$

Which of these two main sources of broadening is dominant? Generally, the damping constant δ'_k is much smaller than the Doppler constant b (both in wavelength units); for example, for the H α line of neutral hydrogen in the Sun, $\delta_k/b \sim 0.001$. However, the dependence of these two types of broadening functions on wavelength is not the same: the probability of the absorption taking place due to Doppler broadening falls off with wavelength from the line centre faster (exponentially, see eq. 6.11) than that due to natural broadening, which has an inverse-square dependence (eq. 6.7). Thus, absorption lines tend to have a Doppler core and damping

wings. Absorption near the line centre is due mostly to the motions of the atoms, while far from the line centre it is the fundamental uncertainty ΔE in the energy of the atomic levels that allows the absorption to take place.

6.4 The Curve of Growth

The full expression for the line optical depth at wavelength λ is obtained by combining eqs. 6.2, 6.4, 6.7, and 6.11 to give:

$$\tau_\lambda = N \sigma_0 \phi_\lambda \otimes \Psi(v) \quad (6.15)$$

where the convolution of the natural broadening and Doppler broadening functions is as given in eq. 6.14, and N (cm^{-2})

$$N = \int_0^s n ds \quad (6.16)$$

is the column density, which measures the number of absorbers in a cylinder of unit cross-section. The expression 6.15 for τ_λ is often referred to as the Voigt function.

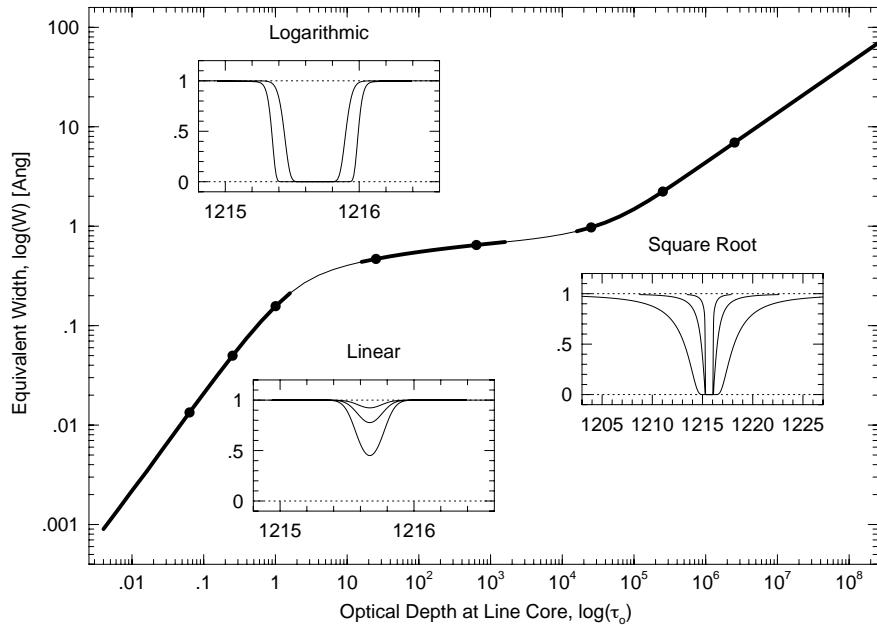


Figure 6.3: Example of a curve of growth. The three regimes discussed in the text, the linear, flat, and damping part of the COG are shown by thicker curves. Corresponding line absorption profiles are shown for each regime and their locations on the COG are marked with filled dots. The wavelength (x -axis) scale in the panel for lines on the damping part of the COG has been expanded relative to the other two panels to illustrate the large extent of damping wings. (Figure courtesy of Chris Churchill).

Integration of 6.3 then gives the sought-after relationship between the equivalent width W_λ of an absorption line and the column density N of absorbing atoms. This relationship, which is illustrated in Fig. 6.3, is known as the Curve of Growth, because it describes how W_λ grows with increasing N .

The precise functional dependence of W_λ on N is sensitive to the optical depth at the line core, τ_0 . Qualitatively, as τ_0 increases from $\tau_0 \ll 1$, the residual intensity I_λ in the line core decreases (the line depth increases) until all the photons at the line core are removed from the incoming beam. At this point, the absorption line is considered to be ‘saturated’. As τ_0 increases further, very little additional light is removed from the beam until we reach a regime where the optical depth becomes significant at wavelengths far from the line centre, where absorption takes place from the natural broadening of the line. The equivalent width can then grow again, as the line develops characteristic ‘damping wings’ [where the absorption coefficient falls off as the inverse square of $(\lambda - \lambda_0)$].

Before proceeding, a note of clarification is required here. The above treatment is strictly only valid for a ‘foreground screen’ of gas illuminated by a background source of light. This situation is an appropriate description of, for example, the interstellar absorption lines seen in the spectra of distant stars and quasars unrelated to the absorbing regions. It is also useful as an approximation to understand the formation of stellar absorption lines. However, recall that at any wavelength we see into a stellar atmosphere from the surface down to $\tau_\lambda = 2/3$. Thus, we see much further into a star in the continuum than in the core of an absorption line. When $\tau_\lambda \gg 1$ in eq. 6.15, and the line profiles intrinsically resemble those shown in the top sub-panel of Figure 6.3, the depth we see into the star approaches zero. But even at the surface of the star, the intensity I_λ is not zero. For LTE, $I_\lambda = B_\lambda \neq 0$, because $T \neq 0$. Thus, in stellar spectra, even strongly saturated absorption lines do not reach down to zero intensity; as a rule, even in the line core, $I_\lambda = I_\lambda(\tau = 2/3)$.

6.4.1 Measuring Column Densities

Returning to Figure 6.3, we can distinguish three portions of the Curve of Growth:

1. **The linear part**, where $\tau_0 < 1$ and $W_\lambda \propto N$. The absorption line is optically thin and W_λ is a sensitive measure of N , irrespectively of the value of the Doppler parameter b .

Using the approximation $e^{-x} \simeq 1 - x$ when $x \ll 1$, it can be readily seen from eq. 6.3 that on the linear part of the COG:

$$W_\lambda = \int_{-\infty}^{\infty} \tau_\lambda d\lambda = N \sigma_0 \int_{-\infty}^{\infty} \Phi_\lambda d\lambda = N \sigma_0 \quad (6.17)$$

since we defined $\int_{-\infty}^{\infty} \Phi_\lambda d\lambda = 1$ (eq. 6.6).

2. **The flat, or logarithmic, part**, where $10 \lesssim \tau_0 \lesssim 10^3$ and $W_\lambda \propto b \sqrt{\ln(N/b)}$. The absorption line is optically thick and W_λ is *not* a good measure of N , but is sensitive to the Doppler parameter b .
3. **The damping, or square root, part**, where $\tau_0 \gtrsim 10^4$ and $W_\lambda \propto \sqrt{N}$. In this regime, the optical depth in the damping wings provides an accurate estimate of N .

Curves of growth tend to be plotted in a variety of ways. Most commonly, they are a plot of W_λ/λ vs. $N f \lambda$, where f is the ‘oscillator strength’ of a transition (available from compilations of atomic data), related to the statistical weights and the transition probability as follows:

$$\lambda g_l f_{lk} = \frac{m_e c \lambda^3}{8\pi^2 e^2} \cdot g_k a_{kl} \quad (6.18)$$

In other applications, COGs are plots of W_λ/λ vs. $A f \lambda$, where A is the abundance of the element of interest relative to H.

6.4.2 Which Curve of Growth?

The transition between the linear and flat parts of the curve of growth occurs at different values of N depending on the Doppler width b , as can be appreciated from inspection of eq. 6.11 and Figure 6.4. In reality, even for a single absorption line, there is a whole family of COGs depending on the value of the Doppler parameter b , which is generally unknown.

Such ambiguity can be resolved by observing many absorption lines formed in the same layer of the stellar atmosphere. Fe is a good example because of the large number of transitions available between its energy levels. The

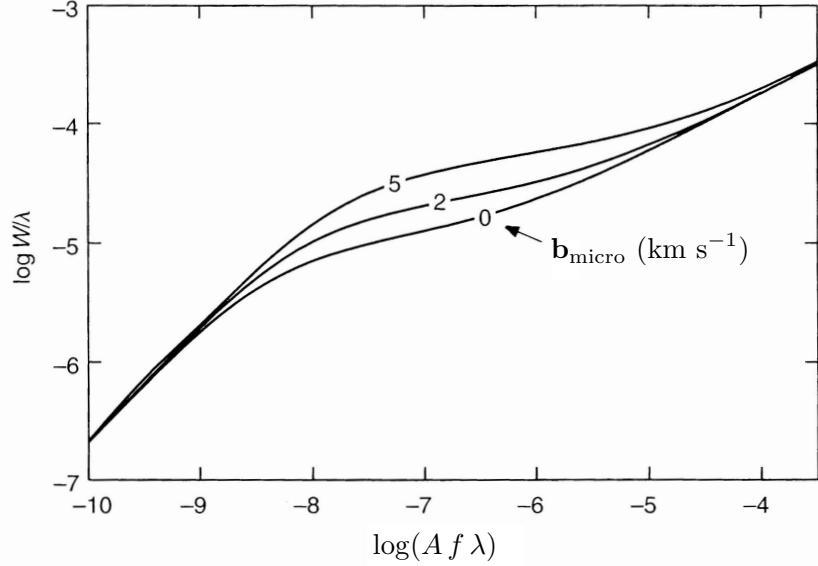


Figure 6.4: Different curves of growth for three values of the microturbulence parameter b_{micro} . At greater values of b , the transition between the linear and flat part of the curve of growth occurs at larger values of W_λ and N .

measured values of W_λ/λ can be fitted with a theoretical COG to determine the equivalent Doppler parameter (which is the sum in quadrature of all the individual b -values, see eq. 6.13), as well as the column density of the ion under consideration, as shown in Figure 6.5.

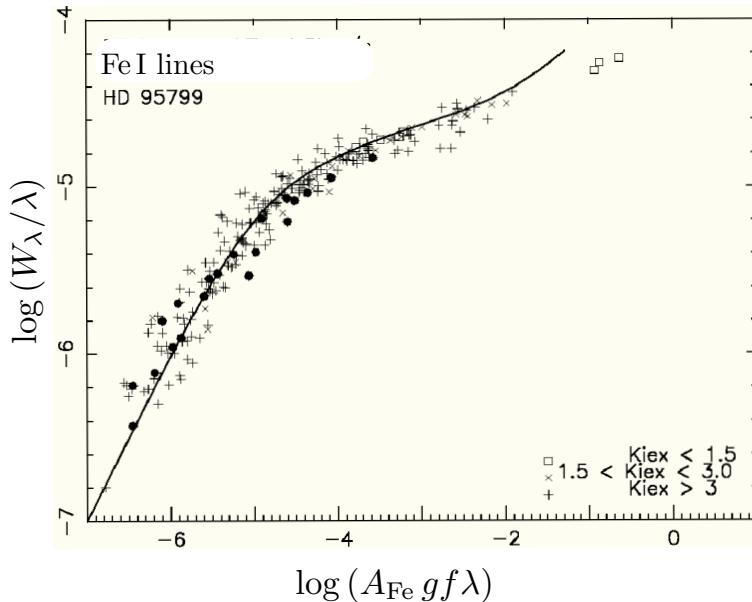


Figure 6.5: Empirical curve of growth for Fe lines in the star HD 95799. Different symbols indicate three ranges of excitation potential for the Fe I lines, while Fe II lines are represented by filled circles. The continuous line is a theoretical curve of growth for $b_{\text{micro}} = 1.5 \text{ km s}^{-1}$ and a solar abundance of Fe: $A_{\text{Fe}} = 3 \times 10^{-5}$ by number.

There can be limitations to the use of COGs, for example if some of the atomic data relevant to the transitions observed are poorly known. Much of the use of COGs is in differential analysis, where the strengths of a set of absorption lines are compared in different stars, to deduce the relative abundances of the elements of interest (an example was shown in Figure 6.1). Similarly, the ratios of the column densities of two (or more) ionisation stages of the same element in a star can be used to infer the ionisation temperature via the Saha equation (eq. 3.4), if the electron density is known. The ratio of the column densities of two excitation stages of the same ion gives the excitation temperature via the Boltzmann equation (eq. 3.1).

6.5 Measurements of Stellar Parameters from Spectral Lines

Nowadays, the analysis of stellar spectra is generally conducted by comparing the observed spectrum with a suite of theoretical ones computed according to a stellar model atmosphere. In such models, each atmospheric layer is involved in the formation of the line profiles and contributes to the final spectrum. Thanks to advances in computational power, models are being developed with increasing levels of sophistication, taking into consideration the full three-dimensional structure of a star and departures from LTE where necessary. Here we give a few examples.

6.5.1 Metal Lines as Temperature Indicators

The sensitivity of metal lines to temperature can vary widely. Generally speaking, lines of neutral metals are most useful in solar-type stars, while lines of ionised metals are better temperature tracers in hotter stars. Many lines can be analysed at once and the curve of growth can be used as a temperature-measuring device by finding the excitation temperature that minimises the scatter in an empirical curve of growth (see Figure 6.5).

Observational precision can be increased by using the ratio of the central depths of two absorption lines when the stellar spectrum has been recorded at high resolution and signal-to-noise ratio, usually with an echelle spec-

trograph. Ideally one would choose two lines from the same element, so as to eliminate any abundance dependence. However, in practice it is advantageous to use pairs of lines at nearby wavelengths, so as to minimise any errors in the continuum normalisation (i.e. in dividing out the underlying stellar continuum radiation against which the absorption appears to take place). In this case, pairs of lines of similar elements, such as Fe, V and Ti, (elements that are not expected to exhibit differential abundance variations) are used, as shown in Figures 6.6 and 6.7.

There are many such pairs of lines in the spectra of cool stars; an example is reproduced in Figure 6.7. Any given pair of lines is limited in the temperature range over which it is a useful measure of T_{eff} . As can be appreciated from the right panel of Figure 6.6, in stars cooler than $T_{\text{eff}} \simeq 3900$ K a different set of lines is needed to measure the temperature.

Application of the line-depth ratio method can yield very high temperature *resolution*, as high as a few K, which is 10–20 times better than the accuracy with which the temperature scale itself can be established.¹ Nevertheless, line-depth ratios are a powerful tool for measuring small temperature differences between stars, or small variations in the temperature of a given star.

¹A good example of a case where the systematic error overwhelms the random error of a physical measurement.

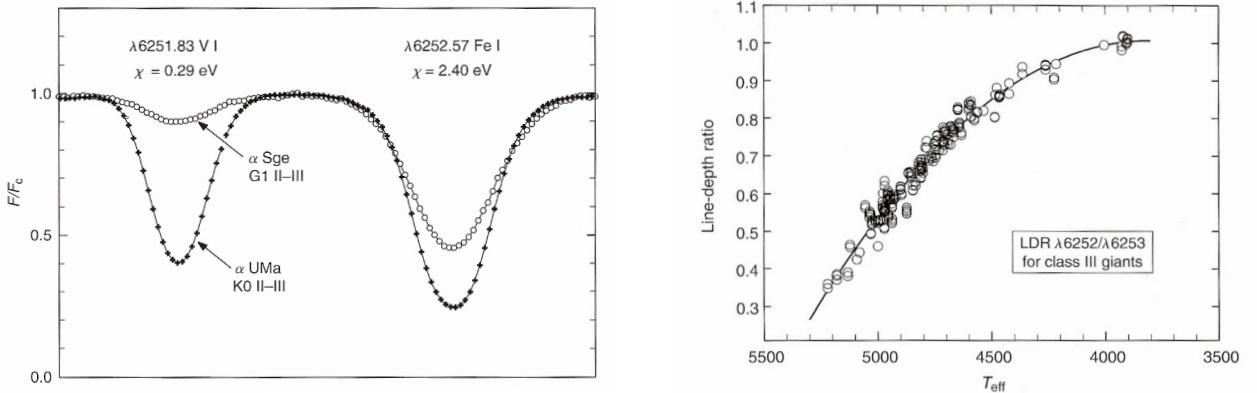


Figure 6.6: *Left:* The ratio of the equivalent widths (or simply the central depths) of two absorption lines originating from atomic levels of different excitation energies, as indicated, can be a sensitive thermometer. *Right:* A typical calibration curve for a line-depth ratio. The slope of the curve reflects the temperature sensitivity of the method. In this case, the line ratio can be used to infer T_{eff} in stars hotter than ~ 4000 K. (Figures reproduced from Gray, D.F., Stellar Photospheres, CUP).

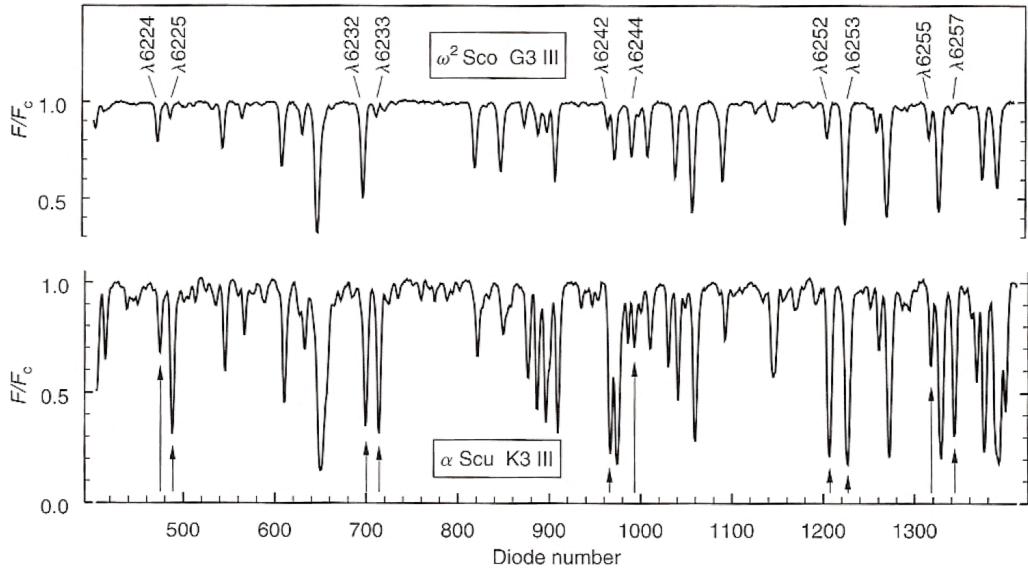


Figure 6.7: Spectral lines having different sensitivity to temperature such as those marked, have been selected in close pairs to avoid continuum setting errors. In these stars, absorption lines of V I vary more strongly than the nearby Fe I lines. (Figure reproduced from Gray, D.F., Stellar Photospheres, CUP).

6.5.2 Spectral Lines as Pressure Indicators

Stars above the main sequence in the HR diagram have higher luminosity and lower photospheric pressure than stars on the main sequence, as a result of their expansion (see Lecture 13). Indeed, pressure effects are often referred to as luminosity effects. Generally speaking, spectroscopic pressure effects are more subtle than temperature effects. On the other hand, while values of T_{eff} span about one order of magnitude between the coolest and the hottest stars, pressure ranges by five orders of magnitude from dwarfs to supergiants. This renders even relatively subtle pressure effects measurable in stellar spectra.

As explained in Section 6.3.1, pressure adds to the natural damping constant to give characteristic broad wings to the absorption lines. From the measurement of these wings, as we shall see in a moment, astronomers deduce the value of *surface gravity*, or $\log g$, where g is the gravitational acceleration (in cm s^{-2}) which would be experienced by a hypothetical test particle of negligible mass at the star's surface. Thus, $\log g = 3$ at the surface of the Earth. Main sequence stars have $\log g \simeq 4.5$, with only a modest dependence on stellar mass. On the other hand, in supergiants $\log g$ varies by three orders of magnitude, from $\log g \simeq 3.5$ in O5 I stars to

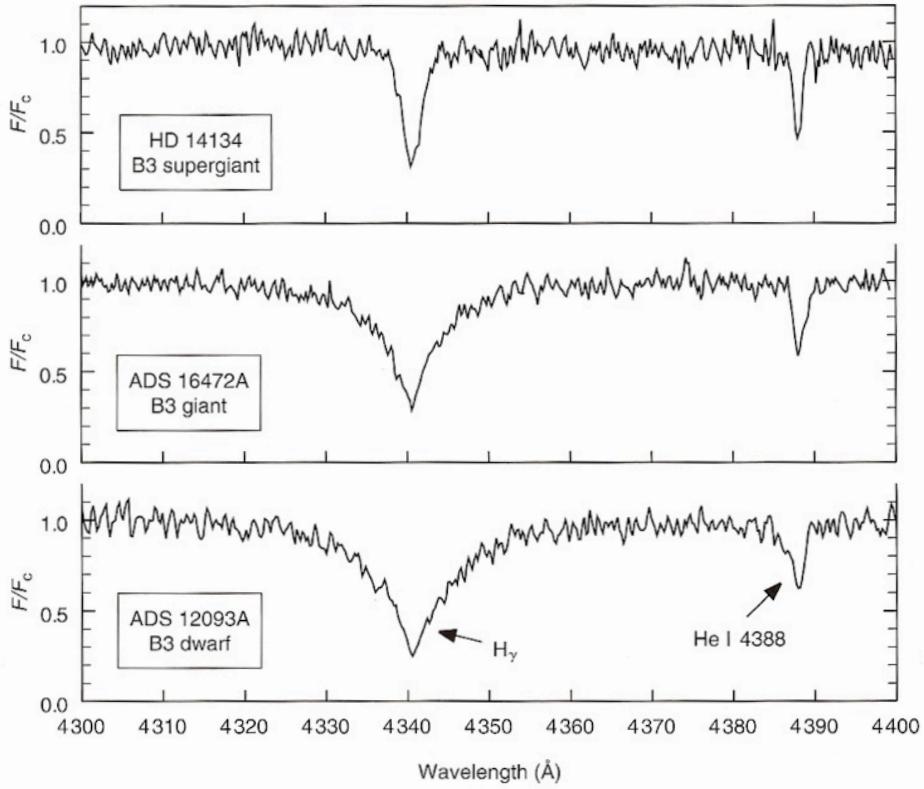


Figure 6.8: Pressure variations in the profile of the $H\gamma$ line in B3 stars of different luminosity classes. The effect arises because of the pressure dependence of the linear Stark effect, the broadening of the $n = 2$ and $n = 5$ energy levels of neutral hydrogen due to presence of an external electric field provided by nearby ions. (Figure reproduced from Gray, D.F., Stellar Photospheres, CUP).

$\log g \simeq 0$ in M2 I red supergiants.

Pressure sensitivity in the hydrogen lines of the Balmer series is one of the classical luminosity indicators in early-type stars (see Figure 6.8). Model profiles are shown in Figure 6.9. The gravity dependence rapidly diminishes

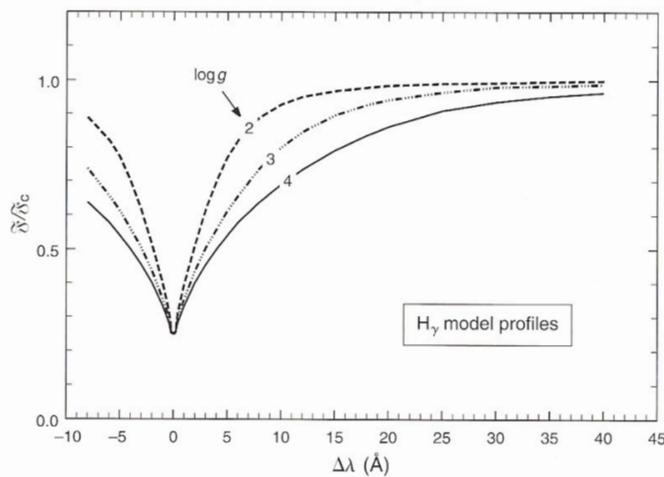


Figure 6.9: Model profiles of the $H\gamma$ line of the Balmer series for different values of surface gravity, as indicated. (Figure reproduced from Gray, D.F., Stellar Photospheres, CUP).

with temperature below $T_{\text{eff}} \simeq 10\,000$ K.

The Balmer lines can be so strong in stars that their equivalent widths can be determined (approximately) from photometric, rather than spectroscopic, measurements. That is, one compares stellar magnitudes measured through two filters of different bandpasses, both centred on the same Balmer line. The most widely used such index is the $H\beta$ index, which measures the magnitude difference between two filters, each centred at the wavelength of the $H\beta$ line, with half-width bandpasses of 29 Å and 129 Å respectively. The $H\beta$ index is the gravity discriminant in the *uvby* β Strömgren photometric system.

Some strong metal lines also show pressure-broadened wings in the spectra of cooler stars; an example is reproduced in the left panel of Figure 6.10. Since pressure diagnostics are always temperature sensitive, it is often useful to make a simultaneous solution for T_{eff} and $\log g$. This can be achieved if we have two lines that have different responses to the two variables, as in the example shown in the right panel of Figure 6.10. Each curve in this gravity-temperature diagram is computed for a fixed Fe abundance, while varying $\log g$ at a given T_{eff} (or vice versa) to recover the observed equivalent width. The crossing point is then the solution, i.e. the correct

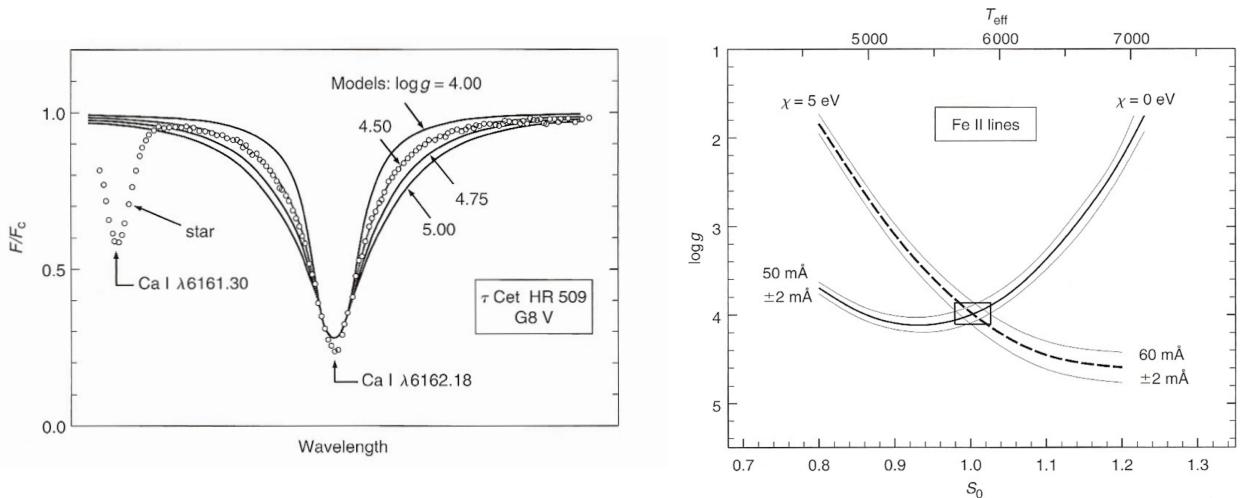


Figure 6.10: *Left:* Fitting theoretical line profiles to the $\text{Ca I} \lambda 6162$ line recorded in the spectrum of the G8 V star τ Ceti gives a measure of the surface gravity: $\log g = 4.50$ *Right:* Gravity-Temperature diagram for the equivalent widths of two Fe II absorption lines of different excitation potential, as indicated. Measurement errors are shown by lighter lines. The solution is localised to the small rectangle where the lines cross. (Figures reproduced from Gray, D.F., Stellar Photospheres, CUP).

effective temperature and surface gravity. The uncertainties in the two parameters can be estimated by mapping the errors ΔW_λ into ranges ΔT_{eff} and $\Delta \log g$, as shown by the rectangle in the Figure.

A number of diagnostics can be placed on the same $\log g - T_{\text{eff}}$ diagram and, in theory, all loci should intersect at the same point within the errors ΔW_λ . The real case differs from the ideal case because of inadequate model stellar atmospheres, incomplete understanding of line formation, systematic errors in the data, and other second-order effects.

6.5.3 Abundance Determinations from Metal Absorption Lines

Once the effective temperature and gravity of a star have been determined by rigorous analysis of absorption lines that are sensitive to these two

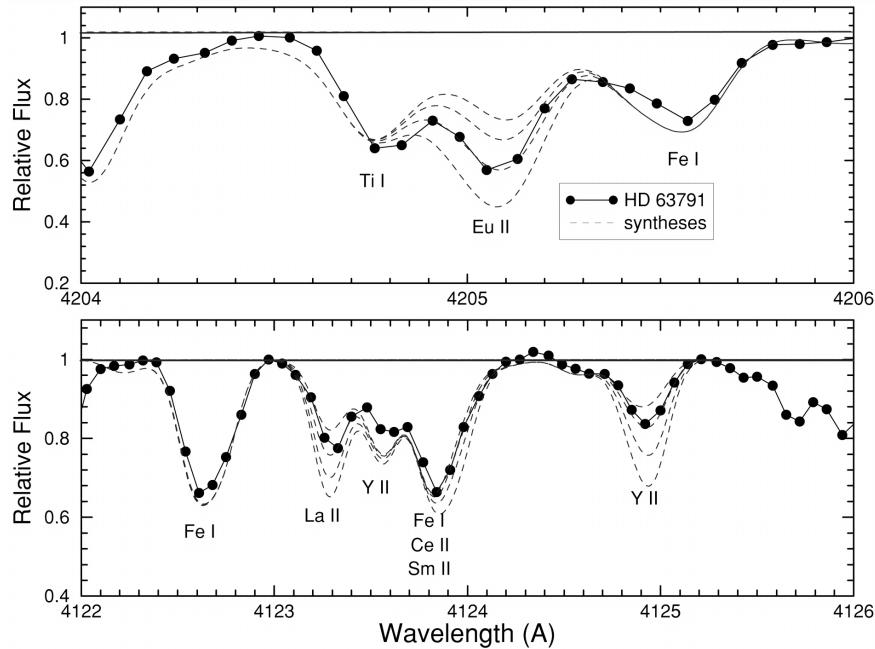


Figure 6.11: Example of spectral synthesis from which element abundances are deduced. The observed spectrum is represented by filled circles, while the lines show model spectra. All the model spectra have the same effective temperature, $T_{\text{eff}} = 4725$ K, the same surface gravity, $\log g = 1.70$, and the same value of microturbulence, $\xi = 1.60 \text{ km s}^{-1}$, but differ in the values of the abundances of the elements of interest, in this case the neutron-capture elements Eu, La, and Y. The solid line is deemed to fit best. The spectrum used for this abundance analysis has resolving power $R \equiv \lambda/\Delta\lambda \simeq 20\,000$ and signal-to-noise ratio $S/N \simeq 100$; such high values are typical of the quality of spectra required for stellar abundance studies. (Figure reproduced from Burris et al. 2000, ApJ, 544, 302).

parameters, one is left with only one free parameter: the abundance of the element under consideration. By comparing the observed spectrum with model spectra computed for a range of values of, for example, Fe/H, the abundance of iron in the stellar photosphere can be deduced. (see Figure 6.11 for an example). Abundances so derived are usually referred to the solar abundance scale, since it is a prerequisite of the models that they reproduce the solar spectrum.

In the Milky Way, stellar metallicities vary from at most twice solar to $\sim 10^{-5.5}$ of solar; Figure 6.12 provides a vivid example of the weakening of stellar absorption lines with decreasing metallicity.

The solar metallicity scale is not without its problems! Some elements have proved difficult to measure accurately, and their abundances have undergone revisions, sometimes by non-negligible factors, in the last 10-20 years. Most people today adopt the Asplund et al. (2009) abundance scale reproduced in Figure 6.13. Note that, for historical reasons, element abundances are normally quoted by number (rather than by mass) on a log scale where H has value 12 (don't blame the messenger or, in this case, your lecturer!).

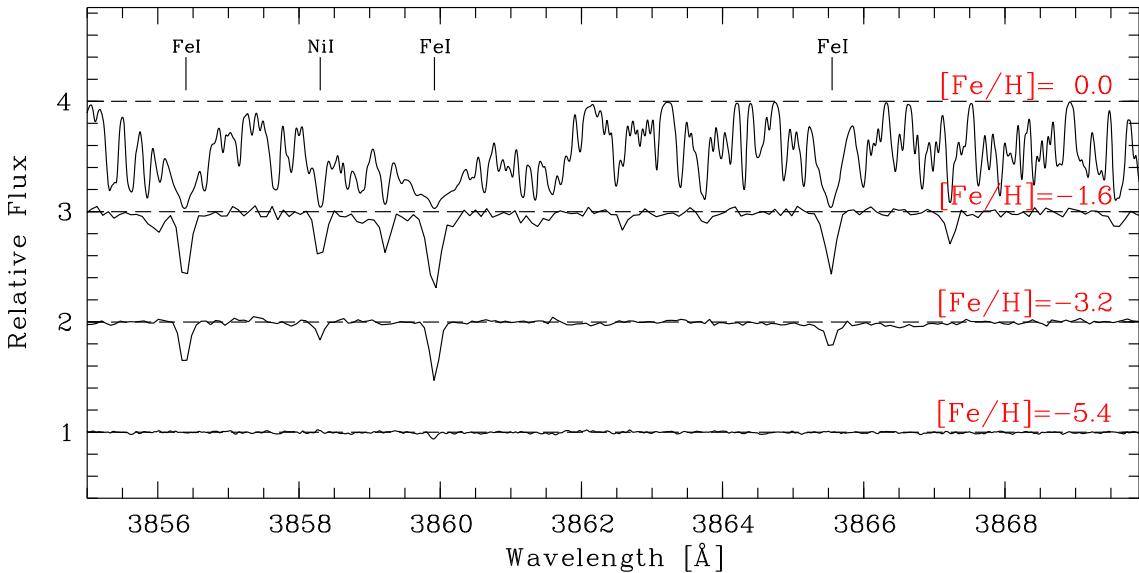


Figure 6.12: Spectral comparison of stars in the main-sequence turn-off region with different metallicities. Several absorption lines are marked. The variations in line strength reflect the different metallicities. From top to bottom: Sun with $[Fe/H] = 0.0$; G66-30 with $[Fe/H] = -1.6$; G64-12 with $[Fe/H] = -3.2$; and HE 1327–2326 with $[Fe/H] = -5.4$. (Figure reproduced from Frebel, A. 2010, Astronomische Nachrichten, 331, 474).

As an example, in the photosphere of the star HE 1327–2326 $[\text{Fe}/\text{H}] = -5.4$ (Figure 6.12). Recalling our definition: $[\text{Fe}/\text{H}] \equiv \log(\text{Fe}/\text{H})_{\text{star}} - \log(\text{Fe}/\text{H})_{\odot}$, we have $\log(\text{Fe}/\text{H})_{\text{star}} = 7.50 - 12 - 5.4 = -9.9$. In other words, in the photosphere of this very metal-poor star there is only 1 atom of Fe for every 10 billion atoms of H!

Table 1: Element abundances in the present-day solar photosphere. Also given are the corresponding values for CI carbonaceous chondrites (Lodders, Palme & Gail 2009). Indirect photospheric estimates have been used for the noble gases (Sect. 3.9).

Elem.	Photosphere	Meteorites	Elem.	Photosphere	Meteorites
1 H	12.00	8.22 ± 0.04	44 Ru	1.75 ± 0.08	1.76 ± 0.03
2 He	$[10.93 \pm 0.01]$	1.29	45 Rh	0.91 ± 0.10	1.06 ± 0.04
3 Li	1.05 ± 0.10	3.26 ± 0.05	46 Pd	1.57 ± 0.10	1.65 ± 0.02
4 Be	1.38 ± 0.09	1.30 ± 0.03	47 Ag	0.94 ± 0.10	1.20 ± 0.02
5 B	2.70 ± 0.20	2.79 ± 0.04	48 Cd		1.71 ± 0.03
6 C	8.43 ± 0.05	7.39 ± 0.04	49 In	0.80 ± 0.20	0.76 ± 0.03
7 N	7.83 ± 0.05	6.26 ± 0.06	50 Sn	2.04 ± 0.10	2.07 ± 0.06
8 O	8.69 ± 0.05	8.40 ± 0.04	51 Sb		1.01 ± 0.06
9 F	4.56 ± 0.30	4.42 ± 0.06	52 Te		2.18 ± 0.03
10 Ne	$[7.93 \pm 0.10]$	-1.12	53 I		1.55 ± 0.08
11 Na	6.24 ± 0.04	6.27 ± 0.02	54 Xe	$[2.24 \pm 0.06]$	-1.95
12 Mg	7.60 ± 0.04	7.53 ± 0.01	55 Cs		1.08 ± 0.02
13 Al	6.45 ± 0.03	6.43 ± 0.01	56 Ba	2.18 ± 0.09	2.18 ± 0.03
14 Si	7.51 ± 0.03	7.51 ± 0.01	57 La	1.10 ± 0.04	1.17 ± 0.02
15 P	5.41 ± 0.03	5.43 ± 0.04	58 Ce	1.58 ± 0.04	1.58 ± 0.02
16 S	7.12 ± 0.03	7.15 ± 0.02	59 Pr	0.72 ± 0.04	0.76 ± 0.03
17 Cl	5.50 ± 0.30	5.23 ± 0.06	60 Nd	1.42 ± 0.04	1.45 ± 0.02
18 Ar	$[6.40 \pm 0.13]$	-0.50	62 Sm	0.96 ± 0.04	0.94 ± 0.02
19 K	5.03 ± 0.09	5.08 ± 0.02	63 Eu	0.52 ± 0.04	0.51 ± 0.02
20 Ca	6.34 ± 0.04	6.29 ± 0.02	64 Gd	1.07 ± 0.04	1.05 ± 0.02
21 Sc	3.15 ± 0.04	3.05 ± 0.02	65 Tb	0.30 ± 0.10	0.32 ± 0.03
22 Ti	4.95 ± 0.05	4.91 ± 0.03	66 Dy	1.10 ± 0.04	1.13 ± 0.02
23 V	3.93 ± 0.08	3.96 ± 0.02	67 Ho	0.48 ± 0.11	0.47 ± 0.03
24 Cr	5.64 ± 0.04	5.64 ± 0.01	68 Er	0.92 ± 0.05	0.92 ± 0.02
25 Mn	5.43 ± 0.05	5.48 ± 0.01	69 Tm	0.10 ± 0.04	0.12 ± 0.03
26 Fe	7.50 ± 0.04	7.45 ± 0.01	70 Yb	0.84 ± 0.11	0.92 ± 0.02
27 Co	4.99 ± 0.07	4.87 ± 0.01	71 Lu	0.10 ± 0.09	0.09 ± 0.02
28 Ni	6.22 ± 0.04	6.20 ± 0.01	72 Hf	0.85 ± 0.04	0.71 ± 0.02
29 Cu	4.19 ± 0.04	4.25 ± 0.04	73 Ta		-0.12 ± 0.04
30 Zn	4.56 ± 0.05	4.63 ± 0.04	74 W	0.85 ± 0.12	0.65 ± 0.04
31 Ga	3.04 ± 0.09	3.08 ± 0.02	75 Re		0.26 ± 0.04
32 Ge	3.65 ± 0.10	3.58 ± 0.04	76 Os	1.40 ± 0.08	1.35 ± 0.03
33 As		2.30 ± 0.04	77 Ir	1.38 ± 0.07	1.32 ± 0.02
34 Se		3.34 ± 0.03	78 Pt		1.62 ± 0.03
35 Br		2.54 ± 0.06	79 Au	0.92 ± 0.10	0.80 ± 0.04
36 Kr	$[3.25 \pm 0.06]$	-2.27	80 Hg		1.17 ± 0.08
37 Rb	2.52 ± 0.10	2.36 ± 0.03	81 Tl	0.90 ± 0.20	0.77 ± 0.03
38 Sr	2.87 ± 0.07	2.88 ± 0.03	82 Pb	1.75 ± 0.10	2.04 ± 0.03
39 Y	2.21 ± 0.05	2.17 ± 0.04	83 Bi		0.65 ± 0.04
40 Zr	2.58 ± 0.04	2.53 ± 0.04	90 Th	0.02 ± 0.10	0.06 ± 0.03
41 Nb	1.46 ± 0.04	1.41 ± 0.04	92 U		-0.54 ± 0.03
42 Mo	1.88 ± 0.08	1.94 ± 0.04			

Figure 6.13: The most recent compilation of solar element abundances by Asplund, Grevesse, Sauval, & Scott, P. 2009, ARAA, 47, 481.

WHAT MAKES A STAR SHINE?

In this lecture, we are going to look in more detail at some of the ideas outlined in Lecture 1, section 1.3. What are the sources of energy available to a star?

7.1 Gravitational Potential Energy

A potential source of energy to a newly born star is the gravitational energy released when the interstellar clouds from which the star formed began contracting. The virial theorem for a system in equilibrium:

$$-2\langle K \rangle = \langle U \rangle \quad (7.1)$$

where K and U are the kinetic and potential energy respectively and the brackets denote time averages, tells us that:

$$\langle E \rangle = \langle K \rangle + \langle U \rangle = \frac{1}{2}\langle U \rangle \quad (7.2)$$

only half the change in gravitational potential energy is available to be radiated away as the protostar contracts; the remaining potential energy supplies the thermal energy that heats the gas.

Referring to Figure 7.1, the gravitational force acting on a test particle of mass dm_i at a distance r from the centre of a spherically symmetric mass distribution is

$$dF_{g,i} = G \frac{M_r dm_i}{r^2} \quad (7.3)$$

where M_r is the mass contained within r , which acts as a point mass M located at the centre of the sphere, and the force is directed towards the centre of the sphere. The corresponding gravitational potential energy of the test mass is:

$$dU_{g,i} = -G \frac{M_r dm_i}{r} \quad (7.4)$$

Summing all the mass within a thin shell of thickness dr , with total mass

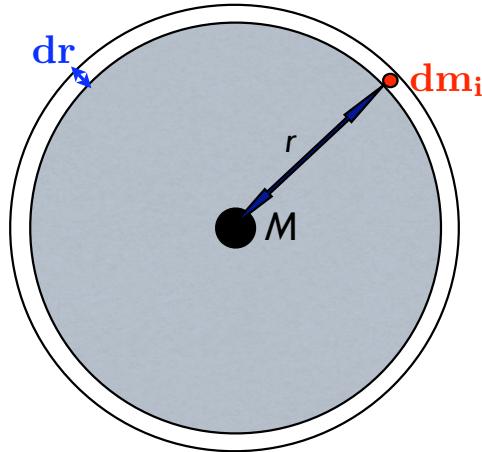


Figure 7.1: Gravitational potential energy of a star.

$dm = 4\pi r^2 dr \rho$ where ρ is the density (mass per unit volume, e.g. g cm⁻³) and $4\pi r^2 dr$ is the volume of the shell, we can rewrite eq. 7.4 as:

$$dU_g = -G \frac{M_r 4\pi r^2 \rho}{r} dr \quad (7.5)$$

To obtain the total gravitational potential energy of a star, we integrate 7.5 over all mass shells from the centre of the star to its surface:

$$U_g = -4\pi G \int_0^R M_r \rho r dr \quad (7.6)$$

To properly evaluate the integral, we need to know $M_r = f(r)$ and $\rho = f(r)$. As an approximation, let us assume that an average density $\langle \rho \rangle$ obtains everywhere within the star; then:

$$M_r \sim \frac{4}{3}\pi r^3 \langle \rho \rangle, \quad (7.7)$$

and the mass of a star of radius R is $M = 4/3\pi R^3 \langle \rho \rangle$. Substituting 7.7 into 7.6 and integrating:

$$U_g \sim -\frac{16\pi^2}{15} G \langle \rho \rangle^2 R^5 = -\frac{3GM^2}{5R}. \quad (7.8)$$

As we saw, applying the virial theorem, tells us that half of this energy could have been radiated away as a protostar contracted from R_{initial} to R_{final} , where $R_{\text{initial}} \gg R_{\text{final}}$.

For the Sun:

$$\Delta E_g = \frac{3}{10} \frac{GM_\odot^2}{R_\odot} = \frac{3}{10} G \frac{(2 \times 10^{33})^2}{7 \times 10^{10}} \frac{\text{cm}^3 \text{g}^{-1} \text{s}^{-2} \text{g}^2}{\text{cm}} \quad \frac{\text{cm}^3 \text{g}^{-1} \text{s}^{-2} \text{g}^2}{\text{cm}}$$

$$\Delta E_g = \frac{3}{10} 6.7 \times 10^{-8} \frac{4 \times 10^{66}}{7 \times 10^{10}} \text{ g cm}^2 \text{ s}^{-2}$$

$$\Delta E_g = 1.2 \times 10^{48} \text{ erg}$$

If the Sun radiated this energy on average at its present luminosity, the corresponding timescale, known as the *Kelvin-Helmholtz timescale*, would be:

$$\tau_{\text{KH}} = \frac{\Delta E_g}{L_\odot} = \frac{1.2 \times 10^{48}}{3.8 \times 10^{33}} \frac{\text{erg}}{\text{erg s}^{-1}}$$

$$\tau_{\text{KH}} \sim 3 \times 10^{14} \text{ s} \simeq 1 \times 10^7 \text{ yr}$$

The Kelvin-Helmholtz timescale is two orders of magnitude smaller than the age of the solar system, as measured by radioactive dating techniques applied to Moon rocks, for instance. Thus, the Sun must now be shining by a different mechanism than converting gravitational potential energy into radiation. Nevertheless, during its initial contraction from a protostellar cloud, the release of gravitational energy must have played an important role.

7.1.1 Average Temperature of the Sun

We can also use the virial theorem to roughly estimate the average temperature of a star. From eqs. 7.1 and 7.8, we have:

$$\langle K \rangle = \frac{3}{10} \frac{GM^2}{R}. \quad (7.9)$$

We now want to know the average temperature $\langle T \rangle$ which corresponds to the average kinetic energy $\langle K \rangle$. We obtain this by recalling that the distribution of particle velocities as a function of T is given by the Maxwell-Boltzmann distribution:

$$f(v) = \left(\frac{2}{\pi}\right)^{1/2} \left(\frac{m}{kT}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2kT}\right). \quad (7.10)$$

We obtain the most probable speed, v_p , by imposing the condition $df(v)/dv = 0$, which gives:

$$v_p = \left(\frac{2kT}{m}\right)^{1/2},$$

and the rms speed:

$$v_{\text{rms}} = \left(\int_0^\infty v^2 f(v) dv \right)^{1/2} = \left(\frac{3kT}{m} \right)^{1/2} = \sqrt{\frac{3}{2}} v_p$$

Thus, the average kinetic energy per particle is:

$$\frac{1}{2} m_p v_{\text{rms}}^2 = \frac{3}{2} k \langle T \rangle \quad (7.11)$$

The total number of particles in a star of mass M is M/m_p ; thus we have:

$$\frac{3}{2} k \langle T \rangle \frac{M}{m_p} = \frac{3}{10} \frac{GM^2}{R}. \quad (7.12)$$

Solving for the average temperature:

$$\langle T \rangle = \frac{1}{5} \frac{GMm_p}{kR}. \quad (7.13)$$

Note the dependence on the stellar mass and radius. Let us work out $\langle T_\odot \rangle$:

$$\begin{aligned} \langle T_\odot \rangle &= \frac{1}{5} \frac{6.7 \times 10^{-8} \cdot 2 \times 10^{33} \cdot 1.7 \times 10^{-24}}{1.4 \times 10^{-16} \cdot 7 \times 10^{10}} \frac{\text{cm}^3 \text{g}^{-1} \text{s}^{-2} \text{g g}}{\text{erg K}^{-1} \text{cm}} \\ \langle T_\odot \rangle &\simeq \frac{0.5 \times 10^1}{1 \times 10^{-6}} \frac{\text{cm}^3 \text{g}^{-1} \text{s}^{-2} \text{g g}}{\text{g cm}^2 \text{s}^{-2} \text{K}^{-1} \text{cm}} \\ \langle T_\odot \rangle &\simeq 5 \times 10^6 \text{ K}. \end{aligned}$$

7.2 Nuclear Fusion

Returning to the question of the main source of radiant energy in stars, another possibility which is readily discounted is chemical reactions. Reactions between ions and atoms involve exchanges of electrons between different energy levels. Since the energies of most electronic levels are of order 1–10 eV, again they provide insufficient energy to power a star like the Sun for $\sim 10^{10}$ years.

On the other hand, the energies that bind protons and neutrons within atomic nuclei are a million times larger, of order MeV. For example, if we wanted to break up the He nucleus (also sometime referred to as an alpha particle) into its constituents two protons and two neutrons, we would need

to supply ~ 27 MeV. It follows that the inverse reaction, whereby four hydrogen nuclei fuse together to form a He nucleus (plus a number of low mass remnants) will release ~ 27 MeV. This energy is the binding energy of the He nucleus, and manifests itself as the mass difference between four H nuclei (4.03 unified atomic mass units or 4.03 u) and one He nucleus (4.00 u). This mass difference $\Delta m = 0.029$ u, or $\sim 0.7\%$ of the rest mass of four H nuclei, corresponds to an energy $E = \Delta m c^2 = 26.7$ MeV, which is released by the nuclear fusion process.

Even if only 10% of the Sun's mass is converted from H into He, the energy generated would be:

$$E = 0.1 M_\odot \cdot 0.007 \cdot c^2 = 7 \times 10^{-4} \cdot 2 \times 10^{33} \cdot (3 \times 10^{10})^2 \quad \text{g} \cdot \text{cm}^2 \text{s}^{-2}$$

$$E = 1.26 \times 10^{2-4+33+20} = 1.26 \times 10^{51} \text{ erg}$$

Thus the nuclear timescale $\tau_{\text{nuclear}} \sim 10^3 \times \tau_{\text{KH}}$, or $\tau_{\text{nuclear}} \sim 10^{10}$ years, consistent with the age of the solar system.

The process of nuclear fusion is not limited to $4 \frac{1}{1}\text{H} \rightarrow 1 \frac{4}{2}\text{He}$ (using the standard notation ${}_Z^A\text{X}$, where A is the mass number (total number of protons + neutrons), Z is the atomic number (i.e. the number of protons, giving the total positive charge of the nucleus in units of e), and X is the chemical symbol for the element under consideration). Three $\frac{4}{2}\text{He}$ nuclei can fuse to give $\frac{12}{6}\text{C}$, in what is known as the triple alpha process. The addition of an alpha particle to $\frac{12}{6}\text{C}$ produces $\frac{16}{8}\text{O}$, and the subsequent incorporation of additional alpha particles into the nucleus produces the so called alpha-capture elements, Ne, Mg, Si, S and so on.

The process can continue so long as it is exothermic, that is so long as the mean mass per nucleon of the final fusion product is lower than that of the fusing nuclei. The binding energy per nucleon curve (see Figure 7.2) reaches a peak near Fe (hence the term Fe-peak). Fusion of Fe-peak elements to form elements of higher mass is an endothermic process requiring an additional supply of energy (and conversely, energy can be released by the fission of these heavier nuclei into lighter ones, as in a nuclear power station). The astrophysical production of these heavier elements is thought to occur during late stages of stellar evolution.

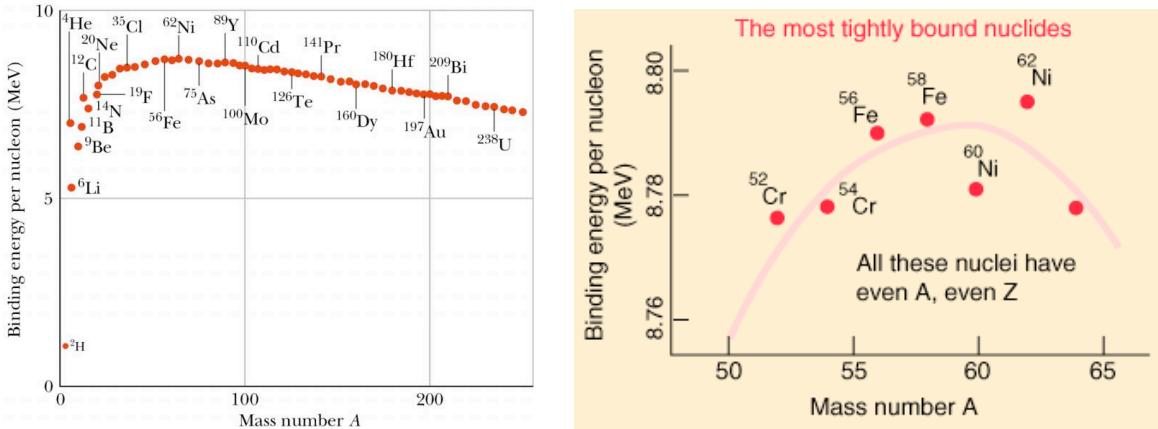


Figure 7.2: Left: Binding energy per nucleon as a function of mass number A . Right: Close-up near the iron peak. ^{62}Ni has the highest binding energy per nucleon of any isotope of any element.

7.2.1 Quantum Mechanical Tunneling

The nucleosynthesis of the common elements of the Periodic Table in stars is a good example of the fundamental relevance of quantum mechanics to our everyday world. In order for two nuclei to fuse, they must come sufficiently close together for the *short-range* strong nuclear force to overcome the Coulomb repulsion between the two positively charged particles. In the classical description, this would require kinetic energies corresponding to temperatures much higher than those found even in the interiors of stars. Only in the quantum-mechanical description of the process, can the reaction take place at stellar temperatures.

We can estimate the temperature required in the classical treatment by equating the mean velocity between two nuclei to the potential energy of the Coulomb barrier at the turn-around point (see Figure 7.3):

$$\frac{1}{2}\mu_m v_{\text{rms}}^2 = \frac{3}{2}kT_{\text{class}} = \frac{1}{4\pi\epsilon_0} \frac{Z_1 Z_2 e^2}{r} \quad (7.14)$$

where ϵ_0 is the permittivity of free space [the constant that relates electric charge to mechanical quantities, defined by Coulombs law which gives the force between two electric charges separated by a distance r as $F_C = (1/4\pi\epsilon_0)(q_1 q_2/r^2)$], $Z_i e$ is the charge of each nucleus, and μ_m is the reduced

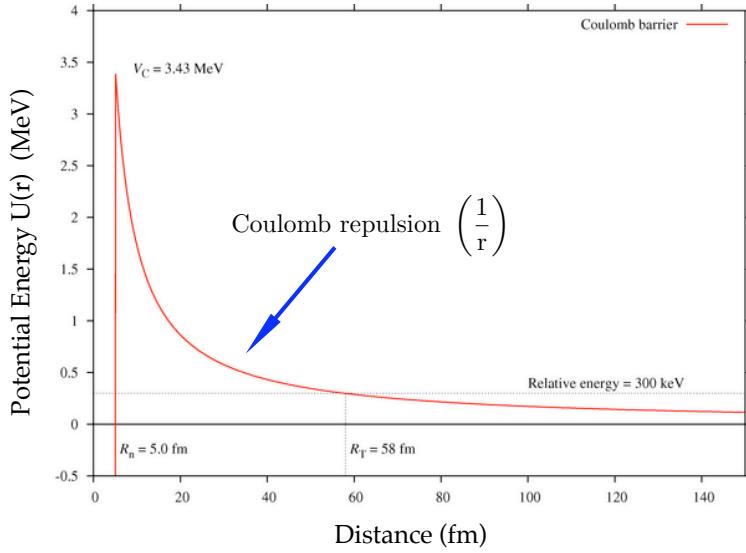


Figure 7.3: The Coulomb barrier classically prevents low-energy particles from approaching each other. The nuclear potential is represented here as a square well. In this example, the values shown are those appropriate to the $^{12}_6\text{C} + ^4_2\text{He} \rightarrow ^{16}_8\text{O}$ fusion reaction. The relative energy of ~ 300 keV corresponds to the Gamow-peak energy described below.

mass of the two colliding particles.¹ Thus,

$$T_{\text{class}} = \frac{Z_1 Z_2 e^2}{6\pi\epsilon_0 k r} \sim 10^{10} \text{ K} \quad (7.15)$$

if $r \sim 10^{-13}$ cm (1 fm), the typical nuclear radius. The value of $\langle T_{\text{class}} \rangle$ is three orders of magnitude higher than the central temperature of the Sun; even allowing for the high velocity tail of a Maxwell-Boltzmann distribution, an insufficient number of particles would penetrate the Coulomb barrier in classical physics to power the Sun's luminosity.

Quantum-mechanically, however, the two nuclei may find themselves within range of the strong nuclear force even if their kinetic energy is insufficient to overcome the Coulomb barrier, due to the inherent uncertainties in their positions and momenta given by Heisenberg uncertainty principle:

$$\Delta p_x \Delta x \geq \frac{\hbar}{2}.$$

The distance of interest here is the de Broglie wavelength associated with a particle, $\lambda = h/p$. Rewriting the mean kinetic energy in terms of the momentum:

$$\frac{1}{2}\mu_m v_{\text{rms}}^2 = \frac{p^2}{2\mu_m}$$

¹We have already encountered the concept of reduced mass for a two-body problem in Lecture 4, where it was defined as $\mu = m_1 \cdot m_2 / (m_1 + m_2)$.

we can equate potential and kinetic energies at $r = \lambda$:

$$\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 \lambda} = \frac{p^2}{2\mu_m} = \frac{(h/\lambda)^2}{2\mu_m} \quad (7.16)$$

or

$$\lambda = \frac{2\pi\epsilon_0 h^2}{\mu_m Z_1 Z_2 e^2}. \quad (7.17)$$

Substituting λ for r in eq. 7.15, we find:

$$T_{\text{quantum}} = \frac{Z_1^2 Z_2^2 e^4 \mu_m}{12\pi^2 \epsilon_0^2 h^2 k}. \quad (7.18)$$

In the case of two protons, $T_{\text{quantum}} \simeq 10^7$ K, comparable to the temperature in the core of the Sun.

7.3 Nuclear Reaction Rates

While nuclear fusion reactions can in principle supply the energy that makes the stars shine, we still need to consider quantitatively the rate at which such reactions occur in stellar interiors. In general terms, the rate at which a given nuclear reaction will proceed (number of reactions per unit volume per unit time) will depend on: (i) the volume density of the particles involved, (ii) their energy distribution (which will determine how closely two particles can approach each other), and (iii) the cross-section, or probability, of interaction.

With the Maxwell-Boltzmann distribution in energy terms:

$$n_E dE = \frac{2n}{\pi^{1/2}} \frac{1}{(kT)^{3/2}} E^{1/2} \exp[-\frac{E}{kT}] dE, \quad (7.19)$$

the reaction rate defined as above can be written as:

$$r_{it} = \int_0^\infty n_i n_t \sigma(E) v(E) \frac{n_E}{n} dE, \quad (7.20)$$

where n_i and n_t are the volume densities of incident and target particles respectively, $v(E)$ is the particle velocity, and $\sigma(E)$ is the cross-section for the interaction.

In order to evaluate this integral, we need to know the functional form of $\sigma(E)$. A great deal of effort is devoted to the calculation and laboratory

measurement of nuclear cross-sections, which can have quite complicated functional forms. But we can still make some general considerations about the overall behaviour of $\sigma(E)$ and r_{it} .

We saw that the quantum-mechanical size of a particle, its de Broglie wavelength, is inversely proportional to its momentum, $\lambda = h/p$. Thus, the cross-sectional *area* for interaction may reasonably be expected to have a $\sigma(E) \propto 1/E$ dependence. But the probability of quantum tunneling is also related to the ratio of the Coulomb barrier potential energy to the particle kinetic energy. This second factor gives rise to an exponential dependence of the cross-section on the energy. Taken together, these effects give rise to the traditional form of the cross-section:

$$\sigma(E) = \frac{S(E)}{E} \exp[-bE^{-1/2}]. \quad (7.21)$$

where $S(E)$ is a slowly varying function of E . Combining 7.21 and 7.20, we find:

$$r_{it} = \left(\frac{2}{kT}\right)^{3/2} \frac{n_i n_t}{(\mu_m \pi)^{1/2}} \int_0^\infty S(E) \exp[-bE^{-1/2}] \exp[-\frac{E}{kT}] dE, \quad (7.22)$$

where b is a constant that includes the charge and the mass of the particles:

$$b = \frac{\pi \mu_m^{1/2} Z_1 Z_2 e^2}{\sqrt{2} \epsilon_0 h}$$

The important thing to appreciate is that there are two competing energy dependencies in the integral at 7.22. The $\exp[-E/kT]$ is the decreasing number of particles with high energies in the tail of the Maxwellian distribution (eq. 7.19). On the other hand, the probability of tunnelling through the Coulomb barrier increases at high energies as indicated by the term $\exp[-bE^{-1/2}]$. The net result of combining these two functions is a strongly peaked dependence of the reaction rate on energy, as shown in Figure 7.4. Thus, the greatest contribution to the reaction rate integral is from a relatively narrow energy range that depends on the temperature of the gas, and the charges and masses of the nuclei involved. By setting $dr_{it}/dE = 0$, we find that the peak of the Gamow curve (so named from the Russian physicist George Gamow who for a period in the late 1920s and early 1930s worked with Rutherford at the Cavendish Laboratory) occurs at energy:

$$E_0 = \left(\frac{bkT}{2}\right)^{2/3}$$

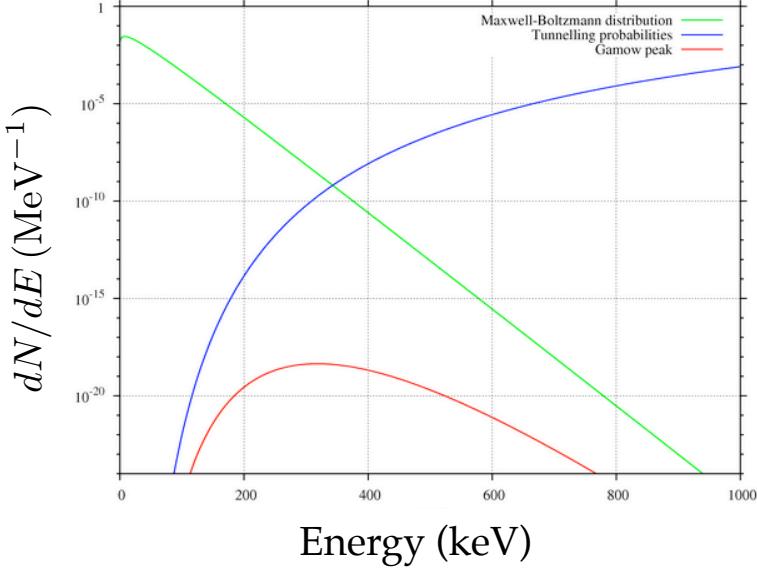


Figure 7.4: The red curve is the product of the Maxwell-Boltzmann distribution (number of particles per unit energy interval, shown in green) with the tunnelling probability of the nuclei through their Coulomb barrier (dimensionless, shown in blue). The curve is strongly peaked: this is the energy range over which the nuclear reaction is most likely to take place. At higher energies the number of particles becomes insignificant, while at lower energies the tunnelling through the Coulomb barrier becomes improbable. This Figure refers to the ${}^1_6\text{C} + {}^4_2\text{He} \rightarrow {}^{16}_8\text{O}$ fusion reaction at $T = 2 \times 10^8 \text{ K}$; the Gamow peak is at an energy $E \simeq 300 \text{ keV}$, much larger than $\langle E \rangle = 3kT/2 = 26 \text{ keV}$.

Before moving on to consider the nuclear reaction rates which are most important for stellar nucleosynthesis, we mention two effects which add to the above treatment of the reaction rates. First, the cross sections of some nuclear reactions exhibit resonances—small energy intervals where the likelihood of interaction is boosted significantly by energy levels within the nucleus. Second, at the high temperatures of stellar interiors, the high densities of free electrons can partially shield the positive nuclear charge. The net effect of this electron screening is to reduce the Coulomb barrier, thereby enhancing the reaction rates.

7.4 Stellar Nucleosynthesis

We now consider the main pathways for energy generation by nuclear reactions in stellar interiors. For each nucleosynthetic reaction, we are interested in the amount of energy released per second per unit mass of nuclear fuel. For this calculation, it is most convenient to express the reaction rate

in power-law form:

$$r_{it} \simeq r_0 X_i X_t \rho^\gamma T^\beta \quad (7.23)$$

where r_0 is a constant, $X_{i,t}$ are the mass fractions of the two particles, the exponent of the density dependence is normally $\gamma = 2$ for two-body collisions, whereas the power-law dependence on the temperature β can range from ~ 1 to $\gtrsim 40$. If \mathcal{E}_0 is the energy released per reaction, then the rate of energy release per unit mass of nuclear fuel is:

$$\mathcal{E}_{it} = \left(\frac{\mathcal{E}_0}{\rho} \right) r_{it} \frac{\text{erg s}^{-1} \text{cm}^{-3}}{\text{g cm}^{-3}} \Rightarrow \text{erg s}^{-1} \text{g}^{-1} \quad (7.24)$$

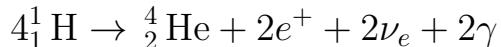
or

$$\mathcal{E}_{it} = \mathcal{E}_0 r_0 X_i X_t \rho^\alpha T^\beta \quad (7.25)$$

where $\alpha = \gamma - 1$.

7.4.1 The Proton-Proton Chain

We begin with the main reaction chain which powers stars on the Main Sequence, hydrogen burning:



with $\mathcal{E}_0 = 26.73 \text{ MeV}$. The two positrons subsequently annihilate with free electrons: $e^+ + e^- = 2\gamma$. The reaction can proceed through three channels (it is quite a common situation in nuclear reactions that different pathways can lead to the same end result), as shown in Figure 7.5. The balance between PP I and PP II varies with temperature, with the former preferred at $T \lesssim 1.5 \times 10^7 \text{ K}$; the values indicated in Figure 7.5 are those appropriate to the central temperature of the Sun, $T = 1.57 \times 10^7 \text{ K}$. PP III is never very important, but it is a source of high energy neutrinos.

Each step in the PP chain (and indeed in any chain) has its own reaction rate, determined by the Coulomb barriers and cross-sections involved. The slowest step in the PP chain is the first one because it is necessary for one of the protons to undergo a β^+ decay: $p^+ \rightarrow n + e^+ + \nu_e$ via the weak nuclear force. On average, a proton in the Sun will undergo one such decay only once in the lifetime of the Sun ($\sim 10^{10} \text{ years}$)! All the subsequent steps are much quicker. For the pp-chain:

$$\mathcal{E}_{pp} \propto X^2 \rho T^4 \quad \text{erg s}^{-1} \text{g}^{-1},$$

with the neutrinos carrying off $\sim 1\%$ of the energy generated.

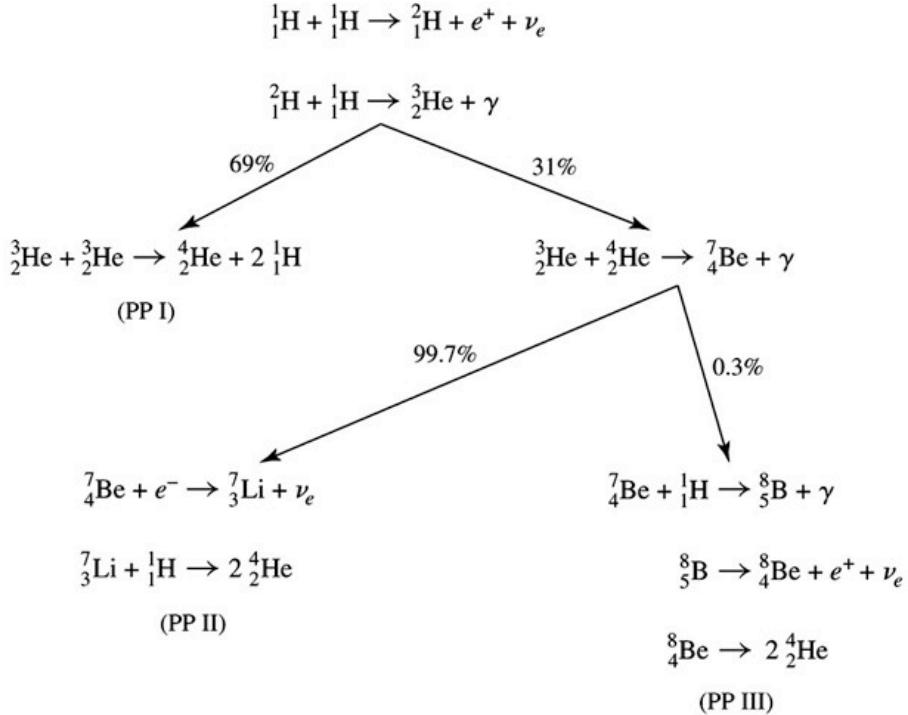
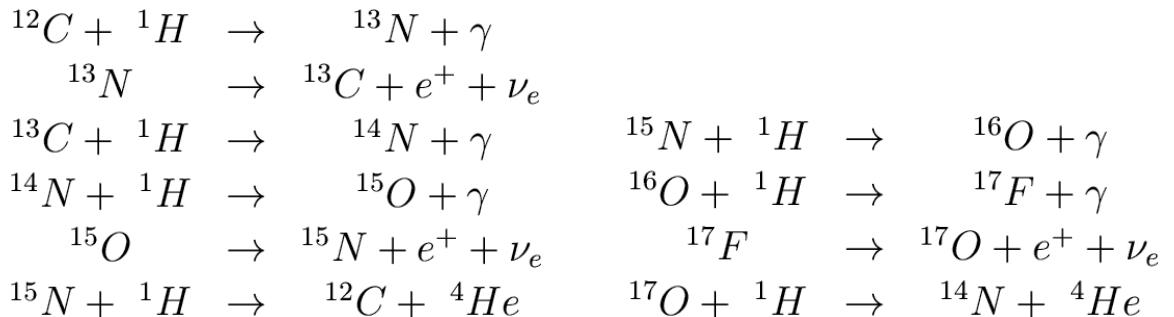


Figure 7.5: The three proton-proton chains. The branching ratios are appropriate to the temperature in the core of the Sun.

7.4.2 The CNO Cycle

When C, N and O are present and the temperature is sufficiently high, He can be synthesised from four H nuclei through a series of pathways known as the CNO cycle. C,N and O act as catalysts: they make He fusion possible through a series of reactions, but their number is conserved in the cycle. The two main pathways of the CNO cycle are as follows:



with the pathway on the right occurring only $\sim 0.04\%$ of the time. For the CNO cycle:

$$\mathcal{E}_{\text{CNO}} \propto X_{\text{CNO}} \rho T^{17} \quad \text{erg s}^{-1} \text{g}^{-1}.$$

Note the much steeper dependence on T than the p-p chain. The two

functions $\mathcal{E} = f(T)$ are shown in Figure 7.6 from where it can be seen that He fusion proceeds mostly via the CNO cycle in stars with $M \gtrsim 2M_{\odot}$, and via the p-p chain in cooler stars.

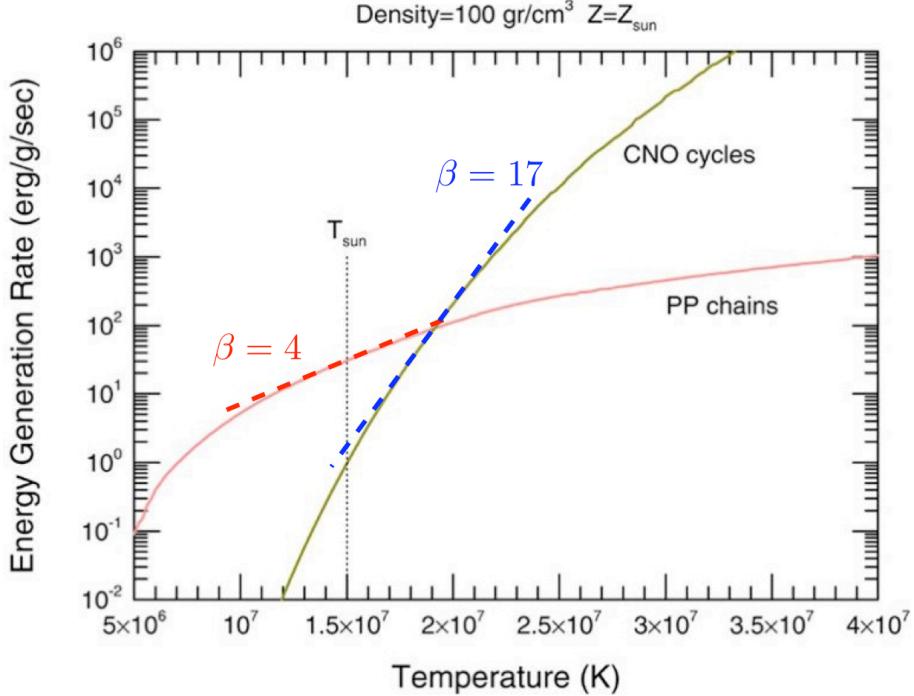


Figure 7.6: The temperature dependence of the energy generation rates of the p-p chain and CNO cycle.

Of course, in stars of lower metallicity, the cross-over point moves to higher temperatures (X_{CNO} is lower). In the first stars, which presumably consisted just of H and He synthesised in the Big Bang, only the p-p chain was operative.

A consequence of the steep T dependence of \mathcal{E}_{CNO} is that in massive stars H fusion must be more concentrated in the inner core of the star than is the case for lower mass stars where the p-p chain is the main channel. A second point of note is the following. As the name implies, the CNO cycle is a cyclic process that quickly reaches equilibrium. This means that: (a) the total number of C+N+O ions is conserved, and (b) each step has to proceed at the same rate. Given that rate for each step is $r_{i \rightarrow j} \propto n_i \cdot \sigma_{i \rightarrow j}$, the requirement $r = \text{constant}$ implies that $n_i \propto 1/\sigma_{i \rightarrow j}$. In other words, steps with smaller cross-sections require proportionally higher concentrations of the ions involved to keep the cycle in equilibrium. In particular, $\sigma(^{14}\text{N}) \simeq 1/30 \sigma(^{12}\text{C})$, so that in equilibrium $n_{^{14}\text{N}} \simeq 30 \times n_{^{12}\text{C}}$. In other words, the CNO cycle piles up ^{14}N at the expense of ^{12}C and ^{16}O ,

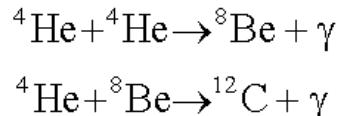
and is in fact the main channel for the nucleosynthesis of N in stars.

7.4.3 Helium Burning

The fusion of four H nuclei to form He, via either process, increases the mean molecular weight of the gas μ (defined so that the mean particle mass in a gas is $\langle m \rangle = \mu m_H$, where m_H is the mass of the hydrogen atom). From the ideal gas law:

$$P = \frac{\rho k T}{\mu m_H}$$

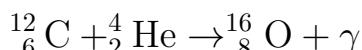
the pressure decreases and is no longer able to support the star against the pull of gravity. As a result, the stellar core contracts, raising both the temperature and the density of the gas (recall the virial theorem at the beginning of this lecture). When the temperature and density become sufficiently high, He nuclei can overcome their Coulomb barrier and combine to form ^{12}C by capturing three ^4He nuclei, as follows:



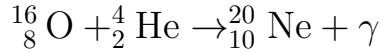
The lifetime of ^8Be is very short, $\sim 3 \times 10^{-16}$ s; consequently, this is really a three-body interaction, with a reaction rate $r \propto (\rho Y)^3$, where Y is the mass fraction of He. This triple alpha reaction was predicted by Fred Hoyle (a previous director of the Institute of Astronomy) in 1954, based on the high abundance of C in the Sun and H II regions like the Orion nebula. It produces C from He nuclei bypassing completely the intermediate elements Li, Be, B, and it explains why C is 10^5 – 10^7 times more abundant than Li, Be and B in the Universe. The energy generation rate of the triple alpha reaction has an extraordinarily steep dependence on T :

$$\mathcal{E}_{3\alpha} \propto Y^3 \rho^2 T^{40} \quad \text{erg s}^{-1} \text{g}^{-1}.$$

Once sufficient quantities of ^{12}C have been synthesised via the triple alpha reaction, heavier elements can be formed from ^{12}C via the capture of additional ^4He nuclei:

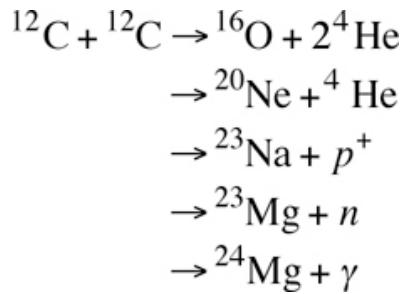


and

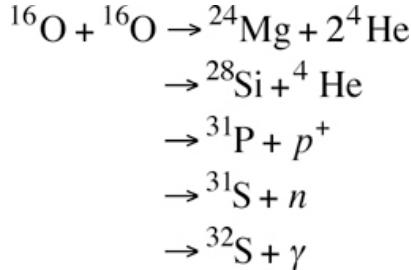


7.4.4 Carbon and Oxygen Burning

Higher and higher temperatures are required to overcome the increasing Coulomb barriers as heavier elements are synthesised. At temperatures $T \geq 6 \times 10^8 \text{ K}$, as found in the cores of stars with mass $M \geq 8M_\odot$, carbon burning can proceed via several reactions:



When $T \geq 1 \times 10^9 \text{ K}$, oxygen can burn:



When $T > 1.5 \times 10^9 \text{ K}$, we find from Wien's law (eq. 2.11), the maximum emission occurs at $\lambda = 2 \times 10^{-2} \text{ \AA}$. The corresponding photon energy, 0.6 MeV, is sufficient to photo-disintegrate heavy nuclei, creating a mixture of massive nuclei, p, n, He nuclei, and photons. These can facilitate reactions (silicon burning) that ultimately lead to the formation of stable elements of the iron group, at the peak of the binding energy per nucleon curve (Figure 7.2).

Each of the steps in the nucleosynthesis chain we have described, starting from H and ending at the Fe-peak, requires progressively higher temperatures, because of the increasing height of the corresponding Coulomb barrier. Thus, the core has to contract before the next step can start. Each successive reaction also has a steeper temperature dependence. Thus, it

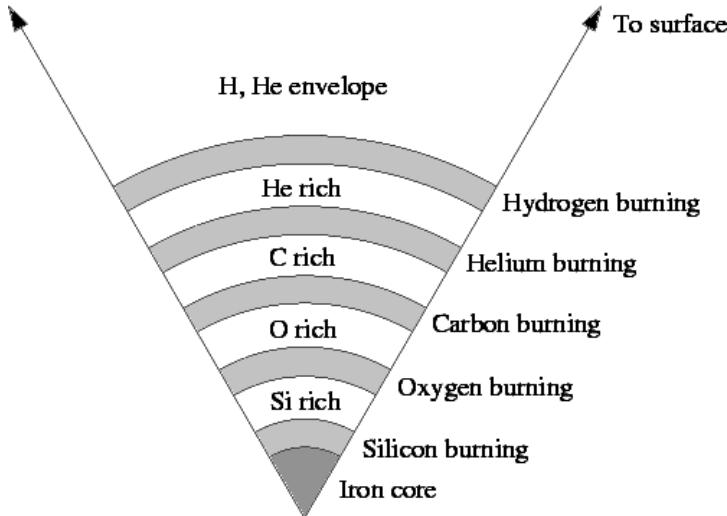


Figure 7.7: The stratified structure of the core of a massive star.

will take place in a more concentrated region within the core of the star, and will involve less mass. Therefore, at the end of the process, the interior of the star is stratified, as shown in Figure 7.7; this structure is sometimes referred to as an onion-skin structure.

Furthermore, each step in the nucleosynthesis of elements from H to Fe extracts less energy from the system (notice that the biggest individual ‘jump’ in Figure 7.2 occurs when He is synthesised). Consequently, the reaction rate has to be higher and fusion faster to provide the luminosity. Another factor is that at $T > 10^9$ K, neutrinos carry away larger and larger fractions of the energy produced. Neutrinos travel right through the star without interacting with the plasma; this has the effect of reducing the net energy production by nuclear burning. Thus, each successive nuclear burning phase lasts a progressively shorter period of time.

Table 7.1 summarises the main nuclear burning processes.

Table 2.1 Main nuclear burning processes

Fuel	Process	T_{thresh}^1 (K)	Products	$E/\text{nucleon}^2$ (MeV)	Timescale ³ (yr)
H	p-p	$\sim 4 \times 10^6$	He	6.55	
H	CNO	1.5×10^7	He	6.25	1×10^7
He	triple- α	1×10^8	C, O	0.61	1×10^6
C	C + C	6×10^8	O, Ne, Na, Mg	0.54	300
O	O + O	1×10^9	Mg, S, P, Si	~ 0.3	0.5
Si	Nucl. equil.	3×10^9	Co, Fe, Ni	$\lesssim 0.2$	0.005 (2 days!)

Notes:

¹ Threshold temperature

² Energy released per nucleon

³ Typical timescale for a $15M_\odot$ star.

7.5 Neutron Capture

We conclude our description of stellar nucleosynthesis with a brief mention of the mechanism whereby elements heavier than Fe are thought to be produced in stars.

As we have seen, at $T \gtrsim 1.5 \times 10^9$ K photodisintegration of nuclei becomes important and creates a mixture of neutrons, protons and other nuclei. Neutrons play an important role here. Since they do not experience a Coulomb barrier, they can easily penetrate the nuclei of even fully ionised heavy elements, such as Fe^{+26} . If they are captured by the nucleus, new neutron-rich isotopes can be produced. Such isotopes can be either stable or unstable.

It is important to distinguish between slow and rapid neutron capture (termed the s-process and the r-process), depending on the relative timescales of β -decay and neutron capture. In the example of the s-process shown in Figure 7.8, ^{56}Fe absorbs a neutron to form ^{57}Fe . Subsequent capture of two more neutrons leads to the formation of ^{59}Fe . Of the four Fe isotopes shown, the three lighter ones are stable, but ^{59}Fe is unstable, with a half-life of 44.5 days. Thus, if the flux of neutrons is not high and the interval between successive n-captures is longer than the half-life of ^{59}Fe , there is time for ^{59}Fe to decay to ^{59}Co by β -decay ($n \rightarrow p + e^- + \bar{\nu}_e$). The process can continue to form higher and higher mass elements, as shown in Figure 7.8.

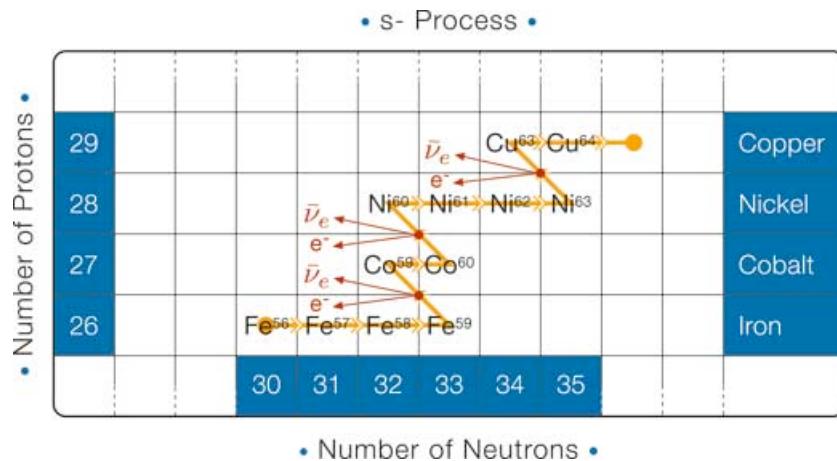


Figure 7.8: Example of the nucleosynthesis of ^{59}Co , ^{60}Ni and ^{63}Cu from ^{56}Fe via slow neutron capture.

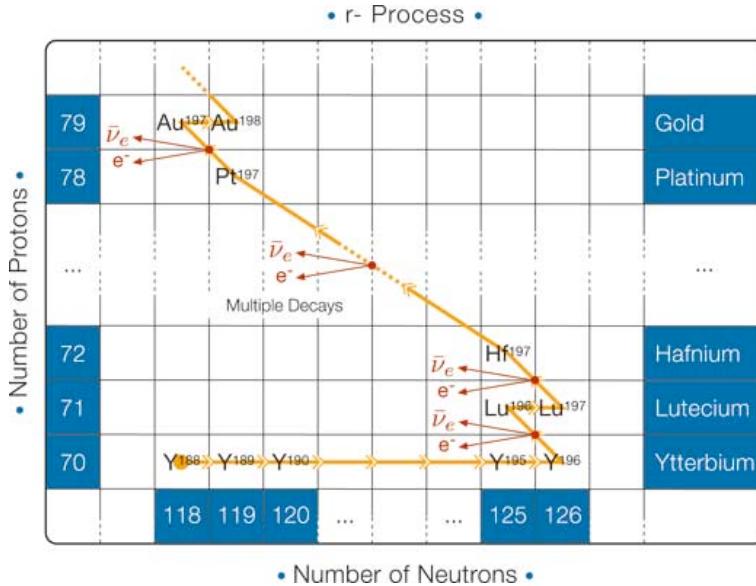


Figure 7.9: Example of the nucleosynthesis of ^{197}Au (stable gold) from ^{188}Yb via rapid neutron capture.

On the other hand, if the flux of neutrons is sufficiently high and the time interval between subsequent neutron captures is small compared to the half-life of the isotopes concerned, super-neutron-rich isotopes can be formed, as in the example of the r-process shown in Figure 7.9. When the neutron flux stops, these super-neutron-rich isotopes will undergo a series of β -decays until a stable isotope is reached.

Trans-Fe-peak elements can be formed by either s- or r-process nucleosynthesis, or both, depending on the stability of their neighbours in the Periodic Table. Typical s-process elements include Cu and Pb, while Eu is the prototypical signature of r-process nucleosynthesis in stellar spectra (see Figure 7.10). It is generally thought that s-process nucleosynthesis takes place primarily in AGB stars, while the r-process occurs mainly in supernova explosions.

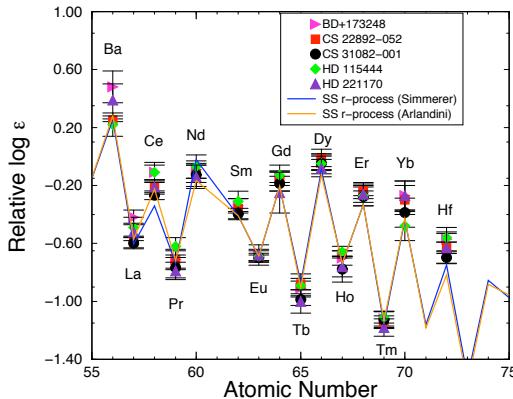


Figure 7.10: Relative abundances of several r-process elements in five Galactic halo stars (Cowan et al 2011).

ENERGY TRANSPORT WITHIN A STAR

8.1 Introduction

Up to now, we have considered how energy is generated within the interior of stars by the processes of gravitational contraction and nuclear fusion. However, for a star to shine, the heat and photons generated deep in the interior of the star have to be transferred to the surface, where they are radiated away as electromagnetic radiation. This is the subject of this lecture.

Three different energy transport mechanisms can operate within the interior of a star. **Radiation** can transfer energy from the core to the surface, as photons are continuously absorbed and re-emitted as they interact with the plasma. **Conduction**, whereby heat is transferred at the microscopic level by collisions between particles. **Convection** is the bulk motion of cells, with hot, buoyant mass elements rising towards the surface, while cooler ones sink.

The efficiency of the first two processes depends on the mean free path of, respectively, photons and particles (i.e. electrons); since the photon mean free path is normally greater than that of electrons, radiative diffusion normally dominates over conduction, although the latter becomes dominant in special physical conditions, such as those prevailing in white dwarfs.

8.2 Radiative Transport

The diffusion of energy from the star's core to the surface via radiation is described by the Eddington equation for radiative equilibrium:

$$\boxed{\frac{dT}{dr} = -\frac{3}{4} \cdot \frac{1}{ac} \cdot \frac{\kappa\rho}{T^3} \cdot \frac{L_r}{4\pi r^2}} \quad (8.1)$$

where a is the radiation constant, $a = 4\sigma/c$, σ is Stefan-Boltzmann constant, κ is the opacity and the other symbols have their usual meanings.

This equation can be derived in different ways; here we are going to use a simple physical argument.

Consider a small cell within the interior of a star, say with volume 1 cm^3 (see Figure 8.1). In LTE, the flux from below is $F_2 \sim \sigma T_2^4$ (blackbody emission). The flux from above is $F_1 \sim \sigma T_1^4$. Hence, the net flux through the element is $F \sim \sigma(T_2^4 - T_1^4)$. Generalising, we have:

$$F \sim -\frac{d}{dr}\sigma T^4. \quad (8.2)$$

The flux through the volume element has to be multiplied by ‘the transparency’ of the layer, which is approximately (for small volume elements) $1/\kappa\rho$, that is the photon mean free path (as we saw in Lecture 5). Thus the total flux through the volume element is:

$$F \sim -\frac{1}{\kappa\rho} \cdot \frac{d}{dr}\sigma T^4. \quad (8.3)$$

The third step in the derivation just equates the flux through a spherical surface of radius r to the total luminosity:

$$F = \frac{L_r}{4\pi r^2} \quad (8.4)$$

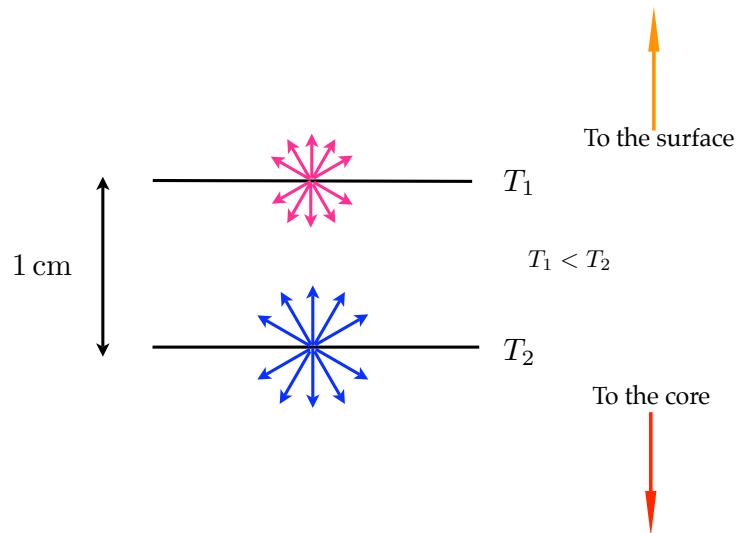


Figure 8.1: Blackbody emission at different distances r from the core of a star.

Combining these three equations we have:

$$L_r \sim -\frac{1}{\kappa\rho} \cdot 4\pi r^2 \cdot \frac{d}{dr} \sigma T^4 = -\frac{1}{\kappa\rho} \cdot 4\pi r^2 \cdot \frac{ac}{4} \cdot 4T^3 \cdot \frac{dT}{dr} \quad (8.5)$$

which can be rearranged in the form:

$$\frac{dT}{dr} \sim -\frac{1}{ac} \cdot \frac{\kappa\rho}{T^3} \cdot \frac{L_r}{4\pi r^2}. \quad (8.6)$$

Eqs. 8.1 and 8.6 differ only in the factor of 3/4 which comes from a proper integration over all angles.

In its 8.5 form (with a 4/3 factor on the right-hand side), the Eddington equation for radiative equilibrium gives the luminosity of star in terms of its radius, temperature, and temperature gradient when energy transport is primarily by radiative diffusion.

In its 8.1 form, the equation gives the temperature gradient required to carry the entire star's luminosity by radiation. A star, or a region within a star, in which this holds is said to be in radiative equilibrium, or simply *radiative*.

The Eddington equation is a valid description as long as the condition of LTE holds. This is clearly *not* the case at the stellar surface, or the photosphere, since this is where the photons escape (and therefore the photon mean free path is no longer small compared to the distance over which dT/dr is small—recall our discussion at 5.2). Thus, near the stellar surface the diffusion approximation is no longer justified and one needs to solve the much more complicated equations of radiative transfer. Fortunately, the LTE and diffusion approximations are valid over almost the entire stellar interior.

Sometimes eq. 8.1 is written in terms of the mass. Recalling (Figure 7.1) that $dm = 4\pi r^2 \rho dr$, we have:

$$\frac{dT}{dm} = -\frac{3}{4} \cdot \frac{1}{ac} \cdot \frac{\kappa}{T^3} \cdot \frac{L_r}{(4\pi r^2)^2} \quad (8.7)$$

8.2.1 The Luminosity of the Sun

We can use the Eddington equation in the eq 8.5 form:

$$L_r = -\frac{4}{3} \frac{1}{\kappa\rho} \cdot 4\pi r^2 \cdot ac \cdot T^3 \cdot \frac{dT}{dr}$$

to estimate the luminosity of the Sun. Integrating the above equation and using the mean temperature of the Sun (as derived in lecture 7.1) as an approximation, we have:

$$L_\odot \simeq \frac{1}{3} \mu \cdot 4\pi r_\odot \cdot ac \cdot \langle T_\odot \rangle^4.$$

$$L_\odot \simeq \frac{4}{3} \cdot 0.1 \cdot 3 \times 10^{10} \cdot 3 \times 10^{10} \cdot 7.6 \times 10^{-15} \cdot (5 \times 10^6)^4 \quad \text{cm}^3 \text{s}^{-1} \text{erg cm}^{-3} \text{K}^{-4} \text{K}^4$$

where $\mu \sim 0.1 \text{ cm}$.

$$L_\odot \simeq 6.4 \times 10 \cdot 1 \times 10^5 \cdot 625 \times 10^{24} \quad \text{erg s}^{-1}$$

$$L_\odot \simeq 4 \times 10^{33} \quad \text{erg s}^{-1}$$

For comparison, $L_\odot = 3.8 \times 10^{33} \text{ erg s}^{-1}$ – not bad!

8.3 Convection

Convection is a familiar phenomenon in our everyday lives: for example, our daily weather is caused by convection in the Earth's atmosphere. The surface of the Sun (Figure 8.2) is not smooth; instead we see bright granules separated by darker intergranular lanes. We know from Doppler velocity measurements that the motion of the bright regions is mostly outwards, while in the dark intergranular regions the gas is moving downwards. The motions and temperature inhomogeneities seen in the granulation pattern are attributed to the hydrogen convection zone just below the solar photosphere. The bulk motions of the gas and associated magnetic fields are thought to be the source of the mechanical energy flux that heats the solar chromosphere and corona.

Referring back to eq. 8.1, it can be readily appreciated that an increase in opacity κ in a stellar atmosphere will lead to a larger temperature gradient

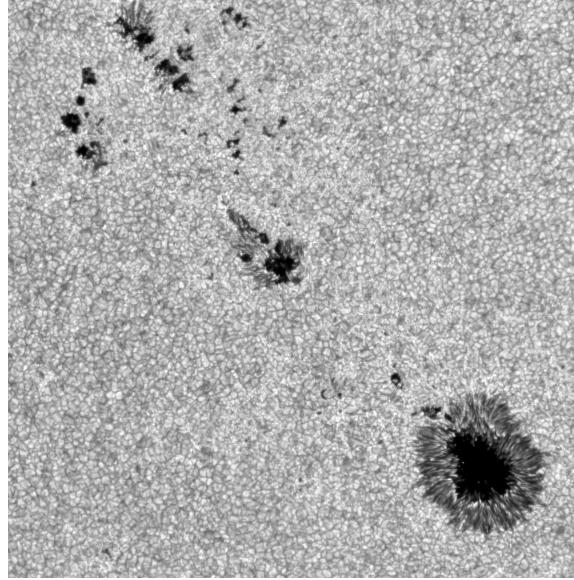


Figure 8.2: High resolution image of the solar photosphere, showing granulation and sunspots.

dT/dr (if the same luminosity is to be maintained). We also saw (Figure 5.5) that as the temperature decreases from $T \sim 10^7$ K, the opacity rises steeply with a Kramer's law $\langle \kappa \rangle \propto T^{-3.5}$. Thus, as we move from the core to the outer regions within a stellar interior, the temperature gradient is expected to become increasingly steep. A very steep temperature gradient is unstable, whether in a star or the Earth's atmosphere.

This can be appreciated by considering the consequences of displacing a volume element of gas, at equilibrium radius r inside a star—where T , P and ρ are the temperature, pressure and density, to a radius $r + dr$, where the ambient parameters are $T + dT$, $P + dP$ and $\rho + d\rho$ (Figure 8.3).

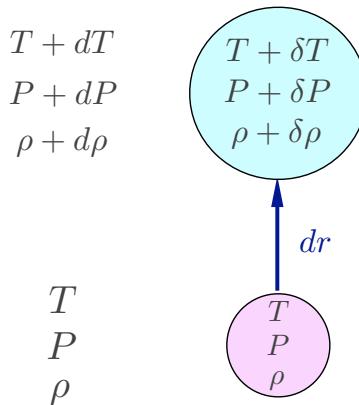


Figure 8.3: Illustration of the onset of convection. If $\rho + \delta\rho > \rho + d\rho$ the element will sink back to its former equilibrium position at radius r . But if $\rho + \delta\rho < \rho + d\rho$, the element will be buoyant and convection will ensue.

In the following treatment, we make two assumptions:

1. Pressure equilibrium: the element maintains the same pressure as its surroundings, and
2. The process is adiabatic, that is there no heat exchange between the volume element and its surroundings.

This is the same as saying that the timescale for removing pressure imbalance is short compared to the timescale for the establishment of thermal equilibrium. The ideal gas law:

$$P = \frac{\rho k T}{\mu m_H} \quad (8.8)$$

where μ is the mean molecular weight, m_H is the mass of the hydrogen atom, and the product μm_H is the mean mass of the gas particles, can be written as:

$$P = K \rho^\gamma \quad (8.9)$$

where K is a constant and γ is the ratio of the specific heats:

$$\gamma = \frac{C_P}{C_V} \quad C_P \equiv \left. \frac{\partial Q}{\partial T} \right|_P, \quad C_V \equiv \left. \frac{\partial Q}{\partial T} \right|_V \quad (8.10)$$

which measure the amount of heat required to raise the temperature of a unit mass of material by a unit temperature interval at constant pressure and at constant volume respectively. Note that $C_P > C_V$ (and therefore $\gamma > 1$), because at constant pressure some of the energy input goes into increasing the volume of the gas and hence more energy is required to raise the temperature by 1 K. From 8.9, we have:

$$P \rho^{-\gamma} = (P + \delta P) (\rho + \delta \rho)^{-\gamma} \quad (8.11)$$

What happens next depends on the difference $\Delta \rho \equiv (\rho + \delta \rho) - (\rho + d\rho)$. In the gravitational field g of the star, our rising volume element will experience a force $f = -g \Delta \rho$. Thus, if $\Delta \rho$ is +ve, the element will sink back to its equilibrium position at radius r but, if $\Delta \rho$ is -ve, the element will be buoyant and will rise further. This is the onset of convection.

Thus, the condition for convective instability is set by the density gradient within the star: if the gradient is less than that experienced by the volume

element rising adiabatically, the star will be unstable against convection. Since we have assumed pressure equilibrium with its surrounding, we could also state the same instability criterion in terms of the temperature gradient (see eq. 8.8) as:

$$\left| \frac{dT}{dr} \right|_{\text{rad}} > \left| \frac{dT}{dr} \right|_{\text{ad}}. \quad (8.12)$$

Eq. 8.12 is known as the Schwarzschild criterion for convective instability (see Figure 8.4). It tells us that if the temperature profile within a star is super-adiabatic, the star is unstable against convection.

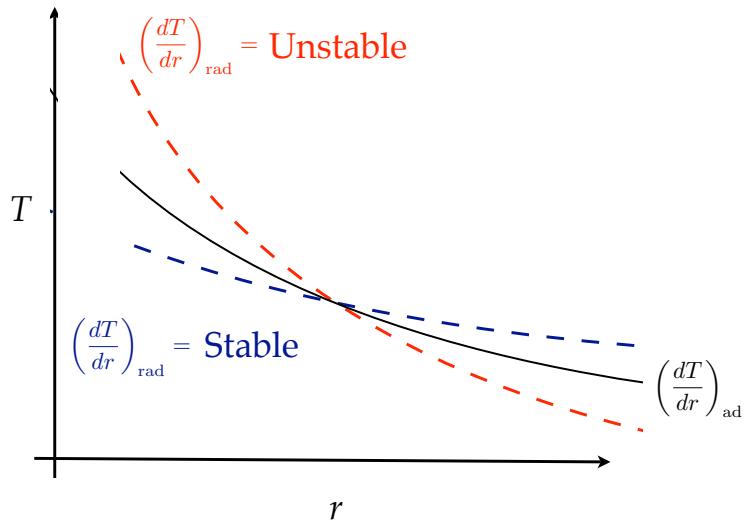


Figure 8.4: The Schwarzschild criterion for convective instability.

It is also instructive to express the convective instability criterion in terms of the parameter γ . We can do this as follows. From eq. 8.12 we have:

$$\left| \frac{d \ln T}{dr} \right|_{\text{rad}} > \left| \frac{d \ln T}{dr} \right|_{\text{ad}}, \quad (8.13)$$

dividing through by $d \ln P / dr$,

$$\left| \frac{d \ln T}{d \ln P} \right|_{\text{rad}} > \left| \frac{d \ln T}{d \ln P} \right|_{\text{ad}} \quad (8.14)$$

or

$$\left| \frac{d \ln P}{d \ln T} \right|_{\text{rad}} < \left| \frac{d \ln P}{d \ln T} \right|_{\text{ad}}. \quad (8.15)$$

Eqs. 8.8 and 8.9 can be combined to give the adiabatic relation between pressure and temperature:

$$P T^{\frac{\gamma}{1-\gamma}} = K', \quad (8.16)$$

so that

$$\frac{dP}{P} = \frac{\gamma}{\gamma - 1} \frac{dT}{T} \quad (8.17)$$

or

$$\left| \frac{d \ln P}{d \ln T} \right|_{\text{ad}} = \frac{\gamma}{\gamma - 1}, \quad (8.18)$$

The Schwarzschild instability criterion then becomes:

$$\left| \frac{d \ln P}{d \ln T} \right|_{\text{star}} < \frac{\gamma}{\gamma - 1}. \quad (8.19)$$

Sometimes, this condition is also given as:

$$\left| \frac{dT}{dr} \right| > \frac{\gamma - 1}{\gamma} \frac{T}{P} \left| \frac{dP}{dr} \right| \quad (8.20)$$

obtained from 8.17 by dividing both sides by dr and rearranging.

The above treatment emphasises the importance of the ratio of the specific heats in determining the stability of a star against convection. Recall that $C_P > C_V$. Their difference is given by:

$$C_P - C_V = \frac{k}{\mu m_H}$$

so that:

$$\frac{\gamma}{\gamma - 1} = C_P \frac{\mu m_H}{k}.$$

Thus, the higher the specific heats, the closer their ratio is to 1; the closer γ is to 1, the higher is the value of the adiabatic $|d \ln P / d \ln T|$ gradient (eq. 8.18), leading to instability (eq. 8.19). Under what conditions are C_P and C_V high? One example is partially ionised gas, where some of the heat supplied to the system may go into further ionizing atoms instead of increasing their kinetic energy. Similarly, the presence of molecules in the mix, would increase the specific heats since some of energy supplied would go into breaking the molecular bonds.

8.3.1 Which Stars Are Convective?

From the discussion above we can develop some physical understanding of when a star will develop a convective layer in its core. From eq. 8.1 we saw

that the temperature gradient is proportional to the opacity κ ; thus, we expect that in layers where the opacity is very high, the temperature gradient required for radiative energy transport becomes unachievably steep. We also saw from Figure 5.5 that stellar opacities increase dramatically as T decreases from 10^6 to 10^5 K; at $T \sim 10^5$ K the gas is only partly ionised (at typical stellar densities); the rise in κ is produced by the availability of many bound-bound and bound-free transitions.

Hand in hand with this is the increase in the specific heats, and therefore the increase in γ as the gas becomes partly, as opposed to fully, ionised. As we have discussed, this will increase the adiabatic $|d \ln P / d \ln T|$ gradient, leading to convection. For both reasons, convection will occur in the outer layers of cool stars. In a G0 V star the convective layer is thin, while main sequence M stars are almost fully convective. Red giants and supergiants are also convective over most of their interiors. Figure 8.5 shows the onset of convection in the Sun at a radius $r \simeq 0.7R_\odot$.

Convection is also important in stellar layers where the ratio $L / 4\pi r^2$ is high (cfr. eqs. 8.1 and 8.12), that is where large luminosities are generated over small volumes. This is the situation in the cores of massive stars, given

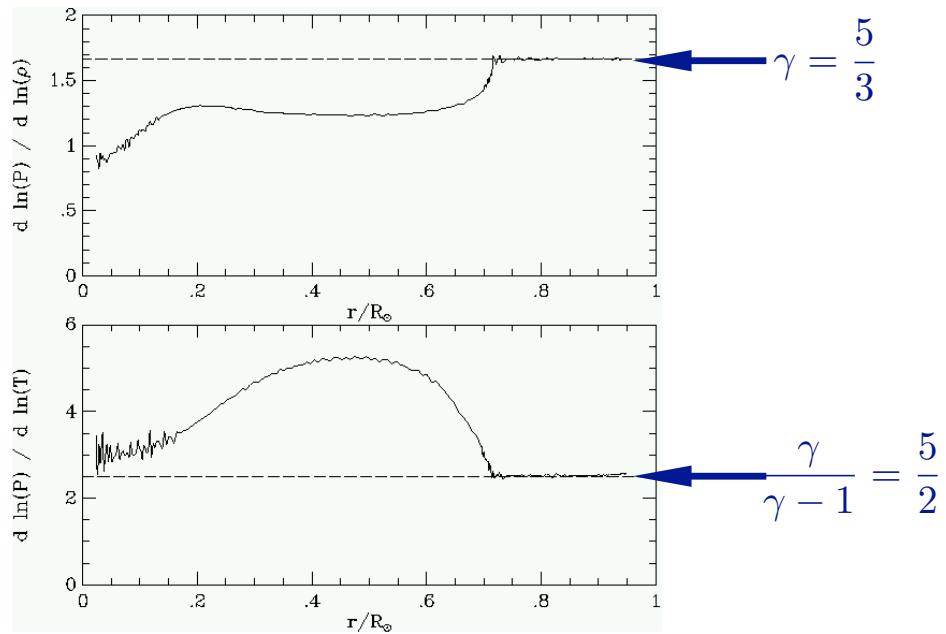


Figure 8.5: Plots of $d \ln P / d \ln \rho$ (top) and $d \ln P / d \ln T$ for the Standard Solar Model. For a monoatomic gas $\gamma = 5/3$.

the steep temperature dependencies of the CNO cycle and the triple-alpha process (Lecture 7.4.2 and 7.4.3).

Figure 8.6 summarises pictorially the points made in this section.

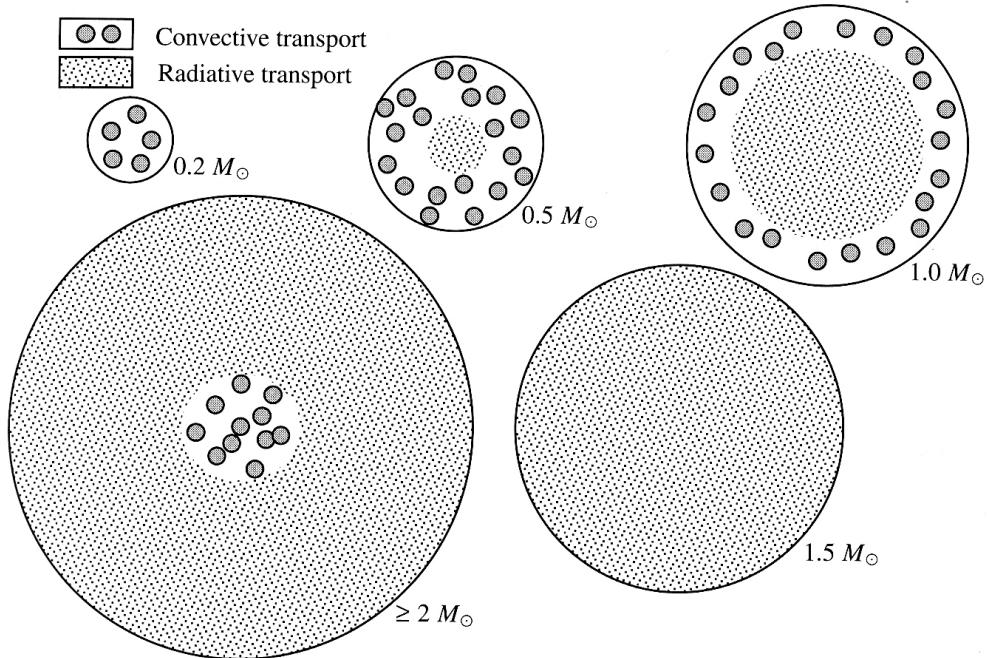


Figure 8.6: Zones of convection and radiation in main-sequence stars of various masses. The lowest mass stars are completely convective. A radiative core develops at $M \simeq 0.4M_{\odot}$, and a star is fully radiative at $M \simeq 1.5M_{\odot}$. The core region is again convective for masses $M \gtrsim 2M_{\odot}$. The relative sizes of the stars shown here are approximately correct, while on the main sequence. (Figure reproduced from Bradt, H., *Astrophysics Processes*, CUP).

8.4 Hydrostatic Equilibrium

Having established the conditions under which convection will occur, it may be of interest to estimate the convective flux F_c , that is the energy transported by convective cells per unit time through a unit surface area of the star. In order to do so, we need to first consider the concepts of hydrostatic equilibrium and pressure scale height.

8.4.1 Hydrostatic Equilibrium

What stops a star from collapsing under its own gravity? The inward force of gravity must be balanced by an equal force of opposite sign; this force is provided by pressure. Considering a small cylindrical element of mass dm and height dr located at a distance r from the centre of a spherically symmetric star (Figure 8.7), we have:

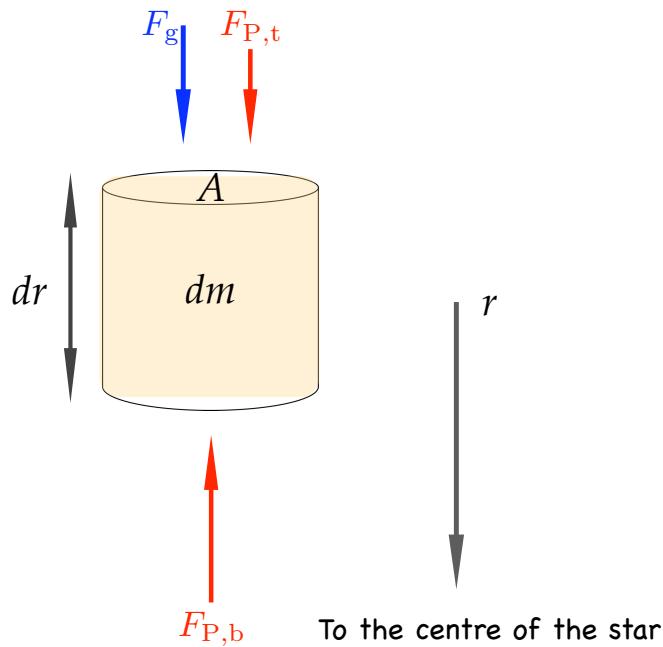


Figure 8.7: In hydrostatic equilibrium $|F_g + F_{P,t}| = |F_{P,b}|$.

$$A dP = -G \frac{M_r dm}{r^2} \quad (8.21)$$

where: (i) M_r is the mass enclosed within radius r ; (ii) A is the area of the base of the cylinder; and (iii) dP is the difference in pressure (defined as force per unit area, i.e. $P \equiv F_p/A$) between the top and bottom faces of the cylinder. Since we are considering an infinitesimally small cylinder, we can assume that the density is constant within the cylinder and express the mass in terms of the density: $dm = \rho A dr$:

$$A dP = -G \frac{M_r \rho A dr}{r^2}. \quad (8.22)$$

Dividing both sides by the volume of the cylinder $dV = A dr$, we obtain the equation of hydrostatic equilibrium:

$$\boxed{\frac{dP}{dr} = -G \frac{M_r \rho}{r^2}} \quad (8.23)$$

Since $G M_r / r^2 \equiv g$, the local acceleration of gravity at radius r , we can also write:

$$\frac{dP}{dr} = -\rho g. \quad (8.24)$$

In order for a star to be in hydrostatic equilibrium, a *negative pressure gradient* must exist within the star, with the pressure being larger in the interior than near the surface.

8.4.2 Dynamical Timescale

What would happen if pressure suddenly vanished and the only force acting on a star were gravity? The star would collapse on a free-fall timescale given by:

$$t_{\text{ff}} = \sqrt{\frac{2R}{g}} \quad (8.25)$$

where R is the star radius (recall $r = \frac{1}{2}at^2$ for a body experiencing acceleration a). With the above definition of g , and expressing M_R in terms of the mean density $\langle \rho \rangle$, $M_R = \frac{4}{3}\pi R^3 \langle \rho \rangle$, we have:

$$t_{\text{ff}} = \sqrt{\frac{2R R^2}{GM_R}} = \sqrt{\frac{2R R^2 3}{G 4\pi R^3 \langle \rho \rangle}} = \sqrt{\frac{3}{2\pi}} \sqrt{\frac{1}{G \langle \rho \rangle}} \quad (8.26)$$

The last term in the equation is often referred to as the *dynamical timescale*:

$$\boxed{t_{\text{dyn}} \sim \sqrt{\frac{1}{G\rho}}} \quad (8.27)$$

Although we have derived it by using the unrealistic example of the free-fall time in a star, the dynamical timescale is an important concept often used in dimensional treatment of astrophysical situations. It describes the

time taken for changes in one part of a body to be communicated to the rest of that body (and thus $t_{\text{dyn}} \approx t_{\text{sc}}$ where t_{sc} is the sound crossing time). Another way to think about the dynamical timescale is as the time required for a system to move from one equilibrium state to another after a sudden change. Thus, for example, an interstellar cloud cannot collapse to form stars over a timescale shorter than t_{dyn} ; similarly, a sudden burst of star formation in a galaxy cannot take place over a timescale shorter than t_{dyn} .

8.4.3 Pressure Scale Height

Returning to the equation of hydrostatic equilibrium, we define the pressure scale height H_p as:

$$\frac{1}{H_p} \equiv -\frac{1}{P} \frac{dP}{dr}, \quad (8.28)$$

so that we can express the variation in pressure with radius as:

$$P = P_0 e^{-r/H_p}. \quad (8.29)$$

In other words, H_p is the radial distance over which the pressure drops by a factor of e . Recalling eq. 8.24, it can be seen that:

$$H_P = \frac{P}{\rho g}. \quad (8.30)$$

8.5 Mixing Length Theory

Let us now return to the question of the energy transport by convection. A buoyant bubble will rise until the temperature inside gradually adjusts to the (lower) temperature of its surroundings. The distance that a hot cell rises, or a cold one sinks, is referred to as the *mixing length*, and is normally expressed in terms of the pressure scale height:

$$\ell = \alpha H_P$$

where $\alpha \sim 1$ is a free parameter (meaning simply that we do not have a physical theory to relate ℓ to H_P). In other words, the mixing length is of the order of the pressure scale height.

As the cell travels towards the stellar surface by one mixing length, the excess heat flow per unit volume from the bubble into its surroundings can be written as:

$$\delta q = (C_P \delta T) \rho, \quad (8.31)$$

where C_P is the specific heat at constant pressure and

$$\delta T = \delta \left(\frac{dT}{dr} \right) dr \quad (8.32)$$

is the temperature difference between the surroundings and the rising bubble. The convective flux is simply:

$$F_c = \delta q \langle v_c \rangle \quad (8.33)$$

where $\langle v_c \rangle$ is the average speed of the convective bubble. We can obtain an expression for $\langle v_c \rangle$ by equating the kinetic energy of the rising bubble to some fraction β ($0 < \beta < 1$) of the work per unit volume done by the buoyant force over the distance ℓ :

$$\frac{1}{2} \rho \langle v_c \rangle^2 = \beta \langle f_{\text{net}} \rangle \ell \quad (8.34)$$

so that:

$$\langle v_c \rangle = \left(\frac{2\beta \langle f_{\text{net}} \rangle \ell}{\rho} \right)^{1/2}. \quad (8.35)$$

where β is another free parameter.

The net force on the bubble is just

$$f_{\text{net}} = -g \delta \rho \quad (8.36)$$

We can express $\delta \rho$ in terms of δT using the ideal gas law

$$\frac{\delta P}{P} = \frac{\delta \rho}{\rho} + \frac{\delta T}{T} \quad (8.37)$$

or, since we have pressure equilibrium between the bubble and its surroundings (i.e. $\delta P = 0$),

$$\delta \rho = -\frac{\rho}{T} \delta T. \quad (8.38)$$

Thus,

$$f_{\text{net}} = \frac{\rho g}{T} \delta T \quad (8.39)$$

and

$$\langle f_{\text{net}} \rangle = \frac{1}{2} \frac{\rho g}{T} \delta T \quad (8.40)$$

taking the average of f_{net} over the distance ℓ travelled by the convective cell. Substituting 8.40 into 8.35, we have:

$$\langle v_c \rangle^2 = \beta \frac{\rho g}{\rho T} \ell \delta T \quad (8.41)$$

$$\langle v_c \rangle^2 = \beta \frac{g}{T} \alpha H_p \delta T \quad (8.42)$$

$$\langle v_c \rangle^2 = \beta \frac{g}{T} \alpha \frac{P}{\rho g} \delta T \quad (8.43)$$

Using the ideal gas law: $P = (\rho kT)/(\mu m_H)$, we have:

$$\langle v_c \rangle^2 = \beta \alpha \frac{k}{\mu m_H} \delta T \quad (8.44)$$

Using 8.32 with $dr = \ell = \alpha H_p = \alpha P/\rho g$ and proceeding as before for P , we have:

$$\langle v_c \rangle^2 = \beta \alpha \frac{k}{\mu m_H} \delta \left(\frac{dt}{dr} \right) \alpha \frac{kT}{\mu m_H} \frac{1}{g} \quad (8.45)$$

$$\langle v_c \rangle^2 = \beta \alpha^2 \left(\frac{k}{\mu m_H} \right)^2 \frac{T}{g} \delta \left(\frac{dt}{dr} \right) \quad (8.46)$$

Substituting 8.46 into 8.33 and using again 8.31 and 8.32, we finally obtain an expression for the convective flux:

$$F_c = \rho C_P \left(\frac{k}{\mu m_H} \right)^2 \left(\frac{T}{g} \right)^{3/2} \beta^{1/2} \left[\delta \left(\frac{dT}{dr} \right) \right]^{3/2} \alpha^2 \quad (8.47)$$

What we have described here is an example of mixing-length theory. The unsatisfactory aspects of it are the ‘fudge-factors’ α and β which are generally adjusted to fit observations. However, it is a significantly simpler treatment than the full 3-D (magneto-)¹hydrodynamical calculations of convective flows (see Figure 8.8).

In closing, convection is a complicated topic and remains an active area of research, in particular because it can have considerable impact on a wide

¹Yes, we have not mentioned them yet, but magnetic fields are clearly important when we are dealing with bulk motions of ionised gas!

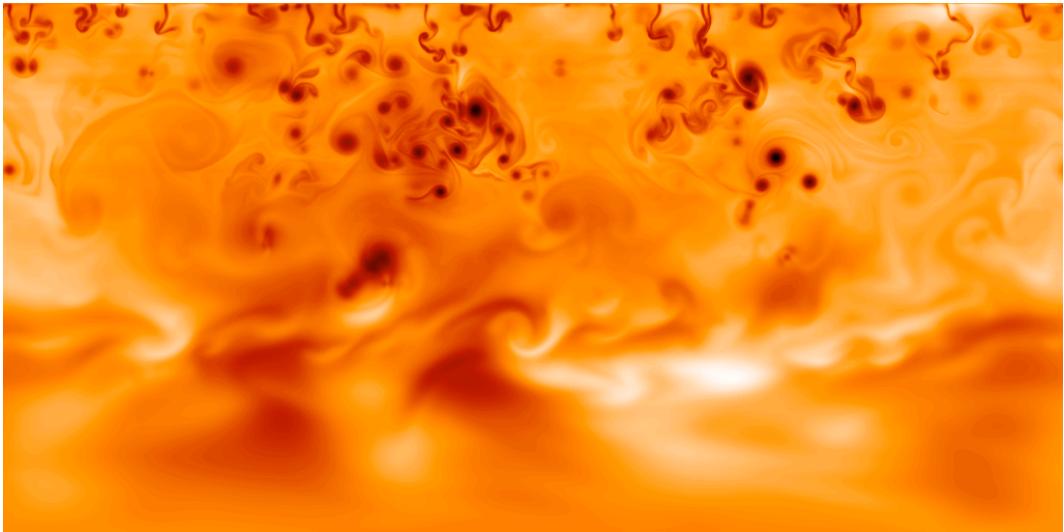


Figure 8.8: Computer simulation of convection.

range of stellar properties and on stellar evolution. The lack of a good theory of convection, and of the amount of energy that can be transferred by convection, is at present an important limitation in our understanding of stellar structure.

However, convection remains an important phenomenon in the interiors of stars, and not only as a way of transporting energy to the surface where it is radiated away. Convection is also important as a means of *mixing* between different layers of the star. Mixing happens because of *convective overshooting*: at the top and bottom boundaries of the convection zone, even though there is no net force, a rising or falling cell will arrive with a finite velocity and will overshoot. This was first realised in the 1980s from observations of the nuclear products at the surface of massive stars, indicative of some degree of mixing to layers above the convective core boundary. Stellar rotation also greatly facilitates mixing between different stellar layers (see Figure 8.9). The mixing of freshly synthesised elements into the outer layers of stars has important consequences for their evolution.

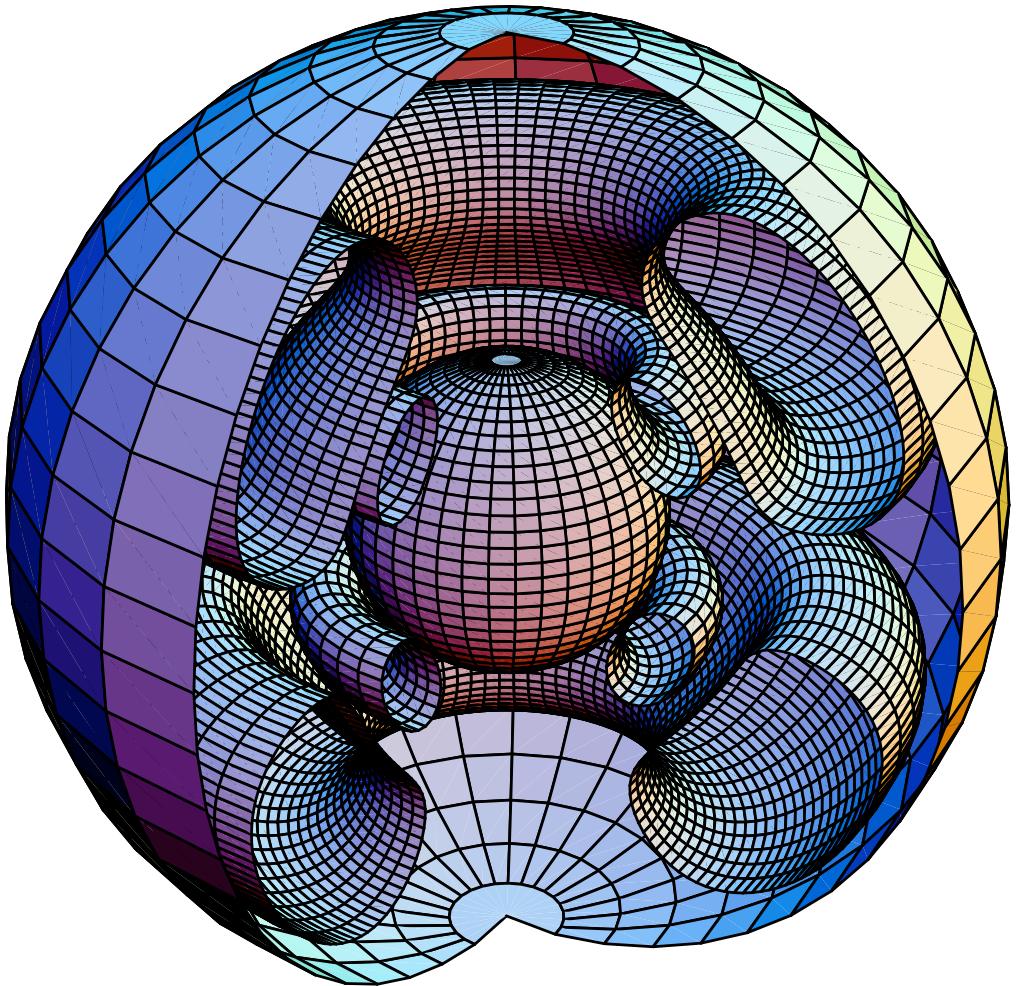


Figure 8.9: Stream lines of meridional circulation in a rotating $20M_{\odot}$ model with solar metallicity and $v_{\text{rot}} = 300 \text{ km s}^{-1}$ at the beginning of the H-burning phase. The streamlines are in the meridian plane. In the upper hemisphere on the right section, matter is turning counterclockwise along the outer stream line and clockwise along the inner one. The outer sphere is the star surface and has a radius equal to $5.2R_{\odot}$. The inner sphere is the outer boundary of the convective core. It has a radius of $1.7R_{\odot}$. (Figure reproduced from Meynet & Maeder 2002, A&A, 390, 561).

STELLAR MODELS

9.1 Introduction

Stars are complex physical systems, but not too complex to be modelled numerically and, with some simplifying assumptions, analytically. This is not always the case in astrophysics. For example, the interstellar medium in galaxies is a very complex environment and we are still a long way from being able to satisfactorily model it in our computers. On the other hand, the evolution of the whole Universe can be described with the Friedmann equations and once the fundamental cosmological parameters are known with sufficient accuracy, its past history can be deduced and its future evolution predicted.

The aim of stellar models is to construct a representation of a star which is physically accurate to a sufficient degree to reproduce the observed properties of stars, and thereby obtain a physical understanding of what determines those properties. For example, in the first few lectures of the course we saw that most stars fall in narrow strip on the colour-magnitude diagram (Figure 2.8); with the aid of stellar models we now interpret this Main Sequence as the locus of stars during their hydrogen burning phase. Stellar models can help us understand how stars move onto the Main Sequence when they are born out of an interstellar cloud, and how they evolve off the Main Sequence—that is the tracks they follow in the colour-magnitude plane during the late stages of their evolution (see Figure 9.1). In Lecture 4 we also saw that stars obey a mass-luminosity relation (Figure 4.8) and that there is a strong dependence of stellar lifetimes on stellar mass (Figure 4.9); again, a successful stellar model should be able to reproduce these relationships and thereby provide physical insight into their origin.

9.2 Static Models

In this lecture we are going to consider the simplest class of stellar models. We assume that stars:

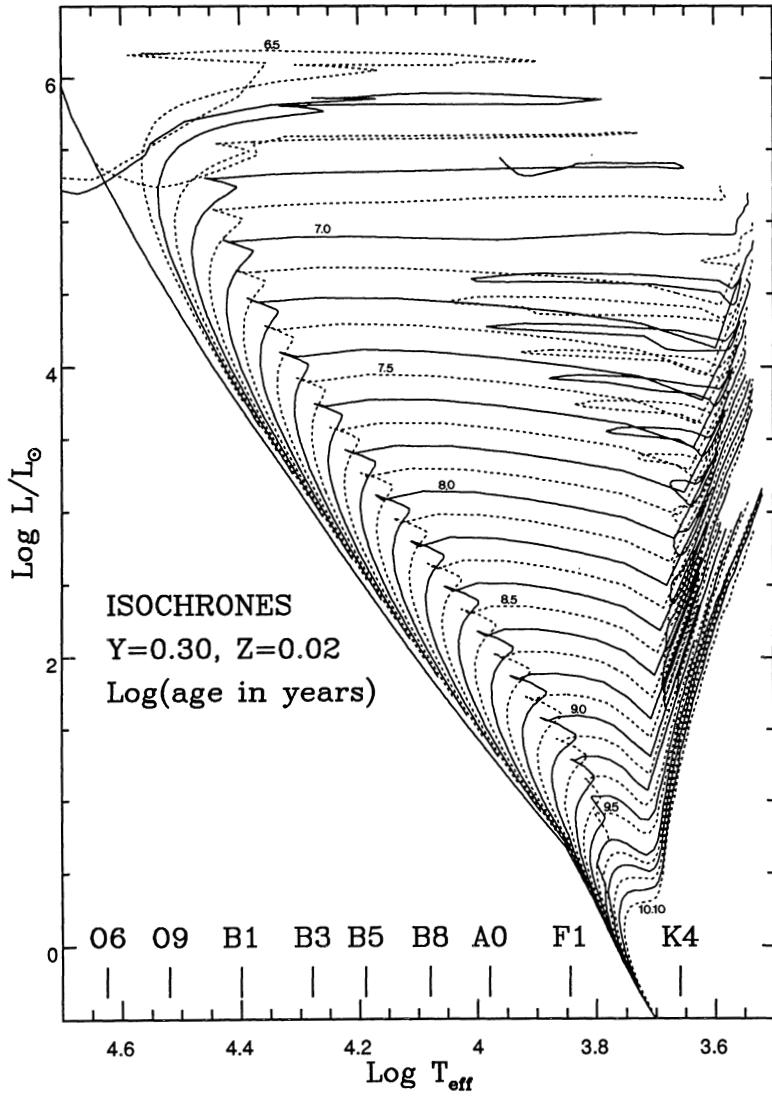


Figure 9.1: Theoretical H-R diagram computed with the Geneva stellar models which include moderate overshooting. (Reproduced from Meynet et al. (1993), A&A Supp., 98, 477).

- are spherically symmetric,
- are static,
- have no magnetic field.

None of these conditions apply to real stars. Real stars rotate, and as a consequence they are slightly ‘flattened’ at the poles compared to the equator. In neglecting this slight asymmetry, we make the assumption that the additional centripetal force acting on a test mass at the equator of the star is small compared to the gravitational force. This inequality can be also expressed in terms of timescales: departures from spherical

symmetry can be neglected when the period of rotation is much longer than the dynamical time, i.e. $P \gg \tau_{\text{dyn}}$. This condition is certainly satisfied in the Sun, where $P \sim 30$ days, while $\tau_{\text{dyn}} \sim 2000$ s. (Incidentally, in the Sun the flattening at the poles is only $1 \times 10^{-5} R_{\odot}$). However, some stars rotate much faster than the Sun and in some cases rotational effects must be included in the structure equations—and can have a big effect on the output of the models.

Similar arguments can be used to argue for static models, by which we mean models which do not include time as a variable. So long as the timescale for evolutionary changes is much longer than the dynamical timescale, time-dependent effects can be neglected. This is certainly the case while stars are on the Main Sequence, but the approximation may no longer be valid during periods when the star expands or contracts (or pulsates). The inclusion of magnetic fields in the stellar models is beyond the scope of these lectures.

To build a stellar model for a star in hydrostatic and thermal equilibrium¹, we begin with four differential equations. We have already encountered three of these, but we shall recapitulate now.

1. The equation of Mass Continuity:

$$\boxed{\frac{dm}{dr} = 4\pi r^2 \rho} \quad (9.1)$$

(see Figure 7.1).

2. The equation of Hydrostatic Equilibrium:

$$\boxed{\frac{dP}{dr} = -G \frac{M_r \rho}{r^2}} \quad (9.2)$$

(see Figure 8.7).

3. The equation of Thermal Equilibrium:

$$\boxed{\frac{dL_r}{dr} = 4\pi r^2 \rho \mathcal{E}} \quad (9.3)$$

to which we shall return in a moment, and

¹That is, a star for which τ_{dyn} (Lecture 8.4.2) $\ll \tau_{\text{KH}}$ (Lecture 7.1) $\ll \tau_{\text{nuclear}}$ (Lecture 7.2).

4. The equations of Energy Transport:

$$\boxed{\frac{dT}{dr} = -\frac{3}{4} \cdot \frac{1}{ac} \cdot \frac{\kappa\rho}{T^3} \cdot \frac{L_r}{4\pi r^2}} \quad (9.4)$$

i.e. the Eddington equation for radiative transport, or

$$\boxed{\frac{dT}{dr} = -\frac{\gamma-1}{\gamma} \frac{\mu m_H}{k} \frac{GM_r}{r^2}} \quad (9.5)$$

for energy transport by adiabatic convection.

Eq. 9.3 follows directly from 9.1. We can write the contribution to the total luminosity of a star by an infinitesimal mass dm as:

$$dL = \mathcal{E} dm$$

where \mathcal{E} is the rate of energy release per unit mass ($\text{erg s}^{-1} \text{ g}^{-1}$) by all processes. In a static model, $\mathcal{E} = \mathcal{E}_{\text{nuclear}}$ but, more generally, if the layer of the star is expanding or contracting, $\mathcal{E} = \mathcal{E}_{\text{nuclear}} + \mathcal{E}_{\text{gravity}}$, where the second term is negative if the star is expanding. Rigorously, even in the static case, one should also include the energy lost from escaping neutrinos, i.e. $\mathcal{E} = \mathcal{E}_{\text{nuclear}} - \mathcal{E}_\nu$. In any case, using eq. 9.1, we have:

$$\frac{dL_r}{dr} = 4\pi r^2 \rho \mathcal{E} .$$

So, we have five coupled differential equations describing the run of mass, pressure, luminosity and temperature within a star with r , the distance from the centre of the star, as the independent variable. Note that the first two equations describe the *mechanical* structure of the star, while the last three describe the *energy and thermal* structure. The equations are coupled to each other through the fact that, for a general equation of state, P is a function of both ρ and T .

More specifically, in order to tackle these five equations, we need expressions for the pressure P , the opacity κ , and the energy generation rate \mathcal{E} , in terms of the fundamental physical characteristics of the plasma, that is density ρ , temperature T , and chemical composition. Thus we have three additional relations, which are sometimes referred to as the ‘Constitutive Relations’:

1. The nuclear energy production rate:

$$\mathcal{E} \propto \mathcal{E}_0 \rho^\alpha T^\beta \quad (9.6)$$

where the exponents α and β and the constant of proportionality depend on the specific nuclear reaction rate (for hydrogen burning stars, the p-p chain or the CNO cycle), as we saw in Lecture 7. Thus, implicit in eq. 9.6 is a dependence on the chemical composition of the star.

2. As we saw in Lecture 5, **the opacity** κ is the sum of several processes: $\langle \kappa \rangle = \langle \kappa_{\text{bb}} + \kappa_{\text{bf}} + \kappa_{\text{ff}} + \kappa_{\text{es}} + \kappa_{\text{H-}} \rangle$. Of these, two of the main sources of opacity, κ_{bf} and κ_{ff} have functional form:

$$\langle \kappa_{\text{bf}} \rangle = \kappa_{0,bf} \rho T^{-3.5} \quad (9.7)$$

where the constant of proportionality $\kappa_{0,bf}$ depends on the composition of the gas.

3. **The Pressure** has two components: gas pressure and radiation pressure:

$$P = P_g + P_{\text{rad}} = \frac{\rho k T}{\mu m_H} + \frac{1}{3} a T^4 \quad (9.8)$$

We discuss each component in the following two subsections.

9.2.1 Mean Molecular Weight

The P_g term in eq. 9.8 is simply the ideal gas law, $PV = NkT$, expressed in terms of the density $\rho = N \langle m \rangle / V$, where N is the number of particles, and the average mass of a particle, $\langle m \rangle = \mu m_H$. μ is the *mean molecular weight* which we have already encountered several times. Let us consider its value for fully ionised gas (a reasonably good assumption in the *interiors* of stars, where the most abundant elements, H and He, are fully ionised. However, the assumption is no longer valid in cool stellar atmospheres).

Recall our definition (Lecture 1) of X, Y, Z as the mass fractions, respectively, of H, He and ‘everything else’. Thus, in a unit volume of density ρ , the mass of H is $X\rho$, that of He is $Y\rho$ and so on. In a fully ionised plasma, H will contribute two particles (one proton and one electron) per m_H . He will contribute 3/4 particles per m_H : two electrons and one alpha particle

(the He nucleus) which weighs $4m_H$. Heavier elements will in general give $1/2$ particles per m_H : fully ionised C, $7/12$, fully ionised O, $9/16$ and so on. Thus, the total number of particles per unit volume is:

$$n = \frac{2X\rho}{m_H} + \frac{3Y\rho}{4m_H} + \frac{Z\rho}{2m_H} \quad (9.9)$$

or

$$n = \frac{\rho}{4m_H} \cdot (8X + 3Y + 2Z). \quad (9.10)$$

Recalling that $X + Y + Z = 1$,

$$n = \frac{\rho}{m_H} \cdot \frac{6X + Y + 2}{4}. \quad (9.11)$$

and that the density is just $\rho = n \langle m \rangle = n m_H \mu$, we have:

$$\mu = \frac{4}{6X + Y + 2} \simeq 0.6 \quad (9.12)$$

for a fully ionised plasma of solar composition, where $X = 0.75$, $Y = 0.235$ and $Z = 0.015$.

9.2.2 Radiation Pressure

Einstein's relativistic energy equation

$$E^2 = p^2 c^2 + m^2 c^4 \quad (9.13)$$

expresses the total energy of a particle as the sum of its momentum p and rest-mass mc^2 contributions. Even though photons are massless particles, they do have a momentum $p = E/c = h\nu/c$ associated with their frequency $\nu = c/\lambda$. This momentum can be transferred to particles during absorption and scattering, that is photons can exert *radiation pressure*. In many astrophysical situations, the second term on the right-hand side of eq. 9.8 can be greater than the first term. In some cases, the force exerted by photons can even exceed the gravitational force, resulting in an overall expansion of the system. Radiation pressure is, for example, believed to be the force that triggers mass-loss through stellar winds in the most luminous stars.

For an isotropic radiation field of intensity I_λ , the radiation pressure per unit wavelength interval is simply:

$$P_{\text{rad},\lambda} d\lambda = \frac{1}{c} \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} I_\lambda d\lambda \cos^2 \theta \sin \theta d\theta d\phi = \frac{4\pi}{3c} I_\lambda d\lambda \quad (9.14)$$

and, for blackbody radiation:

$$P_{\text{rad}} = \frac{4\pi}{3c} \int_0^\infty B_\lambda(T) d\lambda \quad (9.15)$$

or

$$P_{\text{rad}} = \frac{4\sigma T^4}{3c} = \frac{1}{3} a T^4 \equiv \frac{1}{3} u. \quad (9.16)$$

Thus, the blackbody radiation pressure is one third of the energy density (see Lecture 5). For comparison, the pressure of an ideal monoatomic gas is two thirds of its energy density.

9.2.3 The Equations of Stellar Structure in Terms of Mass

As we saw earlier, eqs. 9.1–9.5 describe the run of mass, pressure, luminosity and temperature within a star with r , the distance from the centre of the star, as the independent variable. This is sometimes referred to as the formulation of the stellar structure equations in Euler coordinates.

Alternatively, we could express these equations in terms of the mass $m = M(r)$, sometimes referred to as Lagrange coordinates. A formulation in terms of mass makes more sense when we consider that the mass of a star remains approximately constant during most of the star's lifetime, whereas its radius can change by orders of magnitude.

The Lagrangian equations can be derived from the Euler equations by using:

$$\frac{dx}{dm} = \frac{dx}{dr} \cdot \frac{dr}{dm}$$

and

$$\frac{dr}{dm} = \frac{1}{4\pi r^2 \rho}$$

Thus we have:

1a. **Mass Continuity:**

$$\boxed{\frac{dr}{dm} = \frac{1}{4\pi r^2 \rho}} \quad (9.17)$$

2a. **Hydrostatic Equilibrium:**

$$\boxed{\frac{dP}{dm} = -G \frac{M_r}{4\pi r^4}} \quad (9.18)$$

3a. **Thermal Equilibrium:**

$$\boxed{\frac{dL}{dm} = \mathcal{E}} \quad (9.19)$$

4a. **Energy Transport:**

$$\boxed{\frac{dT}{dm} = -\frac{3}{4ac} \cdot \frac{\kappa}{T^3} \cdot \frac{L_r}{(4\pi r^2)^2}} \quad (9.20)$$

for radiative transport, and

$$\boxed{\frac{dT}{dm} = -\frac{\gamma-1}{\gamma} \frac{\mu m_H}{k} \frac{GM_r}{4\pi r^4 \rho}} \quad (9.21)$$

for convection.

9.3 Solving the Equations of Stellar Structure

9.3.1 Boundary Conditions

In order to solve these four differential equations (for either radiative or convective energy transport), we need to specify four boundary conditions. Boundary conditions are physical constraints to the mathematical equations defining the limits of integration. Two of the boundary conditions are obvious:

$$\left. \begin{array}{l} M_r = 0 \\ L_r = 0 \end{array} \right\} \text{ at } r = 0 \quad (9.22)$$

Physically, this is simply saying that there isn't a core of infinite density or negative luminosity at the centre of a spherically symmetric star.

Unfortunately, we do not have similar boundary conditions for pressure and temperature in the stellar core. However, we can place some boundary conditions at the stellar surface. If stars had clear-cut surfaces, then appropriate boundary conditions might be:

$$\left. \begin{array}{l} P = 0 \\ T = 0 \end{array} \right\} \text{ at } r = R_* \quad (9.23)$$

where R_* is the stellar radius, and indeed this is what is often adopted. The justification for this simplification is that $P(r = R_*) \ll \langle P \rangle$ and similarly, $T(r = R_*) \ll \langle T \rangle$. A more appropriate boundary condition for the temperature might be (eq. 2.15):

$$T(r = R_*) = \left[\frac{L}{4\pi\sigma R_*^2} \right]^{1/4} \quad (9.24)$$

An even better treatment would include a proper model atmosphere for the outer layers of the star.

Irrespectively of whether they are formulated in Eulerian or Lagrangian coordinates, the four independent equations of stellar structure cannot be solved analytically without making some simplifying assumptions. The reasons are that:

- (i) the equations are very non linear. The energy generation rate \mathcal{E} in the equation of thermal equilibrium (9.3 and 9.19) depends very steeply on temperature, $\mathcal{E} \propto \rho^\alpha T^\beta$ with $\beta \gg 1$, as we saw in Lecture 7.4. The opacity κ in the equation of radiative energy transport (9.4 and 9.20) is also a complicated function of ρ and T , as we saw in Lecture 5.3.
- (ii) The four differential equations are coupled and have to be solved simultaneously.
- (iii) There are two boundary conditions at $r = 0$ and two boundary conditions at $r = R_*$.

Thus, unless one makes some simplifying assumption,² the equations of stellar structure are normally integrated numerically. This is accomplished

²One such family of approximate solutions is known as polytropes, obtained by assuming that pressure and density are simply related through an equation of the form $P = K\rho^\gamma$, appropriate to an adiabatic gas.

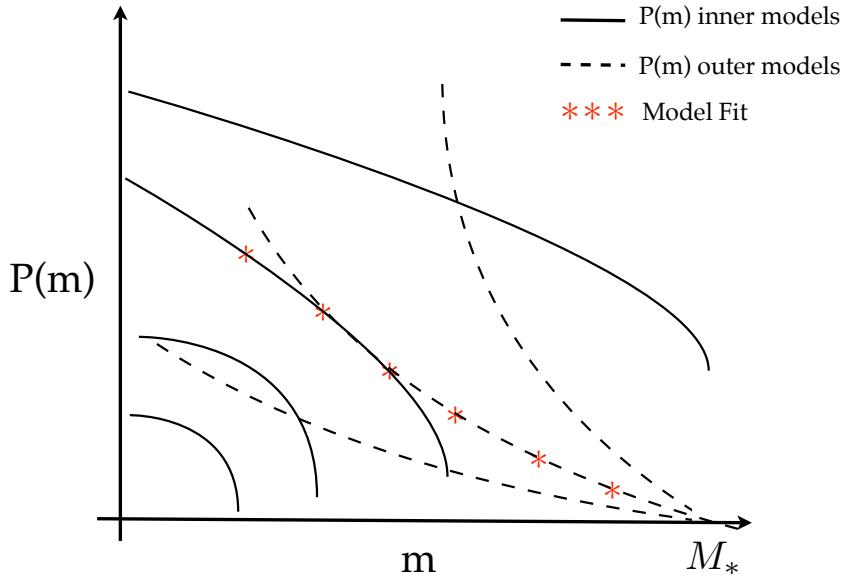


Figure 9.2: Illustration of M. Schwarzschild’s method to solve iteratively the equations of stellar structure.

by replacing the differential equations with difference equations and approximating the internal structure of a star by a series of concentric shells separated by a small, but finite, radial distance δr . For example, if the pressure in shell i is P_i , the pressure in the next shell is $P_{i+1} = P_i + (\Delta P / \Delta r) \delta r$.

The numerical integration is carried out from the centre to the surface using some initial estimates of $P(r = 0)$ and $T(r = 0)$ and, simultaneously, from the surface towards the centre adopting some initial values of L and R_* . Out of a large number of possible models, the condition for a satisfactory model is a smooth transition in the values of all the quantities, P , T , L and their derivatives at some transition radius r' —see Figure 9.2. This is the method developed by Martin Schwarzschild in 1958.

With modern computing facilities, the equations are solved iteratively with multidimensional matrices. Figures 9.3, 9.4 and 9.5 are examples of stellar models computed with modern numerical methods for stars on the Main Sequence of masses, respectively 1 , 3 , and $15M_\odot$. Comparison between the different curves gives us insight into the internal structure of stars of different masses.

Figure 9.6 is an example of an application of some of the best current stellar models, showing a fully theoretical Main Sequence, covering a wide range of stellar masses from $0.4M_\odot$ to $120M_\odot$, as indicated.

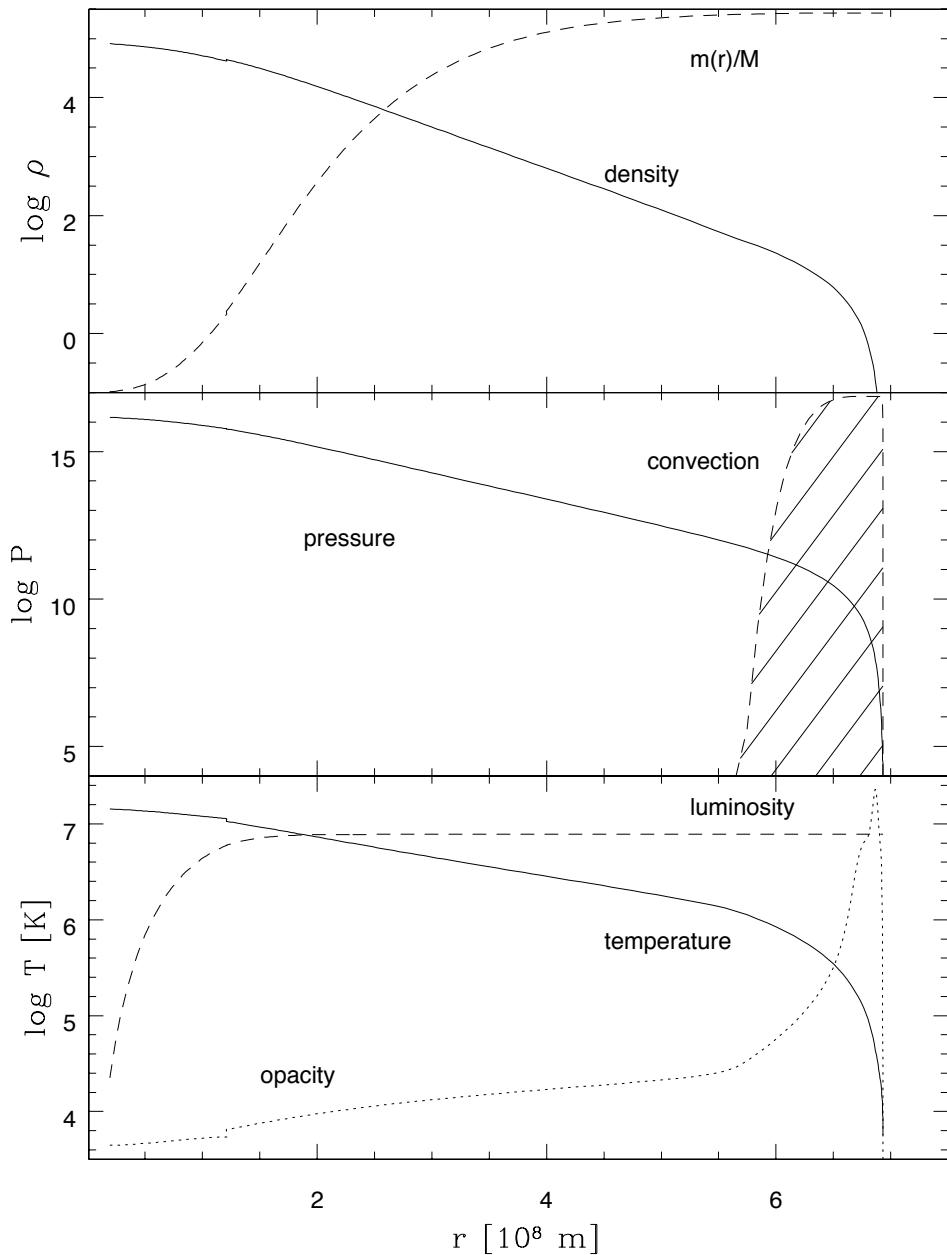


Figure 9.3: Internal structure of a $1M_{\odot}$ star computed with modern stellar models. (A. Zijlstra, private communication).

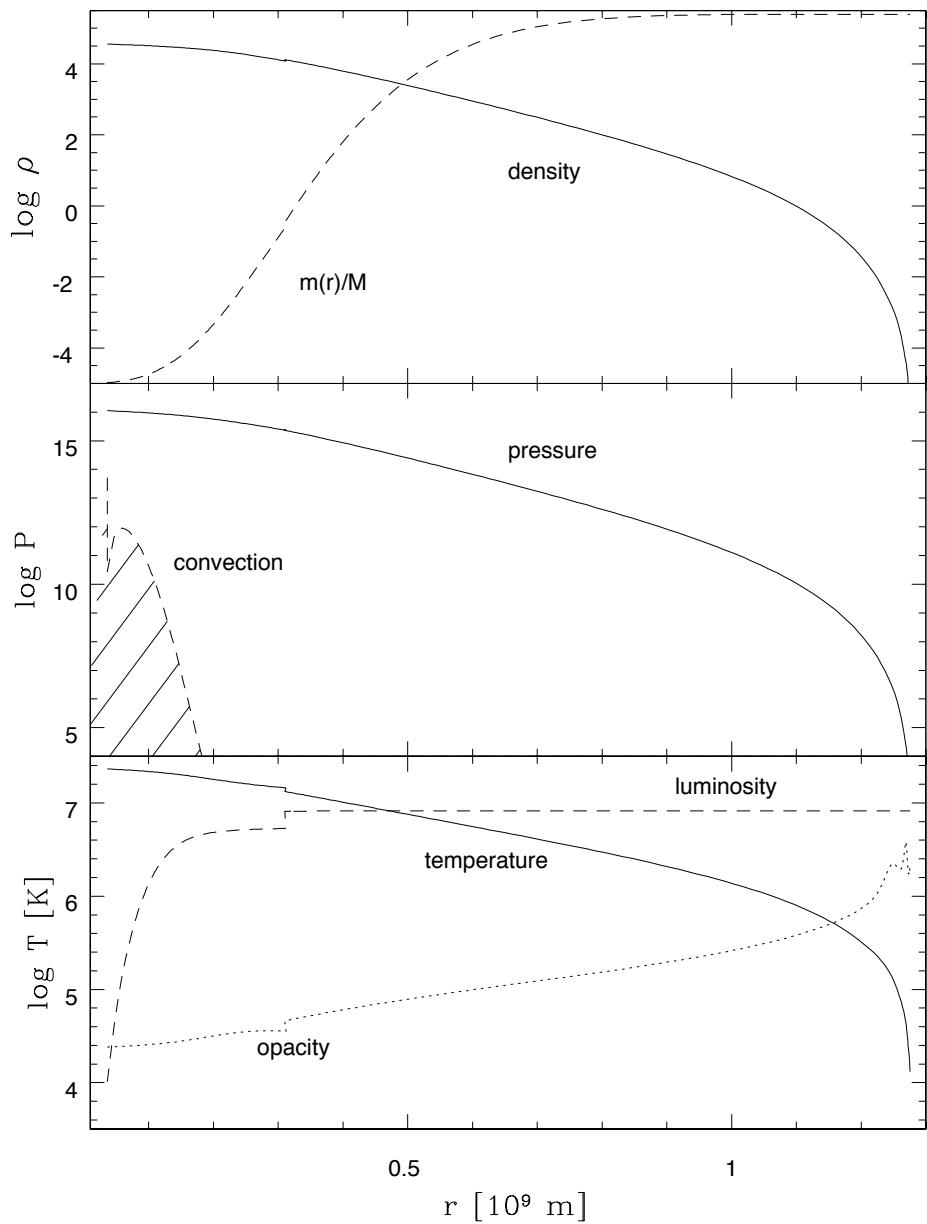


Figure 9.4: Internal structure of a $3M_{\odot}$ star computed with modern stellar models. (A. Zijlstra, private communication).

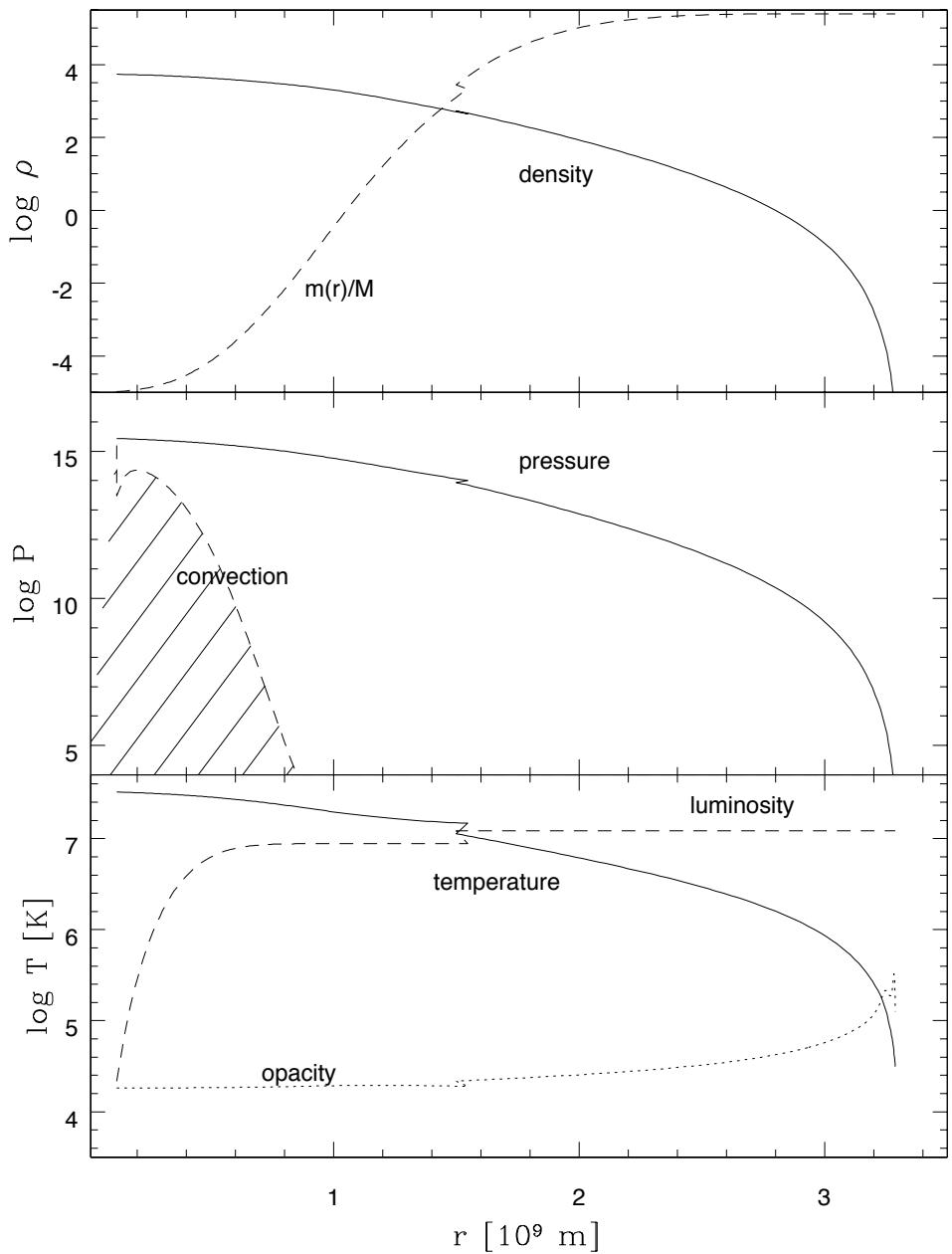


Figure 9.5: Internal structure of a $15M_{\odot}$ star computed with modern stellar models. (A. Zijlstra, private communication).

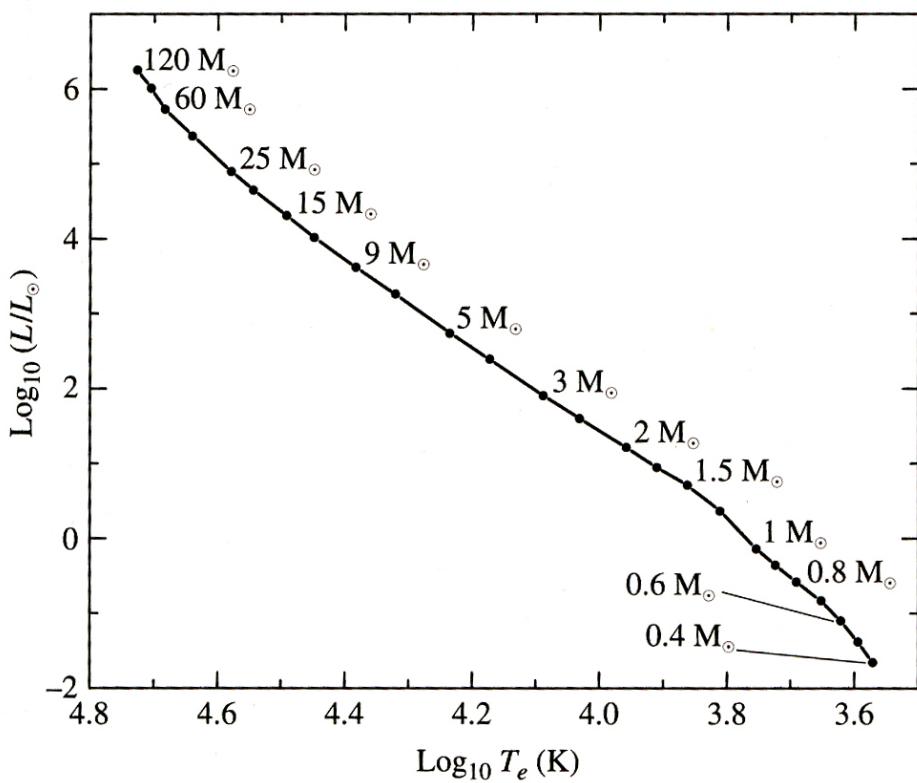


Figure 9.6: Theoretical Main Sequence computed with modern stellar models (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

STELLAR MODELS II: HOMOLOGY

10.1 Introduction

As we have seen in the previous lecture, there are no straightforward analytic solutions to the equations of stellar structure. Rather, the equations are solved iteratively with numerical methods, a process that can be quite consuming in terms of computing resources. However, one important property of the equations is that they are *homologous*. Homology means that, given a solution to the equations—whereby the quantities (P , T , L , and ρ) are specified as a function of the radial coordinate r (in the Eulerian formulation) or as a function of m (Lagrangian formulation) for a star of total mass M and of a given chemical composition (this is what constitutes a stellar model), then we can find a new solution for a star of a different total mass M' simply by multiplying the other physical variables by appropriate scaling factors.

This approach is adequate if, for example, we are interested in reproducing the stellar mass-luminosity relationship (Figure 4.8), or the Main Sequence of hydrogen burning stars in the luminosity-temperature diagram (Figure 3.6). Of course, the implicit assumption in homologous stellar models is that stars have the same chemical composition, but this is approximately the case for most stars on the Main Sequence in the solar neighbourhood.

Physically, in two stars related by a homology transformation we assume that the *way* in which a physical quantity varies from the centre of a star to its surface is the same for all stars, irrespectively of the total stellar mass. For example, the increase of luminosity with radius reduces to a single curve (Figure 10.1) when plotted as a function of the *fractional* mass $x = m(r)/M$. In this way, any stellar model is related to an initial (or reference) one by a simple change in scale.

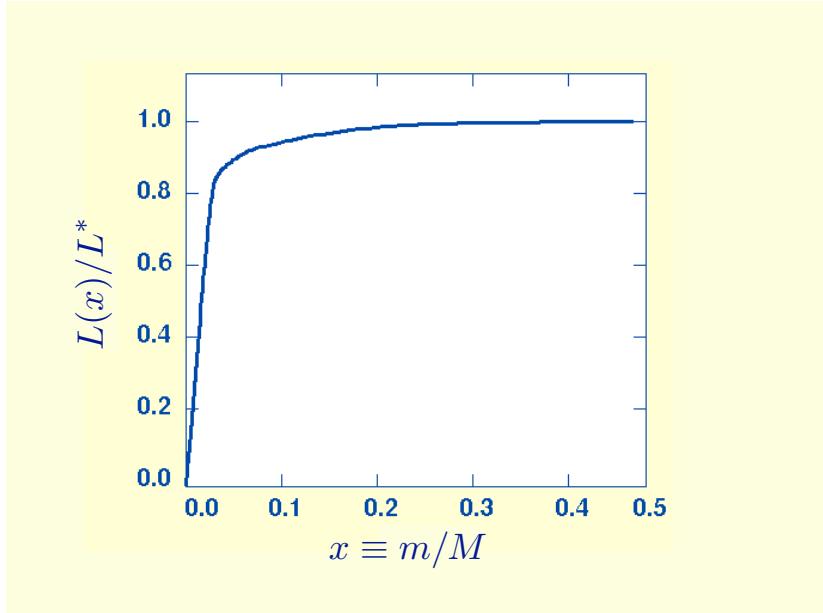


Figure 10.1: In homologous stellar models the growth of fractional luminosity with fractional stellar mass is the same for all stars, irrespectively of a star’s total luminosity L and total mass M .

10.2 An example

Consider two stars with masses M_1 and M_2 and radii R_1 and R_2 . We define:

$$x' = \frac{r_1}{R_1} = \frac{r_2}{R_2}. \quad (10.1)$$

Then,

$$x \equiv \frac{m_1(r_1)}{M_1} = \frac{m_2(r_2)}{M_2} \quad (10.2)$$

expresses the fact that each star contains the same fraction x of its mass within the same fraction x' of its radius.

Let us now consider the homology transformation of the first equation of stellar structure, the equation of mass continuity. In Euler coordinates, we have for star 1:

$$\frac{dm_1}{dr_1} = 4\pi r_1^2 \rho_1 \quad (10.3)$$

which may be transformed for star 2 as:

$$\frac{dm_2}{dr_2} = \frac{M_2}{M_1} \frac{dm_1}{dr_2} = \frac{M_2}{M_1} \frac{R_1}{R_2} \frac{dm_1}{dr_1}. \quad (10.4)$$

But we also have:

$$\frac{dm_2}{dr_2} = 4\pi r_2^2 \rho_2 = 4\pi \left(\frac{R_2}{R_1}\right)^2 r_1^2 \rho_2. \quad (10.5)$$

Equating the r.h.s. of the last two equations, we have:

$$\frac{dm_1}{dr_1} = 4\pi r_1^2 \frac{M_1}{M_2} \left(\frac{R_2}{R_1}\right)^3 \rho_2 \quad (10.6)$$

Comparing 10.3 and 10.6, it can be seen immediately that the density of star 2 at x can be obtained by scaling the density at x of star 1:

$$\rho_2(x) = \left(\frac{M_2}{M_1}\right) \left(\frac{R_1}{R_2}\right)^3 \rho_1(x). \quad (10.7)$$

10.3 Homologous Transformations

More generally (and working now in Lagrangian coordinates), we can replace the solutions to the four equations of stellar structure plus the constitutive equation for density, $r(m)$, $P(m)$, $L(m)$, $T(m)$ and $\rho(m)$, with five relations as follows:

$$\begin{aligned} r &= f_1(x) \cdot R^* \\ P &= f_2(x) \cdot P^* \\ L &= f_3(x) \cdot L^* \\ T &= f_4(x) \cdot T^* \\ \rho &= f_5(x) \cdot \rho^* \end{aligned} \quad (10.8)$$

where R^* , P^* , L^* , T^* , ρ^* are dimensional coefficients for, respectively, the radius, pressure, luminosity, temperature and density, and the $f_n(x)$ are the (dimensionless) scaling functions. By definition, $f_n(x) = 0$ –1; with f_1 and $f_3 = 1$ when $x = 1$, and f_2 , f_4 , and $f_5 = 1$ when $x = 0$.

Consider the equation of hydrostatic equilibrium:

$$\frac{dP}{dm} = -\frac{Gm}{4\pi r^4} \quad (10.9)$$

With the substitutions: $dP = df_2 \cdot P^*$, $dm = dx \cdot M$, $m = x \cdot M$, and $r = f_1 \cdot R^*$, it becomes:

$$\frac{df_2}{dx} \cdot \frac{P^*}{M} = -\frac{GMx}{4\pi f_1^4 (R^*)^4}$$

or

$$\frac{df_2}{dx} = -\frac{x}{4\pi f_1^4} \cdot \frac{GM^2}{(R^*)^4 P^*}. \quad (10.10)$$

The advantage of this formulation is that it clearly separates the stellar structure aspect (the first term on the r.h.s. of the equation) from the scaling between stars of different masses (the second term on the r.h.s.). Recalling that the f_n functions are dimensionless, all the dimensions are in the second term on the r.h.s., which therefore implies:

$$P^* = \frac{GM^2}{(R^*)^4}$$

Repeating this procedure for the four equations of stellar structure (ignoring convection and radiation pressure), we have five pairs of equations, as follows:

$$\frac{df_1}{dx} = \frac{1}{4\pi f_1^2 f_5}; \quad \rho^* = \frac{M}{(R^*)^3} \quad (10.11)$$

$$\frac{df_2}{dx} = -\frac{x}{4\pi f_1^4}; \quad P^* = \frac{GM^2}{(R^*)^4} \quad (10.12)$$

$$\frac{df_3}{dx} = f_5 f_4^\beta; \quad L^* = \mathcal{E}_0 \rho^* (T^*)^\beta M \quad (10.13)$$

$$\frac{df_4}{dx} = -\frac{3f_3}{4f_4^3 (4\pi f_1^2)^2}; \quad L^* = \frac{ac}{\kappa} \frac{(T^*)^4 (R^*)^4}{M} \quad (10.14)$$

$$f_5 = \frac{f_2}{f_4}; \quad T^* = \frac{\mu m_H}{k} \frac{P^*}{\rho^*} \quad (10.15)$$

Note that the equations on the l.h.s. constitute a set of nonlinear differential equations which are independent of M . On the r.h.s. we have five algebraic equations which relate the dimensional coefficients R^* , P^* , ...; these coefficients can be manipulated to obtain the dependence on M of the property of interest. Some pertinent examples follow.

10.3.1 Mass-Luminosity Relation

Substituting the expressions for ρ^* and P^* from 10.11 and 10.12 into 10.15, we obtain:

$$T^* = \frac{\mu m_H}{k} \frac{GM^2}{(R^*)^4} \frac{(R^*)^3}{M} = \frac{\mu m_H}{k} \frac{GM}{R^*} \quad (10.16)$$

Using this expression for T^* in eq. 10.14, we have:

$$L^* = \frac{ac}{\kappa} \left(\frac{\mu m_H G}{k} \right)^4 \cdot \frac{M^4 (R^*)^4}{M(R^*)^4} \quad (10.17)$$

The above equation is valid at any value of x (the fractional mass). But since $f_3(x) = 1$, when $x = 1$, we obtain the mass-luminosity relationship:¹

$$L \propto M^3 \quad (10.18)$$

Equation 10.18 is not a bad approximation to the real $M - L$ relationship we found in Lecture 4 and shown again here in Figure 10.2, considering that we have ignored convection and radiation pressure.¹

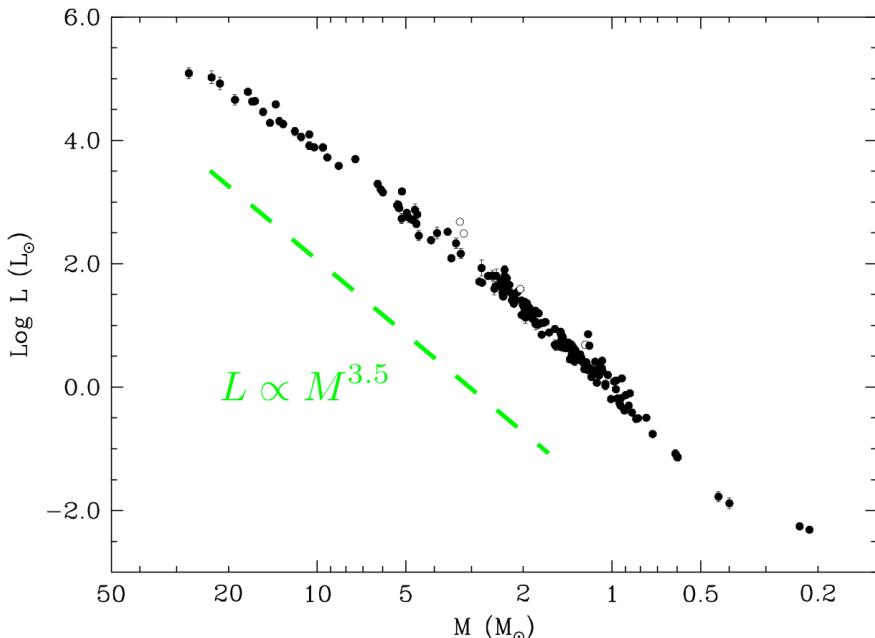


Figure 10.2: The empirical stellar mass-luminosity relation constructed from observations of 190 binary stars with well-determined parameters. (Reproduced from Torres et al. 2010).

Note that there is a real spread to the M-L relation, better seen in Figure 10.3: at each value of mass there a spread in luminosity which is much larger than the observational error. The spread is thought to be due to stellar evolution *during the main sequence lifetime*, as well as to the fact that not all the stars in the plot have the same chemical composition.

¹Note also that in going from 10.17 to 10.18 we have treated the opacity κ as a constant. For simplicity, we shall assume $\kappa \approx \text{const}$ throughout section 10.3 even though, as we saw in Lecture 5, the mean opacity follows a Kramers' law: $\langle \kappa \rangle \propto \rho T^{-3.5}$ (see Figure 5.5). A more rigorous treatment would include the density and temperature dependence of the opacity in the scaling relations.

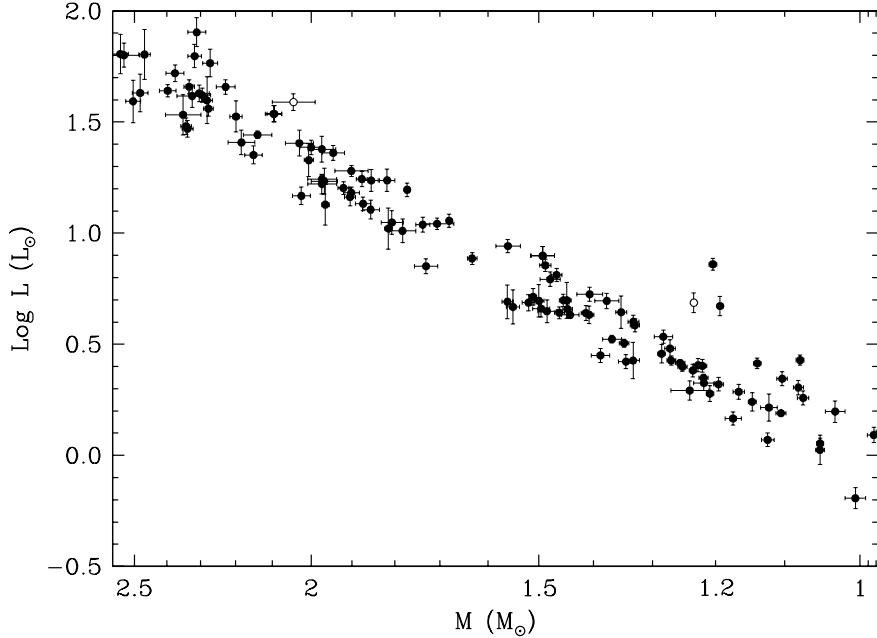


Figure 10.3: Close-up of the $1-2.5M_{\odot}$ range of the mass-luminosity relation of Torres et al. 2010. The very significant (in the sense that it is many times the typical error) scatter in $\log L$ at each mass value is due to the combined effects of stellar evolution and abundance differences. Stars classified as giants are shown with open circles. (Reproduced from Torres et al. 2010).

10.3.2 Mass-Radius Relation

We have written down two equations for the luminosity:

$$L^* = \frac{ac}{\kappa} \left(\frac{\mu m_H G}{k} \right)^4 M^3; \quad L^* = \mathcal{E}_0 \rho^* (T^*)^{\beta} M; \quad (10.19)$$

therefore:

$$\rho^* = \frac{M^2}{(T^*)^{\beta}} \cdot \frac{ac}{\kappa \mathcal{E}_0} \cdot \left(\frac{\mu m_H G}{k} \right)^4, \text{ or} \quad \rho^* \propto \frac{M^2}{(T^*)^{\beta}} \quad (10.20)$$

We also have:

$$T^* \propto \frac{P^*}{\rho^*}, \quad T^* \propto \frac{M^2}{(R^*)^4} \cdot \frac{(R^*)^3}{M}, \quad T^* \propto \frac{M}{R^*} \quad (10.21)$$

using 10.11 and 10.12. Also, from 10.11, we have:

$$(R^*)^3 \propto \frac{M}{\rho^*}, \quad (R^*)^3 \propto M \cdot \frac{(T^*)^{\beta}}{M^2}, \quad (R^*)^3 \propto \frac{(T^*)^{\beta}}{M} \quad (10.22)$$

Thus,

$$(R^*)^3 \propto \frac{M^{\beta}}{M} \cdot \frac{1}{(R^*)^{\beta}} \quad (10.23)$$

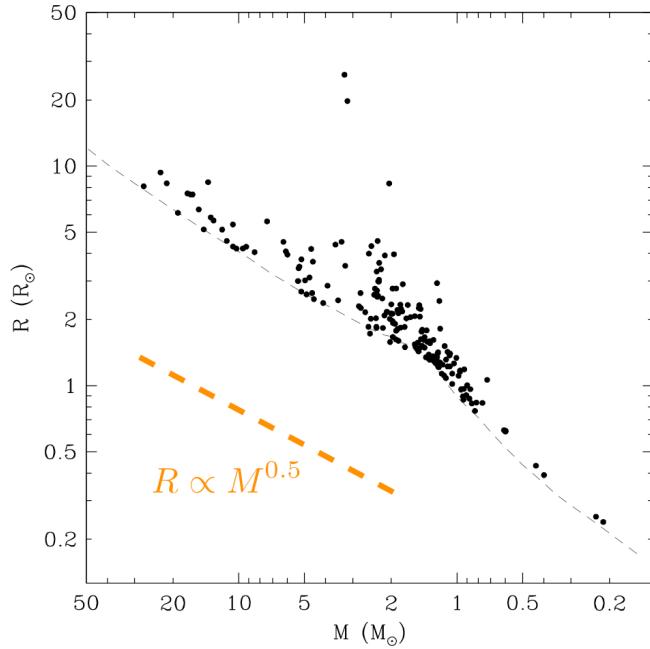


Figure 10.4: Mass-radius relation for 190 stars whose masses and radii are known to better than 3%. The faint dashed line shows a theoretical M - R relation for a zero-age Main Sequence (ZAMS) of solar metallicity, computed from a full set of stellar models. (Reproduced from Torres et al. 2010).

which gives us the mass-radius relation:

$$R \propto M^{(\beta-1)/\beta+3} \quad (10.24)$$

Equation 10.24 tells us that the scaling of stellar radius with mass depends on the exponent β of the power-law dependence on temperature of the rate of energy generation per unit mass of nuclear fuel (eq. 7.25). We saw in Lecture 7.4.1 that for stars burning hydrogen into helium via the p-p chain, $\beta = 4$. Thus for these stars:

$$R \propto M^{3/7}$$

High mass stars burn H into He via the CNO cycle with $\beta = 17$, leading to:

$$R \propto M^{16/20}$$

These scaling are not too dissimilar from the empirical mass-radius relation derived for the same set of well observed stars in binary systems already considered in section 10.3.1, from the compilation by Torres et al. 2010 (A&ARv, 18, 67), as can be seen from Figure 10.4.

Since the density is just mass over volume, we have the scaling of density with mass:

$$\rho^* = \frac{M}{(R^*)^3}, \quad \rho^* \propto M^{(6-2\beta/\beta+3)}$$

from which we see that more massive stars have lower densities for $\beta > 3$, i.e. even for hydrogen burning via the p-p chain.

10.3.3 Luminosity-Temperature Relation

With the mass-luminosity and mass-radius relations just derived we can now make an approximate prediction for the slope of the Main Sequence in the Luminosity-Temperature relation (i.e. the H-R diagram). Recalling that:

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4 \quad (10.25)$$

we obtain, combining 10.25, 10.18 and 10.24:

$$L^{1-\frac{2(\beta-1)}{3(\beta+3)}} \propto T_{\text{eff}}^4 \quad (10.26)$$

which implies

$$\log L = 5.6 \log T_{\text{eff}} + C \quad (10.27)$$

for stars burning H into He via the p-p chain, and

$$\log L = 8.6 \log T_{\text{eff}} + C' \quad (10.28)$$

for more massive stars burning H via the CNO cycle.

Again, these are in reasonable agreement with the approximate double power-law slope of the Main Sequence for stars with masses $M \gtrsim 1M_\odot$ (see Figure 10.5).

10.4 Minimum Stellar Mass

We can use the homologous transformations developed in Section 10.3 to estimate the mass limits of the Main Sequence. The low mass limit is determined by the minimum temperature required to ignite hydrogen fusion via the p-p chain, while the upper mass limit is set by radiation pressure.

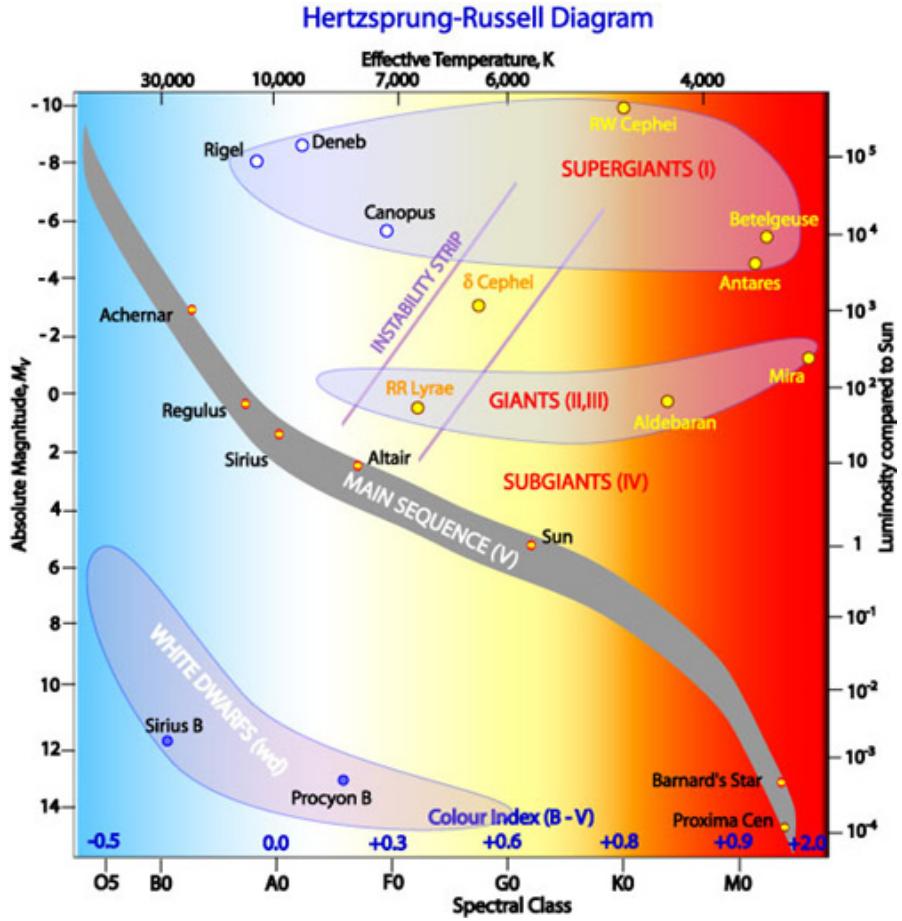


Figure 10.5: Schematic representation of the H-R diagram.

The approximate minimum temperature required for nuclear reactions to take place is $T \simeq 4 \times 10^6$ K; at lower temperatures, the protons kinetic energy is too low to bring sufficient numbers of them close enough for quantum mechanical tunneling of the Coulomb barrier to take place (Section 7.2.1).

Details models of the Sun indicate that its central temperature, $T_c \equiv T(x = 0) = 1.5 \times 10^7$ K. Using eq. 10.16 together with our mass-radius relation (10.24), we have:

$$T^* = \frac{\mu m_H}{k} \frac{GM}{R^*}; \quad T^* \propto \frac{GM}{M^{(\beta-1)/\beta+3}}; \quad T^* \propto M^{1-\frac{\beta-1}{\beta+3}}$$

or

$$T^* \propto M^{4/7}$$

for the p-p chain ($\beta = 4$). Thus,

$$M_{\min} = M_{\odot} \cdot \left(\frac{4}{15} \right)^{7/4} \simeq 0.1 M_{\odot} \quad (10.29)$$

Such a star will have a luminosity of only $L_{\min} \simeq 10^{-3}L_{\odot}$, given the $L \propto M^3$ scaling (Section 10.3.1). Furthermore, using the mass-radius relation $R \propto M^{3/7}$ appropriate to the p-p chain (eq. 10.24), together with eq. 10.25, we have:

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4 \Rightarrow M^3 \propto M^{6/7} T_{\text{eff}}^4 \quad (10.30)$$

Hence:

$$T_{\text{eff}} \propto M^{15/28} \quad (10.31)$$

and

$$T_{\text{eff,min}} = T_{\text{eff},\odot} \cdot \left(\frac{M_{\min}}{M_{\odot}} \right)^{15/28} \simeq 5770 \times 0.1^{15/28} \simeq 1700 \text{ K} \quad (10.32)$$

In summary, scaling from the solar parameters, the low mass end of the Main Sequence of hydrogen burning stars is expected to occur at $L \sim 10^{-3}L_{\odot}$, $T_{\text{eff}} \sim 1700 \text{ K}$. The value of L_{\min} thus derived matches observations, while that of $T_{\text{eff,min}}$ is about a factor of two too low (see Figure 10.5). One reason for the discrepancy may be the fact that, as we saw in Lecture 8.3.1, such low mass stars are almost entirely convective, and we have not included convection in our homologous transformations.

An example of a star near the low mass limit of the hydrogen burning Main Sequence is the red dwarf star Wolf 359, in the constellation of Leo (see Figure 10.6). At a distance of only 2.4 pc, Wolf 359 is the fifth closest star to the Sun, after the three stars that make up the α Centauri system and Barnard's star. Its spectral type is M6.0 V and its luminosity is $0.001L_{\odot}$;

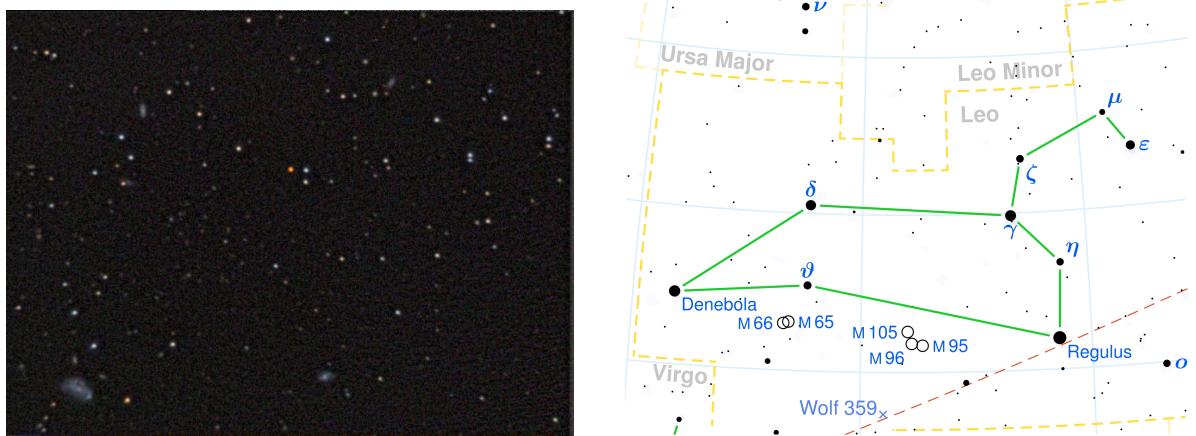


Figure 10.6: The red dwarf star Wolf 359 in the constellation of Leo is the orange-red object just above centre in the left-hand image.

if it were placed at the Sun’s distance, it would be only about ten times brighter than the full Moon. The mass of Wolf 359 is $M = 0.09M_{\odot}$, just above the limit of $0.08M_{\odot}$ (derived more carefully than the homologous transformation above) for hydrogen fusion via the p-p chain. Its effective temperature is $T_{\text{eff}} = 2800 \pm 100$ K. Despite its proximity, Wolf 359 is a faint object at visible wavelengths, although it is brighter in the near-infrared: $V = 13.5$, but $J = 7.1$.

10.5 Eddington Luminosity

We saw in Section 9.2.2 that photons can exert pressure on the gas, so that the total pressure is the sum of two components:

$$P = P_g + P_{\text{rad}} = \frac{\rho k T}{\mu m_H} + \frac{1}{3} a T^4 \quad (10.33)$$

This equation shows that if the temperature is sufficiently high and the density is sufficiently low, radiation pressure can dominate over the gas pressure—such a situation can occur in the outer layers of very massive stars. We also saw in Lecture 8.4.1 that it is the pressure gradient within a star that counteracts the force of gravity and stops a star from collapsing onto itself.

The Cambridge astrophysicist Arthur Eddington, Plumian professor and director of the Cambridge Observatory from 1914, understood that there is a limit to the luminosity that a star can attain, beyond which the radiation pressure would exert an outward force *greater* than the inward force of gravity, rendering the star highly unstable. This upper limit to a star’s luminosity is called the Eddington limit.

For a star to maintain hydrostatic equilibrium, we require:

$$\left| \frac{dP_{\text{rad}}}{dr} \right| < G \frac{M_r \rho}{r^2} \quad (10.34)$$

If radiation pressure dominates:

$$\begin{aligned} P_{\text{rad}} &= \frac{1}{3} a T^4 \\ \frac{dP_{\text{rad}}}{dT} &= \frac{4}{3} a T^3 \end{aligned} \quad (10.35)$$

Recalling the Eddington equation for radiative equilibrium:

$$\frac{dT}{dr} = -\frac{3}{4} \cdot \frac{1}{ac} \cdot \frac{\kappa\rho}{T^3} \cdot \frac{L_r}{4\pi r^2} \quad (10.36)$$

we have:

$$\frac{dP_{\text{rad}}}{dr} = -\frac{\kappa\rho}{c} \frac{L}{4\pi r^2} \quad (10.37)$$

But we also have, from the equation of hydrostatic equilibrium:

$$\frac{dP}{dr} = -G \frac{M_r \rho}{r^2} \quad (10.38)$$

Equating 10.37 and 10.38, we find:

$$L_{\text{edd}} = 4\pi c G \frac{M}{\kappa} \quad (10.39)$$

A star of this luminosity is supported by radiation pressure alone! And if this luminosity is exceeded, material can be peeled off the surface of the star by the transfer of momentum from photons.

The influence of radiation pressure is often expressed in terms of the *Eddington factor*:

$$\Gamma_{\text{edd}} \equiv \frac{L}{L_{\text{edd}}} = \frac{\kappa L}{4\pi c G M} \quad (10.40)$$

10.6 Maximum Stellar Mass

If the gas is fully ionised, the main source of opacity is electron scattering (Lecture 5.4). In Lecture 5.5, the point was made that electron scattering has no wavelength, density, nor temperature dependence, so that κ_{es} takes a particularly simple form: $\kappa_{es} = 0.2(1 + X) \text{ cm}^2 \text{ g}^{-1}$, where X is the mass fraction of hydrogen. Entering the values of the physical constants in eq. 10.39, we now have:

$$\begin{aligned} \frac{L_{\text{edd}}}{L_{\odot}} &= \frac{1.3 \times 10^4}{\kappa_{es}} \frac{M}{M_{\odot}} \\ &= \frac{1.3 \times 10^4}{0.20(1 + 0.7)} \frac{M}{M_{\odot}} \\ &\simeq 3.8 \times 10^4 \frac{M}{M_{\odot}} \end{aligned} \quad (10.41)$$

We can now use our homology transformation $L \propto M^3$, to find the maximum stellar mass:

$$M_{\max} = \sqrt{3.8 \times 10^4} L_{\odot} \simeq 200 M_{\odot} \quad (10.42)$$

Such a star will have a luminosity:

$$L_{\max} = 7.4 \times 10^6 L_{\odot} \quad (10.43)$$

Empirically, however, it appears that $M \simeq 100\text{--}120 M_{\odot}$ is a more realistic upper limit to the stellar mass (see Figure 10.7). The observed luminosity limit appears to *decrease* with decreasing effective temperature for stars with $T_{\text{eff}} \gtrsim 10\,000$ K and then remain approximately constant at $\log L/L_{\odot} \simeq 5.7$. Lamers and Fitzpatrick (1988) understood this behaviour in terms of the “Photospheric Eddington Limit”. The limits deduced above apply to the case when electron scattering is the only source of opacity. But in the atmospheres of even the hottest stars, not all elements are fully ionised. Including the full effects of metal line opacities (bound-bound transitions), shows that $\langle \kappa \rangle$ *increases* from $T_{\text{eff}} = 50\,000$ K to 10 000 K, and then drops steeply to lower T_{eff} . Once this behaviour of the full opacity is taken into account, there is excellent agreement between the Humphreys-Davidson empirical limit and the modified (or atmospheric) Eddington limit (see Figure 10.8).

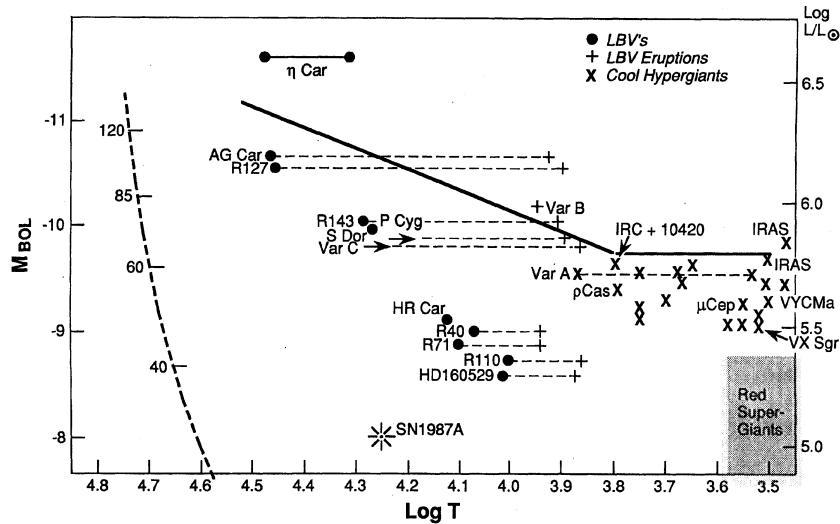


Figure 10.7: A schematic HR diagram for the most luminous known stars. The continuous black line shows the empirical upper luminosity boundary known as the Humphreys-Davidson limit. (Figure reproduced from Humphreys & Davidson 1994, PASP, 106, 1025).

As one might expect, the boundary set by the photospheric Eddington limit changes with metallicity.

10.6.1 Luminous Blue Variables

Luminous Blue Variable, or LBV, is a term coined by Peter Conti in 1984, to describe a class of very rare (only about 20 are known), extremely luminous blue stars which undergo outbursts, increasing their visual brightness by 1–2 magnitudes and experiencing significant mass loss, at a rate that can be as high as 10^{-2} – $10^{-1} M_{\odot} \text{ yr}^{-1}$. Such outbursts occur on timescales of a few decades or, in some cases, centuries, making our census of such stars very incomplete. This, coupled with their presumably very short lifetimes, explains the paucity of objects in this rare class.

During their quiescent periods, LBVs are found just on the hotter side (to the left in Figure 10.8) of the Humphreys-Davidson limit, unstable because they are near their (atmospheric) Eddington limits. However, as these stars cool and approach the Eddington limit they are not totally

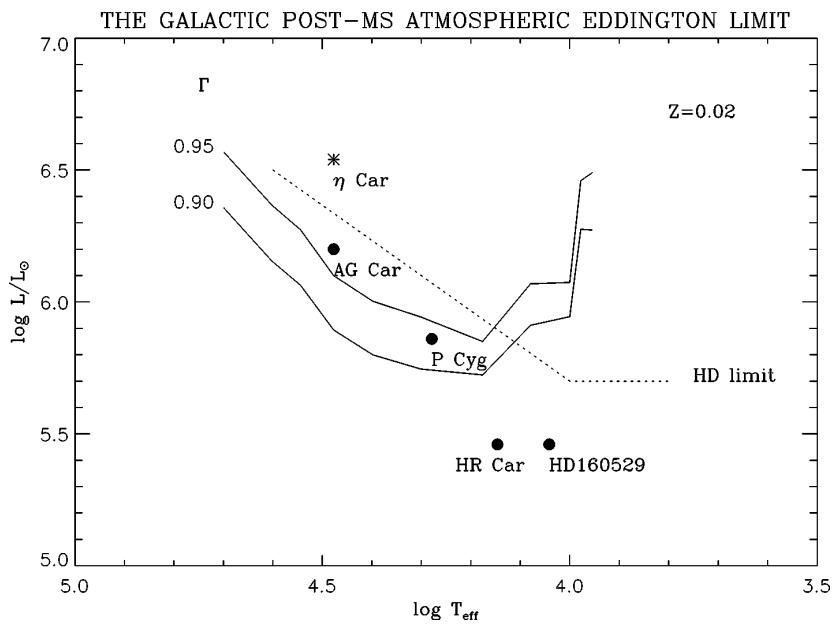


Figure 10.8: The empirical luminosity limit for hot stars found by Humphrey & Davidson, 1984 (dotted line) is well matched by the atmospheric Eddington limit (continuous lines). The parameter Γ is the Eddington factor given in eq. 10.40. The locations of a few well-known Luminous Blue Variables are indicated. (Figure reproduced from Massey 2003, ARA&A, 41, 15)

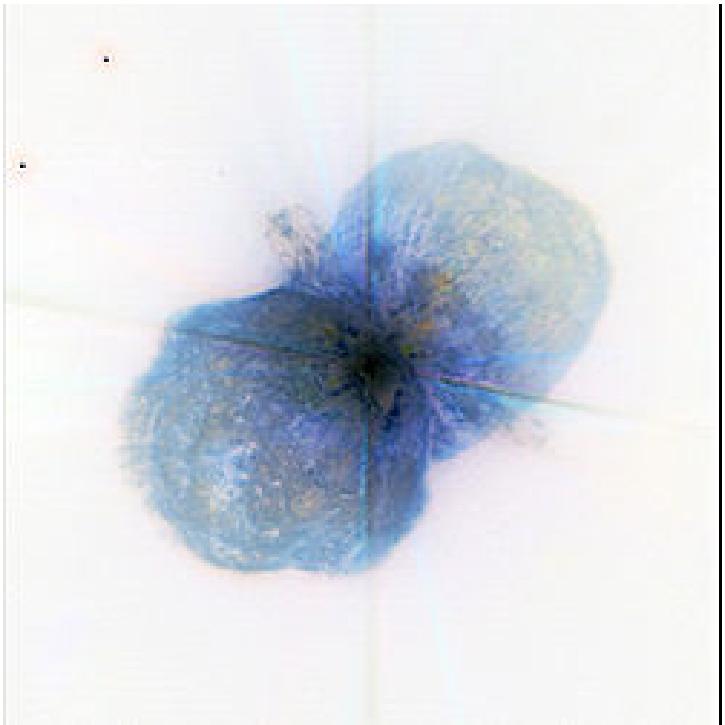


Figure 10.9: The star η Carinae is thought to be the closest example of a Luminous Blue Variable. It has recently been realised that it actually consists of at least two stars. η Carinae has undergone several periods of relative brightening and dimming in historical times; the last major outburst occurred in the 1840s, when the star reached visible magnitude $V = -0.8$, second only to Sirius (α CMa) which is ~ 1000 times nearer to the Earth! The 1940s outburst, during which the star shed more than $1M_{\odot}$, is thought to have created the Homunculus Nebula, seen in this spectacular image. Note the bipolarity of the outflow and the jets bisecting the lobes emanating from the central star. η Car still undergoes unexpected outbursts; its high mass and volatility make it a candidate to explode as a supernova sometime in the next few million years.

disrupted (as would be the case if they exceeded their classical Eddington limit), because it is only in the outer layers of the star that line opacity is effective in transferring momentum from the radiation field to matter.

The frequent (on an astronomical timescale) eruptions undergone by LBVs are recorded in the nebulae that surround them. The closest, and best studied, example is the star η Carinae in the Great Carina nebula, one of the regions of most intense star formation in the Milky Way, containing some of the most massive stars known (see Figure 10.9).

This concludes the part of the course which deals with stellar structure. In the next lectures we shall look in more detail at how stars form and at how they evolve during their lifetimes.

HOW DO STARS FORM?

11.1 Introduction

The last decades have seen great strides forward in our understanding of the processes that govern the formation and evolution of stars. Progress has been made on both observational and theoretical fronts, driven by improvements in instrumentation (primarily infrared) and in computational resources. Indeed, it is now possible to ‘follow’ the process of star formation in sophisticated numerical simulations of increasing spatial resolution.



Figure 11.1: At a distance of 7.4 Mpc from the Milky Way, M101 (the Pinwheel galaxy) is a ‘grand-design’ spiral galaxy seen nearly face on ($i = 18^\circ$). This picture is a composite of three images, each taken with one of the NASA’s ‘Great Observatories’. Visible light (yellow) captured by the Hubble Space Telescope traces the general stellar population. Infrared light (red) recorded with the Spitzer Space Telescope is emitted by hot dust in molecular clouds. X-ray light (blue) seen by the Chandra X-ray Observatory marks the locations of high-energy gas in the vicinity of young, massive stars. Note the close correspondence between X-ray emitting regions of recent star formation and dust clouds.

Despite such advances, star formation is still considered to be a poorly-understood problem in astrophysics. The reason is that we do not yet have a *predictive* theory of star formation; for example, given a set of initial conditions, we do not yet have the means to predict reliably important properties, such as the star formation efficiency (that is, what fraction of the gas is turned into stars), or the Initial Mass Function which describes the relative numbers of stars of different masses. On the other hand, once a star has formed and nuclear burning starts, all the uncertain details of the star formation process are no longer relevant to its evolution. In this lecture, we explore some of the basic ideas concerning the way stars form.

Observationally, we know that there is a strong spatial association between clusters of newly formed stars and interstellar gas. Both are concentrated in spiral arms in external galaxies (see Figure 11.1 for an example); in the Milky Way, regions of recent star formation such as the Great Carina Nebula (Figure 11.2) are also some of most spectacular visualisations of the interstellar medium.

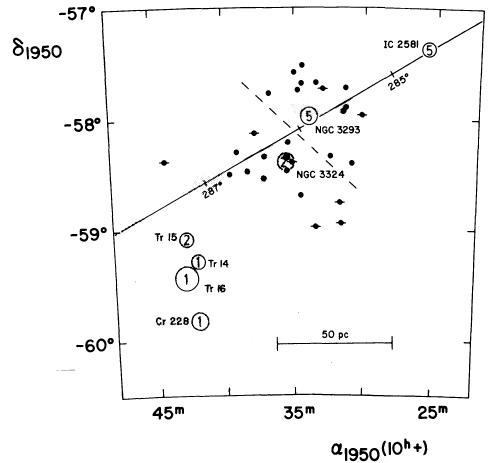


Figure 11.2: Located ~ 2.5 kpc from the Sun, the Giant Carina Nebula is one of the Galaxy's largest star-forming regions. It is rich in clusters of newly formed O and B stars, harbouring a significant fraction of the most massive stars known in the Milky Way, including several O3 V stars and the LBV η Carinae. The picture on the left highlights the interstellar matter in the nebula, some glowing in the light of newly-formed stars, while other regions are dense clouds of molecules and dust. The sketch on the right, reproduced from Turner et al. (1980), shows the locations and ages (in millions of years) of the most prominent star clusters. Star formation is propagating from the north-west to the south-east of the region at a speed of ~ 30 km s $^{-1}$, attesting to the interaction between massive stars and the ambient interstellar medium. The continuous line is the Galactic equator with intervals of Galactic longitude marked.

The interstellar medium is a complex environment, worthy of its own lecture course. For our present purposes, suffice to say that among the various components of the ISM, it is the Giant Molecular Clouds that are associated with sites of star formation. Images taken at near-infrared wavelengths ($\lambda = 1\text{--}2.2\,\mu\text{m}$), show very young clusters and protostellar objects still embedded deep into such clouds, before the ultraviolet radiation emitted by the most massive stars evaporates the cloud away.

Giant molecular clouds, of which there are thousands in the Milky Way, are enormous complexes of dust and gas, sufficiently dense to be self-shielding from the diffuse interstellar ultraviolet radiation field. Typical parameters are temperatures $T \sim 15\,\text{K}$, volume densities $n = 100\text{--}300\,\text{cm}^{-3}$, and masses $M = 10^5\text{--}10^6 M_\odot$. The low temperatures and high densities (and the presence of solid particles known as interstellar dust which acts as a catalyst) favour the formation of molecules; thus most of hydrogen in GMCs is in the form of H_2 rather than $\text{H}\,\text{I}$. Dust makes up about 1% of the material, and it is the dust that renders the clouds very opaque at visual wavelengths.

With typical sizes of 10s of parsecs, GMCs show considerable structure on a variety of scales. On scales of 10 pc, $M \sim 10^4 M_\odot$, densities can be $n \sim 500\,\text{cm}^{-3}$ and the extinction at visible wavelengths is $A_V \sim 5\,\text{mag}$; such regions are sometimes referred to as Dark Cloud Complexes. Smaller clumps on scales of 1–2 pc with $M \sim 30 M_\odot$ can have $n \sim 1000\,\text{cm}^{-3}$, $A_V \sim 10\,\text{mag}$, and $T \sim 10\,\text{K}$. Dense Cores on scales of 0.1 pc, can have $M \sim 10 M_\odot$, $n \sim 10\,000\,\text{cm}^{-3}$, and $A_V > 10\,\text{mag}$. Finally, in some localised regions of GMCs observations have revealed ‘Hot Cores’ with $T = 100\text{--}300\,\text{K}$, and extreme values of density and extinction, as high as $n \sim 10^9\,\text{cm}^{-3}$ and $A_V \sim 50$; masses can be in the range $10\text{--}3000 M_\odot$. It is in these hot cores that near-IR imaging¹ has revealed the presence of embedded young O and B-type stars, strongly suggesting a causal connection. As we shall see presently, giant molecular clouds are indeed unstable to gravitational collapse.

¹Recalling that $A_V = 10\,\text{mag}$ corresponds to a dimming of the visible light by a factor of 10 000, it can be easily appreciated that these dense regions are opaque to light of visible wavelengths. Longer wavelength light, in the near-IR regime, suffers considerably less extinction.

11.2 Brief Overview

It is useful to distinguish six stages in the star formation process:

1. The initial free-fall collapse of the parent interstellar cloud.
2. Cloud fragmentation, leading to a range of stellar masses.
3. Formation of a protostellar core. The star appears on the H-R diagram.
4. Accretion of the surrounding gas, generally through an accretion disk.
5. Dissociation of molecules and ionisation of H and He.
6. Pre-main sequence phase.
7. Star formation is considered to be completed once the star appears on the “Zero Age Main Sequence” (ZAMS).

In the following sections we shall consider these different stages in turn.

11.3 The Jeans Criterion for Gravitational Collapse

The condition for a cloud of interstellar gas to collapse and form stars was first formulated by James Jean in 1902 while working at Trinity College, Cambridge. His treatment considers only gravitation and thermodynamics and ignores other important effects such as rotation, turbulence and magnetic fields. Nevertheless, it provides a useful insight into the development of protostars.

As we saw in Lecture 7.1, the condition for equilibrium of a stable, gravitationally bound system is given by the virial theorem:

$$2K + U = 0 \tag{11.1}$$

where K is the kinetic energy and U is the potential energy. Consider now what happens if such a system is perturbed from equilibrium. If $2K > |U|$, the force due to gas pressure will dominate the force of gravity

and the cloud will expand. On the other hand, if $2K < |U|$, the reverse will be true and the cloud will collapse under the force of gravity.

We showed in Lecture 7.1 (eqs. 7.1–7.8) that the gravitational potential energy can be written as:

$$U \simeq -\frac{3}{5} \frac{GM_c^2}{R_c}, \quad (11.2)$$

where M_c and R_c are, respectively, the mass and the radius of the cloud under consideration.

We also saw (eqs. 7.9–7.11) that the average kinetic energy per particle is

$$K = \frac{3}{2}kT \quad (11.3)$$

where k is Boltmann constant. Thus, the total internal kinetic energy of the cloud is just:

$$K = \frac{3}{2}\mathcal{N}kT \quad (11.4)$$

where \mathcal{N} is the total number of particles. We can write \mathcal{N} in terms of the mass and the mean molecular weight (Lecture 9.2.1):

$$\mathcal{N} = \frac{M_c}{\mu m_H} \quad (11.5)$$

We can therefore rewrite the condition for gravitational collapse ($2K < |U|$) as:

$$\frac{3M_c k T}{\mu m_H} < \frac{3}{5} \frac{GM_c^2}{R_c} \quad (11.6)$$

We can eliminate the radius R_c from the above equation using:

$$R_c = \left(\frac{3}{4} \frac{M_c}{\pi \rho_0} \right)^{1/3} \quad (11.7)$$

where ρ_0 is the initial density of the cloud prior to collapse, with the assumption that the cloud is a sphere of constant density.

Substituting the above expression into eq. 11.6, we obtain the important concept of the *Jeans mass*:

$$M_J \simeq \left(\frac{5kT}{G\mu m_H} \right)^{3/2} \left(\frac{3}{4\pi\rho_0} \right)^{1/2} . \quad (11.8)$$

If the mass of a cloud exceeds the Jeans mass, the cloud will be unstable against gravitational collapse.

Note that for a given chemical composition (i.e. for a given μ), the Jeans mass depends only on temperature and density: the lower the temperature and the higher the density, the smaller the cloud mass which is unstable against gravitational collapse. Note also that this is an oversimplification of the real situation, not only because we have ignored rotation, internal macroscopic velocity gradients and magnetic fields, but also because we have neglected any external pressure on the cloud (such as the external pressure of an encompassing GMC on an embedded dense core).

The same criterion for gravitational collapse can also be expressed in terms of the Jeans length:

$$R_J \simeq \left(\frac{15kT}{4\pi G\mu m_H \rho_0} \right)^{1/2}, \quad (11.9)$$

where the condition for gravitational collapse is $R_c > R_J$, or the Jeans density:

$$\rho_J \simeq \frac{3}{4\pi M_c^2} \left(\frac{5kT}{G\mu m_H} \right)^3, \quad (11.10)$$

where the condition for gravitational collapse is $\rho_c > \rho_J$.

Let us consider the Jeans mass of some interstellar structures. In diffuse hydrogen clouds, typical values of temperature and density are $T \sim 100$ K and $n_{\text{HI}} \sim 10^7 \text{ m}^{-3}$. Thus, $\rho_0 = n_{\text{HI}} m_H \sim 1 \times 10^7 \cdot 1.7 \times 10^{-27} \sim 1.7 \times 10^{-20} \text{ kg m}^{-3}$. With the gas fully neutral, $\mu = 1$. Entering these values in eq. 11.8, we have (using S.I. units throughout):

$$M_J \simeq \left(\frac{5 \cdot 1.4 \times 10^{-23} \cdot 100}{6.7 \times 10^{-11} \cdot 1 \cdot 1.7 \times 10^{-27}} \right)^{3/2} \left(\frac{3}{4 \cdot 3.14 \cdot 1.7 \times 10^{-20}} \right)^{1/2} \quad (11.11)$$

or

$$M_J \simeq (6.1 \times 10^{16})^{3/2} (1.4 \times 10^{19})^{1/2} \quad (11.12)$$

$$M_J \simeq 5.7 \times 10^{34} \text{ kg} \sim 30\,000 M_\odot \quad (11.13)$$

This value of the Jeans mass is two orders of magnitude higher than the typical mass of diffuse interstellar clouds, which are therefore very stable against gravitational collapse.

On the other hand, in the dense core of a giant molecular cloud, typical values are $T \sim 10\text{ K}$, and $n_{H_2} \sim 10^{10}\text{ m}^{-3}$. The density is thus $\rho_0 \sim 2n_{H_2}m_H \sim 2 \times 10^{10} \cdot 1.7 \times 10^{-27} \sim 3.4 \times 10^{-17}\text{ kg m}^{-3}$, a factor of 2000 greater than in a diffuse hydrogen cloud. With $\mu = 2$, the Jeans mass will therefore be lower by a factor of $(10 \cdot 2)^{3/2} \cdot \sqrt{2000} = 4000$, i.e. $M_J \simeq 7.5M_\odot$.

Now we see that, with typical masses $M \sim 10M_\odot$, the dense cores of GMCs are indeed unstable to gravitational collapse, consistent with being the sites of star formation. While a GMC may originally be in pressure equilibrium with the surrounding interstellar medium, a small perturbation can initiate its contraction and trigger gravitational collapse. Such a perturbation can be provided by cloud-cloud collisions, or by a passing interstellar shock wave originating from a nearby region of star formation, where some of the most massive stars have already exploded as supernovae. This can lead to a ‘forest fire’ type of star formation, as has been suggested for the Carina nebula (see Figure 11.2), whereby star formation propagates from one side of a giant molecular cloud complex. Galaxy interactions are another type of process that can trigger star formation by tipping GMCs past the Jeans mass limit.

11.4 Free-fall Timescale

Although the Jeans criterion provides the necessary condition for the onset of collapse of a gas cloud, such collapse involves the release of gravitational energy. If this energy were converted directly to thermal energy, the temperature would rise and, according to eq. 11.8, the Jeans mass would increase, halting the collapse. However, in the early stages of collapse, the cloud is transparent to far-infrared radiation and can cool efficiently by converting kinetic energy of its molecules and atoms into infrared photons which can escape the cloud. Thus, the early stages of the collapse are isothermal and the cloud is essentially in free-fall collapse.

We can obtain an estimate of the free-fall timescale as follows. We begin

by writing the equation of gravitational acceleration:

$$\frac{d^2r}{dt^2} = -G \frac{M_r}{r^2} \quad (11.14)$$

where as usual M_r denotes the mass enclosed within radius r (assuming spherical symmetry). Writing the mass in terms of the initial density and radius, and multiplying both sides by the velocity at the surface of the spherical cloud, we have:

$$\frac{dr}{dt} \frac{d^2r}{dt^2} = - \left(\frac{4\pi}{3} G \rho_0 r_0^3 \right) \frac{1}{r^2} \frac{dr}{dt} \quad (11.15)$$

which can be integrated with respect to time to give:

$$\frac{1}{2} \left(\frac{dr}{dt} \right)^2 = \left(\frac{4\pi}{3} G \rho_0 r_0^3 \right) \frac{1}{r} + C_1. \quad (11.16)$$

The integration constant C_1 can be evaluated with the boundary condition that the infall velocity be zero at the onset of collapse, that is $dr/dt = 0$ when $r = r_0$. This gives:

$$C_1 = - \frac{4\pi}{3} G \rho_0 r_0^2. \quad (11.17)$$

Substituting 11.17 into 11.16 and solving for the velocity at the surface, we have:

$$\frac{dr}{dt} = - \left[\frac{8\pi}{3} G \rho_0 r_0^2 \left(\frac{r_0}{r} - 1 \right) \right]^{1/2}, \quad (11.18)$$

where we have chosen the negative root because the cloud is collapsing. In order to obtain an expression for the position of the spherical cloud surface as a function of time we need to integrate 11.18. With the substitutions:

$$\theta \equiv \frac{r}{r_0}$$

and

$$\chi \equiv \left(\frac{8\pi}{3} G \rho_0 \right)^{1/2}$$

eq. 11.18 can be re-written as:

$$\frac{d\theta}{dt} = -\chi \left(\frac{1}{\theta} - 1 \right)^{1/2}. \quad (11.19)$$

With the further substitution:

$$\theta \equiv \cos^2 \xi,$$

eq. 11.19 can in turn be re-written as

$$\cos^2 \xi \frac{d\xi}{dt} = \frac{\chi}{2} \quad (11.20)$$

which can now be integrated with respect to time to give:

$$\frac{\xi}{2} + \frac{1}{4} \sin 2\xi = \frac{\chi}{2} t + C_2 \quad (11.21)$$

Again, we can evaluate the constant of integration C_2 by considering that at $t = 0$, $r = r_0$. Hence, $\theta = 1$ and $\xi = 0$. Therefore, $C_2 = 0$. We thus arrive at the equation of motion for the gravitational collapse of the cloud:

$$\xi + \frac{1}{2} \sin 2\xi = \chi t \quad (11.22)$$

We define the free-fall timescale as the time taken by a cloud in free-fall to collapse from $r = r_0$ to $r = 0$. (In reality, of course, this final condition is never reached, but the concept is still valid so long as $r_{\text{final}} \ll r_0$). When $r = 0$, $\theta = 0$ and $\xi = \pi/2$. Then, from 11.22 we have:

$$t_{\text{ff}} = \frac{\pi}{2\chi} \quad (11.23)$$

or

$$t_{\text{ff}} = \left(\frac{3\pi}{32} \frac{1}{G\rho_0} \right)^{1/2}. \quad (11.24)$$

Note that the free-fall timescale is independent of the initial radius of the sphere, and depends only on the initial density ρ_0 . Thus, in a spherical molecular cloud of uniform density (admittedly a rather implausible simplification!), all parts of the cloud will take the same length of time to collapse and the density will increase at the same rate everywhere within the cloud. This behaviour is known as *homologous collapse*.

11.5 Cloud Fragmentation

From the above treatment, one would be justified in concluding that the entire mass of GMC exceeding the Jeans limit would collapse to form a

single star. This is clearly not the case: most early-type (i.e. young) stars are found in clusters, and maybe as many as three quarters of all stars are in binary or multiple systems.

During the free-fall collapse phase, the density within the cloud increases by many orders of magnitude. If the temperature remains approximately constant, then the Jeans mass criterion (eq.11.8) implies that the mass limit for instability decreases dramatically. Any initial density inhomogeneities which may have been present within the cloud will cause individual regions within the GMC to cross the instability threshold independently and collapse locally. This could lead to the formation of a large number of smaller objects.

This is still only part of the story, however. The fragmentation process must stop at some point because: (i) the process of star formation is not 100% efficient—it is actually rather inefficient, with only about 1% of the gas mass of the parent cloud being turned into stars; and (ii) the most common type of star in the Galaxy has a mass of $\sim 1M_{\odot}$, and lower mass stars are actually *less* common. Something must intervene to limit the fragmentation process.

Fragmentation stops when the assumption of isothermal contraction breaks down. The increasing density of the collapsing cloud fragment eventually renders the gas opaque even to infrared photons. As a result, radiation is trapped within the central part of the collapsing cloud, leading to heating and an increase in gas pressure. When radiation can no longer escape the cloud, the collapse turns from isothermal to adiabatic. (Of course, in reality the collapse is never totally isothermal nor adiabatic, but somewhere between these two limits. As is often the case, we can gain insight into the relevant physical processes by considering the two limiting cases.)

We can appreciate the transition from isothermal to adiabatic collapse by recalling (Lecture 8) the adiabatic relationship between pressure and density:

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

where γ is the ratio of specific heats and K is a constant. Combined with the ideal gas law

$$P = \frac{\rho k T}{\mu m_H},$$

it implies an adiabatic relationship between temperature and density:

$$T = K' \rho^{\gamma-1}. \quad (11.25)$$

Substituting 11.25 into the expression for the Jeans mass (eq. 11.8), we have:

$$M_J \propto \rho^{(3\gamma-4)/2}. \quad (11.26)$$

For atomic hydrogen² $\gamma = 5/3$, giving $M_J \propto \rho^{1/2}$; in other words, the Jeans mass *increases* as the density increases (for a perfectly adiabatic collapse). This behaviour results in a minimum fragment mass determined by the transition from a predominantly isothermal to a predominantly adiabatic collapse. In turn, the transition is driven by the opacity of the interstellar mix of molecules, gas and dust to infrared radiation.

We can have a stab at estimating the lower mass limit of the fragmentation process as follows. As we have already seen in lecture 7.11, the energy released during the collapse of a protostellar cloud is half its potential energy:

$$\Delta E_g \simeq \frac{3}{10} \frac{GM_J^2}{R_J} \quad (11.27)$$

for a spherical cloud satisfying the Jeans instability criterion. Averaged over the free-fall time, the luminosity produced in the collapse is:

$$L_{\text{ff}} = \frac{\Delta E_g}{t_{\text{ff}}} \simeq \frac{3}{10} \frac{GM_J^2}{R_J} \cdot \left(\frac{3\pi}{32} \frac{1}{G\rho_0} \right)^{-1/2} \quad (11.28)$$

Using $\rho_0 = M_J / (\frac{4}{3}\pi R_J^3)$, we have:

$$L_{\text{ff}} \sim G^{3/2} \left(\frac{M_J}{R_J} \right)^{5/2}. \quad (11.29)$$

Now, if the cloud were optically thick and in thermodynamic equilibrium, this energy would be emitted as blackbody radiation, with its luminosity given by the familiar expression:

$$L_{\text{bb}} = 4\pi R^2 \sigma T^4.$$

²At the low temperatures of interstellar clouds, molecular hydrogen behaves like a monoatomic gas. This observation was extremely puzzling in the early part of the twentieth century and engaged the minds of some of the most prominent physicists of the time. It took the development of the quantum theory to show that diatomic hydrogen, with its tiny rotational inertia, requires a large amount of energy to excite its first excited molecular rotation quantum state. Since it can not get that amount of energy at low temperatures, it acts like a monoatomic gas.

However, the collapsing cloud is not in thermodynamic equilibrium, so we introduce an efficiency factor e ($0 < e < 1$) into the above equation to give:

$$L_{\text{rad}} = 4\pi R^2 e \sigma T^4. \quad (11.30)$$

Equating the two expressions for the cloud luminosity, we have:

$$M_J^{5/2} = \frac{4\pi}{G^{3/2}} R_J^{9/2} e \sigma T^4. \quad (11.31)$$

Expressing the radius in terms of the mass and density (eq. 11.7), and then the density in terms of the Jeans mass (eq. 11.8), we finally arrive at an expression for the minimum obtainable Jeans mass corresponding to when adiabatic effects become important:

$$M_{J_{\min}} = 0.03 \left(\frac{T^{1/4}}{e^{1/2} \mu^{9/4}} \right) M_{\odot} \quad (11.32)$$

where T is in kelvins. With $\mu \sim 1$, $e \sim 0.1$, and $T \sim 1000$ K at the time when adiabatic effects may start to become significant, we have $M_{J_{\min}} \sim 0.5 M_{\odot}$; that is, fragmentation ceases when individual fragments are approximately solar mass objects. Note that the parameters which determine $M_{J_{\min}}$ in eq. 11.32 are all to relatively low powers. Thus the conclusion that fragmentation does not continue much beyond masses comparable to a solar mass is not highly dependent on the values of e and T . For example, if $e = 1$ (fully adiabatic), then $M_{J_{\min}} \sim 0.2 M_{\odot}$.

11.6 Protostars

Once the density of a collapsing fragment has increased sufficiently for the gas to become opaque to infrared photons, 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 is nearly in hydrostatic equilibrium and the dynamical collapse is slowed to a quasistatic contraction. At this stage we may start to speak of a protostar.

The mass of such a ‘protostar’ is still only a small fraction of the mass it will have once it reaches the Main Sequence. The surrounding gas continues to free-fall onto the protostellar core, so that the next phase is dominated by

accretion. In general, the infalling gas will form an accretion disk around the protostar, reflecting the fact that the contracting gas cloud has a net angular momentum. Such accretion disks are commonly seen around very young stars.

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:

$$L \sim L_{\text{acc}} = \frac{1}{2} \frac{GM\dot{M}}{R} \quad (11.33)$$

where \dot{M} is the mass accretion rate. The core heats up almost adiabatically since the accretion timescale t_{ff} is much smaller than the thermal timescale, which is just the Kelvin-Helmholtz timescale, t_{KH} , we have already encountered in Lecture 7.1.

11.6.1 Dissociation and Ionisation

When the temperature of the core reaches $T \sim 2000$ K, the average particle energy is comparable to the dissociation energy of molecular hydrogen. The energy produced by contraction is now absorbed by this process, rather than providing the pressure gradient necessary to maintain hydrostatic equilibrium. As a result, the core becomes dynamically unstable, and a second collapse occurs during which the gravitational energy released is absorbed by the dissociating molecules without a significant rise in temperature. When H_2 is completely dissociated into atomic hydrogen, hydrostatic equilibrium is restored and the temperature rises again. Somewhat later, further dynamical collapse phases follow when first H and then He are ionised at temperatures $T \sim 10^4$ K.

When ionization of the protostar is complete, it settles back into hydrostatic equilibrium at a much reduced radius. We can estimate the radius R_p of a protostar after the dynamical collapse phase is over by equating the change in potential energy to the sum of the energies of: (i) dissociation of molecular hydrogen ($\chi_{\text{H}_2} = 4.48$ eV per H_2 molecule), (ii) ionisation of atomic hydrogen ($\chi_{\text{HI}} = 13.6$ eV per H atom), and (iii) ionisation of helium ($\chi_{\text{He}} = 79$ eV per He atom: 24.6 eV for $\text{He}^0 \rightarrow \text{He}^+$ and 54.4 eV for $\text{He}^+ \rightarrow$

He^{++}). Thus we have:

$$\Delta E_g \simeq \frac{3}{10} \frac{GM^2}{R_p} \approx \frac{M}{m_H} \left(\frac{X}{2} \chi_{\text{H}_2} + X \chi_{\text{H}} + \frac{Y}{4} \chi_{\text{He}} \right) \equiv \frac{M}{m_H} \chi, \quad (11.34)$$

where we have taken the collapse to start from infinity because $R_p \ll R_{\text{init}}$. With $X = 0.74$ and $Y = 1 - X$ (ignoring the metals), $\chi = 16.9 \text{ eV}$ per baryon. Therefore we have:

$$R_p \approx \frac{3}{10} \frac{GMm_H}{\chi} \approx 35R_{\odot} \left(\frac{M}{M_{\odot}} \right). \quad (11.35)$$

We can use the virial theorem to estimate the average temperature of the protostar, $\langle T_p \rangle$. Following the same steps as in Lecture 7.1.1 (eqs. 7.9–7.13), we find:

$$\langle T_p \rangle = \frac{1}{5} \frac{\mu m_H}{k} \frac{GM}{R_p} = \frac{2}{3} \frac{\mu}{k} \chi \approx 8 \times 10^4 \text{ K}, \quad (11.36)$$

for $\mu \simeq 0.6$ appropriate to a fully ionised plasma (Lecture 9.2.1). There are two things of note about eq. 11.36. First, the temperature is independent of the mass of the protostar. Second, the temperature reached once the protostar settles back into hydrostatic equilibrium (following H_2 dissociation and He and He ionisation) is still much lower than the temperature necessary to ignite nuclear reactions. At these ‘low’ temperatures, the opacity is very high and is dominated by the H^- ion (Lecture 5.4.1). Under these circumstances, radiative energy transport is very inefficient, making the protostar *convective* throughout, except for the outermost layer (photosphere). Thus a new star is chemically homogeneous. This will change as the star ages.

11.7 The Hayashi Track and the Final Approach to the Main Sequence

Fully convective stars of a given mass occupy an almost vertical line in the H-R diagram. The line is called the *Hayashi line* from the Japanese astrophysicist who worked on stellar evolution during the pre-main sequence stages in the early 1960s. A protostar of a given mass moves along its Hayashi line as it approaches the main sequence (see Figure 11.3).

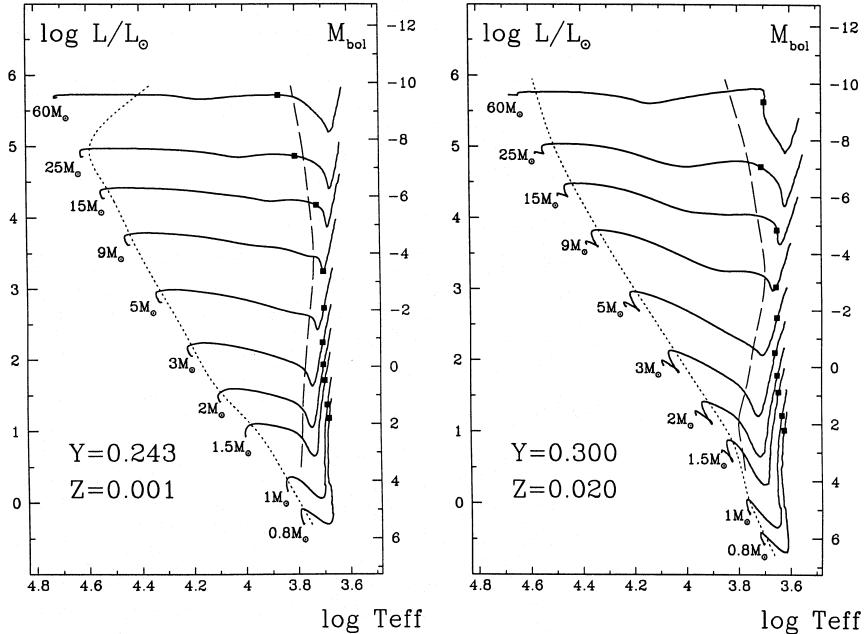


Figure 11.3: Theoretical pre-main sequence evolutionary tracks computed for stars of various masses for two different metallicities, as indicated. A pre-main sequence star evolves along its track from right to left. The black square on each track indicates the onset of D burning. (Reproduced from Bernasconi & Maeder 1996, A&A, 307, 829).

The Hayashi track is actually a boundary. It represents the minimum effective temperature for a star in hydrostatic equilibrium. To the right of the Hayashi track, there is no mechanism that can adequately transport the luminosity out of the star at such low effective temperatures; hence no stable star can exist there and this region is sometimes referred to as the *forbidden region* of the H-R diagram. On the other hand, stars to the left of the Hayashi line (at higher T_{eff}) cannot be fully convective but must have some portion of their interior in radiative equilibrium.

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. 11.35. 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. One such evolutionary track is given approximately by:

$$\log L = 10 \log M - 7.24 \log T_{\text{eff}} + C \quad (11.37)$$

This track is steeply descending, and shifts upwards with increasing initial mass.

As the central temperature rises, we move beyond the peak in the opacity curves shown in Figure 5.5, into a region where the opacity follows a Kramers law, $\langle \kappa \rangle \propto T^{-3.5}$ (Lecture 5.5). A radiative core develops, growing with time to encompass more and more of the star's mass. The point of minimum luminosity in the evolutionary tracks shown in Figure 11.3 corresponds to the development of a radiative core which allows energy to escape more readily into the convective envelope, causing the luminosity of the star to increase. The effective temperature continues to increase since the star is still shrinking.

Contraction continues, as dictated by the virial theorem, until the central temperature becomes sufficiently high to ignite nuclear fusion. Once the energy generated by H 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 $0.08M_{\odot}$ (Lecture 10.4).

Before thermal equilibrium on the ZAMS is reached, however, several nuclear reactions have already set in. In particular, a small quantity of deuterium ($D \equiv {}^2_1H$) is present in the interstellar gas out of which stars form, with mass fraction of order 10^{-5} . D is a very fragile nucleus that reacts easily with atomic hydrogen. What little is left over from Big-Bang nucleosynthesis [$(D/H)_{\text{prim}} = 2.5 \times 10^{-5}$ by number] is destroyed by the second reaction in the p-p chain (see Figure 7.5):



which destroys all D present in a star at $T \sim 1 \times 10^6$ K, while the protostar is still on the Hayashi line (see Figure 11.3).³

The reaction (11.38) is exothermic and produces 5.5 MeV of energy, sufficient to halt the contraction of the PMS star for $\sim 10^5$ yr. (A similar, but much smaller effect, occurs somewhat later when lithium is burned at a higher T). Furthermore, the ${}^{12}C(p, \gamma){}^{13}N$ reaction is already activated

³This process is known as the *astration* of deuterium, and makes the calculation of the abundance of deuterium as star formation progresses in a galaxy particularly simple. For the same reason, the D/H ratio measured in the present-day interstellar medium is a *lower limit* to the primordial abundance of deuterium.

at a temperature below that of the full CNO cycle, due to the relatively large initial abundance of ^{12}C 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 evolutionary tracks just above the ZAMS location visible in Figure 11.3. Note that this occurs even in low mass stars, with $M < 1M_{\odot}$, even though the p-p chain is their main energy generation mechanism.

Finally, the time taken for a protostar to reach the ZAMS depends on its mass. This time is essentially the Kelvin-Helmholtz contraction timescale which we considered in Lecture 7.1. Since contraction is slowest when both R and L are small (recall from Lecture 7.1 that $\tau_{\text{KH}} = \Delta E_g / L$ and $\Delta E_g \propto 1/R$ for a given mass M), the pre-main sequence lifetime is dominated by the final stages of contraction, when the star is already close to the ZAMS. For stars with $M \geq 1M_{\odot}$, we find $\tau_{\text{KH}} \simeq 5 \times 10^7 (M/M_{\odot})^{-2.5}$ yr, while for lower mass stars the mass dependence is somewhat shallower (see Figure 11.4). 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).

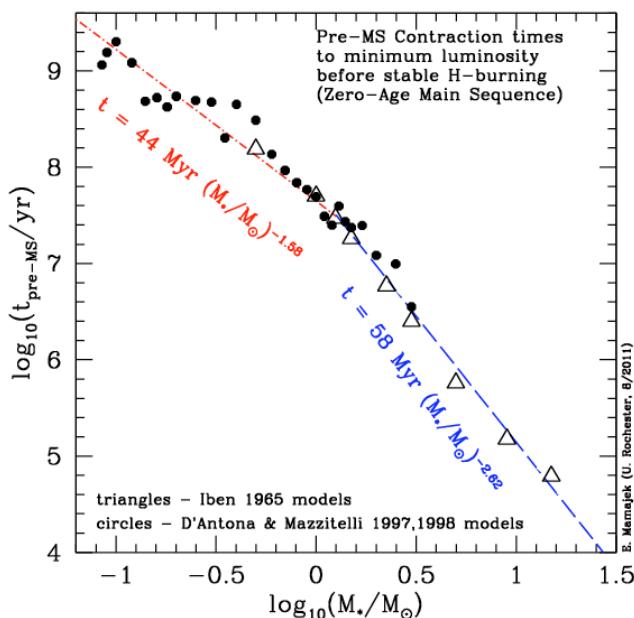


Figure 11.4: Pre-main sequence contraction times as a function of stellar mass, from the evolutionary tracks calculated by D'Antona & Mazzitelli and by Iben, as indicated.

11.8 Nomenclature of Objects Associated with Star Formation

Before concluding the topic of star formation, we briefly mention some of the objects which are associated with this process.

T Tauri stars (named after the first star of their class to be identified as an object of special interest) are an important class of low-mass, pre-main sequence objects. As can be seen from Figure 11.5, many of them lie on the Hayashi tracks. Some of their characteristics [many of them common to Young Stellar Objects (YSO) in general] are:

1. Variability in the light output with timescales of the order of days, as material falls down onto the surface of the star from a residual disk;

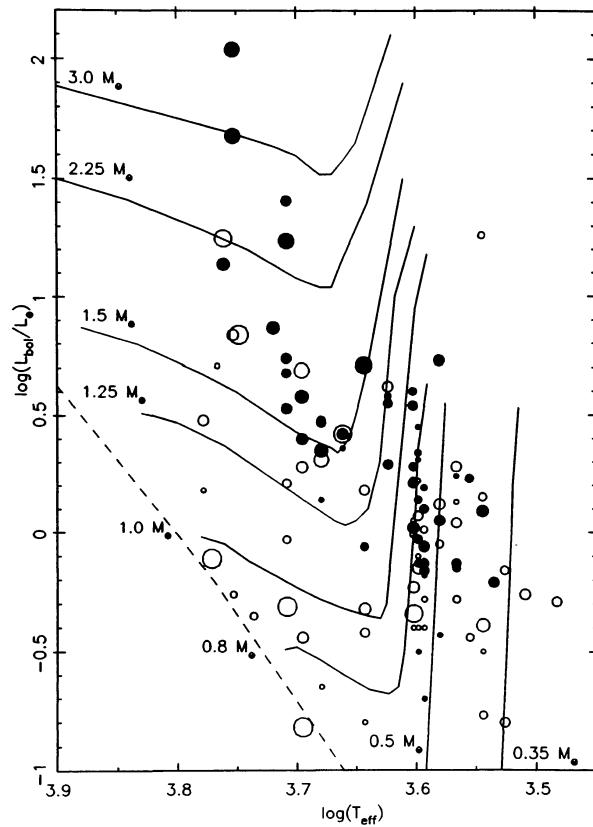


Figure 11.5: Positions of T Tauri stars on the H-R diagram. The sizes of the circles are proportional to the rate of rotation. Filled circles indicate stars with strong emission lines. Theoretical pre-main sequence evolutionary tracks are also shown. (Reproduced from Bertout 1989, ARA&A, 27, 351).

2. Emission lines from both permitted and forbidden transitions, from the disk, or the bipolar outflow, or both. Some emission lines have a characteristic P-Cygni profile, which is a combination of emission and absorption indicative of mass loss.
3. Higher infrared luminosity than main sequence stars of the same mass, because there is more dust in their immediate vicinity.
4. A high level of what is called activity is seen, meaning flares, star spots and emission from a hot corona (including X-ray emission). There seem to be two reasons for this: many T Tauri stars are rapid rotators, with rotation periods from hours to days, as opposed to a month for the Sun.⁴ Second, as we have seen, the convection zone extends deeper into the star when it is on the Hayashi track than when it reaches the main sequence. The combination of rapid rotation and deep convection results in a strong dipole magnetic field which, in turn, drives the activity.

Herbig-Haro objects are apparently associated with the jets produced by T Tauri stars. As the jets expand supersonically into the ambient interstellar medium, collisions excite the gas resulting in bright nebulosities with emission line spectra. Continuous emission is also observed in some protostellar objects as the central star, which is hidden from direct view behind dust in the accretion disk, illuminates the surfaces of the disk. These accretion disks seem to be responsible for many of the characteristics of YSOs, including emission lines, mass loss, jets, variability, and ultimately planet formation.

OB associations are groups of stars dominated by O and B-type main sequence stars (lower mass stars are of course also present, but they are less conspicuous being less luminous). Since the lifetimes of O stars are $< 10^7$ yr, OB associations mark the sites of recent star formation (see Figure 11.6), and indeed they are usually found close to molecular clouds (although not necessarily the parent cloud from which they formed, which is rapidly dispersed following the formation of massive stars—see later). Studies of the kinematics of individual stars in OB associations generally

⁴Stars are born with high rotational velocities which subsequently decline with age. The braking is thought to result from the interaction of the star's magnetic field and the stellar wind emitted from the photosphere causing a steady transfer of angular momentum away from the star.

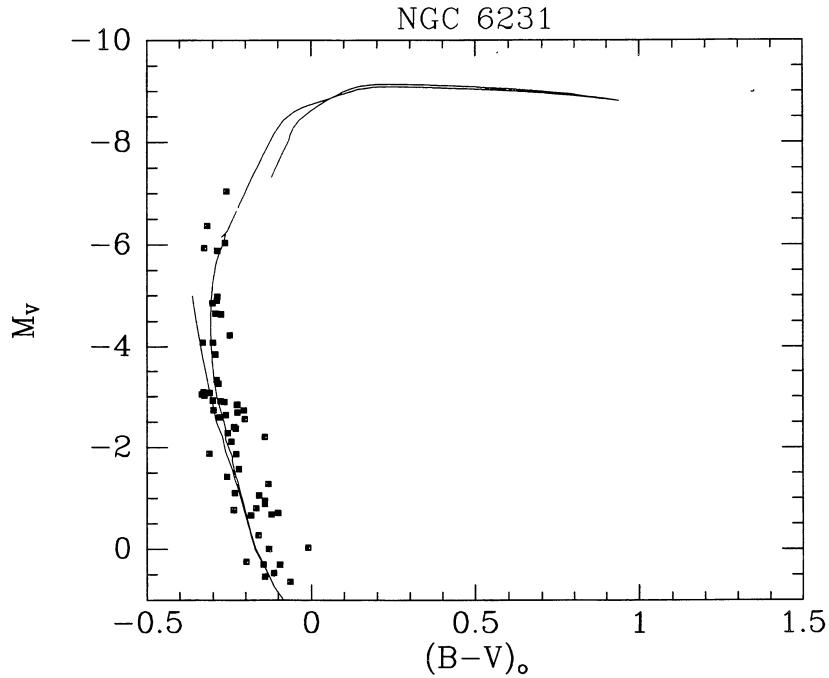


Fig. 10. Same as in Fig. 9 for NGC 6231, $m - M = 12.50$, $E(B - V) = 0.46$, $\log \text{age} = 6.75$

Figure 11.6: Isochrone fitting to the upper main sequence of the open cluster NGC 6231 indicates an age of $\sim 5.6 \times 10^6$ yr. (Reproduced from Meynet et al. 1993, A&A Supp., 98, 477).

lead to the conclusion that they are not gravitationally bound and they will eventually dissolve into the field stellar population.

On the other hand, **super star clusters**, most commonly found in regions undergoing very intense episodes of star formation—or starbursts, consist of 10^4 to 10^6 stars concentrated within a few pc. Given the high density of stars, they may remain bound even after their massive stars have exploded as Type II supernovae and may also survive other disruptive processes later on. If so, after several billion years, they would evolve into objects similar to the old globular clusters in the halo of the Milky Way. Thus, super star clusters may simply be young globular clusters.

11.8.1 H II Regions

Whether isolated or in a cluster, massive stars of spectral type O and B have a profound effect on their surroundings once they arrive on the main sequence. With effective temperatures $T_{\text{eff}} > 30\,000$ K, the peak of their blackbody curve is at ultraviolet wavelengths. Photons with $\lambda < 912$ Å are sufficiently energetic to ionise ground-state hydrogen atoms (with

ionisation potential IP = 13.6 eV) in the surrounding interstellar medium, producing an H II region around the star(s).

We can calculate the size of an H II region provided we know the number of ionising photons emitted by the star and the ambient density. In a steady state, the ionisation rate (number of ionisations per unit time) must balance the recombination rate (the inverse process). If such an equilibrium did not develop, the size of the H II region would continue to grow until Lyman continuum photons are diluted enough for equilibrium to be established.

The recombination rate per unit volume is given by:

$$\mathcal{R}_{\text{rec}} = \alpha(T) n_H n_e$$

where n_H and n_e are the volume densities of H^+ ions and electrons respectively, and $\alpha(T)$ is the temperature dependent radiative recombination coefficient. Since hydrogen is the most abundant element and is fully ionised, $n_H \simeq n_e$. Assuming spherical symmetry, we therefore have:

$$Q_* = \mathcal{R}_{\text{rec}} \frac{4}{3} \pi r_{\text{HII}}^3 \quad (11.39)$$

where Q_* is the number of ionising photons emitted by the star per unit time and r is the radius of the H II region, also called the Strömgren radius from the Danish astrophysicist who first carried out the analysis in the late 1930s. Solving for the Strömgren radius, we have:

$$r_{\text{HII}} = \left(\frac{3Q_*}{4\pi\alpha} \right)^{1/3} n_H^{-2/3}. \quad (11.40)$$

The Strömgren sphere of a star marks the sharp transition between fully ionised circumstellar gas and mainly neutral interstellar gas. Within the H II region, hydrogen is continuously being ionised and recombining with electrons. In general, recombination takes place to a high energy level, followed by cascading of the electron through intermediate levels to the ground state. Each step in the process is accompanied by the emission of a photon with energy *lower* than the original 13.6 eV responsible for ionising H from the ground state. The dominant visible wavelength photon produced in this way results from the transition between the $n = 3$ and $n = 2$ levels, corresponding to the $\text{H}\alpha$ line of the Balmer series, at a wavelength of 6563 Å, in the red region of the optical spectrum. It is this process that causes H II regions to fluoresce in red light.

11.8.2 Coeval Star Formation in Clusters?

The effect of massive stars on their environment goes far beyond ionising the gas within the Strömgren sphere. All stars more massive than $\sim 20M_{\odot}$ experience mass loss driven by radiation pressure while still on the main sequence. For the most massive stars, the energy deposited into their surroundings via strong stellar winds is comparable to that associated with the explosion of a Type II supernova, $\sim 10^{51}$ ergs.

The ‘double-whammy’ of stellar winds and supernova explosions can disperse the remainder of the parent molecular cloud on a timescale of only a few million years, depending on the richness of the newly formed cluster of OB stars, effectively shutting down any further star formation. This timescale is shorter than the contraction times of lower mass stars, and yet many clusters are known to contain both early-type and late-type stars. This apparent contradiction, first noted by the American astronomer George Herbig in the 1960s, has led to the view that, when stars in a cluster form, the low- and intermediate-mass stars form first, with the process continuing gradually until the high-mass stars form and quickly halt all subsequent star formation. It may also be the case that the most massive stars form by mergers of smaller stars in dense protostellar environments.

11.9 The Initial Mass Function

The end product of an episode of star formation is a cluster of stars of different masses. Observationally, it is clear that low mass stars are much more common than high mass stars. The distribution of stellar masses immediately after a population of stars arrives on the main sequence is described by the *Initial Mass Function* (IMF). For a young cluster, the IMF can be deduced from the observed *Present-day Mass Function*, after correcting for the fact that the most massive stars have evolved off the main sequence in clusters older than a few million years. In the field, we need to combine the PDMF with a knowledge (or model) of the past history of star formation to reconstruct the IMF. In both cases, models are required to convert the measured stellar luminosities to stellar masses.

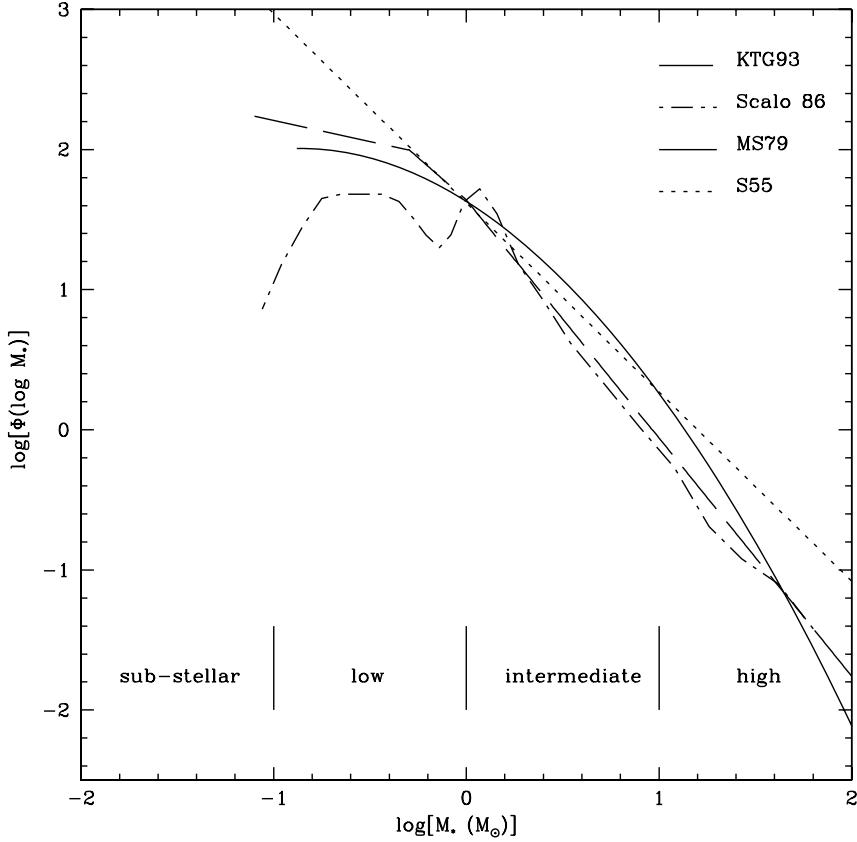


Figure 11.7: Different realisations of the stellar Initial Mass Function. S55 is for Salpeter (1955), and KTG93 is for Kroupa, Tout & Gilmore (1993). (Figure reproduced from Meyer et al. 1999, astro-ph/9902198).

The simplest form of the IMF is a single power law of the form:

$$N(M) dM \propto M^{-\alpha} dM \quad (11.41)$$

where $N(M) dM$ is the number of stars per unit volume with mass between M and $M + dM$.

Edwin Salpeter suggested in 1955 that, in the solar neighbourhood, the frequency distribution of stars more massive than the Sun follows eq. 11.41 with $\alpha = 2.35$. Other formulations have been proposed, usually consisting of a combination of different power laws in different mass intervals (see Figure 11.7). Thus, for example, the Kroupa, Tout & Gilmore (1993) IMF is a combination of three power laws:

$$N(M) dM \propto \begin{cases} M^{-2.7} dM & \text{if } M > 1M_\odot \\ M^{-2.2} dM & \text{if } 1 \geq M \geq 0.5M_\odot \\ M^{-(0.70 \text{ to } 1.85)} dM & \text{if } 0.5 \geq M \geq 0.08M_\odot \end{cases} \quad (11.42)$$

The faintness of low mass stars, the rarity and short lives of high mass stars,

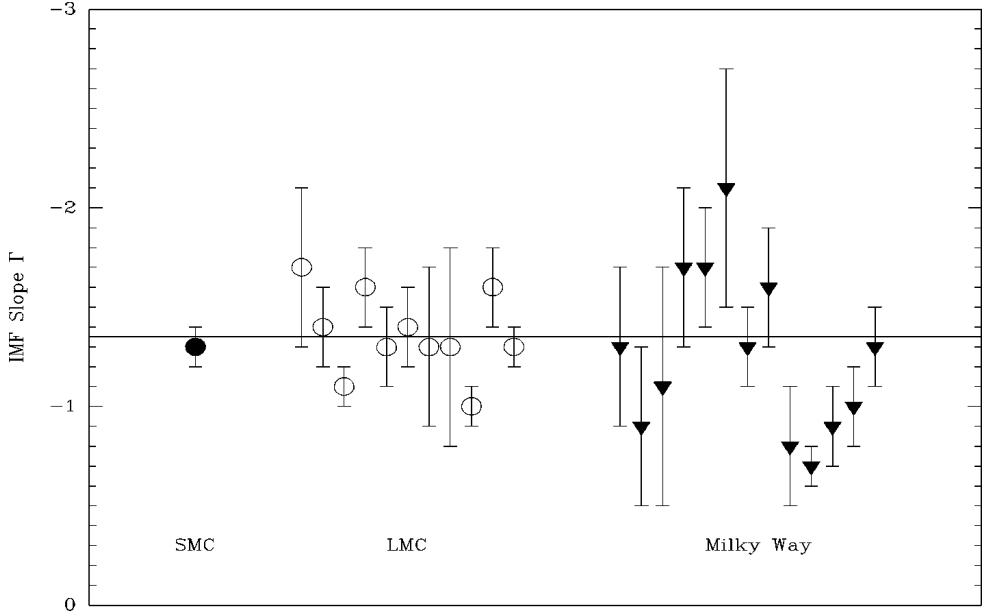


Figure 11.8: The slope of the IMF (in its integral, rather than differential, formulation) measured in OB associations and clusters in the Milky Way and our companion galaxies, the Large and Small Magellanic Clouds, is consistent with a Salpeter slope of $\alpha + 1 = -1.35$. (Figure reproduced from Massey 2003, ARA&A, 41, 15).

and the frequency of binaries are all issues affecting the determination of the IMF.

A much debated question is whether the IMF is ‘universal’, or whether it varies with local conditions, such as metallicity, or with redshift. Opinions are divided as to the reality of claimed IMF variations, although it seems to be well established that the same IMF applies to stars in the Milky Way and Large and Small Magellanic Clouds (our companion galaxies—see Figure 11.8). Most astronomers consider it likely that the IMF of the First Stars that formed in the Universe was ‘top-heavy’, i.e. lacking in low mass stars.

The stellar initial mass function is a crucial factor in many astrophysical problems. Examples are the determination of the star formation rate (SFR) in galaxies and studies of the chemical evolution of galaxies. In the former case, the SFR is usually determined from a tracer—such as the H α emission line—which is produced by stars in a limited mass range (for H α the most massive stars dominate), and then the result is extrapolated to all stellar masses. In the latter, the relative proportions of different chemical elements depend on the slope and mass range of the IMF because stars of different masses synthesise different elements in different proportions. Typical examples are Oxygen which is produced and ejected into

the ISM mainly by massive stars, and Iron which is due mostly to low- and intermediate-mass stars. Thus, altering the IMF would affect the resultant (O/Fe) ratio following an episode of star formation. The determination of the mass-to-light ratio in galaxies is another example of a problem where the slope and mass range of the IMF are crucial: how much mass is ‘hidden’ in very low mass stars and ‘brown dwarfs’ depends sensitively on the poorly known IMF in the subsolar regime.

STELLAR EVOLUTION. I: LIFE ON THE MAIN SEQUENCE

12.1 Introduction

In lectures 9 and 10 we developed stellar models for stars in hydrostatic equilibrium, and found that a star of a given mass and composition has a unique, fully determined structure. However, the nuclear reactions which power a star by fusing four hydrogen nuclei into a helium nucleus deep in the stellar core change the initial chemical composition of the star over time. Furthermore, convection may set in at some radius, and mix processed and unprocessed gas. As we saw, the equations of pressure, opacity and nuclear power density all depend on the chemical composition of the

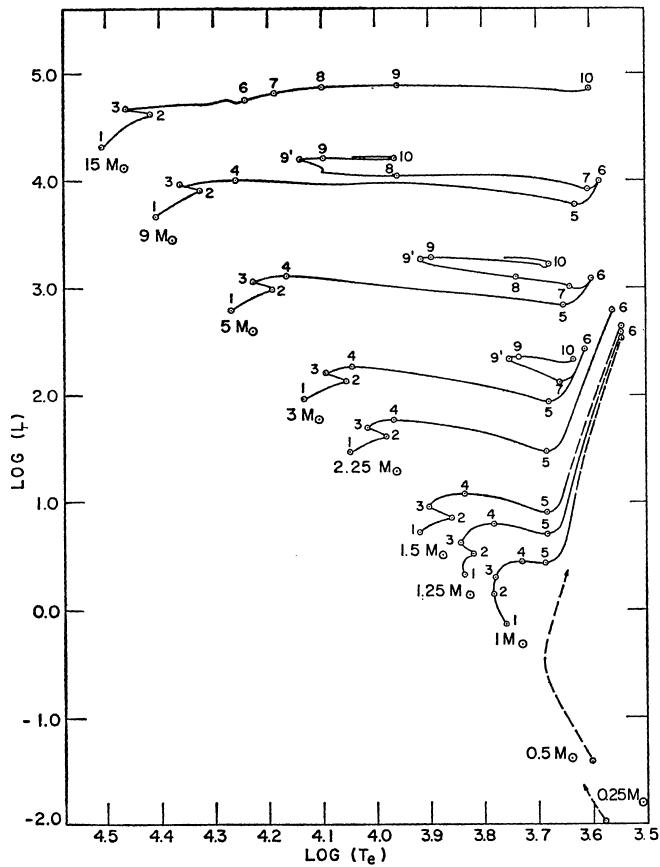


Figure 12.1: Theoretical evolutionary tracks for stars of different masses calculated by Icko Iben in 1967 (ARA&A, 5, 571). The numbered stages are referred to in the text. The y -axis units are L_\odot .

gas. It is therefore inevitable that stars evolve with time. In the next few lectures we shall look in some detail at the process of stellar evolution.

Stellar evolution, as opposed to equilibrium, can be reproduced in our computers by solving a series of equilibrium stellar models—normally referred to as a stellar evolution track—in which one updates in steps the gradual enrichment by elements heavier than hydrogen at different radii within the star. One such set of tracks is shown in Figure 12.1.

The net effect of stellar evolution is to move a star away from the main sequence onto other regions in the H-R diagram. Observations immediately tell us that stars spend most of their lives on the main sequence, and that the subsequent evolution is relatively short-lived. This is evident as soon as one plots on the H-R diagram all of the stars within a given volume (see Figure 12.2): the number of stars away from the main sequence is much smaller than the number of stars on the main sequence.

How do we know that our computer models of stellar evolution are realistic? The requirement of course is that they must match the data, in this case provided by the H-R diagrams of clusters, where all the stars are assumed to

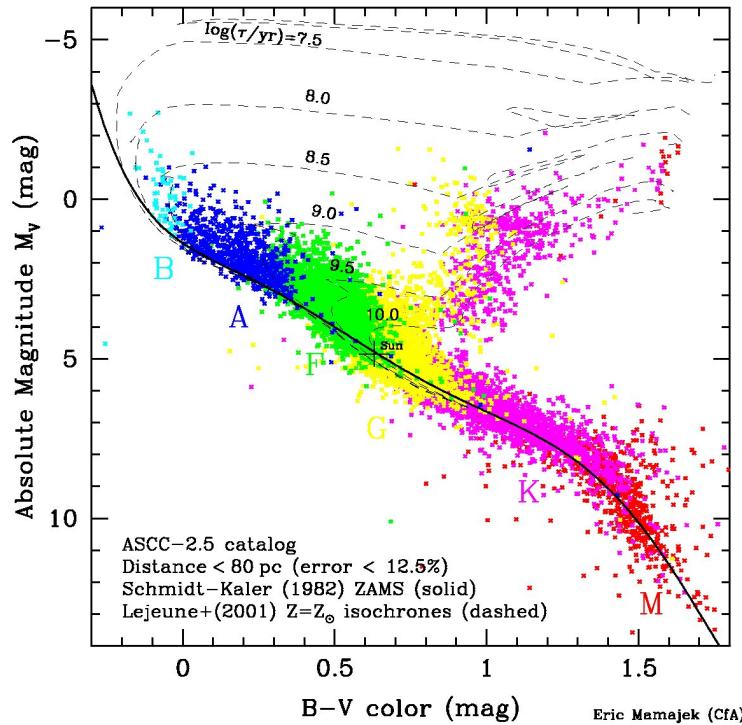


Figure 12.2: Colour-magnitude diagram ($B - V$ vs. M_V) for stars within 80 pc of the Sun, with colour-coding by spectral type. Most stars are found on the main sequence.

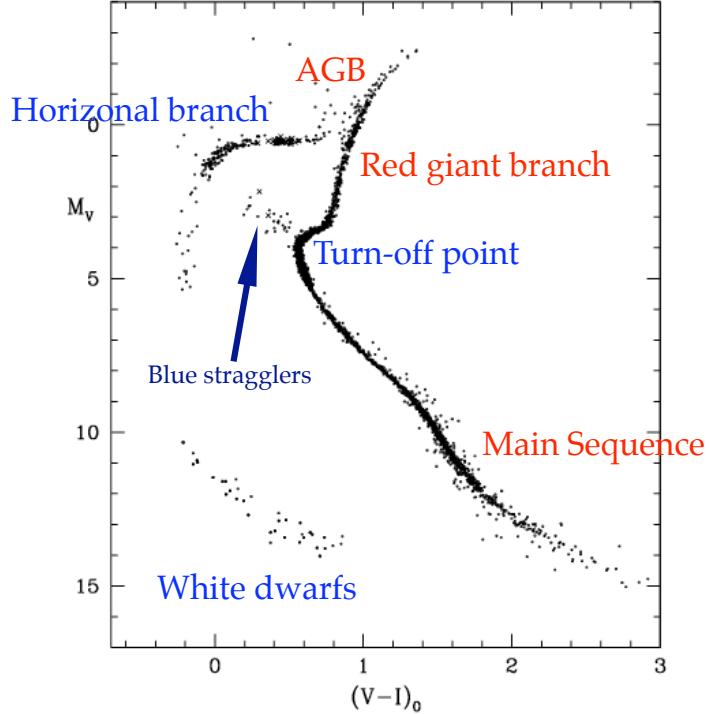


Figure 12.3: Hertzsprung–Russell diagram of a typical globular cluster.

have formed more or less at the same time. This is an adequate assumption in most cases, given the long timescales of stellar evolution, $t = \tau_{\text{nuclear}}$, compared to the star formation timescales, $t \approx \tau_{\text{KH}}$, as we saw in Lecture 7.1 and 7.2.

Particularly useful in this respect are the H-R diagrams of halo globular clusters which, being among the oldest stellar systems known, give us a view of the late stages in the evolution of long-lived stars with masses comparable to that of the Sun (see Figure 12.3).

The combination of computer modelling and observations has led to a well defined picture of stellar evolution. As you may well imagine (at least if you have been following these lectures so far!), mass is the dominant parameter determining not only the length of time a star spends on the main sequence, but also the path that it follows once it evolves off the main sequence. Figure 12.4 gives a schematic summary of the most important evolutionary stages of stars of different masses together with some of their observational characteristics.

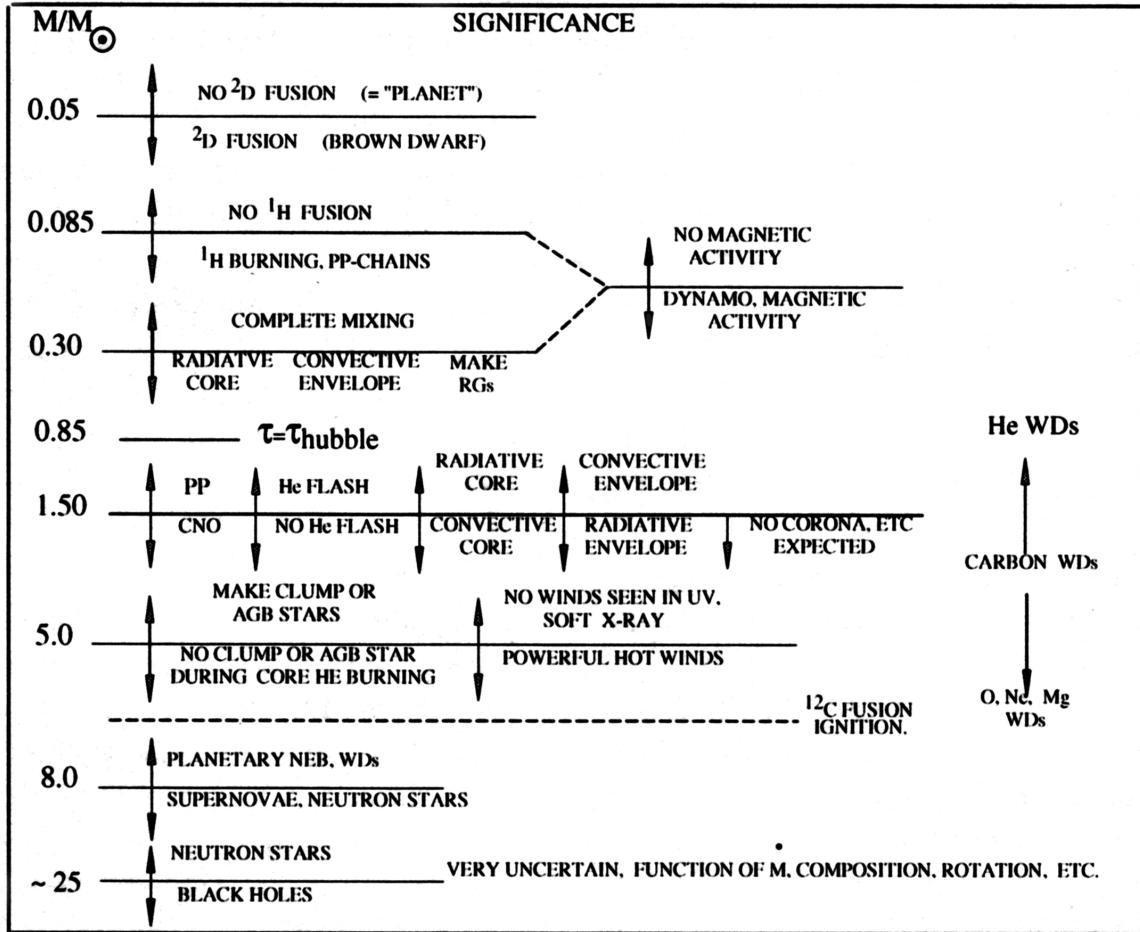


Figure 12.4: Diagram showing the main evolutionary stages of stars in various mass ranges. (Reproduced from Hansen et al. 2004, Stellar Interiors, Springer).

12.2 Life on the Main Sequence

The main sequence has an intrinsic width, even after accounting for errors in the photometry, reddening correction, distance and other factors which may be of importance. The finite width arises because the internal structure of a star does change (slowly) even while it is on the hydrogen-burning main sequence. We look at the factors involved using a $1M_{\odot}$ star as an example.

As H is converted into He deep in the stellar core via the p-p chain, the mean molecular weight μ :

$$\mu = \frac{4}{6X + Y + 2}$$

of the plasma increases (as the He fraction, Y , increases at the expense of the H fraction, X). As a consequence, the pressure would also decrease in

accordance with the ideal gas law:

$$P = \frac{\rho k T}{\mu m_H} \equiv \mathcal{R} \frac{\rho T}{\mu} \quad (12.1)$$

and become insufficient to support the overlying layers of the star, unless either ρ or T , or both increase. The result therefore is that the core will shrink, and its density will increase. In the process, gravitational energy is released, half of which will go into heating the plasma while half of it is radiated away (recall the virial theorem).

In Lecture 7 we saw that rate of energy release via the p-p chain per unit mass of nuclear fuel is:

$$\mathcal{E}_{pp} \propto X^2 \rho T^4 \quad \text{erg s}^{-1} \text{ g}^{-1}.$$

The higher values of T and ρ more than offset the decrease in X ; the resulting increase in \mathcal{E}_{pp} in turn increases the radius of both the core and the envelope, and therefore the luminosity L , of the star. As a consequence, our $1M_\odot$ star will move upwards on the H-R diagram. Referring to Figure 12.1, the portion of the $1M_\odot$ track between points 1 and 2 corresponds to the evolution just described, which takes place over about half of the main sequence lifetime; thus point 2 corresponds roughly to the location of the Sun today, $\sim 30\%$ more luminous than it was on its arrival on the ZAMS.

12.2.1 H shell burning

Figure 12.5 shows the depletion of H from the core of a $1M_\odot$ star with time. At $t \simeq 1 \times 10^{10}$ yr, most of the H burning has become confined to a thick shell surrounding a small but growing He core (point 3 on the $1M_\odot$ track in Figure 12.1). The development of a chemically inhomogeneous structure marks the beginnings of the end of the main sequence life of the star.

The temperature of the He core is insufficient at this stage to ignite He burning. Without nuclear reaction taking place in the He core, the luminosity of the core is $L_{\text{He core}} \simeq 0$. Therefore, according to Eddington's equation for radiative equilibrium (eq. 8.1), the temperature gradient vanishes: $dT/dr \simeq 0$ —that is, the He core is now essentially isothermal.

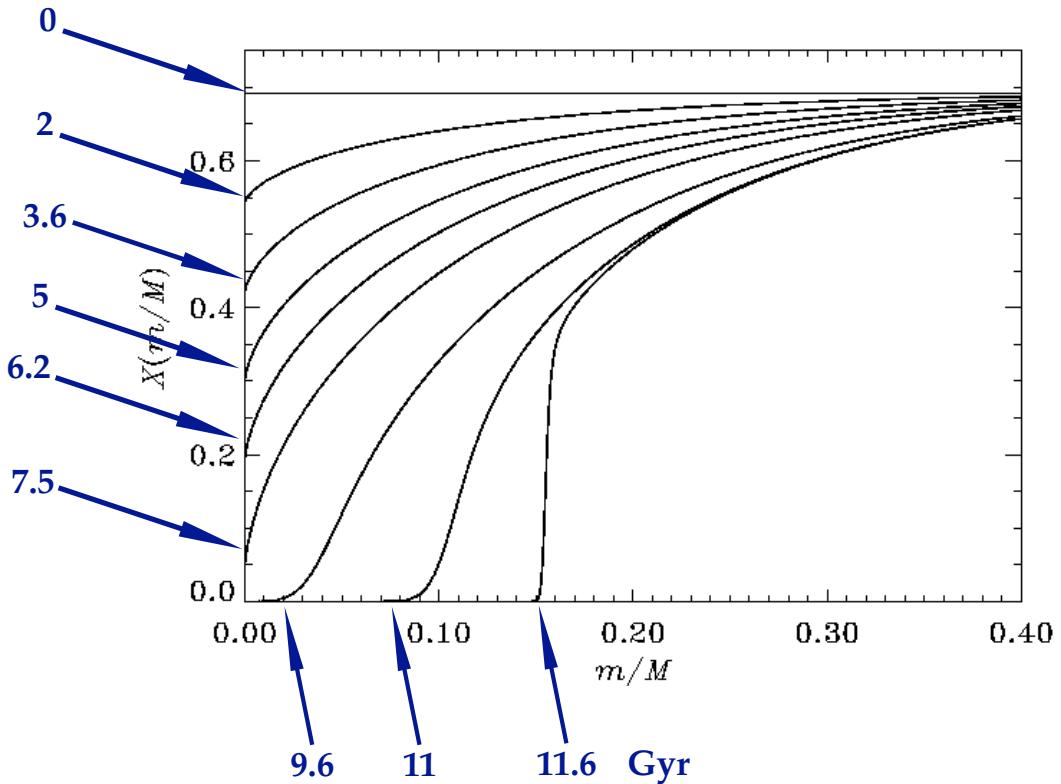


Figure 12.5: Hydrogen profiles showing the gradual exhaustion of H in the core of a $1M_{\odot}$ star. The homogeneous initial model consists of a mixture of H, He and metals with H mass fraction $X = 0.69$. The plot shows X as a function of stellar mass fraction m/M for nine models at different ages in Gyr after the onset of H burning, as indicated. The model at 5×10^9 yr corresponds roughly to the present-day Sun; the last two models are in the shell hydrogen burning phase. (Reproduced from Christensen-Dalsgaard: Lecture Notes on Stellar Structure and Evolution).

Recall our conclusion in Lecture 8.4.1: in order for a star to be in hydrostatic equilibrium, a negative pressure gradient must exist within the star:

$$\frac{dP}{dr} = -\rho g. \quad (12.2)$$

Differentiating eq. 12.1, we find:

$$\frac{dP}{dr} = \frac{\mathcal{R}\rho}{\mu} \frac{dT}{dr} + \frac{\mathcal{R}T}{\mu} \frac{d\rho}{dr}. \quad (12.3)$$

If $dT/dr \simeq 0$, the entire pressure gradient necessary to support the star in pressure equilibrium must be provided by the density gradient. This can work, provided the isothermal core contains only a small fraction of the total stellar mass. We'll return to this point presently.

At this stage, energy generation is taking place in a thick hydrogen-burning shell. The temperature in the shell increases as the core contracts, leading

to higher rates of energy generation. Some of this energy reaches the surface as radiation, and some goes into a slow expansion of the envelope, causing the effective temperature to decrease (recall that the luminosity of a blackbody is given by $L = 4\pi R^2 \sigma T^4$). The net effect is that the evolutionary track bends to the right in the H-R diagram (see Figure 12.1).

12.3 Main Sequence Evolution of Massive Stars

The evolution of more massive stars on the main sequence differs from that of solar mass stars in one important aspect: stars more massive than the Sun have convective cores, as we saw in Lecture 8.3.1. The timescale for convection, defined as the time it takes for a convective element to travel one mixing length (Lecture 8.5), is considerably shorter than the nuclear timescale. Consequently, material is well mixed in the convective core, keeping its composition nearly homogeneous. So, while hydrogen is being converted to helium more quickly in the center of the star where the temperature is greatest (recall the steep temperature dependence of the energy generation rate in the CNO cycle), hydrogen is depleted more or less uniformly throughout the core. In the case of massive stars with convective cores, the core actually *shrinks* as hydrogen is depleted (see Figure 12.6 and compare it with Figure 12.5).

Why does the core shrink? It is primarily because the *opacity* within the core is decreasing, and the lower the opacity, the easier it is for radiation to escape and convection is no longer needed to transfer energy efficiently. The opacity in the core is due primarily to electron scattering which, as we saw in Lecture 5.5, has the form:

$$\kappa_{es} = \kappa_{0,es} \frac{1}{\mu_e} \simeq \kappa_{0,es} \frac{1}{2}(1 + X) \quad (12.4)$$

where $1/\mu_e$ is the number of electrons per nucleon. As hydrogen is converted into helium, positrons are produced. These annihilate with electrons, reducing the number of electrons available for scattering and thus reducing the opacity.

It can also be realised from inspection of Figure 12.1 that T_{eff} changes differently in massive stars and low mass stars, as the stars evolve on the main sequence. In massive stars T_{eff} *decreases* while in stars with

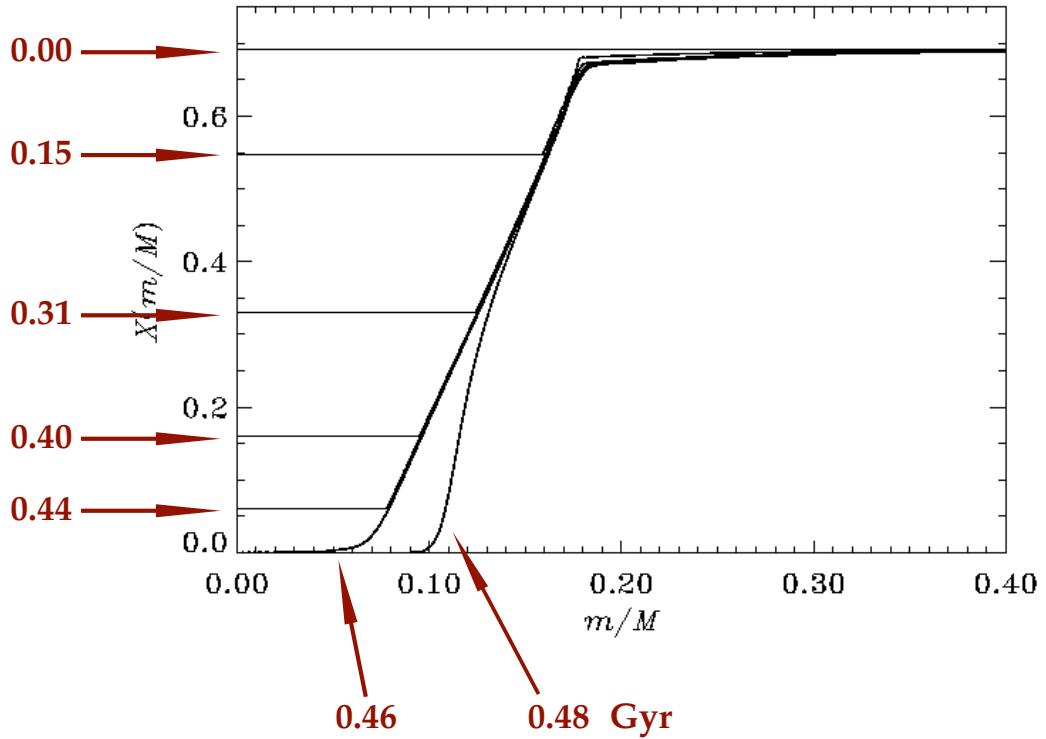


Figure 12.6: Hydrogen profiles showing the gradual exhaustion of H in the core of a $2.5M_{\odot}$ star. The homogeneous initial model consists of a mixture of H, He and metals with H mass fraction $X = 0.69$. The plot shows X as a function of stellar mass fraction m/M for nine models at different ages in Gyr after the onset of H burning, as indicated. The last model is in the shell hydrogen burning phase, the helium core having grown substantially beyond the smallest extent of the convective core. (Reproduced from Christensen-Dalsgaard: Lecture Notes on Stellar Structure and Evolution).

$M \leq 1M_{\odot}$ T_{eff} increases during their main sequence lifetimes. All stars experience expansion of their outer layers throughout their main sequence evolution. Low mass stars experience only a modest expansion, so the energy produced by changes in the core leads to an overall increase in luminosity and surface temperature. More massive stars experience a more rapid expansion. The energy goes into moving the outer layers of the star and results in an overall decrease in effective temperature. Recalling the expression for blackbody luminosity, if the luminosity increases only slowly and the radius increases quickly, the effective temperature must decrease.

In the final stages of H fusion in massive stars, the whole star contracts in an attempt to maintain the energy production by increasing the core temperature. This produces the ‘left hook’ in the H-R diagram visible in Figures 12.1 and 12.7.

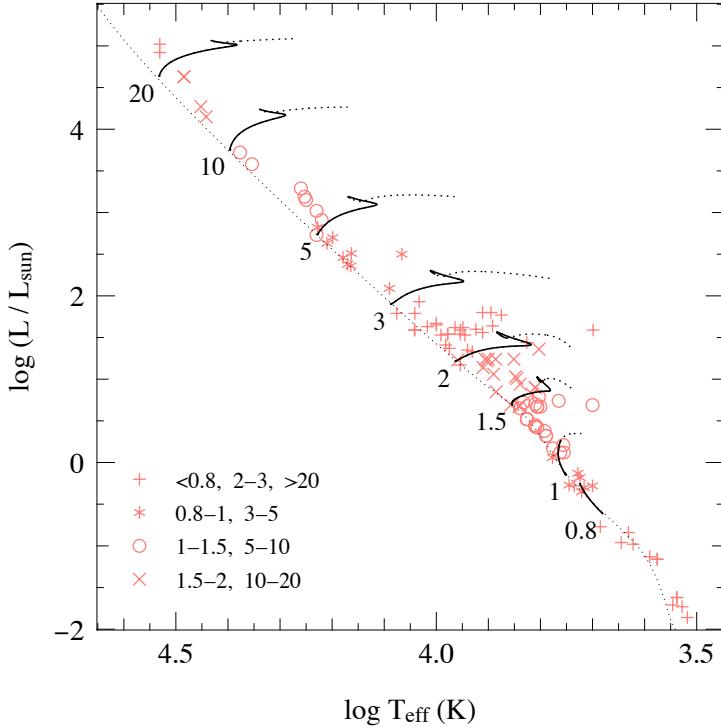


Figure 12.7: Evolutionary tracks in the HR diagram during central hydrogen burning for stars of various masses, as labelled in the legend (in units of M_{\odot}). The thin dotted line is the ZAMS, while the dotted portion of each track shows the continuation of the evolution after central hydrogen exhaustion. The evolution of the $0.8M_{\odot}$ star is terminated at an age of 14 Gyr. Symbols show the locations of binary star components with accurately determined masses, luminosities and radii. Note that each symbol is used to indicate more than one mass range.

12.4 The Mirror Principle of Stars with Shell Burning

Whenever a star has a shell-burning source, it appears to act like a mirror:

“ When the region within a burning shell contracts, the region outside the shell expands; and when the region inside the shell expands, the region outside the shell contracts.”

This is also known as the “shell burning law”. It is not a physical law as such, but an empirical observation, supported by the results of numerical simulations. Although the detailed physics is complicated, we can try and understand it in simple terms as follows. If the virial theorem holds and the total stellar energy remains constant, then the gravitational and thermal energy are each conserved. Any contraction of the core must be accompanied by the expansion of the envelope to conserve the gravitational

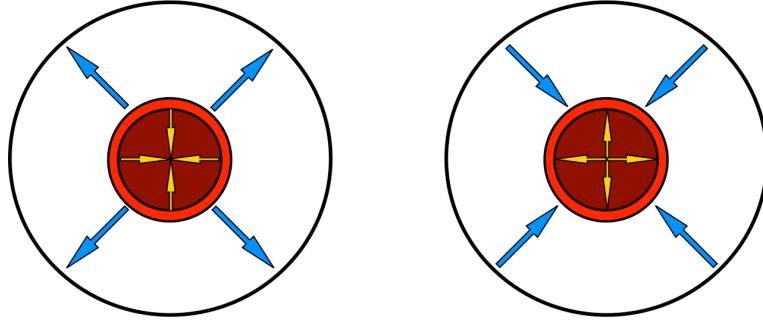


Figure 12.8: The ‘mirror principle’ of shell-burning stars.

potential energy. At the same time, the heating of the core must result in a cooling of the envelope for the thermal energy to be conserved. If the total energy does not remain constant, but $L_{\text{nuc}} > L$, then the envelope will expand considerably on core contraction.

We can state the above a little more quantitatively, as follows. If we consider timescales shorter than the Kelvin-Helmholtz timescale, then both energy conservation

$$\langle U \rangle + \langle K \rangle = \text{constant}$$

and the virial theorem

$$\langle U \rangle + 2\langle K \rangle = \text{constant}$$

must be satisfied, which means that $\langle U \rangle$ and $\langle K \rangle$ must be conserved separately. For $M_c \gg M_{\text{env}}$, we have:

$$|U| \approx \frac{GM_c^2}{R_c} + \frac{GM_c M_{\text{env}}}{R} \quad (12.5)$$

where R is the radius of the star and we have assumed that the binding energy of the envelope is dominated by the gravity of the core. If we take M_c to be constant (and therefore M_{env} too), then:

$$-\frac{GM_c^2}{R_c^2} \frac{dR_c}{dt} - \frac{GM_c M_{\text{env}}}{R^2} \frac{dR}{dt} = 0 \quad (12.6)$$

which implies:

$$\frac{dR}{dR_c} = - \left(\frac{M_c}{M_{\text{env}}} \right) \left(\frac{R}{R_c} \right)^2 \quad (12.7)$$

i.e. the envelope expands as the core contracts and vice versa.

12.5 The Schönberg-Chandrasekhar Limit

We mentioned earlier that once an isothermal He core has developed, hydrostatic equilibrium is maintained by the density gradient alone (eq. 12.3). In 1942, Schönberg and Chandrasekar (ApJ, 96, 161) showed that there is a limit to the fraction of stellar mass that can be supported by an isothermal core, given by:

$$\left(\frac{M_{\text{ic}}}{M}\right)_{\text{SC}} \simeq 0.37 \left(\frac{\mu_{\text{env}}}{\mu_{\text{ic}}}\right)^2 \approx 0.1 \quad (12.8)$$

where μ_{env} and μ_{ic} are the mean molecular weights of the envelope and core respectively. If the mass of the core exceeds the Schönberg-Chandrasekhar limit, pressure support from the core can no longer hold the overlying stellar envelope. As a result, the core collapses on a Kelvin-Helmholtz timescale and the star evolves very rapidly relative to the nuclear timescale of main sequence evolution. This occurs at the points labelled 4 in Figure 12.1.

The mass of an isothermal core may exceed the Schönberg-Chandrasekhar limit if an additional source of pressure can be found to supplement the ideal gas pressure. In stars of relatively low mass ($M \lesssim 2M_{\odot}$), where the helium core is relatively cool and dense, the electrons in the gas start to become *degenerate* and degeneracy pressure can become an important additional source of pressure.

12.6 Electron Degeneracy

Recall that electrons are *fermions*, particles with half-integer spin (angular momentum in units of \hbar), as opposed to *bosons* which are particles with integer spin. Fermions and bosons obey very different statistical rules on how they occupy quantum mechanical states.

The state of a free particle in 1-D space may be described by its position x and its momentum p_x ; together, these two quantities define a position in a two-dimensional phase space (see Figure 12.9). An area in this 2-D phase space has the dimensions of energy \times time (equivalent to momentum \times length). Quantum mechanics tells us that, on tiny scales, this phase space is quantized, that is it consists of discrete cells or states, each of area equal

to Planck constant:

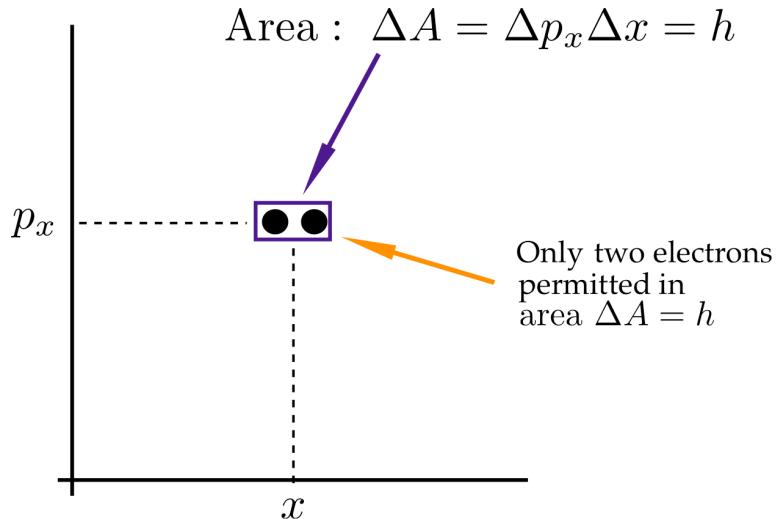


Figure 12.9: Two-dimensional phase space (x, p_x) for a one-dimensional gas showing an area of $h = 6.6 \times 10^{-27}$ erg s (Planck constant). This area is a quantum state that can contain no more than two electrons.

The *Pauli exclusion principle* dictates that no more than one fermion of a given spin can occupy a given state in the x, p_x phase space. Thus, at most two electrons, with opposite spins, can occupy a given state of area h . In 3-D space, phase space has six dimensions: x, y, z, p_x, p_y, p_z , so that a phase space quantum state has a 6-D volume:

$$\Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z = h^3$$

with only two half-spin particles permitted to occupy such a volume element.

A gas of fermions with sufficiently low temperature and/or sufficiently high density will fill all the lowest momentum states. In such degenerate conditions, the momentum (or energy) distribution differs markedly from the Maxwell-Boltzmann distribution of an ideal gas, which describes the situation when the particles occupy only a very small portion of the lowest states. The electrons are forced to occupy higher momentum states than they would normally occupy, resulting in an abnormally high pressure.

The situation is illustrated in Figure 12.10 where panel (a) corresponds to non-degenerate gas. In this case, the greatest density of electrons is in

the lowest momentum state, in accord with the 1-D Maxwell-Boltzmann distribution:

$$\mathcal{P}(p_x) dp_x = \left(\frac{1}{2\pi mkT} \right)^{1/2} \exp \left(-\frac{p_x^2}{2mkT} \right) dp_x \quad (12.9)$$

shown on the right side of panel (a). Also indicated in panel (a) is the momentum p_x at which the kinetic energy is at its average value, $p_x^2/2m = kT/2$. The key to the non-degenerate character is that the lowest-momentum states are not all fully occupied—some are empty and some contain only one electron. In reality, in a truly non-degenerate situation, the electrons would occupy only < 1% of these states, owing to their thermal content.

Consider now what happens if the density is increased significantly (panel b). The gas is ‘squeezed’ by reducing the length of physical space Δx

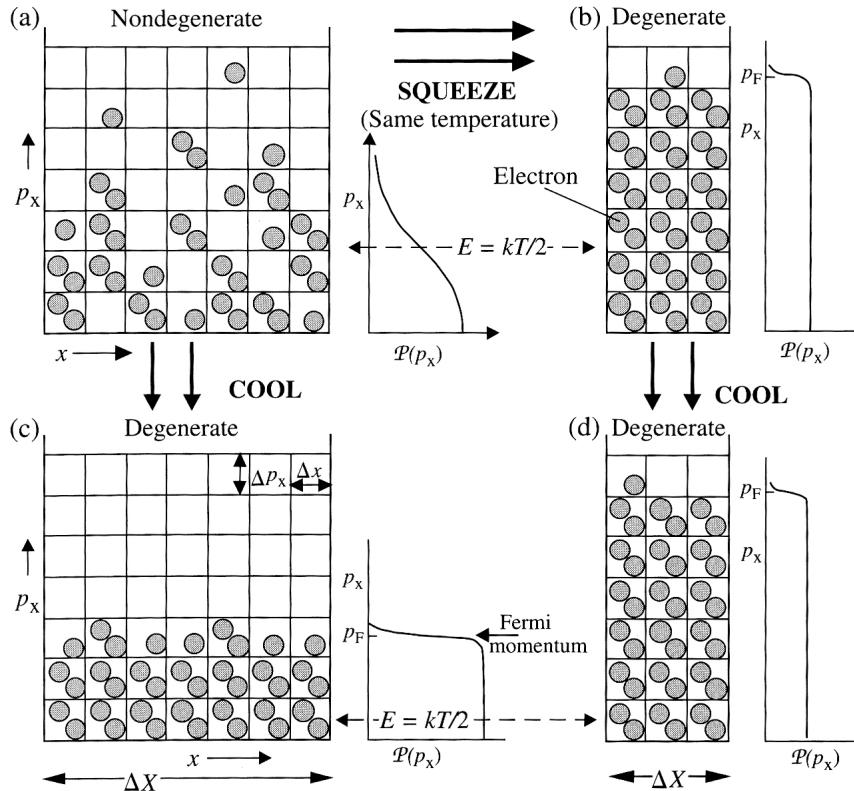


Figure 12.10: The two dimensions of phase space, x and p_x , for a one-dimensional gas of fermions containing 37 particles, for different temperatures and physical lengths ΔX . (a): non-degenerate gas; (b), (c) and (d): fully degenerate gas. The electrons in a degenerate gas are forced to much higher average energies than $kT/2$ because of a shortage of available states at lower energies. (Figure reproduced from Bradt, H., Astrophysics Processes, CUP).

available to it. In the compressed state, very few particles can occupy the lowest-momentum states, because there are not many of them. Electrons are therefore forced to abnormally high energies for a given temperature and the electron gas is said to be *degenerate*. Note that the distribution of particle momenta, $\mathcal{P}(p_x)$, shown on the right of panel (b), is no longer a Gaussian as in eq. 12.9. Rather, electrons are packed solidly up to a maximum value of momentum called the *Fermi momentum*, p_F . It is these high momenta that give rise to degeneracy pressure.

12.6.1 Pressure of a Degenerate Gas

What would be the result if the gas in panel (b) ‘cools down’ to a lower temperature as in panel (d)? The electrons cannot move to lower energy states, because they are all occupied; they stay very nearly in the same quantum states as before, exerting the same pressure as before. Thus, in a completely degenerate gas, the pressure is independent of temperature and is a function of density only:

$$P \propto \rho^{5/3} \quad (12.10)$$

What does ‘cooling down’ mean in these circumstances? The answer lies in the presence of heavier particles: protons, neutrons and atomic nuclei. When the gas is non-degenerate, equipartition of energy dictates that the average electron energy, $p_x^2/2m$, will be the same as the average proton energy. But, the mass of a proton is 1836 times greater than the mass of an electron. Consequently, p_x will be $\sqrt{1836} \simeq 43$ times higher for a proton than for an electron. In the core of a star near the end of its main sequence lifetime, p_x will be two orders of magnitude higher for the He nuclei than for the electrons.

The protons are also constrained by the Pauli exclusion principle, but they occupy much higher p_x states than the electrons. Consequently, as cooling or squeezing takes place, the electrons reach degeneracy long before the protons do. The normal Maxwell-Boltzmann distribution of proton energies serves to define the temperature. Thus, a *partly* degenerate gas does in fact have a well-defined thermodynamic temperature: it is simply the temperature reached by a non-degenerate gas (a ‘thermometer’) that is in contact with it. In the case just described, the thermometer will

indicate a thermodynamic temperature much lower than the high electron momenta would otherwise imply.

Eq. 12.10 applies to a fully degenerate gas. Returning to our stars moving away from the main sequence, if the helium core is only partially degenerate, then some temperature dependence of the pressure remains. The isothermal core of a $1M_{\odot}$ star between the points labelled 3 and 4 in Figure 12.1 is partially degenerate; consequently the core mass can reach $\sim 13\%$ of the entire stellar mass before it begins to collapse. Less massive stars exhibit even higher levels of degeneracy on the main sequence and may not exceed the Schönberg-Chandrasekhar limit at all before the next stage of nuclear burning.

We shall return to degenerate gas later in the course, when discussing white dwarfs.

12.7 Low Mass Stars and Brown Dwarfs

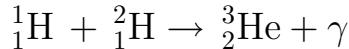
As we saw in Lecture 4 (Figure 4.10), the main sequence lifetimes of stars scale approximately as $t_{\text{MS}} \propto M^{-2.5}$. Thus, while the Sun's main sequence lifetime is approximately 10^{10} yr, for stars with masses $M < 0.85 M_{\odot}$ $t_{\text{MS}} > t_H$, where $t_H = 13.7$ Gyr is the current best estimate for the age of the Universe (see Figure 12.4). Therefore, stars with $M < 0.85 M_{\odot}$ have not yet evolved off the main sequence, with the possible exception of stars in binary systems where the stellar evolution may have been affected by mass transfer from a companion.

We also saw in Lecture 8.3.1 that the size of the convective envelope increases with decreasing stellar mass below $1M_{\odot}$. Stars with $M \lesssim 0.3M_{\odot}$ are fully convective; they keep fusing and mixing until *all* H is converted to He after 10^{12} yr.

In stars with $M \lesssim 0.085 M_{\odot}$, the core never reaches the temperature $T \simeq 4 \times 10^6$ K required to ignite hydrogen burning via the p-p chain (Lecture 10.4). Therefore, this is considered to be the minimum mass of an object for it to be called a ‘star’.¹

¹The minimum stellar mass is a function of composition. The limit $M_{\min} = 0.085 M_{\odot}$ applies to gas of solar composition; in a very metal-poor star with less opacity, and therefore more transparent to escaping

$M \simeq 0.085 M_{\odot}$ is thus taken to be the limit between H-fusing stars and ‘brown dwarfs’, which should more appropriately be called ‘infrared dwarfs’ because their peak emission is at near-IR wavelengths. Indeed brown dwarfs have been discovered primarily by near-IR surveys such as 2MASS (2 Micron All-Sky Survey). While brown dwarfs are too cool to ignite H, they can generate energy via the nuclear reaction:



which is the second step in the p-p chain (Figure 7.5). ${}^2\text{H}$ is deuterium, the heavy isotope of hydrogen, created in Big Bang nucleosynthesis with a primordial abundance $(\text{D}/\text{H})_p = 2.5 \times 10^5$ by number which has not changed much to the present day. Even such trace amounts of D are sufficient to slow the contraction and fading of a brown dwarf (see Figure 12.11).

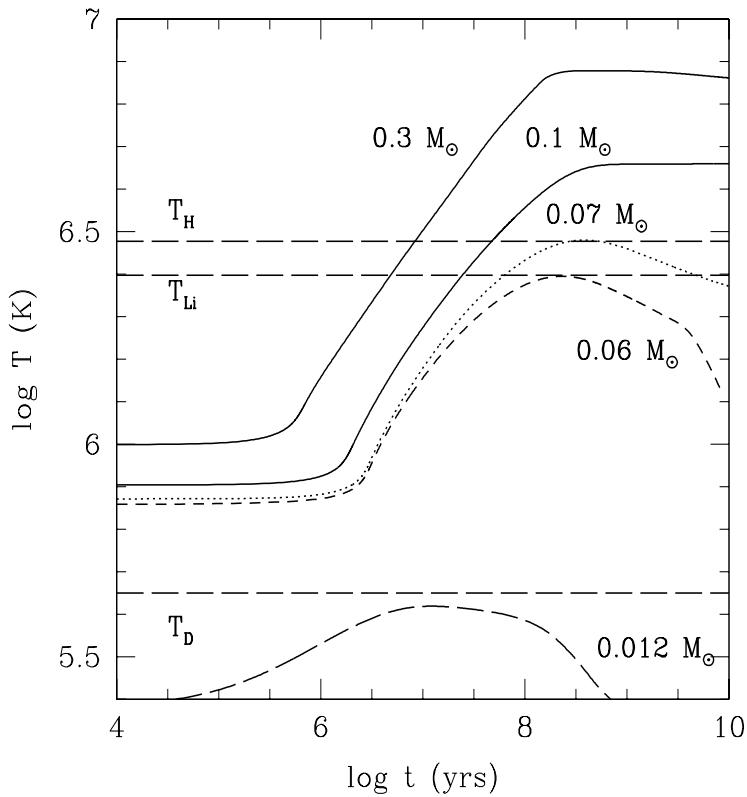


Figure 12.11: Core temperature as a function of age for different masses. T_{H} , T_{Li} , and T_{D} indicate the hydrogen, lithium, and deuterium burning temperatures respectively. (Figure reproduced from Chabrier & Baraffe 2000, ARA&A, 38, 337.)

The first brown dwarfs were discovered in 1995 (Rebolo et al. 1995, Nature, 377, 129; Oppenheimer et al. 1995, Science, 270, 1478). Coincidentally,

radiation, the limit increases to $M_{\min} > 0.085 M_{\odot}$.

this was also the year when the discovery of the first extrasolar planet was announced. Hundreds of brown dwarfs are now known. An ‘acid’ test of the sub-stellar nature of these objects is the presence of methane (CH_4) absorption features in their near-IR spectrum. H-burning stars are far too hot to form methane in their atmospheres, although this is a common molecule in the atmospheres of gas giant planets in our Solar System. Similar test based on the presence of Li and D have also been proposed, although a knowledge of the object age may be required (see Figure 12.11).

Three spectral types have been proposed to classify brown dwarfs: late M (later than M6.5), L and T. Gliese 229B, the prototype T dwarf and the first brown dwarf where methane absorption features were detected, has $T_{\text{eff}} = 950\text{ K}$. Ultracool brown dwarfs of spectral class Y have also been hypothesised. In August 2011 (two weeks before these lecture notes were written!) NASA announced the discovery of six Y dwarfs with temperatures as low as the human body, from data collected with the Wide-field Infrared Survey Explorer (WISE) satellite, now decommissioned (see http://www.nasa.gov/mission_pages/WISE/news/wise20110823.html).

POST-MAIN SEQUENCE EVOLUTION. I: SOLAR MASS STARS

13.1 Introduction

In this and subsequent lectures, we are going to follow the evolution of stars as they progress towards the end of their lives, having concluded core hydrogen burning. The evolution is a sequence of stages which may involve nuclear burning in the core and in concentric mass shells. At various times, either core or shell burning may cease, leading to a readjustement of the internal structure of the star, with expansion or contraction of the core, and subsequent reaction of the envelope. Convection, mixing, and mass loss all play important parts.

All of these processes have been studied extensively with the aid of ever-more sophisticated computer simulations of the internal structure of stars. Unfortunately, they are too complex to be reduced to even illustrative analytical representations. So, this lecture will be largely descriptive, and our principal aim will be to explain the origin and understand the interpretation of different groups of evolved stars in the HR diagram.

As we have emphasised before, the details of the ways stars evolve off the main sequence and their ultimate fate all depend on the stellar mass. In this Lecture we consider low mass stars, with $M \sim 1M_{\odot}$. Figure 13.1 summarises the main evolutionary stages.

13.2 The Red Giant Branch

In the previous lecture, we left our solar mass star at the end of its main sequence evolution, burning H in a shell encompassing an isothermal He core. Because of the mirror action of the shell, the outer layers expand and cool and the star moves to the right in the H-R diagram. During this phase, which for a $1M_{\odot}$ star lasts ~ 2 Gyr, the star moves along the “sub-

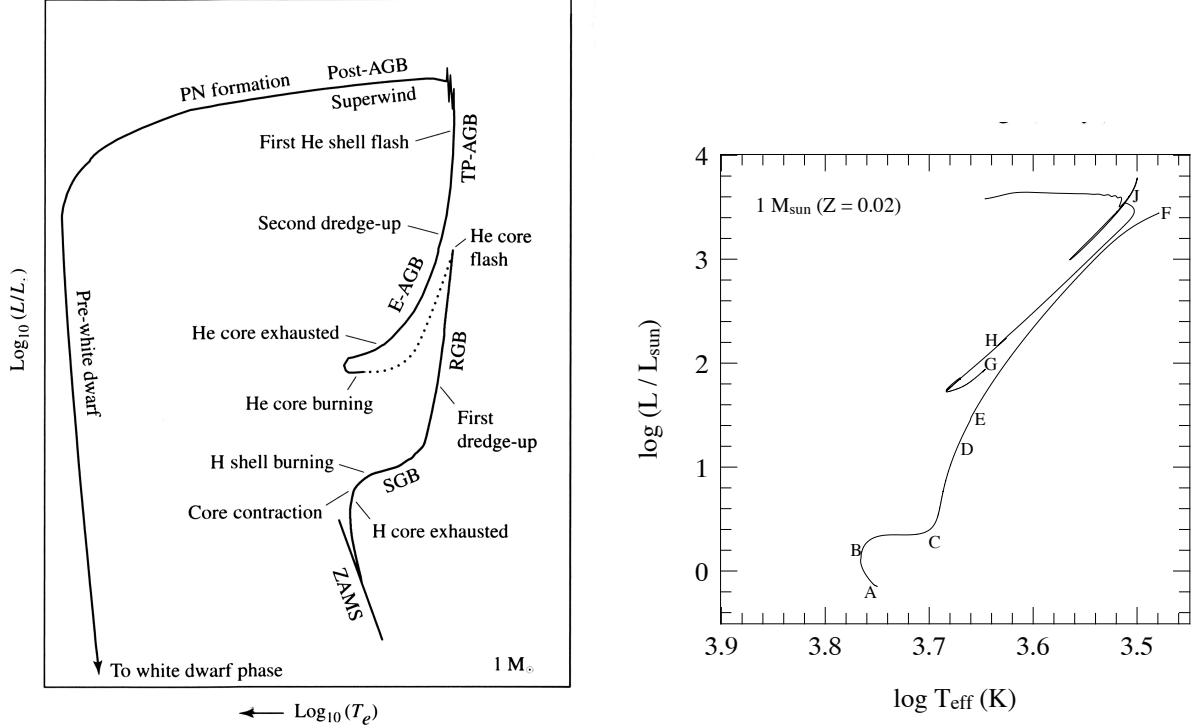


Figure 13.1: Schematic diagrams of the evolution of a $1M_{\odot}$ star of solar metallicity, from the main sequence to a white dwarf (*left*), and from the main sequence to a planetary nebula (*right*). The various acronyms and the lettering indicating different evolutionary stages are explained in the text.

giant branch" (SGB), between points B and C in Figure 13.1. At point C the star's He core has become degenerate.

As the star expands, however, the effective temperature cannot continue to fall indefinitely. With the expansion of the stellar envelope and the decrease in effective temperature, the photospheric opacity increases due to the additional contribution from H^- ions. When the temperature of the outer layers of the star falls below ~ 5000 K, they become fully convective. This enables a greater luminosity to be carried by the outer layers and hence abruptly forces the evolutionary track to travel almost vertically upwards to the *red giant branch* (RGB).

The star now moves along the same path, but in reverse, followed by a fully convective pre-main-sequence star on its approach to the main sequence, which, as we discussed in Lecture 11.7, is a nearly vertical line in the H-R diagram known as the Hayashi track. We also explained then that this track represents a boundary, in the sense that the region to the right of the track is 'forbidden': there is no mechanism that can adequately transport the luminosity out of the star at such low effective temperatures.

A $1M_{\odot}$ star will spend ~ 0.5 Gyr on the RGB, moving from point C to point F in Figure 13.1 at an accelerating evolutionary pace, driven by what is occurring in the core. As H-fusion in the shell deposits more He onto the core, the mass of the core increases. For a fully degenerate gas, it can be shown that this results in a *contraction* of the core. As the core contracts, the density of the H-burning shell directly on top of it increases; this in turns leads to higher fusion efficiency and higher luminosity. This is a runaway process. By the end of the RGB, at point F in Figure 13.1, the degenerate He core has reached a mass of $\sim 0.5M_{\odot}$, and has contracted sufficiently to achieve the temperature required to ignite He fusion.

Table 13.1 Radii and Luminosities of Red Giants

	G0	G5	K0	K5	M0	M5
$\log(R/R_{\odot})$	0.8	1.0	1.2	1.4	1.6	1.9
$\log(L/L_{\odot})$	1.5	1.7	1.9	2.3	2.6	3.0

Table 13.1 gives some representative values for the sizes and luminosities of red giant stars; a main sequence G V star may end up as a high-K or low-M luminosity class III giant. Note that the values in Table 13.1 depend largely on the spectral type, and not on the mass: stars of a wide range of masses follow similar tracks on the RGB, becoming redder and more luminous as the core grows.

For a star with a degenerate core, the density contrast between the core and the envelope is so large that the two are practically decoupled. The pressure at the bottom of the extended envelope is very small compared to the pressure at the edge of the core and in the H-burning shell separating core and envelope. This implies that the efficiency of the shell burning is completely determined by the mass of the He core and not by the envelope. Thus, there is a strong and steep relationship between the He core mass and the luminosity of a red giant, which is entirely due to the hydrogen shell-burning source:

$$L \simeq 2.3 \times 10^5 \left(\frac{M_c}{M_{\odot}} \right)^6 L_{\odot} \quad (13.1)$$

Therefore, the evolutionary tracks of stars of different masses all converge onto the Hayashi line that is the RGB; from the position of a star on the RGB we can deduce the value of M_c , but the total mass is more difficult.

13.2.1 Metallicity Dependence of the RGB

The red giant branch does however exhibit a metallicity dependence. As we discussed in earlier lectures, fully convective stars are on the Hayashi line which is the locus of the lowest values of T_{eff} at which a star of a given luminosity can shine. Convection is related to the opacity (Lecture 8.3.1), and the opacities of stellar atmospheres depend on metallicity. This is the case even when H^- is the main source of opacity because the metals provide the free electrons that form H^- .

With a higher metallicity, an optical depth $\tau \simeq 2/3$ is reached sooner, or at lower density, as we travel from the stellar ‘surface’ to the core (recall the treatment of the optical depth in Lecture 5 where we took $\tau \simeq 2/3$ as

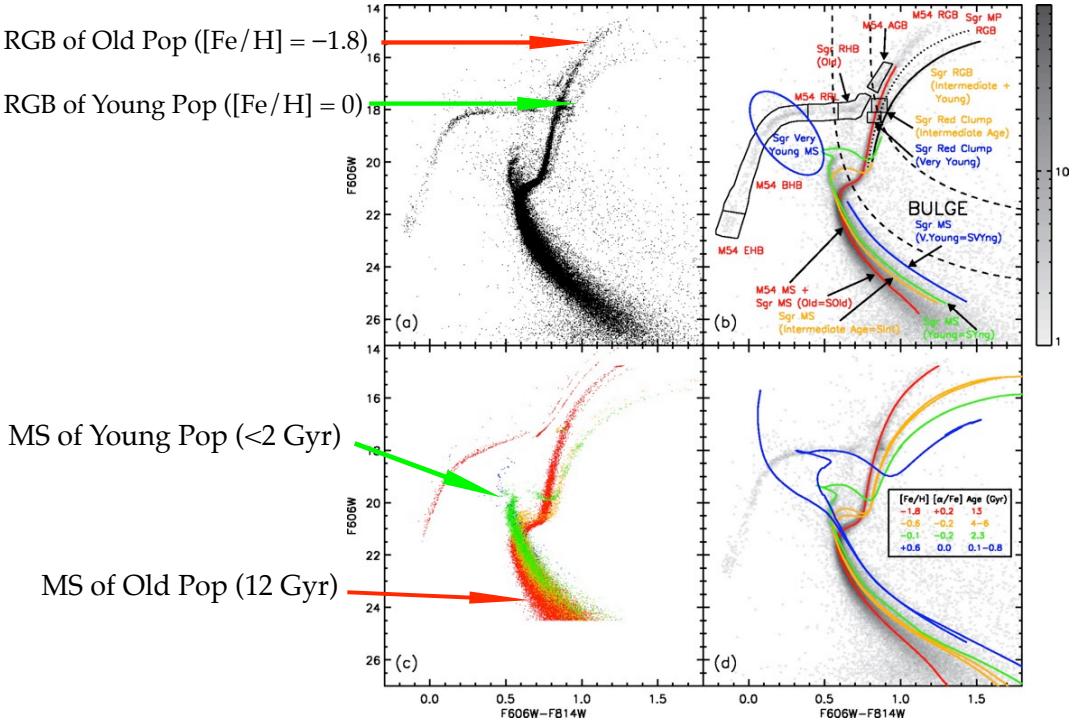


Figure 13.2: Colour-magnitude diagram (CMD) of 60 000 stars in the field of the globular cluster M54, which also includes the dwarf spheroidal galaxy Sagittarius (Sgr), which is in the process of merging with our own Galaxy. Multiple populations, of different ages and metallicities, can be distinguished in this complex CMD, allowing the past history of star formation of this companion galaxy to the Milky Way to be reconstructed. Highlighted in panel (a) are two well-separated red giant branches, whose metallicities differ by two orders of magnitude. The photometric data were measured from high spatial resolution images obtained with the Advanced Camera for Surveys on the *Hubble Space Telescope*, and published by Siegel et al. 2007, ApJ, 667, L57.

the definition of the stellar photosphere). Thus, metal-rich stars of a given mass have slightly larger radii and lower effective temperatures than stars of the same mass but lower metallicity. For the same reason, the RGB of metal-rich stars runs at slightly lower temperatures than that of metal-poor stars. A vivid demonstration is provided by the colour-magnitude diagrams of stellar systems consisting of multiple populations (see Figure 13.2).

The very steep temperature dependence of the opacity at the effective temperatures of red giants, $\kappa \propto T^9$, provides an intuitive explanation for the fact that the Hayashi line is close to vertical on a $L-T_{\text{eff}}$ diagram. Suppose that a cool star of constant L could increase its radius, even by a small amount. This would lower the value of T_{eff} and therefore the opacity of the outer layers. As a result, we would be able to see deeper into the star, down to a depth where $\tau \simeq 2/3$, at nearly constant T_{eff} .

13.2.2 Mass Loss on the RGB

An important process experienced by stars while they are in the red giant phase is mass loss. As the stellar luminosity and radius increase while 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 via radiation pressure (Lecture 9.2.2) on atoms and grains.

Grains are microscopic solid particles that can condense out of the gas phase at the values of temperature and pressure typical of the extended atmospheres of late-type giant and supergiant stars. Their presence in these environments is indicated by a number of infrared spectral features, such as the $9.7\,\mu\text{m}$ band due to silicates, which can appear in emission or absorption in the spectra of red giants and supergiants. The winds from these stars are responsible for distributing grains into the interstellar medium, where they can subsequently grow through accretion of atoms. Interstellar grains, or *dust* as they are often referred to, are an important constituent of the diffuse interstellar medium. They regulate the heating and cooling of the ISM, act as a catalyst in the formation of H_2 molecules, and of course are responsible for interstellar extinction, the process that reddens the light of all stars.

Returning to mass loss on the RGB, red giant stars are observed to lose

mass in the form of a slow wind ($v_{\text{wind}} \simeq 5\text{--}30 \text{ km s}^{-1}$) at a rate $\dot{M} \simeq 10^{-8} M_{\odot} \text{ yr}^{-1}$. A $1M_{\odot}$ star loses $\sim 0.3M_{\odot}$ of its envelope mass by the time it reaches the tip of the giant branch.

When calculating the effect of mass loss in evolution models an empirical formula due to Reimers is often used:

$$\dot{M} = -4 \times 10^{-13} \eta \frac{L}{L_{\odot}} \frac{R}{R_{\odot}} \frac{M_{\odot}}{M} M_{\odot} \text{ yr}^{-1} \quad (13.2)$$

where the efficiency factor $\eta \simeq 0.25\text{--}0.5$. However, this relation is based on observations of only a handful of stars with well-determined stellar parameters. Note that eq. 13.2 implies that a fixed fraction of the stellar luminosity is used to lift the wind material out of the gravitational potential well of the star.

13.2.3 The First Dredge-up

As the star climbs up the RGB, its convection zone deepens until the base reaches down into regions where the chemical composition has been modified by nuclear processes. This transports processed material from the deep interior to the surface in what is referred to as the first *dredge-up* phase. This phase provides us with the first opportunity to verify empirically our ideas about nuclear burning which, up to this point, has been completely hidden from view.

For example, Li is destroyed by collisions with protons at relatively low temperatures, $T \gtrsim 2.7 \times 10^6 \text{ K}$; as a consequence of the first dredge-up the atmospheres of evolved stars exhibit a Li deficiency compared to the Li abundance of the proto-stellar nebula. Indeed, the Li abundance is often used as a test to decide whether the atmospheric abundances can be trusted to represent the composition of the gas from which the star formed.

Similarly, the surface He abundance increases and the H abundance decreases while a star ascends the RGB. In intermediate mass stars ($M \sim 5M_{\odot}$), the convective envelope brings material processed by the CNO cycle to the surface. The C-N cycle reaches equilibrium before the O-N cycle, and thus CN-processed material (N enriched, C depleted) is first exposed on the surface. The N abundance increases by a factor of ~ 2 , C is decreased by 30% and O is unchanged. Many red giants are observed to have

CN-processed material in their atmospheres.

13.3 The Red Giant Tip and the Helium Flash

At the tip of the RGB (point F in the right panel of Figure 13.1), the central temperature and density have finally become high enough ($T > 10^8$ K) for quantum tunnelling to overcome the Coulomb barrier between He nuclei, allowing the triple-alpha process to begin. Some of the resulting ^{12}C is further processed into ^{16}O via capture of an alpha particle (see Lecture 7.4.3). This is the onset of the helium burning phase of evolution. Unlike H-burning, the reactions involved in He-burning (Lecture 7.4.3) 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).

The electrons in the core of a $1M_\odot$ star are completely degenerate by the time the star reaches point F in Figure 13.1. Ignition in a degenerate core results in an explosive start of the fusion known as the “*Helium Flash*”. The ignition of He-fusion raises the temperature of the core, but this does not raise the pressure, because in a degenerate gas $P \neq f(T)$. Thus, as T increases the core does *not* expand, and the density remains the same. As we saw in Lecture 7.4.3, the energy generation rate of the triple alpha reaction has an extraordinarily steep dependence on T : $\mathcal{E}_{3\alpha} \propto Y^3 \rho^2 T^{40}$. Thus, the rise in T leads to more efficient fusion, which in turn raises the T , and so on: a degenerate core that is ignited acts like a bomb!

The thermonuclear runaway leads to an enormous overproduction of energy: at maximum, the local luminosity in the helium core is $L_c \sim 10^{10} L_\odot$, comparable to the luminosity of a small galaxy! However, this only lasts for a few seconds. All the nuclear energy released is absorbed by expansion of the non-degenerate layers surrounding the core, so none of this luminosity reaches the surface. The short duration, and the presence of a very extended convective envelope that can absorb the energy created by the flash explain why the He flash has never been observed, other than in our computers (see Figure 13.3).

Since the temperature increases at almost constant density, degeneracy is eventually lifted when $T_c \simeq 3 \times 10^8$ K. Further energy release increases the

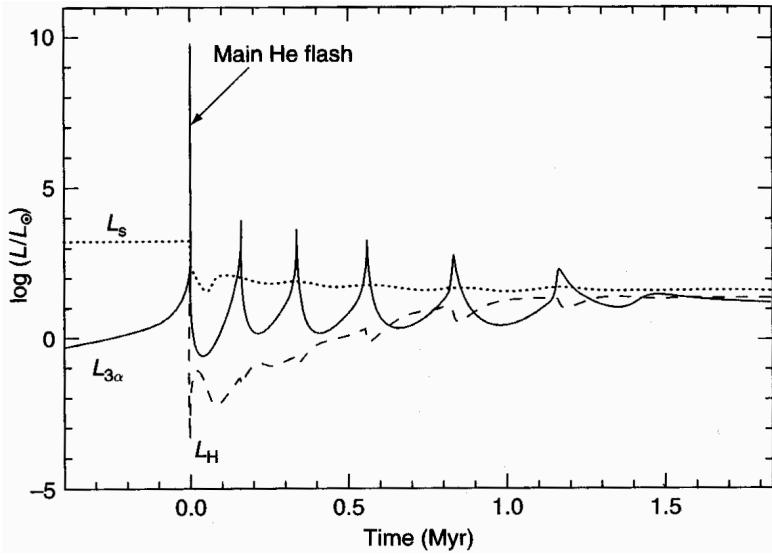


Figure 13.3: The helium flash. Evolution with time of the surface luminosity (L_s), the He-burning luminosity ($L_{3\alpha}$) and the H-burning luminosity (L_H) during the onset of He burning at the tip of the RGB in a low-mass star. Time $t = 0$ corresponds to the start of the main helium flash. (Figure from Salaris & Cassisi, Evolution of Stars and Stellar Populations, Wiley).

pressure when the gas starts behaving like an ideal gas and thus causes expansion and cooling. This results in a decrease of the energy generation rate until it balances the energy loss rate and the core settles in thermal equilibrium at $T_c \simeq 10^8$ K. Further nuclear burning is thermally stable.

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 Figure 13.3) for ~ 1.5 Myr, before degeneracy in the centre is completely lifted and further He burning proceeds stably in a convective core.

This is the situation when stars with a *non-degenerate* core reach $T_c \sim 10^8$ K at the tip of the RGB. In this case, the rise in T_c is accompanied by an increase in P . The core expands, T_c and P decrease, the energy production drops, and the core shrinks until it reaches hydrostatic equilibrium again. So, in this case, gravity acts like a regulator and the star does not experience a He flash. The dividing line between stars with degenerate and non-degenerate cores at the tip of the RGB is $\sim 2M_\odot$; stars with $M \lesssim 2M_\odot$ undergo a He flash, while in those with $M \gtrsim 2M_\odot$ He burning is ignited without a thermonuclear runaway event.

13.4 The Horizontal Branch

In our $1M_{\odot}$ star, the He flash occurs at point F in Figure 13.1. Evolution through the helium flash was not calculated for the model shown in this figure, because the evolution is very fast and hard to follow. The evolution of the star is resumed at point G when the star has settled into a new equilibrium configuration with an expanded non-degenerate core which is hot enough to burn He. The star now has two sources of energy generation: core He fusion and shell H fusion. However, the H-burning shell has also expanded and now has lower temperature and density; consequently, it generates less energy than when the star was at the upper end of the RGB. The lower total luminosity is insufficient to keep the star in its distended red giant state; the star shrinks in size, dims and settles on the *horizontal branch*.

At this point the luminosity and radius of the star have decreased by more than one order of magnitude from their values 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.

The horizontal branch is the core He-burning equivalent of the core H-burning main sequence. However, while a $1M_{\odot}$ star spends $\sim 1 \times 10^{10}$ yr on the main sequence, its core He-burning phase on the HB lasts only ~ 120 Myr, or $\sim 1\%$ of its main sequence lifetime, because of the much higher luminosity of the He-burning phase.

13.4.1 The Horizontal Branch Morphology

For the $1M_{\odot}$ star of solar composition considered in Figure 13.1, He burning occurs between points G and H. The location of the star in the H-R diagram does not change very much during this period, always staying close to (but somewhat to the left of) the red giant branch. Its luminosity is $\sim 50L_{\odot}$ for most of the time, a value determined mainly by the core mass. Since the core mass at the start of helium burning is $\sim 0.45M_{\odot}$ for all low-mass stars, irrespectively of stellar mass, the luminosity at which He burning occurs is also almost independent of the total stellar mass. Thus, it is only the

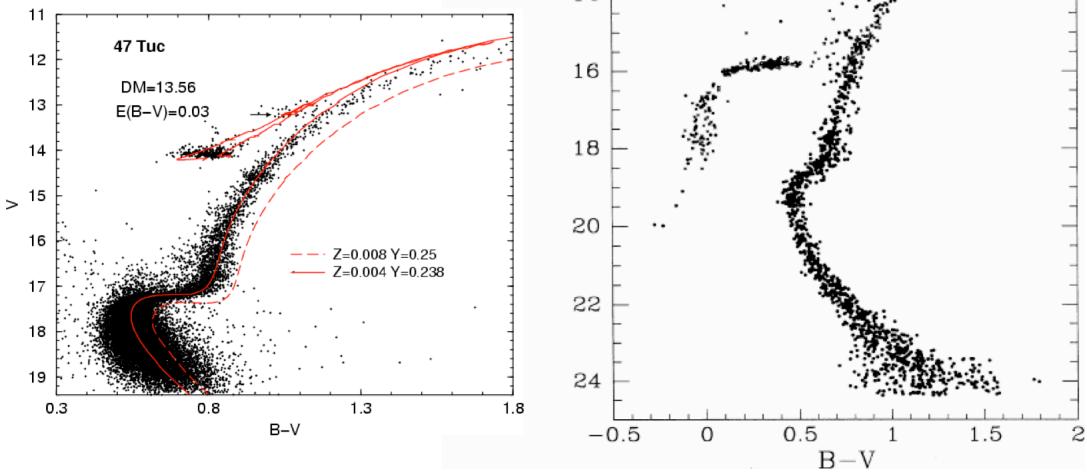


Figure 13.4: The Fe abundances of the stars in these two Galactic globular clusters differ by a factor of ~ 40 . In the more metal-poor globular cluster, M15 on the right, the horizontal branch extends much further to the blue (hotter effective temperatures, implying smaller radii) than in the more metal-rich one (47 Tuc, on the left).

envelope mass that varies from star to star, either because of differences in mass on the ZAMS, or as a result of different amounts of mass loss on the RGB.

At solar metallicity, all core He-burning stars occupy a similar locus in the H-R diagram, which is referred to as the ‘*red clump*’. However, in metal-poor globular clusters these stars are found to be spread out over a range of effective temperatures at the same approximate luminosity—hence the ‘*horizontal branch*’ nomenclature. It is thought that the location of a star on the HB is a reflection of its envelope mass: stars with smaller envelopes (and hence radii) are bluer (and are therefore found on the left of the HB).

The extent of the horizontal branch in globular clusters seems to be related to their metallicity. The more metal-rich globular clusters tend to have a red HB, while in the more metal-poor ones the HB extends further to the blue (see Figure 13.4). In this respect, the red clump of solar metallicity stars may be simply the red extreme of the HB. But metallicity is not the only factor at play here, because there are known cases of globular clusters with a blue HB and others with a red HB even though their metallicities are similar! This has led astronomers to invoke a ‘second parameter’, a label that acknowledges that some other unknown physical effect is responsible for HB morphology differences in clusters that seem to be very similar in most of their physical properties. Age, He content and rotation have been

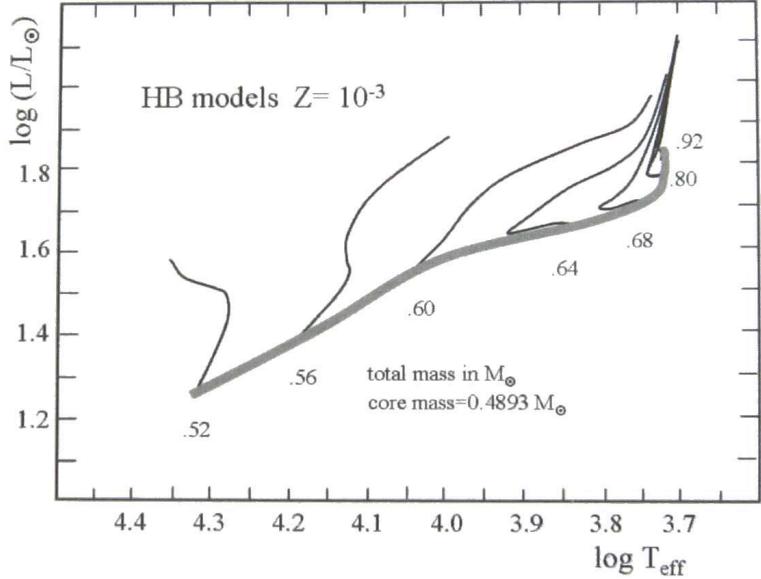


Figure 13.5: Location of the zero-age horizontal branch (thick grey line) for a metallicity $Z = 0.001$ which is typical of Galactic globular clusters. The models shown all have the same core mass ($M_c = 0.489 M_\odot$) but varying total (i.e. envelope) mass, which determines their position in the H-R diagram. Evolution tracks during the HB phase for several total mass values are shown as thin solid lines. (Figure from Maeder, A., Physics, Formation and Evolution of Rotating Stars, Springer-Verlag).

proposed, but the underlying cause of different HB morphologies remains a long-standing problem in stellar astrophysics.

Once a star has entered the HB (on the left, the right or in between), evolution moves it to the right during the core He fusion phase, due to the increasing depth of the convection zone (see Figure 13.5).

13.5 The Asymptotic Giant Branch (AGB)

Returning to Figure 13.1, by the time the star has reached stage H, it has exhausted its supply of He in the core which now consists of C and O. The core contracts again, but in stars with masses $M < 8M_\odot$ there is insufficient gravitational energy to generate the high temperatures required to fuse C and O into heavier nuclei. Thus, no more core fusion takes place in these stars. However, the core contraction generates sufficient heat for the surrounding layer of He to start fusing in a shell.

The next phase of the evolution is very similar to the evolution we have already discussed following exhaustion of the hydrogen burning core. The contraction of the core leads to a strong expansion of the star's outer

layers, causing its surface temperature to drop and moving the star to the right and upwards in the H-R diagram along the *Asymptotic Giant Branch* (AGB). A $1M_{\odot}$ AGB can reach a luminosity $L \sim 10^5 L_{\odot}$! The AGB is so named because the evolutionary track approaches the line of the RGB asymptotically from the left, and indeed it can be thought of as the shell He-burning analogue of the shell H-burning RGB. At solar metallicities the AGB lies close to the RGB, but in metal-poor globular clusters the AGB and RGB appear well separated.

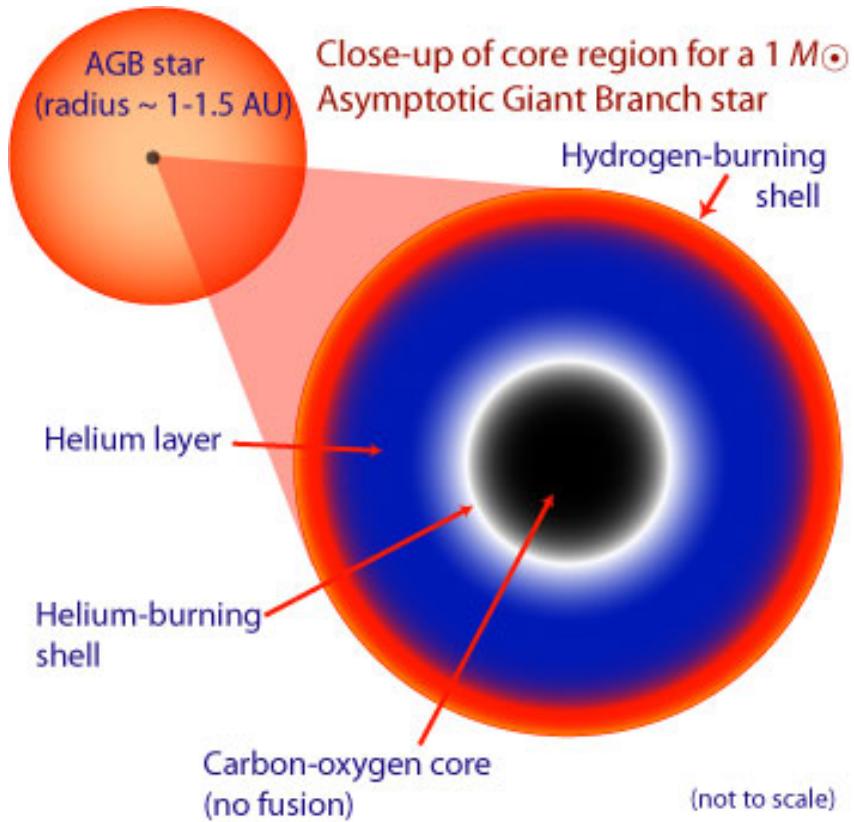


Figure 13.6: Schematic structure of a solar mass star during the RGB phase.

At this point, the star consists of: (i) a degenerate C+O core; (ii) a He-burning shell; (iii) an inert He-shell around it; (iv) a H-burning shell; and (v) an outer H-rich convective envelope (see Figure 13.6).

The evolution is now complex because the huge differences between the two nuclear fusion processes do not allow a steady state to exist. The two shells supply the luminosity of the AGB star alternately in a cyclical process, or a thermal pulsation, which has a period of ~ 1000 yr, with the changes triggered by shell flashes.

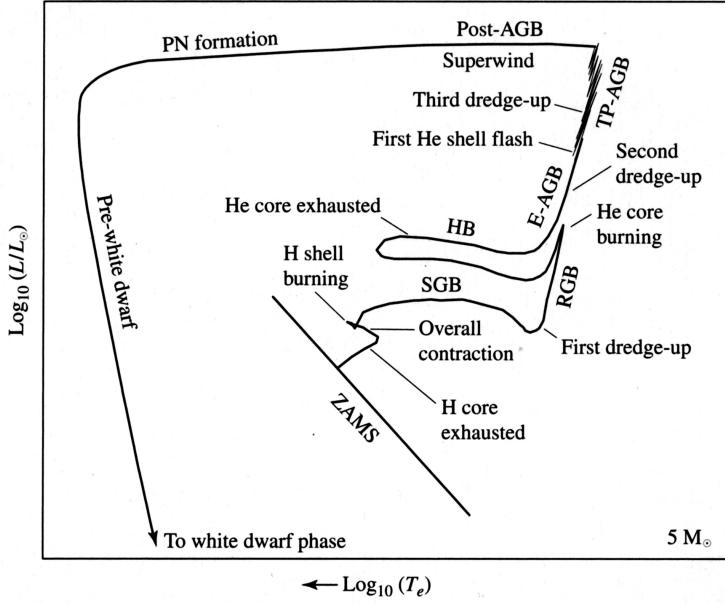


Figure 13.7: Schematic diagram of the evolution of a $5M_{\odot}$ star of solar metallicity, from the main sequence to a white dwarf.

Although brief (a $1M_{\odot}$ star will spend $\sim 5 \times 10^6$ yr on the AGB), the AGB is an interesting and important phase of stellar evolution. The expansion and cooling of the envelope increase its opacity and the depth of the convection zone which can reach down to the chemical discontinuity between the H-rich outer layer and the He-rich region between the two burning shells. The mixing that results during this *second dredge-up* phase increases the He and N content of the envelope.

In stars more massive than $\sim 2M_{\odot}$, there a *third dredge-up* as the tip of the AGB is approached, driven by thermal pulsations (see Figure 13.7). This brings to the surface C-rich material and ‘s-process’ elements (see later) In stars more massive than $\sim 3M_{\odot}$, the base of the convective envelope becomes hot enough for the CN cycle to operate and the dredged-up C is converted to N, in a process called ‘hot bottom burning’.

At the low temperatures of the atmospheres of AGB stars, most of the C and O atoms are bound into CO molecules, since this is the most stable molecule. In the protostellar nebula, $C/O \sim 0.5$ (see Figure 6.13). If this initial abundance has not been changed appreciably and all the C is locked in CO molecules, 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. However, as a result of repeated dredge-up events, at some point the C/O ratio can exceed

unity. In this case all O is locked into CO molecules and the remaining C forms carbon-rich molecules and dust grains, e.g. C₂, CN, C_nH_n, and carbonaceous grains like graphite and SiC. 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 Thermal-Pulse (TP) AGB phase.

13.5.1 Slow Neutron Capture Nucleosynthesis

Direct evidence for active nucleosynthesis in AGB stars was provided in 1953 by the detection of technetium (₄₃Tc), the lowest atomic number element *without* any stable isotopes: every form of it is radioactive. The longest lived isotope, ₄₃⁹⁹Tc, decays on a timescale of only 2×10^5 yr. Spectroscopic observations actually 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 s-process which we have already discussed in Lecture 7.5. 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 a number of reactions. AGB stars are nowadays considered to be major producers in the Universe of carbon, nitrogen and of elements heavier than iron synthesised via the s-process. They also make an important contribution to the production of ¹⁹F, ²⁵Mg, ²⁶Mg and other isotopes.

13.5.2 Mass loss and the post-AGB phase

During the AGB phase, the mass loss increases dramatically from $\dot{M} \simeq 10^{-8}$ to $\simeq 10^{-4} M_\odot \text{ yr}^{-1}$. We can easily see this:

$$\frac{dM_{\text{star}}}{dt} = -\frac{dM_{\text{wind}}}{dt}, \quad (13.3)$$

and

$$\frac{dM_{\text{env}}}{dt} = -\frac{dM_{\text{wind}}}{dt} - \frac{dM_{\text{core}}}{dt}. \quad (13.4)$$

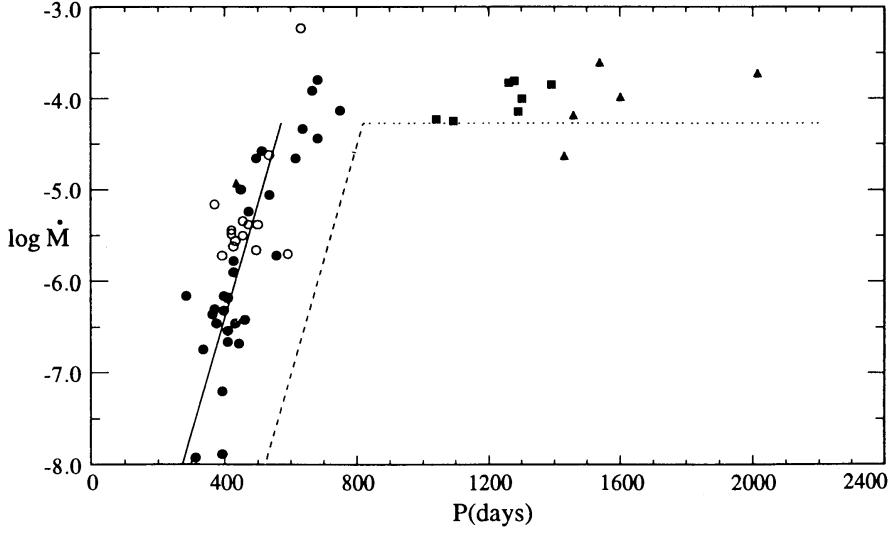


Figure 13.8: Mass loss of AGB stars. The observed correlation between the pulsation period P of Mira variables and their mass-loss rate \dot{M} , in $M_{\odot} \text{ yr}^{-1}$. (Figure reproduced from Vassiliadis & Wood 1993, ApJ, 413, 641).

But:

$$\frac{dM_{\text{wind}}}{dt} = f(L) \quad (13.5)$$

(see eq. 13.2), and

$$L = f(M_{\text{core}}) \quad (13.6)$$

In fact, 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, $t_{\text{TP-AGB}} \sim 1-2 \times 10^6 \text{ yr}$, is essentially determined by the mass-loss rate.

The high mass-loss rate distributes the chemical elements and dust grains found in the outer atmospheres of AGB stars into the surrounding interstellar medium. Many AGB stars (known as OH/IR stars) are completely enshrouded in a dusty circumstellar envelope which renders them invisible at optical wavelengths. The mechanisms driving such strong mass loss are not fully understood, but it is likely that both dynamical pulsations and radiation pressure on dust particles play a role.

AGB stars undergoing strong radial pulsations are known as ‘Mira variables’. Observationally, a correlation is found between the pulsation period and the mass-loss rate, shown in Figure 13.8. As a star evolves towards larger radii along the AGB, the pulsation period increases and so does the mass-loss rate, from $\dot{M} \sim 10^{-8}$ to $\sim 10^{-4} M_{\odot} \text{ yr}^{-1}$ for pulsation periods in excess of about 600 days. 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 marking 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.

The mass loss rate increases until the mass of the remaining envelope has reached some minimum value, $10^{-2}\text{--}10^{-3}M_{\odot}$, such that a convective envelope can no longer be sustained and the envelope starts to contract into radiative equilibrium 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_{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 $\sim 10^4\text{yr}$.

13.6 Planetary Nebulae

As the star gets hotter and T_{eff} exceeds 30 000 K, two effects come into play: (1) the star develops a weak but fast wind ($\dot{M} \simeq 10^{-6}M_{\odot}\text{ yr}^{-1}$, $v_{\text{exp}} \simeq 1000\text{ km s}^{-1}$), driven by radiation pressure in UV absorption lines (similar to the winds of massive OB-type stars which we shall discuss in a subsequent lecture); 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 H II region) and starts radiating in recombination lines: a young *Planetary Nebula* (PN) is born. (PNs have nothing to do with planets, of course. The name has its origin in the fact that, like planets, they are not point-like sources, and therefore did not appear to ‘twinkle’ due to atmospheric turbulence when observed with the naked eye by early astronomers. The misnomer has stuck).

For a long time it was believed that the PN represented the previous AGB

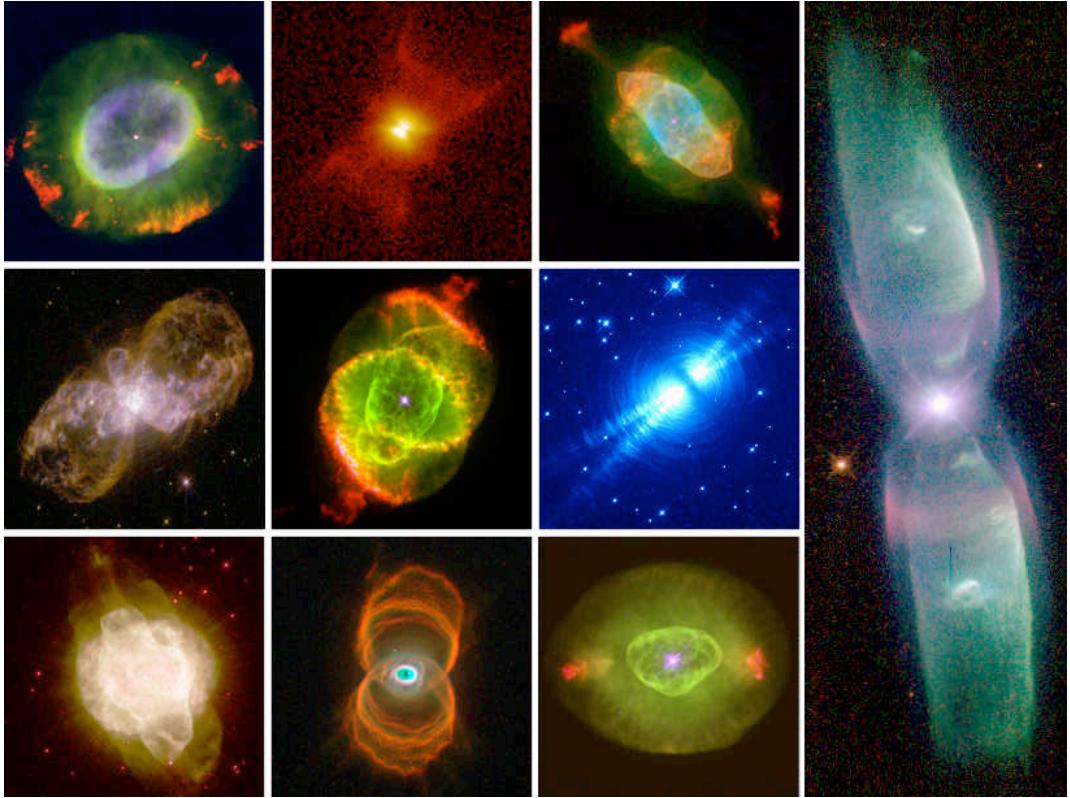


Figure 13.9: Images of Planetary Nebulae taken with the *Hubble Space Telescope*.

wind. However, a puzzle remained: the observed expansion speeds of PNs are typically $v_{\text{exp}} \simeq 50 \text{ km s}^{-1}$, whereas AGB winds are ejected with much lower velocities, only $10\text{--}15 \text{ km s}^{-1}$. How could the AGB material be accelerated? However, once it was realised (in 1975) that the central stars of PNs have fast stellar winds, then the current scenario was proposed (by Sun Kwok and collaborators): planetary nebulae result from the interaction between the slow AGB wind and the fast wind from the central star. The fast wind sweeps up and accelerates the AGB wind, forming a compressed optically thin shell from which the radiation is emitted.

Images of Planetary Nebulae taken with the *Hubble Space Telescope* reveal a wide variety of often very complex shapes (see Figure 13.9). Bruce Balick (1987, AJ, 94, 671) put forward an empirical morphological classification of PNs, ranging from spherical, through elliptical, to ‘butterfly’, and proposed that these different morphologies result from spherical asymmetries in the wind of the AGB progenitor. According to his suggestion, a relatively moderate density contrast in the slow AGB wind between the equatorial and polar directions produces an elliptical shell, while a strong contrast results in a ‘butterfly’ shape. Such a density contrast may be related to rotation. Soker & Livio (1989, ApJ, 339, 268) extended this idea to PNs

containing binary nuclei. Here common envelope evolution of the binary system can naturally produce the density contrast between the equatorial (orbital) plane and polar directions upon which the interacting wind model can operate. Some of the more complex PN morphologies may be due to the influence of magnetic fields.

13.6.1 The spectra of planetary nebulae

The diffuse gases which constitute a planetary nebula are heated and ionised by the hard UV radiation emitted by the hot central star with $T_{\text{eff}} = 30\,000\text{--}100\,000\,\text{K}$. Recalling the discussion in Lecture 3 (Figure 3.1), the nebula will emit a spectrum which is dominated by emission lines, with very little or no continuum; an example is reproduced in Figure 13.10.

The strongest emission lines are H and He recombination lines, emitted following capture of free electrons by ionized H and He atoms, and collisional lines excited by inelastic electron impacts with heavy elements ions. The latter are usually indicated with square brackets because they are ‘forbidden’ transitions, by which we mean that they cannot proceed via the most efficient (electric dipole) route, and thus have much lower transition

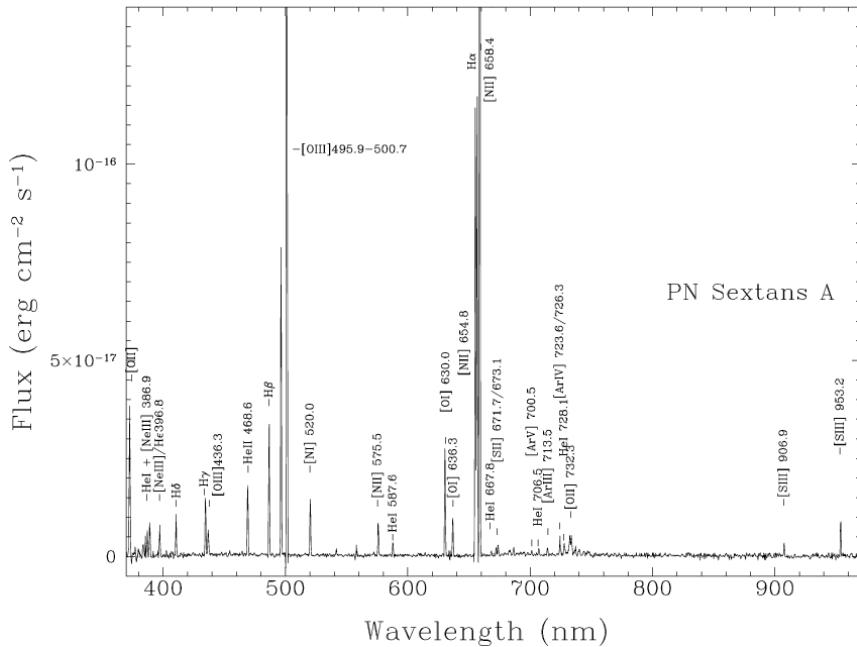


Figure 13.10: Emission line spectrum of a Planetary Nebula recorded with the Very Large Telescope of the European Southern Observatory at Cerro Paranal, Chile. (Figure reproduced from Magrini et al. 2005, A&A, 443, 115).

probabilities than transitions allowed by the selection rules of quantum mechanics.

These emission lines are a veritable treasure-trove of information on the physical conditions within the nebula; from their relative strengths it is possible to deduce values of the electron temperature and density, as well as the relative abundances of different elements. Planetary nebulae have played a key role in the development of nebular diagnostic techniques. They are one of the best astrophysical laboratories for studying the physical processes operating within ionized nebulae in general, and for testing the atomic data that are central to stellar and nebular models.

Because their spectra are dominated by strong and narrow emission lines, PNs can be recognised in external galaxies at distances where individual stars are too faint to be detected. An extra bonus is that their radial velocities can be measured very precisely. These characteristics have been exploited in the recently developed ‘Planetary Nebula Spectrograph’, an instrument especially designed with the goal in mind of using PNs as probes of galaxy kinematics. Specifically, by measuring the projected velocities of large numbers of PNs within a single elliptical galaxy, it is possible to reconstruct its gravitational potential, and thereby probe the distribution of both luminous and dark matter. Traditionally, this has only been possible in gas-rich spiral galaxies via the 21 cm emission line of neutral hydrogen. Elliptical galaxies, on the other hand, have little gas left, so that alternative strategies had to be devised.

The distribution of [O III] $\lambda 5007$ emission line strengths in PNs appears to have a sharp bright-end cut-off. This has led to attempts to use PNs as ‘standard candles’ in determining cosmological distances.

<http://web.williams.edu/astronomy/research/PN/nebulae/> is a great site for further images, spectra and general information on PNs.

13.7 White Dwarfs

When the stellar envelope mass has decreased to $10^{-5} M_{\odot}$, the H-burning shell is finally extinguished and from this point on the luminosity of the star decreases as it cools from $T_{\text{eff}} \simeq 10^5$ K. All stars with initial masses

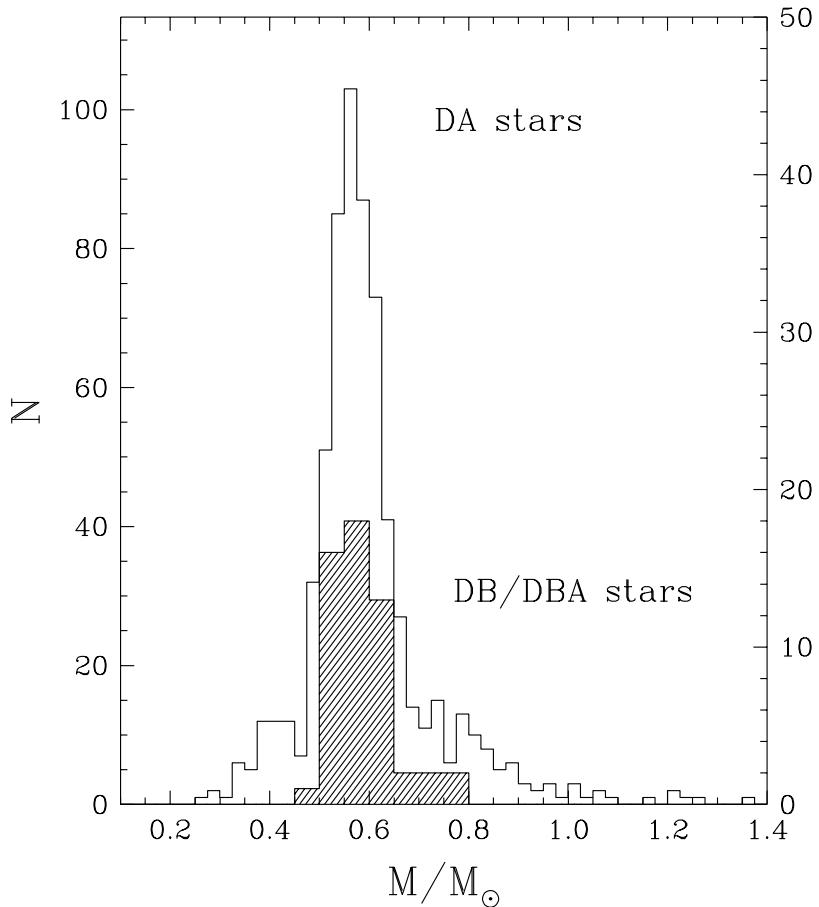


Figure 13.11: Observed mass distribution of white dwarfs, for a large sample of DA white dwarfs and a smaller sample of DB white dwarfs. In both samples, there is a sharp peak between $M = 0.55M_{\odot}$ and $0.6M_{\odot}$, as expected if most white dwarfs come from low-mass progenitors with masses $M_{\text{ZAMS}} \leq 2M_{\odot}$. (Figure reproduced from Bergeron et al. 2007, ASPC, 372, 29).

$M_{\text{ZAMS}} \lesssim 8M_{\odot}$, develop electron-degenerate cores, lose their envelopes during the AGB phase, and end their lives as white dwarfs (WD). Nuclear fusion no longer provides energy: white dwarfs shine by radiating the thermal energy stored in their interiors, cooling at almost constant radius and decreasing luminosity. The faintest white dwarfs detected have $L \approx 10^{-4.5}L_{\odot}$. Observed WD masses are mostly in a narrow range around $\sim 0.6M_{\odot}$ (see Figure 13.11), which corresponds to the CO core mass of low-mass ($M_{\text{ZAMS}} \lesssim 2M_{\odot}$) AGB progenitors. This sharply peaked mass distribution is further evidence that AGB mass-loss is very efficient at removing the stellar envelope.

We shall return to White Dwarfs in a subsequent lecture.

13.8 Summary

Table 13.2 summarises the timescales relevant to the different evolutionary stages of a $1M_{\odot}$ star.

Table 13.2 Approximate Timescales in the Evol. of a $1M_{\odot}$ Star

Phase	t (yr)
Main Sequence	1×10^{10}
Subgiant	2×10^9
Red Giant Branch	5×10^8
Horizontal Branch	1×10^8
Asymptotic Giant Branch	5×10^6
Planetary Nebula	1×10^5
White Dwarf (cooling) ^a	$> 8 \times 10^9$

Notes:

^a Age of the Galactic disk.

STELLAR REMNANTS. I: WHITE DWARFS

14.1 Introduction

We do not have to look very far to find a white dwarf. Sirius (α CMa) is the brightest star in the sky, and the fifth closest to the Sun, at a distance of only 2.6 pc. In 1862 the telescope maker Alvan Clark, while testing an 18-inch refractor lens, discovered that Sirius has a very faint companion: Sirius B. Sirius B always appears very close to Sirius A but is 10 magnitudes (10,000 times!) fainter. This makes it a very difficult object to recognise visually (see Figure 14.1). Sirius B has the same mass as our Sun but only one millionth its volume. That's because Sirius B is a White Dwarf.

Its existence had been predicted nearly twenty years earlier by the German astronomer Bessel (of Bessel functions fame), who correctly deduced the orbital period of the binary star system to be ~ 50 yr from observations of the motion of Sirius A relative to fixed stars over a period of ten years (Figure 14.2).

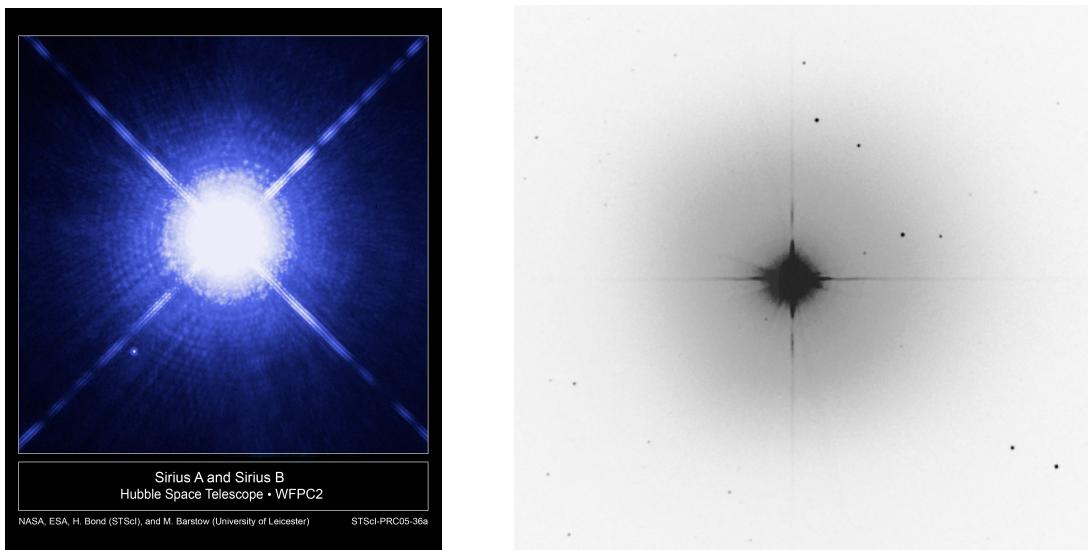


Figure 14.1: Sirius A and B, as seen by the 2.5 m *Hubble Space Telescope* (left) and a ground-based 18-in Celestron (right).

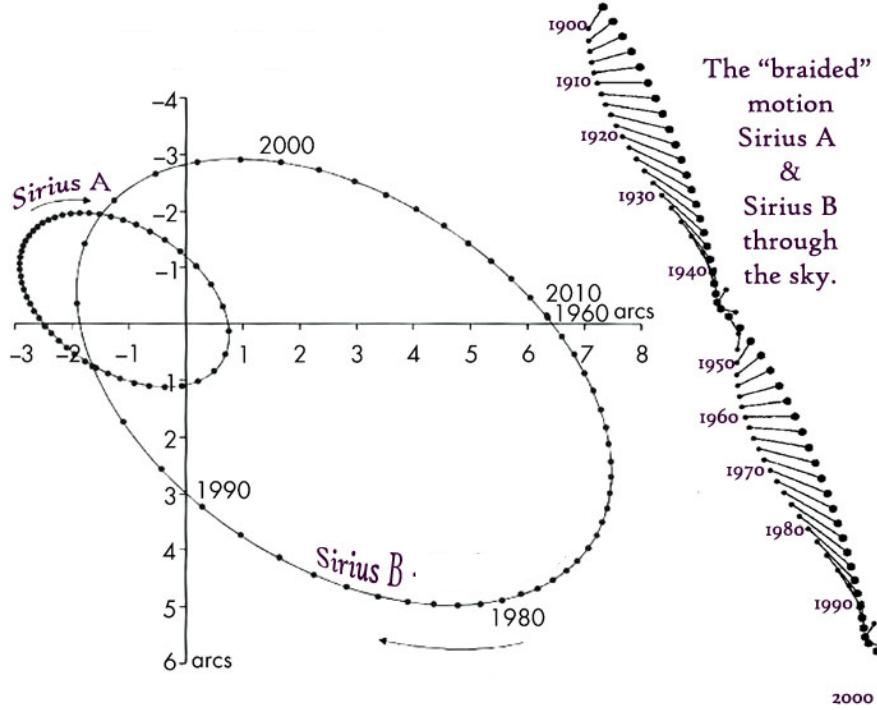


Figure 14.2: The orbits of Sirius A and B about the common centre of mass of the binary system, and their projection on the sky.

The physical parameters of this binary system are collected in Table 14.1. Sirius B has the mass of the Sun confined within a volume smaller than the Earth! The acceleration due to gravity at the surface of Sirius B is $\sim 4 \times 10^5$ times greater than on Earth (where $\log g \simeq 3$). Such strong gravity is reflected in highly pressure-broadened (Lecture 6.5.2) H I Balmer lines superposed on otherwise featureless blue continuum (see Figure 14.4). White dwarf ‘material’ is so dense that a teaspoon would weight 16 tons on Earth, and ~ 6.4 million tons on the surface of the white dwarf. These are extreme physical conditions indeed!

Table 14.1 Physical Parameters of the Sirius A B Binary System.

Property	Sirius A	Sirius B
Spectral type	A1V	DA2
M_V (mag)	1.4	11.2
Mass (M_\odot)	2.0	0.98
Radius (R_\odot)	1.7	0.0084
Surface gravity ($\log g$)	4.3	8.57
Luminosity (L_\odot)	25	0.026
Temperature (K)	9940	25 200

As we have already discussed in Lecture 13.7, white dwarfs are the stellar remnants of low and intermediate mass stars. Their name is related to the fact that they are hot and compact, occupying a narrow sliver in the H-R diagram that is roughly parallel to and well below the Main Sequence (Figure 14.3). The name itself is, however, something of a misnomer since white dwarfs come in all colours, with surface temperatures ranging from $T_{\text{eff}} < 5\,000\,\text{K}$ to $T_{\text{eff}} > 80\,000\,\text{K}$.

The core of a white dwarf is mostly He, or C and O. Stars with $M \lesssim 0.5M_{\odot}$ do not have sufficient gravitational energy to heat up their core to the temperature required to ignite He fusion, and they will end up as He white dwarfs. However, since the lifetime of such stars is greater than the current age of the Universe, such He white dwarfs should not exist yet! Another channel for the creation of a He-core WD may be through binary evolution, whereby the outer layers of a star in the process of becoming a red giant may be stripped by the gravitational pull of its binary companion. If the mass transfer happens before He ignition, further evolution of the star is halted, leaving a white dwarf made up mostly of He.

Stars with masses $M \simeq 1\text{--}8M_{\odot}$ evolve through the stages outlined in Lecture 13, leaving a white dwarf with a CO core of mass $M \simeq 0.6M_{\odot}$ (Fig-

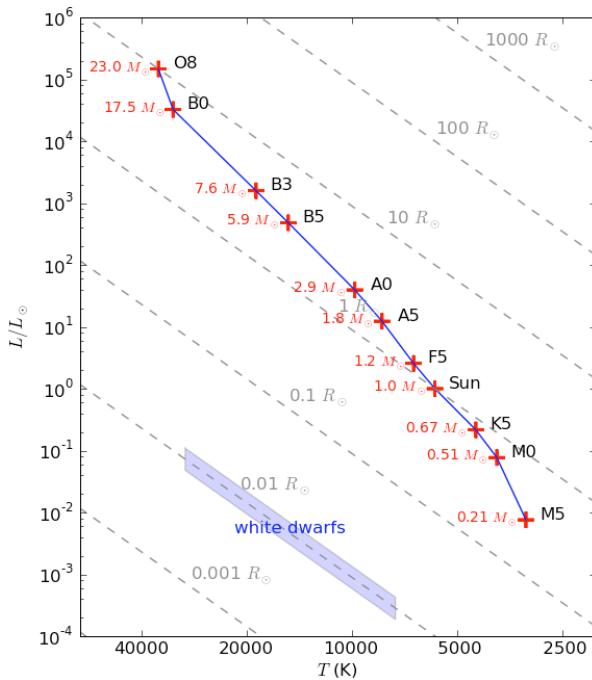


Figure 14.3: Location of white dwarf stars in the H-R diagram.

ure 13.11) following the ejection of the star's outer layers in the planetary nebula stage. In either case, without an internal source of energy, white dwarfs simply cool off at an essentially constant radius as they slowly deplete their supply of thermal energy.

14.1.1 Spectral Classification of White Dwarfs

Although most white dwarfs are thought to be composed of carbon and oxygen, spectroscopy typically shows that their emitted light comes from an atmosphere which is observed to be either H-dominated or He-dominated. (see Figure 14.4). The dominant element is usually at least 1000 times more abundant than all other elements. The cause for this 'purity' is thought to be the high surface gravity which leads to a stratification of the atmosphere, with heavy elements on the bottom and lighter ones on top. This vertical stratification takes only ~ 100 years to become established.

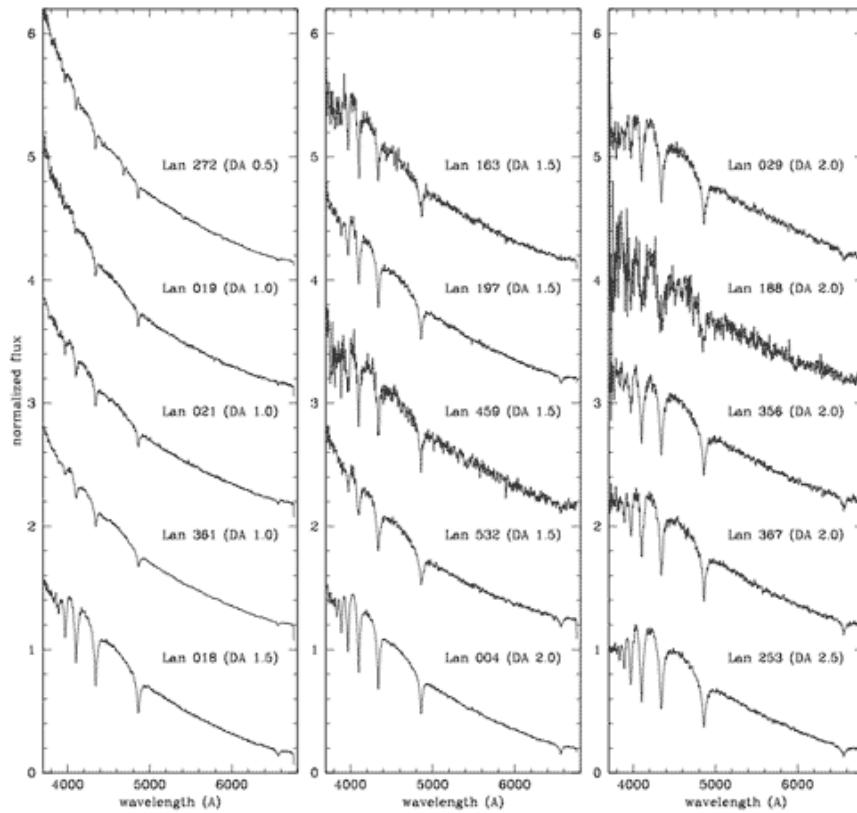


Figure 14.4: Optical spectra of white dwarfs of spectral type DA. Note the blue continuum, indicative of high effective temperatures ($T_{\text{eff}} > 10\,000$ K), and the pressure-broadened (Lecture 6.5.2) H I Balmer absorption lines.

The atmosphere, which is the only part of the white dwarf visible to us, is thought to be the top of an envelope which is a residue of the star's envelope in the AGB phase and may also contain material accreted from the interstellar medium. The envelope is believed to consist of a He-rich layer with mass no more than 1/100 of the star's total mass which, if the atmosphere is H-dominated, is overlaid by a H-rich layer with mass approximately 10^{-4} of the star's total mass.

The current classification scheme uses the letter D, followed by another letter to describe the primary spectral feature. The main classes are DA (H-dominated spectrum) and DB (He-dominated spectrum), with DAs making up $\sim 80\%$ of all observed white dwarfs and DB $\sim 16\%$. The remainder includes DC white dwarfs (continuum spectrum, no lines), DQ (carbon lines present) and DZ (metal lines present).

14.2 Electron Degeneracy Pressure

Without an internal supply of energy, the pressure to support white dwarfs against the pull of gravity is provided by electron degeneracy. We have already introduced the general principle of electron degeneracy in Lecture 12.6; we will now calculate in detail the pressure that the electrons can provide, P_e , as a function of their density ρ_e . We shall use some of the concepts that you have already encountered in the *Principles of Quantum Mechanics* and in the *Statistical Mechanics* Part II Astrophysics courses.

14.2.1 Density of States

To derive $P_e = f(\rho_e)$, we start by considering the density of states for free electrons: how many free-electron states fit into a box of volume $V = L^3$. We should think of this box as filling all space, via periodic replication. This means that the wave vectors of the free-electron quantum states can only take certain values. If the electron wave function is $\psi \propto \exp(i\mathbf{k} \cdot \mathbf{x})$, where $\mathbf{k} = (k_x, k_y, k_z)$, then the requirement of periodicity implies:

$$k_x = n_x \frac{2\pi}{L} \quad \text{where } n_x = 1, 2, \dots \quad (14.1)$$

The allowed states therefore lie on a lattice with spacing $2\pi/L$, and the density of states in k space is

$$dN = g \frac{L^3}{(2\pi)^3} d^3k \quad (d^3k \equiv dk_x dk_y dk_z), \quad (14.2)$$

where g is a degeneracy factor for spin.

The de Broglie relation between the electron's momentum to its wave vector: $\mathbf{p} = \hbar\mathbf{k}$ allows us to convert the density of states in k space to that in momentum space:

$$dN = g \frac{L^3}{(2\pi\hbar)^3} d^3p. \quad (14.3)$$

The number density of particles (per unit volume) with momentum states in the range of d^3p is then:

$$dn = g \frac{1}{(2\pi\hbar)^3} f(p) d^3p \quad (14.4)$$

where $f(p)$ is the occupation number of the mode, i.e. the number of particles in the box with that particular wave function. For bosons (e.g. photons), $f(p)$ is unrestricted. But fermions (such as electrons with spin angular momentum $\hbar/2$) obey Pauli exclusion principle, which states that:

$$f(p) \leq 1.$$

This criterion immediately imposes a restriction on how dense an electron gas can be before it has to be treated in a manner very different from the classical one. Normally, the distribution of momenta of the particles would be treated as a Maxwellian distribution, with each component of velocity having a Gaussian distribution with standard deviation:

$$\Psi(v) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp[-v^2/2\sigma^2] d^3v, \quad (14.5)$$

where v is the particle velocity. The dispersion in velocities, σ , is related to the temperature by equipartition of energy: $\sigma^2 = kT/m$. We can convert this to a number density of particles in a given range of momentum space, by multiplying by the total number density of particles, n , and using $p = mv$:

$$dn = \frac{n}{(2\pi m k T)^{3/2}} \exp[-p^2/2m k T] d^3p. \quad (14.6)$$

Comparing the Maxwellian distribution of particle momenta in eq. 14.6 with the general expression for the number density of particles in momentum space given by eq. 14.4, we deduce that there is a *critical density* at which the classical law would yield $f(p) > 1$ (at $p = 0$ which is where eq. 14.6 peaks):

$$n_{\text{crit}} = \frac{g}{(2\pi)^{3/2}} \frac{(mkT)^{3/2}}{\hbar^3} \quad (14.7)$$

At a fixed density, the gas will be in the classical regime at high values of T , but quantum effects become important as $T \rightarrow 0$.

Integrating eq. 14.4 over all momentum states gives the total number density of particles:

$$n = g \frac{1}{(2\pi\hbar)^3} \int f(p) d^3p. \quad (14.8)$$

In the limit of zero temperature, states are occupied only up to the *Fermi momentum*, p_F . Hence:

$$n = g \frac{1}{(2\pi\hbar)^3} \int_0^{p_F} d^3p = g \frac{1}{(2\pi\hbar)^3} \frac{4\pi}{3} p_F^3. \quad (14.9)$$

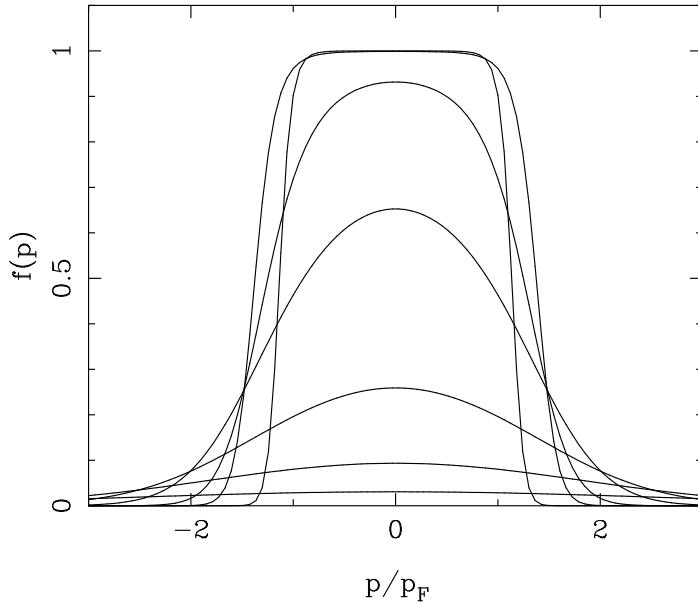


Figure 14.5: The occupation number for a gas of fermions as a function of their density relative to the critical density, ranging from $n/n_{\text{crit}} = 0.03$ to $n/n_{\text{crit}} = 30$. For the lowest densities (or highest temperatures), we have almost exactly the classical Maxwellian velocity distribution. For densities well above critical, the occupation number tends to a ‘top-hat’ distribution: unity for momenta less than the Fermi momentum, and zero otherwise.

Thus, the Fermi momentum is related to the particle density by:

$$p_F = 2\pi\hbar \left(\frac{3}{4\pi g} \right)^{1/3} n^{1/3} \quad (14.10)$$

As the density of the gas goes up, the Fermi momentum increases: the additional particles have to fill higher momentum states because the lower momentum states are fully occupied (see Figure 14.5).

14.2.2 Degeneracy Pressure

We can use the above treatment to work out an expression for the pressure exerted by degenerate electrons. Recall that pressure can be thought of as a flux of momentum. If the number density of electrons is n_e , then the flux of electrons in the x -direction is just the number of electrons crossing a unit area per unit time, or $n_e v_x$. The pressure is then approximately:

$$P_e \simeq p_x n_e v_x . \quad (14.11)$$

The contribution to the total pressure in the x -direction from all electrons with momentum p_x is then just:

$$dP_x = p_x v_x dn_{e,x} \quad (14.12)$$

where $dn_{e,x}$ is the number density of electrons with x -momentum in the range p_x to $p_x + dp_x$. Using eq. 14.4, we obtain for the total pressure in the x -direction (which by isotropy must be equal to the pressure in any direction):

$$P = P_x = g \frac{1}{(2\pi\hbar)^3} \int p_x v_x f(p) d^3p . \quad (14.13)$$

Using spherical polar coordinates in momentum space:

$$\int p_x v_x dp_x dp_y dp_z = \frac{1}{3} \int (p_x v_x + p_y v_y + p_z v_z) dp_x dp_y dp_z = \frac{1}{3} \int p \cdot v 4\pi p^2 dp , \quad (14.14)$$

so that:

$$P = \frac{g}{3} \frac{1}{(2\pi\hbar)^3} \int_0^\infty p \cdot v f(p) 4\pi p^2 dp . \quad (14.15)$$

In the zero-temperature limit (eq. 14.9), if the electron speeds are **non-relativistic** so that $p \cdot v = p^2/m_e$, we have:

$$P_e = \frac{4\pi g}{3(2\pi\hbar)^3} \int_0^{p_F} \left(\frac{p^2}{m_e} \right) p^2 dp = \frac{g}{30\pi^2\hbar^3 m_e} p_F^5. \quad (14.16)$$

What we want is the pressure as a function of density $\rho_e = m_e n_e$. Rewriting eq: 14.10 in terms of ρ_e :

$$p_F = \left(\frac{6\pi^2\hbar^3 \rho_e}{gm_e} \right)^{1/3} \quad (14.17)$$

and substituting into (14.16), we finally arrive at the result stated in Lecture 12.6.1:

$$P_e = \frac{g}{30\pi^2\hbar^3} \left(\frac{6\pi^2\hbar^3}{g} \right)^{5/3} \rho_e^{5/3} m_e^{-8/3} \equiv K_1 \rho_e^{5/3} \quad (14.18)$$

with

$$K_1 = \frac{\pi^2\hbar^2}{5m_e^{8/3}} \left(\frac{6}{g\pi} \right)^{2/3}.$$

Equation 14.10 implies that, at sufficiently high densities, the Fermi momentum can reach relativistic values, with some particles forced into momentum states with velocities approaching the speed of light. This leads to the **relativistic** expression for $P_e = f(\rho_e)$. With $v = c$ and $p \cdot v = pc$, we have:

$$P_e = \frac{4\pi g}{3(2\pi\hbar)^3} \int_0^{p_F} pc p^2 dp = \frac{gc}{24\pi^2\hbar^3} p_F^4 \quad (14.19)$$

or

$$P_e = \frac{gc}{24\pi^2\hbar^3} \left(\frac{6\pi^2\hbar^3 \rho_e}{gm_e} \right)^{4/3} \equiv K_2 \rho_e^{4/3} \quad (14.20)$$

with

$$K_2 = \frac{\pi\hbar c}{4m_e^{4/3}} \left(\frac{6}{g\pi} \right)^{1/3}.$$

What eqs. 14.18 and 14.20 show us is a result that has been stated before: in degenerate gas, the pressure depends only on density and is independent of temperature. Of course, in partially degenerate gas there will be a residual temperature dependence. Note also that the mass of the particle appears on the denominator in these equations. Thus, while electrons may

be degenerate, protons—with a mass $m_p/m_e = 1836$ times larger—exert negligible degeneracy pressure, justifying the fact that we have neglected it in the above treatment.

14.3 The Chandrasekhar Limit to the Mass of a White Dwarf

In July 1930, the bright, not-yet-twenty-year-old, Indian student Subrahmanyam Chandrasekhar who had obtained his B.Sc. degree in Physics the previous month, was awarded a Government of India scholarship to pursue graduate studies at the University of Cambridge, where he was admitted to Trinity College and became a research student of R. H. Fowler. On his journey from India to Cambridge, Chandrasekhar worked out that there is a maximum mass for a white dwarf, now generally referred to as the Chandrasekhar limit and estimated to be $1.44M_\odot$. To see how this limit arises, consider the following.

We can estimate the energy density of degenerate gas as we did the pressure: by integrating over momentum space and including a term, $\epsilon(p)$, to denote the energy per mode:

$$U = g \frac{1}{(2\pi\hbar)^3} \int_0^\infty \epsilon(p) f(p) 4\pi p^2 dp. \quad (14.21)$$

In the zero-temperature limit, $f(p) = 1$ up to the Fermi momentum and $f(p) = 0$ at all other values. In the relativistic case, with $\epsilon(p) = pc$, integration of (14.21) gives:

$$U_e = g \frac{1}{(2\pi\hbar)^3} \frac{1}{4} 4\pi c p_F^4. \quad (14.22)$$

Using (14.10) to express U_e in terms of the number density of electrons, we have:

$$U_e = \frac{3}{4} \left(\frac{6\pi^2}{g} \right)^{1/3} \hbar c n_e^{4/3}. \quad (14.23)$$

In the non-relativistic case, $\epsilon(p) = p^2/2m_e$, which results in:

$$U_e = \frac{3\hbar^2}{10m_e} \left(\frac{6\pi^2}{g} \right)^{2/3} n_e^{5/3}. \quad (14.24)$$

The total kinetic energy supplied by the degenerate electrons is proportional to their energy density times the volume. In the relativistic case, $E_K \propto U_e V \propto n_e^{4/3} V \propto M^{4/3}/R$, where M is the mass and R is the radius. On the other hand, the gravitational energy is proportional to $-M^2/R$. The total energy is the sum of the two terms:

$$E_{\text{tot}} = \frac{(AM^{4/3} - BM^2)}{R} \quad (14.25)$$

where A and B are constants.

Now we see that there is a critical mass for which the two terms in the bracket are equal. If the mass is smaller than this limit, then the total energy is positive and will be reduced by making the star expand until the electrons reach the mildly relativistic regime and the star can exist as a stable white dwarf (see next section). But, if the mass exceeds the critical value, the binding energy increases without limit as the star shrinks: gravitational collapse has become unstoppable. This is believed to be the mechanism responsible for producing Type Ia supernovae which will be discussed in a later lecture.

To find the exact limiting mass, we need to find the coefficients A and B above, where the argument has implicitly assumed the star to be of constant density. In this approximation, the kinetic and potential energies are:

$$E_K = \left(\frac{243\pi}{128g} \right)^{1/3} \frac{\hbar c}{R} \left(\frac{M}{\mu m_p} \right)^{4/3} \quad (14.26)$$

where the mass per electron is μm_p , and

$$E_V = -\frac{3}{5} \frac{GM^2}{R} \quad (14.27)$$

Equating the two terms, we have:

$$M_{\text{crit}} = \frac{3.7}{\mu^2} \left(\frac{2}{g} \right)^{1/2} \left(\frac{\hbar c}{G} \right)^{3/2} m_p^{-2} \simeq \frac{7}{\mu^2} M_\odot \quad (14.28)$$

for $g = 2$. In a star which has burnt most of its initial fuel into elements heavier than Helium $\mu \simeq 2$, giving $M_{\text{crit}} = 1.75 M_\odot$. More precise calculations, which take into account the density profile within the white dwarf, give $M_{\text{crit}} = 1.44 M_\odot$.

14.4 The Mass-Volume Relation of White Dwarfs

As explained in the introduction, white dwarfs are extremely compact objects. To find their radius, we consider the non-relativistic case in which the energy density is $U_e \propto n_e^{5/3}$ (eq. 14.24), so that the total kinetic energy is $E_K = CM^{5/3}/R^2$, where C is a constant. As before, the gravitational potential energy is $E_V = -BM^2/R$, and the total energy is the sum of the two terms. The equilibrium radius is found by imposing the condition $dE_{\text{tot}}/dr = 0$, which gives:

$$R = \frac{2C}{B} M^{-1/3} \quad (14.29)$$

This is a remarkable relation. Since $V \propto R^3$, we have:

$$\boxed{M_{\text{wd}} V_{\text{wd}} = \text{constant}} \quad (14.30)$$

and more massive white dwarfs are actually smaller. This surprising result is a consequence of the star deriving its support from electron degeneracy pressure. The electrons must be more closely confined to generate the larger degeneracy pressure required to support a more massive star. Thus, the mass-volume relation implies that $\rho \propto M_{\text{wd}}^2$.

Note also that eq. 14.30 seemingly implies that the volume of a white dwarf can get infinitely small for an arbitrarily large mass. This is another way of recognising the existence of a critical mass; the former statement is incorrect because it ignores relativistic effects.

We can put numerical values into eq. 14.29 considering the following:

$$n_e = \frac{M}{\mu m_p} \left(\frac{4\pi R^3}{3} \right)^{-1} \quad (14.31)$$

which, when combined with (14.24), gives:

$$C = \frac{3\hbar^2}{10m_e} \left(\frac{6\pi^2}{g} \right)^{2/3} \left(\frac{1}{\mu m_p} \right)^{5/3} \left(\frac{4\pi}{3} \right)^{-2/3}. \quad (14.32)$$

Combined with our previous $B = 3G/5$, we get:

$$R = \frac{3}{2} \left(\frac{6\pi^2}{g^2} \right)^{1/3} \frac{\hbar^2}{Gm_e(\mu m_p)^{5/3}} M^{-1/3}. \quad (14.33)$$

Expressed in terms of the Chandrasekhar mass, this is:

$$R = 3\sqrt{\pi/5} \frac{\hbar}{\mu m_p m_e} \left(\frac{2\hbar}{gcG} \right)^{1/2} \left(\frac{M}{M_{\text{crit}}} \right)^{-1/3} \simeq 5975 \left(\frac{M}{M_{\text{crit}}} \right)^{-1/3} \text{ km}, \quad (14.34)$$

where again we have taken $g = 2$ and $\mu \simeq 2$. In other words, a Chandrasekhar-mass white dwarf is about the size of the Earth. This is an extremely high density, which can be evaluated to be:

$$\rho = \frac{M}{4\pi R^3/3} \simeq 10^{6.6} \left(\frac{M}{M_{\text{crit}}} \right)^2 \text{ g cm}^{-3} \quad (14.35)$$

This are only approximate expressions because they are based on the constant density approximation. If we wanted to do better, we would need to solve for the internal structure of a white dwarf.

In summary, non-relativistic white dwarfs are stable over a range of masses, but as their mass increases, they shrink. By the uncertainty principle, this must force the electrons closer and closer to becoming relativistic, and the limit of this is a star in which the electrons are all relativistic, which has a unique mass. It appears that electron degeneracy pressure has no way of supporting more massive bodies.

14.5 Cooling and the Ages of White Dwarfs

With nuclear reactions over, a white dwarf will slowly cool over time, radiating away its stored heat.¹ Much effort has been directed at understanding the process and timescale of white dwarf cooling, as it potentially provides a chronology of the past history of star formation in our Galaxy.

Recalling the discussion in Lecture 8, in normal stars the mean free path for photons is much greater than that of electrons or atoms; consequently energy transport is mainly by radiative diffusion. This is not the case in white dwarfs, where the degenerate electrons can travel long distances before losing energy in a collision with a nucleus, since the vast majority of lower-energy electron states are already occupied. Thus, in a white dwarf

¹This idea was developed in 1952 by Leon Mestel, who had recently graduated from Cambridge (he was at Trinity College) and was on his first postdoctoral appointment at the University of Leeds. He returned to Cambridge as a lecturer between 1955 and 1966.

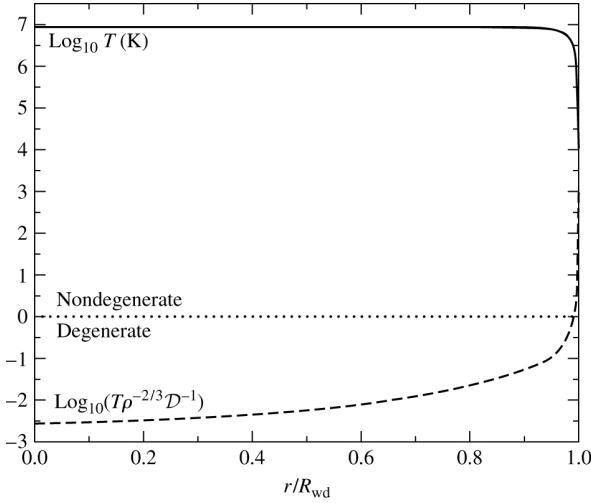


Figure 14.6: Temperature and degree of degeneracy in the interior of a model white dwarf. The condition for degeneracy is defined as $T/\rho^{2/3} < \mathcal{D}$, where $\mathcal{D} = (\hbar^2/3m_e k)(3\pi^2/\mu m_p)^{2/3}$. The smaller $T\rho^{-2/3}$, the more degenerate the gas.

energy is carried by electron conduction (similar to conduction in metals), rather than by radiation.

Electron conduction is so efficient that the interior of a white dwarf is nearly isothermal, with the temperature dropping significantly only in the non-degenerate surface layers (Figure 14.6). The thin ($\sim 1\%$ of the white dwarf radius), non-degenerate envelope transfers heat less efficiently, and acts as an insulating blanket, causing energy to leak out slowly. The steep temperature gradient near the surface creates convection zones that can alter the appearance of a white dwarf’s spectrum as it cools (Section 14.1.1).

We can calculate the initial temperature of a white dwarf by recalling that it forms from the contraction of a thermally unsupported stellar core, of mass M , down to the radius at which degeneracy pressure stops the contraction. The virial theorem tells us that, just before reaching the final point of equilibrium, the thermal energy will equal half of the potential energy:

$$E_{\text{th}} \sim \frac{1}{2} \frac{GM^2}{R}. \quad (14.36)$$

For a pure He composition, the number of nuclei in the core is $M/4m_p$, and the number of electrons is $M/2m_p$. The total thermal energy (which, once degeneracy sets in, will no longer play a role in supporting the star against gravity) is therefore:

$$E_{\text{th}} = \frac{3}{2} NkT = \frac{3}{2} \frac{M}{m_p} \left(\frac{1}{2} + \frac{1}{4} \right) kT = \frac{9}{8} \frac{M}{m_p} kT, \quad (14.37)$$

so that:

$$kT \sim \frac{4}{9} \frac{GMm_p}{R}. \quad (14.38)$$

Substituting the equilibrium radius of a white dwarf from eq. 14.33, we find:

$$kT \propto M^{4/3} \simeq 1.1 \times 10^{-7} \text{ erg} \sim 70 \text{ keV} \quad (14.39)$$

for a $0.5M_\odot$ white dwarf. With $k = 1.38 \times 10^{-16}$ erg K $^{-1}$, the corresponding temperature is $T \sim 8 \times 10^8$ K. Clearly, a just-formed degenerate core is a very hot object, with thermal emission that peaks at X-ray wavelengths. This is why, once the core becomes an exposed white dwarf, its radiation ionises the layers of gas that were blown off during the AGB phase, giving rise to a planetary nebula.

The energy radiated away from the surface of a white dwarf is the thermal energy stored in the still classical gas of *nuclei* within the star's volume. (The degeneracy of the electron gas limits almost completely the ability of the electrons to lose their kinetic energies). We can obtain an *upper limit* to the cooling rate by neglecting the envelope and assuming a uniform temperature throughout. The radiative energy loss is obtained by equating the luminosity of the blackbody, given by the Stefan-Boltzmann law, to the rate of change of thermal energy:

$$L = 4\pi R_{\text{WD}}^2 \sigma T^4 \sim \frac{dE_{\text{th}}}{dt} = \frac{3}{8} \frac{Mk}{m_p} \frac{dT}{dt} \quad (14.40)$$

where we have included only the contribution of nuclei to the thermal energy from eq. 14.37.

Separating the variables T and t and integrating, the cooling time to a temperature T is:

$$\tau_{\text{cool}} \sim \frac{3}{8} \frac{Mk}{m_p} \frac{1}{4\pi R_{\text{WD}}^2 \sigma} \frac{1}{3} \frac{1}{T^3}. \quad (14.41)$$

Using as before eq. 14.33 for R_{WD} and entering the numerical values, we find:

$$\tau_{\text{cool}} = 3 \times 10^9 \text{ yr} \left(\frac{T}{10^3 \text{ K}} \right)^{-3} \quad (14.42)$$

for a $M = 0.5M_\odot$ white dwarf. Writing the temperature as a function of time:

$$\frac{T}{10^3 \text{ K}} \sim \left(\frac{t}{3 \times 10^9 \text{ yr}} \cdot \right)^{-1/3} \quad (14.43)$$

Thus, even with the unrealistically efficient cooling we have assumed, it would take a $0.5M_{\odot}$ white dwarf several Gyr to cool to 10^3 K. In reality, the insulation provided by the non-degenerate envelope results in an effective temperature that is significantly lower than the interior temperature (Figure 14.6), and lowers the cooling rate. Furthermore, as a white dwarf cools, it *crystallises* in a gradual process that starts at the centre and moves outwards. The regular crystal structure is maintained by the mutual electrostatic repulsion of the nuclei: it minimises their energy as they vibrate about their average position in the lattice. As the nuclei undergo this phase change, the latent heat that is released is added to the thermal balance, further slowing down the decline in temperature.

Detailed models that include these and other processes for various masses and chemical compositions of white dwarfs show that over a period of 10 Gyr, comparable to the age of the Universe, white dwarfs cannot cool below ~ 3000 – 4000 K. This explains why most white dwarfs are observed to have high temperatures, and hence their blue to white colours. The coolest white dwarfs known have effective temperatures $T_{\text{eff}} \simeq 3500$ K.

14.5.1 White Dwarfs as Fossil Records of Star Formation

The interest in an accurate calculation of white dwarf cooling stems in part from the fact that, with cooling times comparable to the age of the Universe, it may be possible to use these fossil stars to uncover the history of star formation in our Galaxy by modelling their luminosity function as in Figure 14.7. The model luminosity function shown in the Figure was calculated using the best available models of white dwarf cooling and reproduces well the sudden drop in the population of white dwarfs with luminosities $\log L_{\text{wd}}/L_{\odot} < -4.5$. This decline can best be explained if the first white dwarfs were formed and started cooling 9.0 ± 1.8 Gyr ago. Including the pre-white dwarf evolutionary time, one would arrive at an age of 9.3 ± 2.0 Gyr for the Galactic disk. The model shown in Figure 14.7 assumes that star formation proceeded at the same rate throughout this time. This is an oversimplification, and more realistic models of the past history of star formation in the Milky Way may result in better agreement between calculated and observed luminosity functions.

As a final comment, it is interesting to note that the paper which first

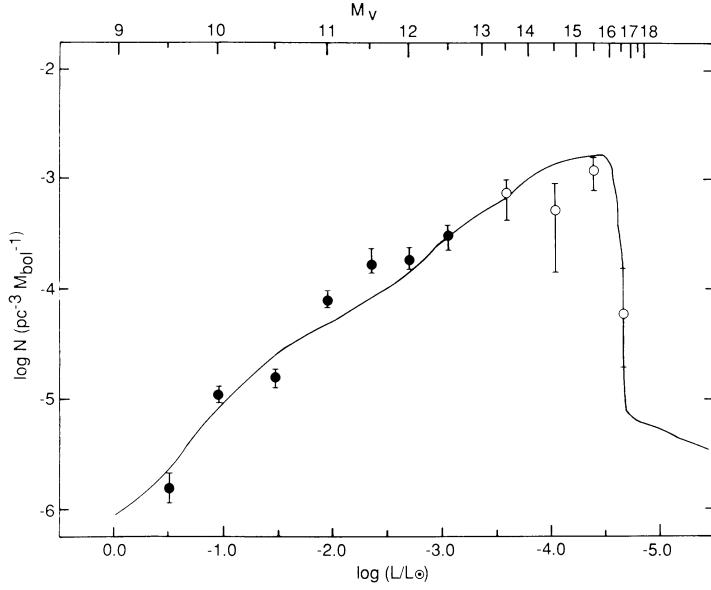


Figure 14.7: Observed (circles) and calculated (continuous line) luminosity function of Galactic disk white dwarfs, reproduced from Winget et al. (1987).

made use of white dwarfs as ‘cosmic clocks’, by Winget et al. (1987, ApJ, 315, L77), was entitled: “*An Independent Method for Determining the Age of the Universe*”. This reflects the fact that the paper predates the current era of ‘precision cosmology’, achieved primarily (but not exclusively) through precise measurement of the spectrum of temperature anisotropies of the cosmic microwave background, which encodes most fundamental cosmological parameters. Today we know the age of the Universe to be 13.69 ± 0.13 Gyr, a precision far superior than any that could possibly be achieved by modelling the luminosity function of white dwarfs. But, by turning the argument around, we can now use this well determined age, together with models of galaxy formation, to refine our understanding of white dwarf cooling.

POST-MAIN SEQUENCE EVOLUTION. II: MASSIVE STARS

15.1 Introduction

We saw in Lecture 13 that low- and intermediate-mass stars (with $M \leq 8M_{\odot}$) develop carbon-oxygen cores that become degenerate after central He burning. The electron degenerate pressure supports the core against further collapse; as a consequence, the maximum core temperature reached in these stars is lower than the $T \simeq 6 \times 10^8$ K required for carbon fusion (Lecture 7.4.4). During the latest stages of evolution on the AGB, these stars undergo strong mass loss which removes the remaining envelope, leaving as their remnants cooling C-O white dwarfs.

The evolution of *massive* stars differs in two important respects from that of low- and intermediate-mass stars:

(1) Massive stars undergo *non-degenerate carbon ignition* in their cores. The minimum CO *core* mass for this to happen at the end of core He burning has been estimated with detailed modelling to be $M_{\text{CO-core}} > 1.06M_{\odot}$. The corresponding ZAMS *stellar* mass is somewhat uncertain mainly due to uncertainties related to mixing (e.g. convective overshooting), but is considered to be $\sim 8M_{\odot}$. This is why $M = 8M_{\odot}$ is conventionally taken as the boundary between ‘intermediate’ and ‘high’ mass stars.

Stars with $M \gtrsim 11M_{\odot}$ achieve core temperatures high enough to ignite and burn elements heavier than carbon up to and including Fe which is near the peak of the binding energy per nucleon curve (Figure 7.2). With no more nuclear energy to be extracted by burning Fe, the Fe core collapses leading to a ‘*core-collapse supernova*’ (Lecture 16). Just before exploding as a supernova, a massive star has an ‘onion skin’ internal structure (Figure 15.1)

(2) Whereas low- and intermediate-mass stars experience mass loss during the late stages of their evolution, when they are on the RGB and AGB (Lecture 13), for stars with masses $M \gtrsim 15M_{\odot}$, mass loss by fast and en-

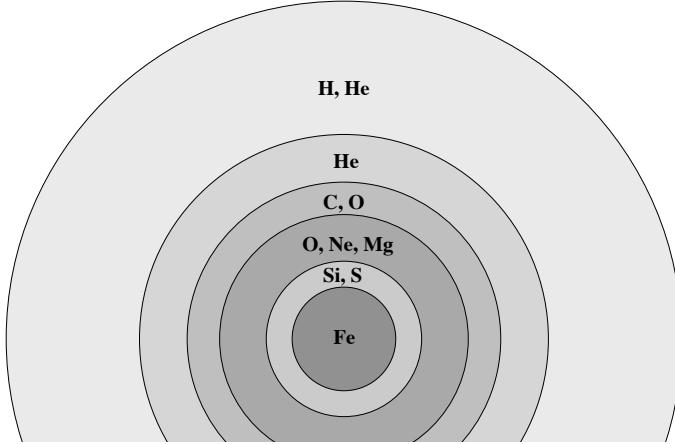


Figure 15.1: Schematic overview of the ‘onion skin’ structure of a massive star at the end of its evolution.

ergetic stellar winds is important during all evolutionary phases, including the main sequence. For $M \gtrsim 30M_{\odot}$, the mass-loss rates \dot{M} are so large that the timescale for mass loss, $t_{\text{ml}} = M/\dot{M}$ because smaller than the nuclear timescale t_{nuc} . Therefore mass loss has a very significant effect on the evolution of these stars, introducing substantial uncertainties in the calculations of massive star evolution.

Mass loss peels off the outer layers of a star, revealing at the stellar surface the products of nuclear burning brought up into higher layers by convective overshooting. Thus the photospheric abundances are changed drastically. The combination of mass reduction in the outer layers on the one hand, and larger core size—and hence increased luminosity—as a result of convective overshooting on the other, makes all massive stars overluminous for their masses. The main sequence lifetime is also increased because of the larger core size.

We look in more detail at the properties of stellar winds in Section 15.3.

15.2 Massive Stars in the H-R Diagram

The evolutionary journey through the H-R diagram of a massive star undergoing the different phases of core and shell burning that lead to the internal structure shown in Figure 15.1 is complicated. Figure 15.2 shows examples of evolutionary tracks for stars of solar metallicity, calculated with computer-generated models that include mass loss and convective

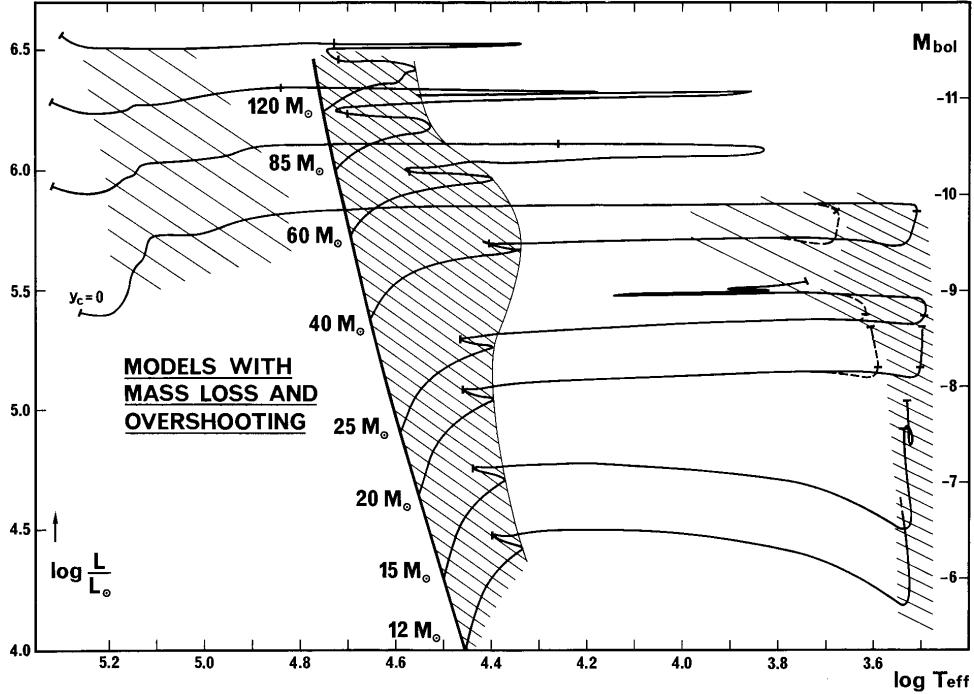


Figure 15.2: Evolutionary tracks of massive stars calculated with mass loss and a moderate amount of convective overshooting. The shaded regions correspond to long-lived evolutionary phases: (i) on the main sequence; (ii) during core He burning as a Red Supergiant (RSG) at $\log T_{\text{eff}} < 4.0$; and (iii) as a Wolf-Rayet (WR) star at $\log T_{\text{eff}} > 4.8$. Stars with $M_{\text{ZAMS}} > 40 M_{\odot}$ are assumed to lose their entire envelope during the Luminous Blue Variable (LBV) phase and never become RSGs. (Figure reproduced from Maeder & Meynet 1987, A&A 182, 243).

overshooting. More recent developments of the code used by the same authors also include the effects of stellar rotation and consider a range of metallicities.

As the core and shell energy sources vary in relative strength, the star makes a number of excursions to and fro across the HR diagram. In high-mass stars, these rightward (core exhaustion) and leftward (core ignition) excursions, between the red and blue (supergiant) branches respectively, occur with only a slight systematic increase in luminosity. Thus, the evolutionary tracks of high-mass stars are close to horizontal in the H-R diagram. In very high-mass stars, the nuclear evolution in the central regions of the star occurs so quickly that the outer layers have no time to respond to the successive rounds of core exhaustion and core ignition, and there is only a relatively steady drift to the right on the H-R diagram before the star arrives at the pre-supernova state shown in Figure 15.1.

The path followed by evolving massive stars in the upper part of the H-R

diagram gives rise to a rich nomenclature. On the main sequence, stars with $M_{\text{ZAMS}} \gtrsim 20M_{\odot}$ are of spectral type O. They evolve off the main sequence as blue and red supergiants (BSG and RSG). Of stars are very massive O supergiants whose spectra show pronounced emission lines. The most massive stars evolve into Luminous Blue Variables (LBVs) which have already encountered in Lecture 10.6.1.

Stars with $M \lesssim 40M_{\odot}$ 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 stellar wind, exposing the helium core of the star as a Wolf-Rayet (WR) star. WR stars are extreme objects which continue to attract a great deal of attention by stellar astronomers. We describe their main characteristics below.

15.2.1 Wolf-Rayet Stars

Spectroscopically, WR stars are spectacular in appearance: their optical and UV spectra are dominated by strong, broad emission lines instead of the narrow absorption lines that are typical of ‘normal’ stars (Figure 15.3). The emission lines are so strong that they were first noticed as early as 1867 by... Charles Wolf and Georges Rayet (!) using the 40 cm Foucault telescope at the Paris Observatory. Nowadays, this characteristic is exploited to identify WR stars in external galaxies with narrow-band imaging. After some debate, which lasted into the 1970s, WR are now recognised as the evolved descendent of O-type stars, whose H or He-burning cores have been exposed as a result of substantial mass loss.

From the earliest stages of the subject, it was clear that WR come in two flavours: those with strong emission lines of He and N (WN subtypes), and those with strong He, C and O lines (WC and WO subtypes), in which the products of, respectively, the CNO cycle and triple- α nuclear reactions are revealed. Each class is further subdivided in subclasses. WN2 to WN5 are early WN, or WNE, WN7 to WN9 are late WN or WNL, with WN6 being referred to as either early or late. The classification is based on the relative strengths of emission lines of N III to N V and He I to He II (see Figure 15.3). Similarly, WC and WO subtype classification is based on the ratios of highly ionised C and O emission lines.

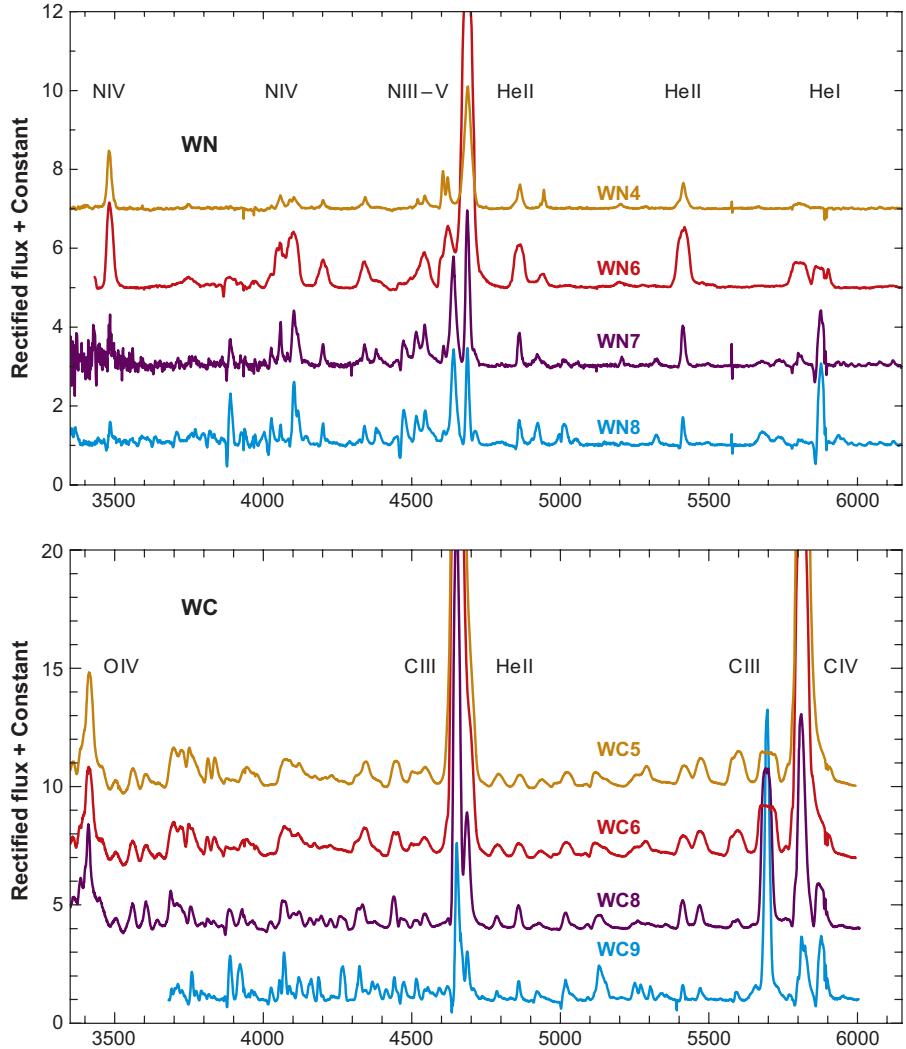


Figure 15.3: Examples of optical spectra of Galactic Wolf-Rayet stars of the WN and WC classes. The x -axis is wavelength is \AA . (Figure reproduced from Crowther 2007, ARAA, 45, 177).

When these spectral characteristics are interpreted in terms of the surface abundances, the following picture emerges:

WNL stars have H present on their surfaces (with $X < 0.4$) 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 lack H ($X = 0$);

WC stars have no H, 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).

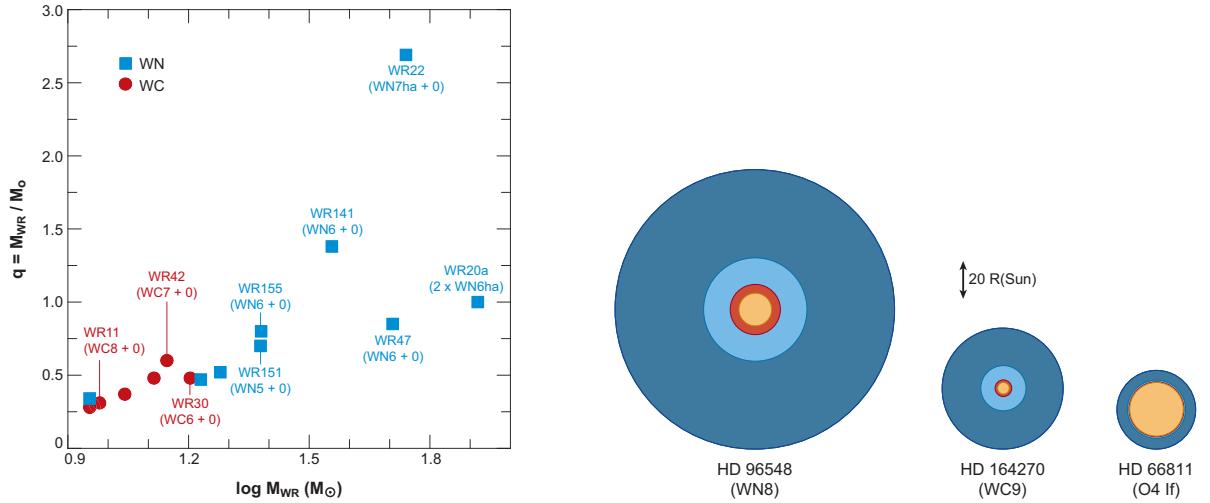


Figure 15.4: *Left:* Stellar masses of Galactic Wolf-Rayet stars (M_{WR}) deduced from the analysis of binary orbits. The y -axis is the ratio of the WR mass to that of its binary companion, usually an O-type star. *Right:* Stellar radii at Rosseland optical depths of 20 (orange) and $2/3$ (red) for an O4 If, a WC9, and a WN8 star, shown to scale. The primary optical emission line region is shown in dark blue, with light blue indicating higher density wind material. WR stars have much more extended winds than Of supergiants. (Figures reproduced from Crowther 2007, ARAA, 45, 177).

Approximately 40% of WR stars in the Milky Way occur in binary systems, allowing their masses to be measured (Lecture 4). As can be seen from Figure 15.4, WCs span a relatively narrow mass range, from 9 to $16M_{\odot}$, whereas WN masses are found between ~ 10 and $83M_{\odot}$, and in some cases exceed the mass of their OB companions, i.e. $M_{\text{WR}}/M_{\text{O}} > 1$.

The strong, broad emission lines characteristic of WR stars are due to their powerful stellar winds, with terminal velocities as high as $v_{\infty} \simeq 2000 \text{ km s}^{-1}$ and mass loss rates as high as $\dot{M} \simeq 10^{-5} M_{\odot} \text{ yr}^{-1}$. WR winds are sufficiently dense that an optical depth of unity in the continuum arises in the outflowing material (rather than in a stationary stellar photosphere, as in ‘normal’ stars). The emission lines are formed far out in the wind; both line- and continuum-emitting regions are much larger than the conventional stellar radius (Figure 15.4, right panel), and their physical depths are highly wavelength dependent.

Some young WR stars, mostly WNs, are surrounded by spectacular ring nebulae, thought to be the result of the interaction between material ejected in a slow wind by the WR precursor and the WR fast wind. The hard radiation from the central WR star photoionises the swept-up circumstellar gas, producing a wind-blown bubble. In OB associations containing

many massive stars, the combined effects of stellar winds and the supernova explosions that mark the ends of the lives of stars more massive than $8M_\odot$ can produce ‘*superbubbles*’.

15.2.2 The Conti Evolutionary Scenario

In 1976, Peter Conti proposed an evolutionary scenario that links the various types of massive stars which, until then, had been classified primarily on the basis of the appearance of their spectra.

$M \lesssim 15 M_\odot$	MS (OB) → RSG (\rightarrow BSG in blue loop? \rightarrow 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

These evolutionary sequences are still being refined. The relation of the final evolutionary stage to the supernova types indicated above will be clarified in Lecture 16.

The limiting masses given above are only indicative, and apply (approximately) to massive stars of solar metallicity. However, mass-loss rates decrease with decreasing Z because, as we shall see presently (Section 15.3 and following), stellar winds are driven by the absorption of photons by metal lines. Thus, the mass limits are higher for stars of lower metallicity.

The metallicity dependence of the winds from massive stars is thought to be the fundamental reason behind the empirical observation that the frequency of WR stars and the breakdown between different WR subclasses are not the same in all galaxies. The best sampled galaxies are the Milky Way ($Z = Z_\odot$), the LMC ($Z \simeq 1/2Z_\odot$) and the SMC ($Z \simeq 1/6Z_\odot$), although with large telescopes the samples are now being extended to other galaxies. In the solar neighbourhood, the ratio of WR stars to their O progenitors is $N(\text{WR})/N(\text{O}) \sim 0.15$, but in the SMC

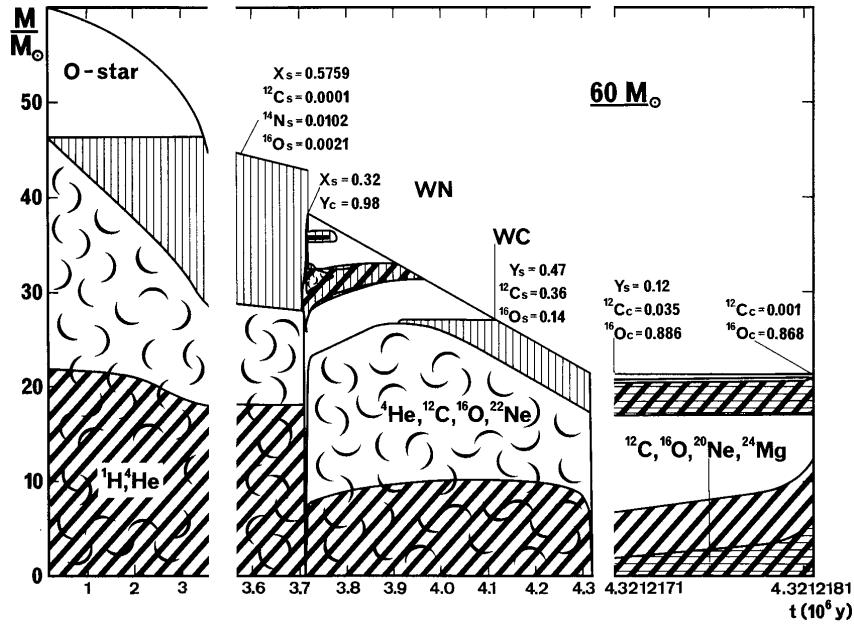


Figure 15.5: Diagram of the evolution of a $60M_\odot$ star of solar metallicity. Cross-hatched areas indicate where nuclear burning occurs, and curly symbols indicate convective regions. (Figure reproduced from Maeder & Meynet 1987, A&A 182, 243).

$N(\text{WR})/N(\text{O}) \simeq 12/1000 = 0.01$. Furthermore, in the solar neighbourhood $N(\text{WN}) \simeq N(\text{WC})$, but in the LMC $N(\text{WN})/N(\text{WC}) \simeq 5$ and in the SMC $N(\text{WN})/N(\text{WC}) \sim 10$.

Figure 15.5 illustrates the evolutionary scenario for a $M = 60M_\odot$ star, indicative of the most massive stars. After ~ 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, especially N, are revealed, and the surface He abundance increases at the expense of H. During the very short phase between core 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 ($t = 3.7$ – 3.9 Myr), the star appears as a WNL star, and subsequently ($t = 3.9$ – 4.1 Myr) as a WNE star, after mass loss has removed the last H-rich layers outside the H-burning shell. After 4.1 Myr, 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. In the last 0.2 Myr of its evolution the star is a WC star.

In general, mass loss rates during all evolutionary phases increase with stellar mass; as a result, there is a convergence of the final (pre-supernova) masses to a range between 5 and $10M_\odot$. However, this effect is diminished in metal-poor stars which experience lower mass loss rates.

Table 15.1 summarises some of the properties of the different nuclear burning stages in a $15M_{\odot}$ star. Note the accelerating timescales as heavier and heavier elements are ignited and the star approaches its final fate as a supernova.

Table 15.1. 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

15.3 Stellar Winds

Most stars lose mass. The existence of the *solar wind*, a stream of high velocity particles moving radially outwards from the Sun and carrying magnetic fields with them, was inferred in the early 1950s from the observation that comet ion tails always point away from the Sun, rather than trailing behind the comet. The total mass flow can be estimated from the particle density and their typical velocity (at the location of the Earth):

$$\dot{M} = n m_H v 4\pi d^2 \sim 10^{-14} M_{\odot} \text{ yr}^{-1} \quad (15.1)$$

where $d = 1 \text{ AU}$, $n \sim 5 \text{ cm}^{-3}$, and $v \sim 500 \text{ km s}^{-1}$. With such a low mass loss rate, the Sun will have lost only $\sim 0.1\%$ of its mass during its entire lifetime of $\sim 10^{11}$ years (if the present mass-loss rate remained constant during the whole lifetime of the Sun, which it won't, as we saw in Lecture 13).

The mechanism by which solar mass stars lose mass is quite different from that which drives the much higher mass loss rates ($\dot{M} = 10^{-7}\text{--}10^{-4} M_{\odot} \text{ yr}^{-1}$, corresponding to $\sim 1/30$ to 30 Earth masses per year) seen in hot stars with $M \gtrsim 15M_{\odot}$ and which we are going to discuss here.

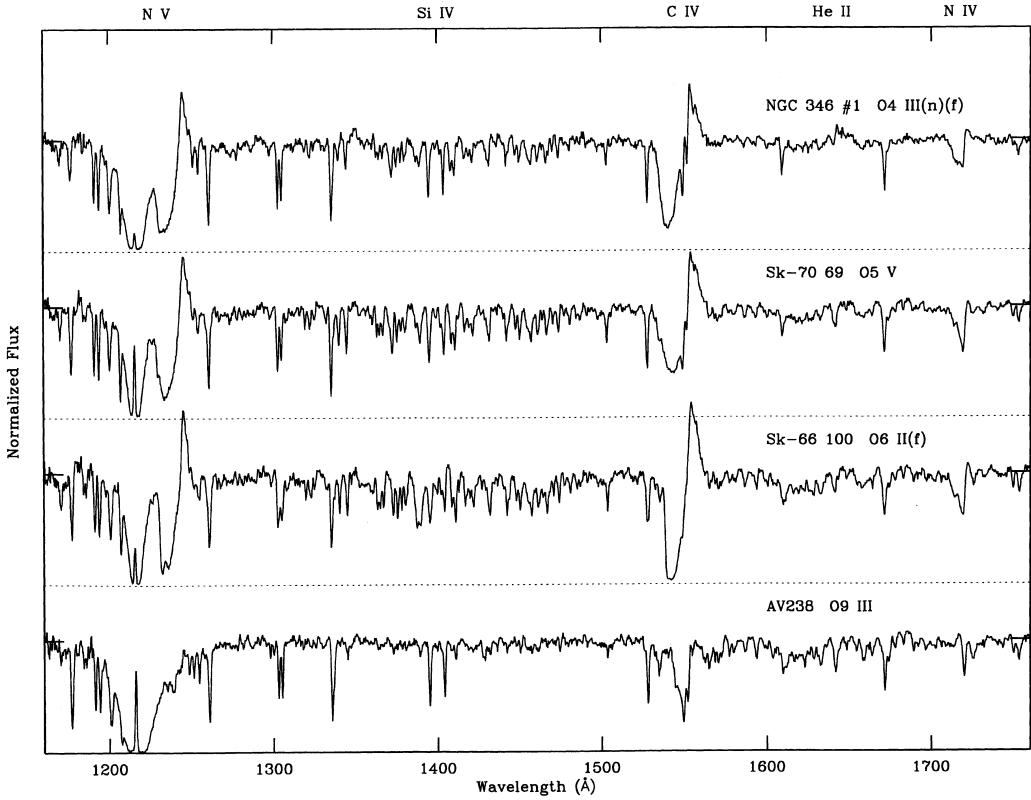


Figure 15.6: Ultraviolet spectra of massive OB stars in the Magellanic Clouds, obtained with the Faint Object Spectrograph on the *Hubble Space Telescope*. The resonance lines of $\text{N V} \lambda\lambda 1238, 1242$ and $\text{C IV} \lambda\lambda 1548, 1550$ display strong P Cygni profiles indicative of high wind terminal velocities and large mass loss rates. (Figure reproduced from Walborn et al. 1995, PASP, 107, 104).

15.3.1 P Cygni Line Profiles

We see *direct* evidence of such mass loss in the profiles of spectral lines of highly ionised species such as C IV , Si IV , N V and O VI in the ultraviolet spectra (from 1549 to 1035 Å) of primarily O and B-type stars (but also WR stars and A supergiants); some examples are shown in Figure 15.6.

These line profiles, which are a mixture of emission and absorption, are called P Cygni profiles from the LBV star in which they were first seen. Figure 15.7 illustrates the basic formation mechanism in an *outflowing extended* stellar atmosphere.

The material in front of the stellar disk absorbs light at the frequencies of these resonance lines. The resulting absorption profile extends from $v = 0 \text{ km s}^{-1}$ (assuming the radial velocity of the star relative to the Earth to be zero) to a maximum *negative* velocity, v_∞ . Thus, the absorption line has a net *blue-shift* relative to the star. Although absorption of starlight

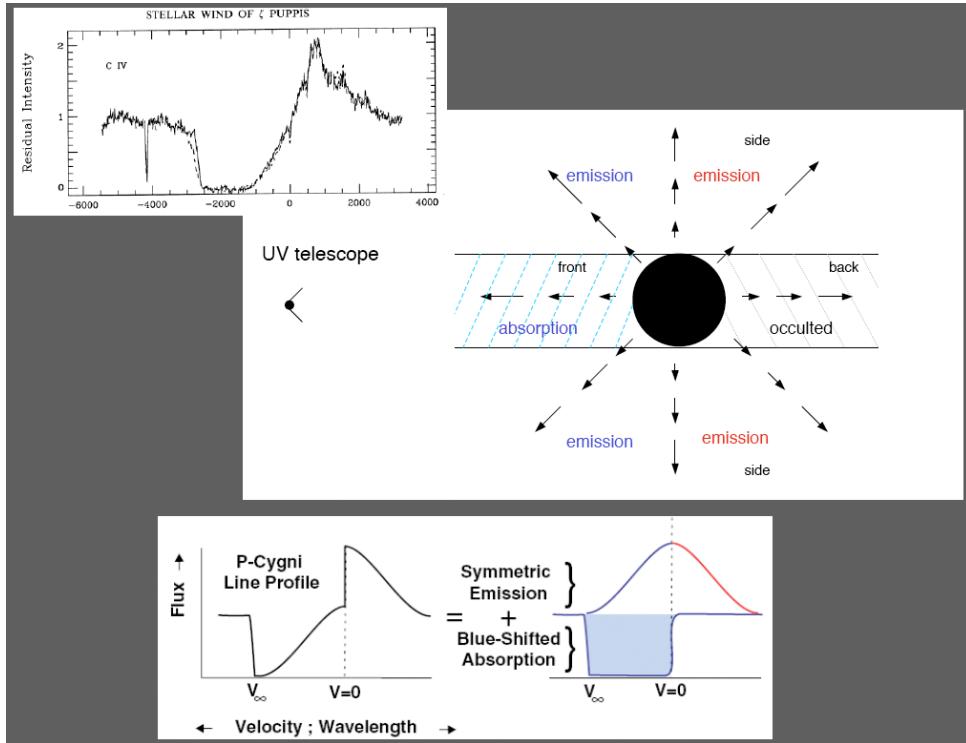


Figure 15.7: Schematic representation of the origin of P Cygni line profiles in an expanding stellar atmosphere.

takes place everywhere within the extended atmosphere, we can only see it if the absorbing ions are located in front of the stellar disk—hence the net blueshift of the absorption component of the P Cygni profile.

Once the outflow reaches its maximum velocity, this velocity will remain approximately constant up to large distances from the star (in the absence of other forces); hence the term ‘*terminal velocity*’. Typical values range from $v_\infty \simeq -200 \text{ km s}^{-1}$ in A-type supergiants to $v_\infty \simeq -3000 \text{ km s}^{-1}$ in early O-type stars. The sound speeds in the atmospheres of these stars are typically $10\text{--}30 \text{ km s}^{-1}$; thus the winds are highly *supersonic*.

If the material in front of the star is optically thick over the full velocity range, it will produce strongly saturated, almost rectangular, absorption line profiles, as in the example shown in Figure 15.7. This is not necessarily always the case, and weaker, unsaturated absorption components to the P Cygni composite profile are possible.

Following absorption in any of the above resonant (from the ground state to the first excited level of the ion under consideration) lines, a new photon will be re-emitted as the electron returns to the ground state. Overall, an observer would see an emission profile resulting from a multitude of such

re-emission processes at different velocities. Note that the wind emission provides additional radiation over that due to the stellar photosphere's blackbody emission; that is, we see an emission line superimposed on the stellar continuum.

Unlike the absorption case, an observer sees photons emitted from gas behind, as well as in front of, the star—gas that is moving with, respectively, positive and negative velocities relative to the star. Such an emission profile has a maximum at $v = 0$ and falls off to zero at v_∞ . This is simply a geometrical effect, with the largest emitting areas having net zero velocity (projected along the line of sight to the observer). The larger the projected wind velocity, the smaller the corresponding emitting area. Since there are no absorbing and re-emitting ions at velocities larger than v_∞ , the emission profile is restricted to this velocity range.

Note that if the radius of the stellar photosphere (where the continuum is produced) is not small compared to the physical extent of the wind, the emission profile can be intrinsically asymmetric, with the approaching, blue-shifted, portion in front of the star emitting more flux (as viewed from a given direction) than the receding back portion which is occulted by the star.

The overall P Cygni profile, as observed for example in the C IV $\lambda\lambda 1548, 1550$ lines in the ultraviolet spectrum of the O5 Iaf star ζ Puppis (Figure 15.7), is the superposition of three components: stellar blackbody continuum, blueshifted absorption, and (possibly asymmetric) emission.

15.3.2 Diagnostics of P Cygni Profiles

There are three main physical quantities that can be determined from the analysis of the P Cygni profiles of UV absorption lines in the spectra of hot, massive stars:

(i) The terminal velocity, v_∞ . Provided there is a sufficiently high column density of absorbing ions over the full velocity range of the wind, v_∞ can be readily deduced by measuring the extreme blue wavelength of the absorption profile and applying the familiar Doppler formula:

$$v_\infty = \frac{\lambda_{\min} - \lambda_0}{\lambda_0} c$$

where λ_0 is the laboratory wavelength of the transition under consideration and c is the speed of light.

- (ii) The ion column density, N . Provided the absorption profile is not saturated (recall the discussion of this point in Lecture 6.4), it may be possible to deduce the distribution of ion column densities at different velocities in the wind. This is normally done by comparing a family of computer-generated theoretical line profiles with the observed one, and minimising the difference between the two. Under favourable circumstances, the ion column densities thus deduced can be interpreted to infer the mass loss rate and relative element abundances.
- (iii) The shape of the velocity field, $v(r)$. Particularly when the absorption profile is saturated, the exact shape of the absorption+emission composite is sensitive to the velocity field—changing the way v varies with distance r from the stellar ‘surface’ (for example, a steep or a shallow velocity gradient) can produce recognisable changes in the overall P Cygni profile. We shall show some examples in the next section.

15.4 Theory of Radiatively Driven Winds

In this section we shall take a closer look at radiative line acceleration in hot star winds. In particular, we’ll see how this acceleration mechanism leads to certain scaling relations for some of the wind properties, such as \dot{M} and v_∞ . We consider the simplest possible model: *a homogeneous, time-independent,¹ spherically symmetric, stellar wind free of magnetic fields*. We further assume that the photons emitted from the stellar photosphere and which are driving the wind are propagating only radially. None of these simplifying assumption actually applies to real hot star winds, of course. However, it turns out that the majority of observations can be reproduced satisfactorily with such a simple model, except for the assumption of homogeneity (that is, introducing *clumping* in the wind turns out to be important).

¹Confusingly, in the literature the assumption that the wind properties are time-independent is referred to as a ‘stationary’ wind.

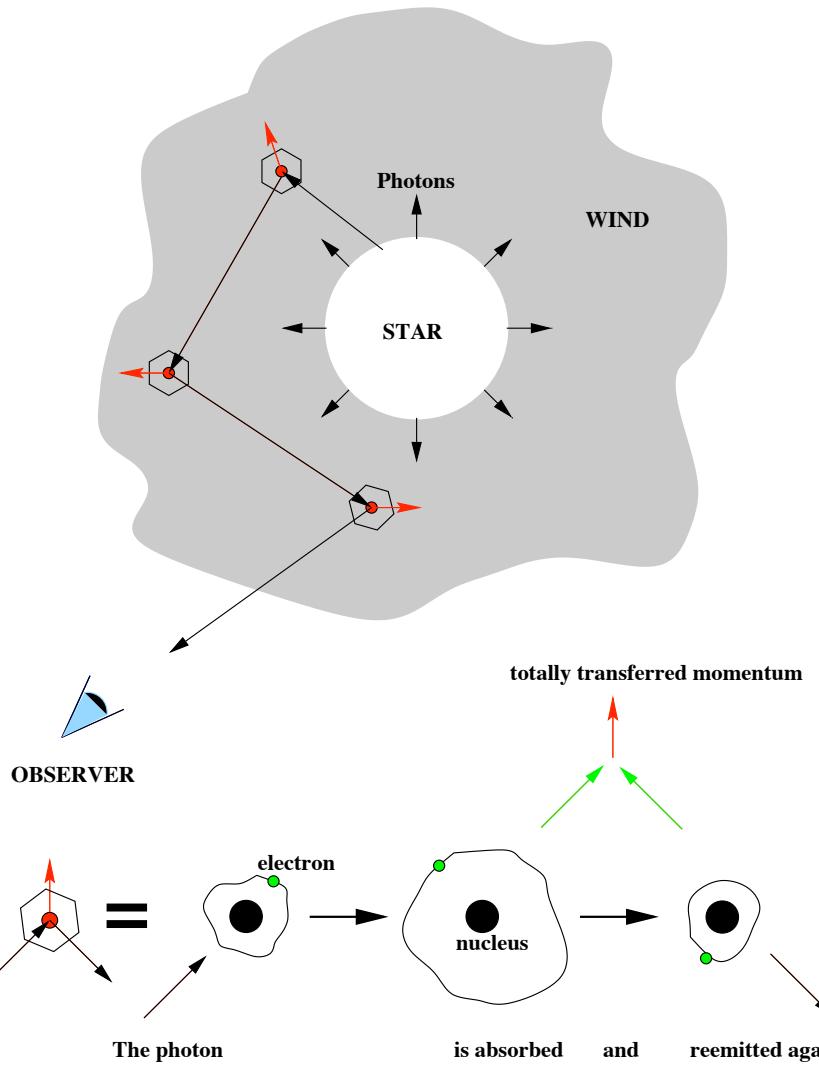


Figure 15.8: Schematic representation of the principle of radiative line-driven winds. Each absorbed photon transmits a momentum $h\nu/c$ to the absorbing ion in the direction of propagation of the photon. The emitted photons transmit a momentum in the direction opposite to the propagation direction. The momenta transmitted by the isotropically re-emitted photons cancel on average. The momenta of the absorbed photons add.

15.4.1 Momentum Transfer via Line Absorption and Re-emission

The basic principle is illustrated in Figure 15.8. The absorption and re-emission of stellar photons in a spectral line with frequency ν_i in the atomic rest frame, result in a net transfer of momentum in the radial direction

$$\Delta P_{\text{radial}} = \frac{h}{c} (\nu_{\text{in}} \cos \theta_{\text{in}} - \nu_{\text{out}} \cos \theta_{\text{out}}) \quad (15.2)$$

to the absorbing and re-emitting ion. Here θ is the angle between the direction of the photon and the radial unit vector (parallel to the velocity vector) of the ion. Since the absorbed photons are re-emitted isotropically,

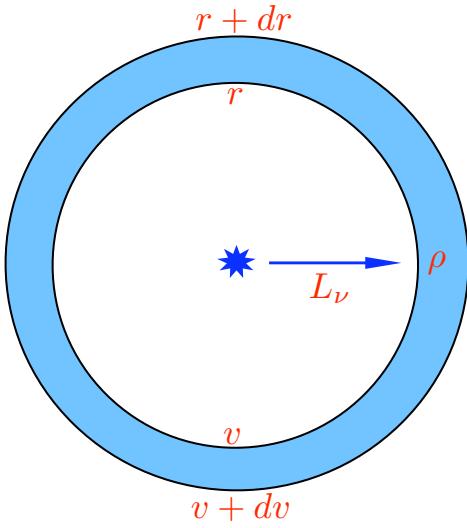


Figure 15.9: Sketch of a blue supergiant star irradiating a thin shell of wind material. L_ν is the star luminosity at frequency ν , v is the wind velocity at radius r , and ρ is the local density within the shell. The shell has mass $m = 4\pi r^2 \rho dr$.

we have:

$$\langle \cos \theta_{\text{out}} \rangle = 0. \quad (15.3)$$

On the other hand, prior to absorption the stellar photons are approaching from the direction of the star, that is their direction of propagation is parallel to the ion velocity vector; thus:

$$\langle \cos \theta_{\text{in}} \rangle \approx 1. \quad (15.4)$$

Therefore,

$$\langle \Delta P_{\text{radial}} \rangle = \frac{h\nu_{\text{in}}}{c} \quad (15.5)$$

Let us now consider a thin shell in the wind, as in Figure 15.9. Within the shell, the wind velocity increases by an amount dv on a scale dr . Photons emitted from the stellar surface (photosphere) with observer's frame frequency ν_{obs} can be absorbed by an ion if their frequency in the *ion* frame equals the transition frequency ν_i . Assuming that the photons propagate radially, the two frequencies are related via the Doppler formula:

$$\begin{aligned} \nu_i &= \nu_{\text{obs}} - \frac{\nu_i}{c} v \\ \nu_i &= (\nu_{\text{obs}} + d\nu_{\text{obs}}) - \frac{\nu_i}{c} (v + dv). \end{aligned} \quad (15.6)$$

In other words, a possible absorption and re-emission process (which from now on we shall call a 'scattering' event) by an ion moving outwards with the wind requires stellar photons which have left the stellar surface with

higher frequencies than that of the atomic transition under consideration. The frequency interval $d\nu_{\text{obs}}$ corresponding to the velocity interval dv is just:

$$d\nu_{\text{obs}} = \nu_i \frac{dv}{c}.$$

The radiative acceleration of the shell resulting from a given absorption line can be calculated using the general definition of any acceleration:

$$g_{\text{rad}}^i = \frac{\Delta P}{\Delta t \Delta m}. \quad (15.7)$$

In our case, the line acceleration is obtained by multiplying the average momentum transferred in a scattering event by the number of available photons in the corresponding frequency interval, per unit time and per unit mass of the accelerated shell. The number of photons per unit time is simply:

$$\frac{N_\nu}{\Delta t} = \frac{\Delta (E_\nu/h\nu)}{\Delta t} = \frac{L_\nu \Delta \nu_{\text{obs}}}{h\nu_{\text{obs}}} \quad (15.8)$$

where L_ν is the stellar luminosity (equal to the radiated energy per unit time per unit frequency) at frequency ν . With $\nu_{\text{in}} = \nu_{\text{obs}}$, the radial acceleration of the shell caused by a single atomic transition is therefore:

$$g_{\text{rad}}^i = \frac{N_\nu \langle \Delta P_{\text{radial}} \rangle}{\Delta t \Delta m} = \frac{L_\nu \Delta \nu_{\text{obs}}}{h\nu_{\text{obs}}} \frac{h\nu_{\text{obs}}}{c} \frac{1}{\Delta m} = \frac{L_\nu \nu_i}{c^2} \frac{dv}{dr} \frac{1}{4\pi r^2 \rho}. \quad (15.9)$$

It is noteworthy that the radiative line acceleration of the shell depends on the velocity gradient dv/dr within the shell (can you see why?).

15.4.2 Total Line Acceleration

To obtain the total radiative acceleration we need to take into account two other factors.

First, until now we have assumed a unit probability of interaction between photons (of the appropriate frequency) and ions. We include in eq. (15.9) a factor which reflects the dependence of this probability on the atomic properties:

$$\mathcal{P}_{\text{inter}} = 1 - e^{-\tau}$$

where τ is the optical depth of the observed transition at coordinates r, v, dv .

Recalling again our treatment of absorption lines in Lecture 6, we can reduce $\mathcal{P}_{\text{inter}}$ to two limiting cases. If $\tau \gg 1$, the atomic transition in question is optically thick, and $\mathcal{P}_{\text{inter}} \approx 1$. On the other hand, if $\tau \lesssim 1$, the line is optically thin and the acceleration due to such lines needs to be modified by a factor corresponding to the local optical depth. If $\tau \ll 1$, $\mathcal{P}_{\text{inter}} \approx \tau$. The radiative acceleration due to an optically thin line is a factor of τ smaller than that due to an optically thick line.

Second, in the expanding atmosphere of a hot star, there is not just a single atomic transition capable of driving a wind. On the contrary, there are literally millions of transitions from atoms and ions of the most abundant elements of the Periodic Table which are potentially capable of absorbing radiation and momentum. In practice, ‘only’ some ten thousand lines are relevant for calculating the overall line acceleration, $g_{\text{rad}}^{\text{tot}}$, because the rest have too low an interaction probability or lie in a spectral range where the photon density is small.

In order to calculate the total line acceleration, we have to sum up all individual contributions:

$$\begin{aligned} g_{\text{rad}}^{\text{tot}} &= \sum_{\text{thin}} g_{\text{rad}}^{\text{i}} + \sum_{\text{thick}} g_{\text{rad}}^{\text{i}} \\ &= \frac{1}{4\pi r^2 c^2} \left(\sum_{\text{thin}} L_\nu \nu_i \frac{dv}{dr} \frac{\tau_i}{\rho} + \sum_{\text{thick}} L_\nu \nu_i \frac{dv}{dr} \frac{1}{\rho} \right) \end{aligned} \quad (15.10)$$

(with the above scheme of dividing the lines into just two categories of optical depth).

In the so-called ‘Sobolev’ approximation:

$$\tau_i = \frac{k_i \rho}{dv/dr} \quad (15.11)$$

the optical depth of a line can be expressed as a function of the velocity gradient, density and a ‘line-strength’ parameter k_i which includes all of the atomic and plasma physical details of the transition (most importantly, occupation number of the absorbing level and the interaction cross-section), and remains roughly constant throughout the wind.

The mammoth task of including thousands of atomic transitions in the summations in eq. 15.10 is made very much easier by the fact that the

distribution of line strengths can be satisfactorily approximated by an analytical power-law:

$$dN(\nu, k_i) = -N_0 f_\nu(\nu) k_i^{\alpha-2} d\nu dk_i. \quad (15.12)$$

Here, $dN(\nu, k_i)$ is the number of lines in the frequency interval $\nu, \nu+d\nu$ with line strengths $k_i, k_i + dk_i$, and the exponent α takes the values $0 < \alpha < 1$. Note that the frequency distribution of lines is independent of the line strength distribution. The validity of eq. 15.12 is confirmed by detailed model atmosphere calculations, which also indicate that typically $\alpha \approx 2/3$.

Equations 15.12 and 15.10 then lead to a rather simple expression for the total line acceleration in the wind:

$$g_{\text{rad}}^{\text{tot}} = \mathcal{C} \frac{L}{4\pi r^2} \left(\frac{dv/dr}{\rho} \right)^\alpha \quad (15.13)$$

where $L = \int L_\nu d\nu$ is the stellar luminosity. Note that in this formulation the line acceleration depends only on hydrodynamical quantities apart from the scaling factor \mathcal{C} and the exponent α .²

15.4.3 Solving the Equation of Motion

With the simple analytical form for the total line acceleration given by eq. 15.13, we can begin to examine the hydrodynamical structure of the stellar wind. The following equations apply to stationary, spherically symmetric flows:

1. The equation of Continuity:

$$\boxed{\frac{dM}{dt} = 4\pi r^2 \rho v} \quad (15.14)$$

2. The equation of Momentum:

$$\boxed{v \frac{dv}{dr} = -\frac{1}{\rho} \frac{dp}{dr} - g_{\text{grav}}(1 - \Gamma) + g_{\text{rad}}^{\text{tot}}} \quad (15.15)$$

²The real challenge of computer modelling of hot star winds is the calculation of these two parameters, whose values depend on the occupation numbers of all the contributing atomic energy levels.

3. The equation of State:

$$p = -\rho a^2 \quad (15.16)$$

where M is the stellar mass, p is the pressure, a is the isothermal sound speed, and g_{grav} is the gravitational acceleration of the star. Note that this last quantity is modified by a factor $\Gamma < 1$ to take into account the acceleration due to Thomson scattering of photons off free electrons. The parameter Γ is the same *Eddington factor* already encountered in eq. 10.40:

$$\Gamma = \frac{\kappa_{es} L}{4\pi c GM} \quad (15.17)$$

with the main source of opacity κ provided by electron scattering. Radiation pressure via Thomson scattering reduces gravity by a constant factor. For this reason, the term $M(1 - \Gamma) \equiv M_{\text{eff}}$ is sometimes referred to as the ‘effective mass’. Clearly, $\Gamma < 1$ for a star to be stable against radiation pressure (as already discussed in Lecture 10.6); in early-type supergiants, $\Gamma \simeq 0.5$ typically.

Let us now solve these equations for the major, supersonic portion ($v > a$) of the wind. In this regime, pressure forces can be neglected. Inserting (15.14) into (15.15) and making use of (15.13), the equation of motion of the wind now becomes:

$$r^2 v \frac{dv}{dr} = -GM(1 - \Gamma) + C'L \left(\frac{dM}{dt} \right)^{-\alpha} \left(r^2 v \frac{dv}{dr} \right)^\alpha. \quad (15.18)$$

Equation 15.18 can be readily solved with the substitution $z = r^2 v dv / dr$. The parameter z needs to be constant throughout the wind to allow for a unique solution, since all the other quantities are constant as well. This constrains the mass loss rate $dM/dt \equiv \dot{M}$ to:

$$\dot{M} \propto L^{1/\alpha} [M(1 - \Gamma)]^{1-1/\alpha}. \quad (15.19)$$

Furthermore, from the condition $z = \text{constant}$, the velocity law is obtained via a simple integration, *independently of the mass loss rate*:

$$v(r) = v_\infty \left(1 - \frac{R_*}{r} \right)^{1/2} \quad (15.20)$$

and

$$v_\infty = \left(\frac{\alpha}{1 - \alpha} \right)^{1/2} \left(\frac{2GM(1 - \Gamma)}{R_*} \right)^{1/2} \quad (15.21)$$

where R_* is the stellar radius. The second term on the r.h.s. of eq. 15.21 is the photospheric escape velocity. We mentioned earlier that detailed model atmosphere calculations indicate that values of $\alpha \approx 2/3$ are typical; thus, $v_\infty \simeq \sqrt{2} v_{\text{esc}}$.

The above analytical treatment of the hydrodynamics of stellar winds has of necessity made a number of simplifying assumptions. More detailed analyses, however, do not change the picture dramatically. Most importantly, the scaling relation for \dot{M} remains unaltered and the proportionality between v_∞ and v_{esc} is maintained, although the constant of proportionality changes somewhat. The most severe change concerns the shape of the velocity field:

$$v(r) = v_\infty \left(1 - \frac{R_*}{r} \right)^\beta \quad (15.22)$$

with the exponent $\beta \approx 0.8$ in most cases, rather than the value of $1/2$ deduced above. That is, the wind velocity increases more slowly with distance from the star.

It was stated earlier (Section 15.3.2) that from the shape of the P Cygni profiles it is possible to infer the shape of the velocity field within the wind. We can now illustrate the dependence of the line profiles on the value of the exponent β in eq. 15.22 with the computed P Cygni lines shown in Figure 15.10.

The higher the value of β , the further from the star is v_∞ reached. Figure 15.10 shows that the shallower the velocity field, the stronger is the P Cygni emission. The variation in the line profiles within each panel (i.e. for the same value of β) illustrates the dependence of the emission/absorption on the ion density, which was assumed to be proportional to the total wind density. Each profile differs by a factor of ten in density. Note that the last two cases are essentially indistinguishable, because the lines have become saturated.

15.4.4 The Wind-Momentum Luminosity Relation

The scaling relations we have derived provide a theoretical explanation for the empirical relationship between wind momentum and stellar luminosity first formulated by Rolf Kudritzki and collaborators on the basis of observations of massive stars.

In Galactic supergiants, the wind-momentum luminosity relation takes the form:

$$\dot{M}v_\infty \left(\frac{R_*}{R_\odot} \right)^{1/2} \propto L^{1.46}. \quad (15.23)$$

What this empirical relation tells us is that, for a given stellar radius, the wind-momentum rate depends on some power of the stellar luminosity alone. On the other hand, from our theoretical scaling relations we would

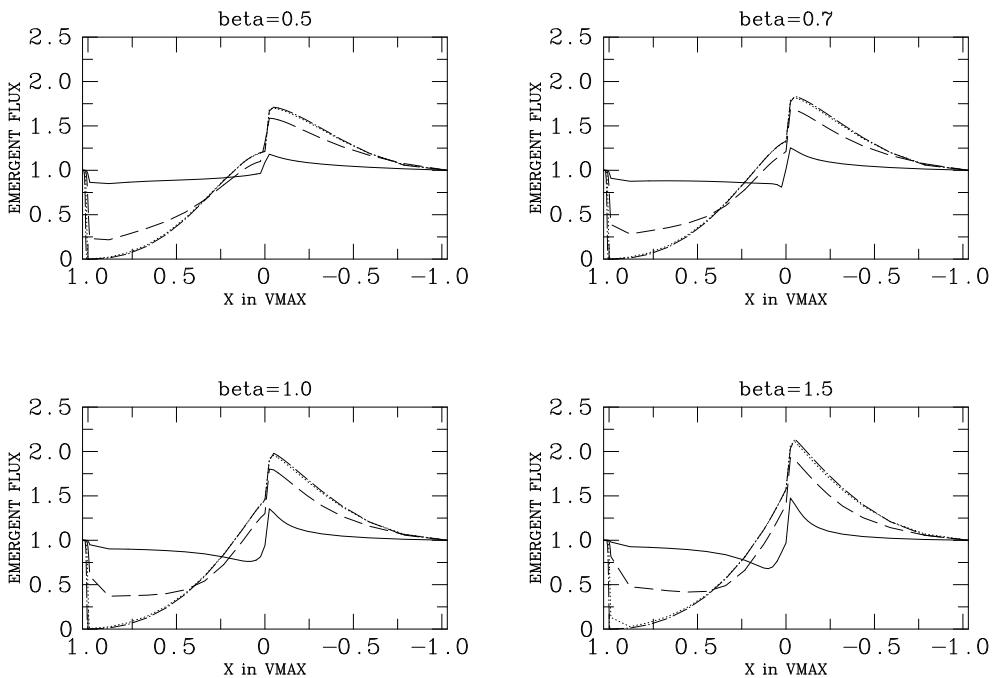


Figure 15.10: Examples of theoretical P Cygni line profiles. The four panels correspond to different values of the exponent β in eq. 15.22, which governs the steepness of the velocity law. In each panel, four profiles are shown for different optical depths of the absorbing ions, as follows. *Continuous line*: low ion column density; *dashed line*: intermediate ion column density; *dotted and dash-dotted lines*: high column densities (these last two lines are superimposed and therefore cannot be readily distinguished in the plots).

predict, combining eqs. 15.19 and 15.21:

$$\dot{M}v_\infty \left(\frac{R_*}{R_\odot} \right)^{1/2} \propto L^{1/\alpha} [M(1 - \Gamma)]^{3/2 - 1/\alpha} \quad (15.24)$$

which at first glance seems quite different in that it includes an additional mass dependence. But, recalling that $\alpha \approx 2/3$, the terms in square brackets in eq. 15.24 becomes unimportant, leaving only the term $L^{1/\alpha} = L^{1.5}$ on the r.h.s. of the equation, in excellent agreement with the empirical relation at (15.23).

Summarizing, the observed wind-momentum luminosity relation can be explained as a consequence of the scaling relations for line-driven winds, plus the exponent of the line-strength distribution function being close $\alpha = 2/3$.

15.5 Metallicity Dependence of Mass Loss

From the discussion in Section 15.4, it should be fairly obvious that the properties of radiatively driven winds are likely to exhibit a metallicity dependence.

Due to its definition, the ‘line-strength’ parameter k_i in the Sobolev approximation (eq. 15.11) scales with metallicity (under the assumption that the ionisation balance is not severely modified) as:

$$k_{i,Z} = k_{i,Z_\odot} \frac{Z}{Z_\odot} \quad (15.25)$$

Thus, the major effect of changing the metallicity is a horizontal shift of the corresponding line-strength distribution function (in the log–log representation) to the ‘left’ (for $Z/Z_\odot < 1$) or to the ‘right’ (for $Z/Z_\odot > 1$), as sketched in Figure 15.11. Such a shift translates to a change in the total number of lines contributing to the wind acceleration, and to the corresponding normalisation constant N_0 in eq. 15.12. For a power-law, the normalisation varies as:

$$N_{0,Z} = N_{0,Z_\odot} \left(\frac{Z}{Z_\odot} \right)^{1-\alpha} \quad (15.26)$$

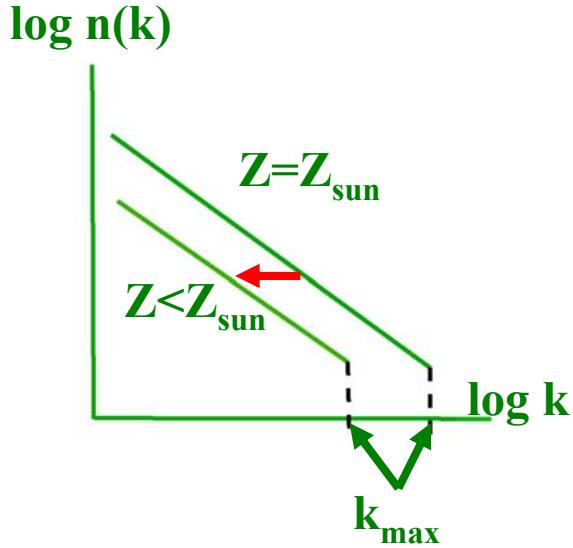


Figure 15.11: Sketch of the effect of decreasing metallicity on the line-strength distribution function.

The overall effect is a metallicity dependence of the mass loss rate of the form:

$$\dot{M}_Z = \dot{M}_{Z_\odot} \left(\frac{Z}{Z_\odot} \right)^{\frac{1-\alpha}{\alpha}} \quad (15.27)$$

or

$$\frac{\dot{M}_Z}{\dot{M}_{Z_\odot}} = \left(\frac{Z}{Z_\odot} \right)^{0.5} \quad (15.28)$$

if $\alpha = 2/3$. Metallicity has a smaller effect on the wind terminal velocity:

$$v_\infty \propto \left(\frac{Z}{Z_\odot} \right)^{0.15} \quad (15.29)$$

More extensive parameterizations of the mass loss rate as a function of L , M , T_{eff} , v_∞ and Z have been computed from ever-more sophisticated numerical modelling (e.g. Vink et al. 2001, A&A, 369, 574). When combined with a theoretical M - L relation for massive stars, we find the typical values listed in Table 15.2. Massive stars with $M > 50M_\odot$ lose more than 10% of their mass while on the Main Sequence.

For Galactic stars, the empirical relationship proposed by Jager et al. (1988):

$$\log \dot{M} = -8.158 + 1.769 \log(L/L_\odot) - 1.676 \log(T_{\text{eff}}/\text{K}) \quad (15.30)$$

is often adopted. Note the inverse temperature dependence. The reason is to be found in the lower opacities at higher temperatures; thus, at a given

bolometric luminosity, an A-type supergiant will have a higher mass loss rate than an O-type star.

Table 15.2 Parameters of Massive Stars

$M_{\text{ZAMS}} (M_{\odot})$	$\log(L/L_{\odot})$	$\log(\dot{M}/M_{\odot} \text{ yr}^{-1})$	$t_{\text{MS}} (10^6 \text{ yr})$	$\Delta M/M^{\text{a}}$
25	4.85	-6.97	6.6	0.028
40	5.34	-6.23	4.5	0.067
60	5.70	-5.68	3.7	0.13
85	5.98	-5.26	3.3	0.21
120	6.23	-4.88	2.8	0.31

^a Fractional mass lost while on the Main Sequence

Both eq. 15.30, scaled for metallicity according to eq. 15.28, and the multi-parameter theoretical formulation of the mass loss rate by Vink et al. (2001) do a reasonably good job of matching the data (Figure 15.12), although there remains significant scatter (~ 0.3 dex), some of it undoubtedly due to uncertainties in the measurements of \dot{M} .

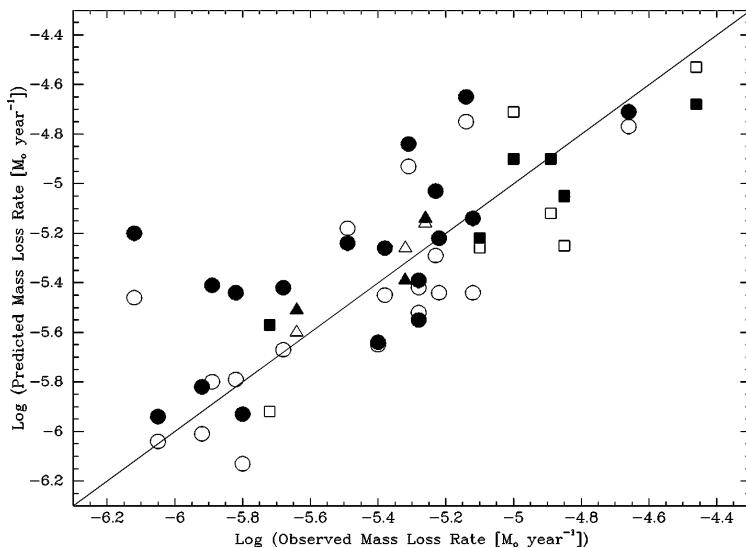


Figure 15.12: Comparison between observed (x -axis) and predicted (y -axis) mass loss rates. Open symbols denote values of \dot{M} calculated from eq. 15.30 with the metallicity scaling of eq. 15.28, while filled symbols are those given by the theoretical formalism of Vink et al. (2001). Circles are Galactic stars, squares LMC stars, and triangles SMC stars. (Figure reproduced from Massey et al. 2003, ARAA, 41, 15).

We still have much to learn regarding how mass loss rates in massive stars vary with metallicity and perhaps other parameters of the host galaxy. Our knowledge of this field is improving with observations, made possible by 8-10 m telescopes, of individual massive stars in galaxies beyond the Local Group, as well as in low-mass, low-metallicity dwarf galaxies in the Local Group. It is important to remember that the metallicity dependence of mass loss rates, as given for example by eq. 15.28, is empirically untested for metallicities $Z \lesssim 1/6Z_\odot$ (the metallicity of SMC O-type stars). This is crucial when we come to consider the evolution of the ‘First Stars’ (also referred to as Population III stars), which formed out of primordial gas—consisting only of H and He—and were likely very massive.

SUPERNOVAE

16.1 A Little History

In 1572 the Danish astronomer Tycho Brahe observed a new star in the constellation of Cassiopeia and described it in his book *De Stella Nova* (Latin for “*Concerning the New Star*”), giving rise to the name *nova*. Novae are now thought to be non-destructive stellar outbursts in one of the members of a binary system (more about this in a subsequent lecture).

In the 1920s, the realisation that many nebulae are located outside the Milky Way Galaxy was taking shape. With the discovery of *A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae* by Edwin Hubble in 1929 (Proceedings of the National Academy of Science, 15, 168), our cosmological horizon expanded significantly.

Only five years later, Baade and Zwicky, working at the Mount Wilson Observatory in southern California, put forward the suggestion that the novae observed in extra-galactic nebulae, the so-called super-novae, were associated with the death of massive stars that produced neutron stars and high energy cosmic rays—a triple correct prediction!

Supernovae are stellar explosions during which the luminosity of a single star quickly reaches $10^9\text{--}10^{10}L_\odot$ and then declines slowly remaining bright for several months. For comparison, the luminosity of the entire Milky Way galaxy is $L_{\text{MW}} \simeq 3 \times 10^{10}L_\odot$, so that a single star which goes supernova can be almost as bright as a whole galaxy for a brief period of time. This makes them ideal candidates for searches by amateur astronomers, as they can easily be spotted with even modest aperture telescopes. Nowadays, hundreds are discovered every year thanks to automatic searches using robotic telescopes. SNe are named with the year of discovery followed by letters; thus, SN2010A through to SN2010Z, followed by SN2010aa through to SN2010az, followed by SN2010ba, and so on, in chronological order of discovery.

16.2 Supernova Remnants

From the rate at which supernovae occur in external galaxies, it has been estimated that the SN rate in our own Galaxy is in the range 2–3 SNe per century. Even allowing for the fact that only about half of the Galaxy may be visible from the Sun’s location (we cannot readily see stars on the opposite side of the Galactic bulge), it appears that we are ‘overdue’ a SN, since the last recorded Galactic supernova exploded in 1604, four years before the invention of the telescope and six years before the Italian astronomer Galileo pointed his home-made telescope at Jupiter and discovered its four main satellites. In other words, there has not been a visible Galactic supernova since the birth of modern astronomy.

Table 16.1. 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

However, at least eight Galactic supernovae have been recorded over the past 2000 years by Chinese and in some cases also by Japanese, Korean, Arabian and European astronomers (see Table 16.1). The remnants of these supernovae are in most cases still visible as luminous expanding nebulae, containing the matter that was expelled in the explosion. Supernova remnants (SNRs) remain observable at optical, radio and X-ray wavelengths for tens of thousands of years, as the stellar ejecta interact with the ambient interstellar medium (see Figure 16.1). David Green’s catalogue of Galactic SNRs (<http://www.mrao.cam.ac.uk/surveys/snrs/>) lists 274 entries.

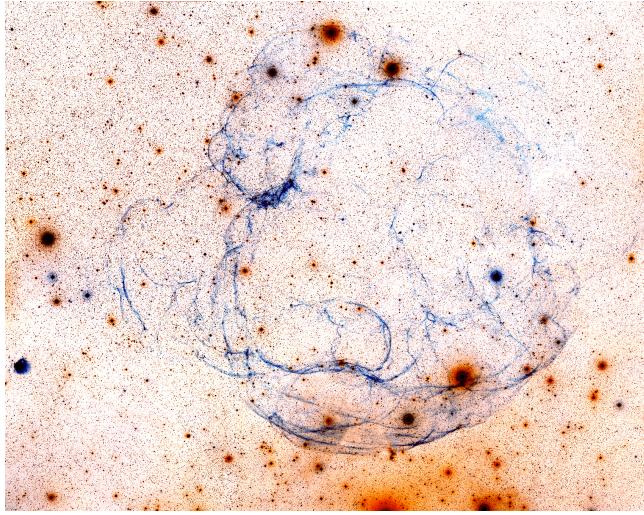


Figure 16.1: Image of the faint supernova remnant Simeis 147. This Type II supernova exploded $\sim 100\,000$ years ago, leaving behind a spinning neutron star or pulsar. Its remnant now extends over nearly three degrees on the sky in the constellation of Taurus.

16.3 Classification of Supernovae

A number of different evolutionary paths can lead to a supernova explosion. Before these paths became understood (and indeed there is still considerable uncertainty regarding some of them, as we shall see presently), an empirical classification for supernovae was developed, based on two observables: (i) their spectra, and (ii) their light curves, by which we mean the way their luminosity decays with time following maximum light. Figure 16.2 summarises the main SN types.

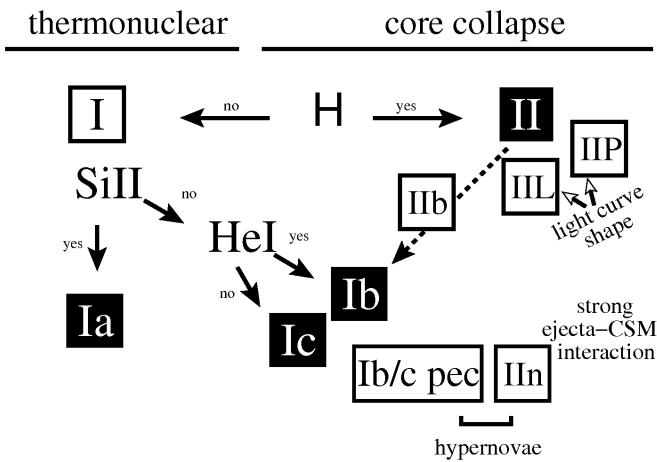


Figure 16.2: Spectral classification of supernovae, based on their spectra and light curve shapes. The main SN types are shown as black squares. (Figure reproduced from Turatto 2003, LNP, 598, 21).

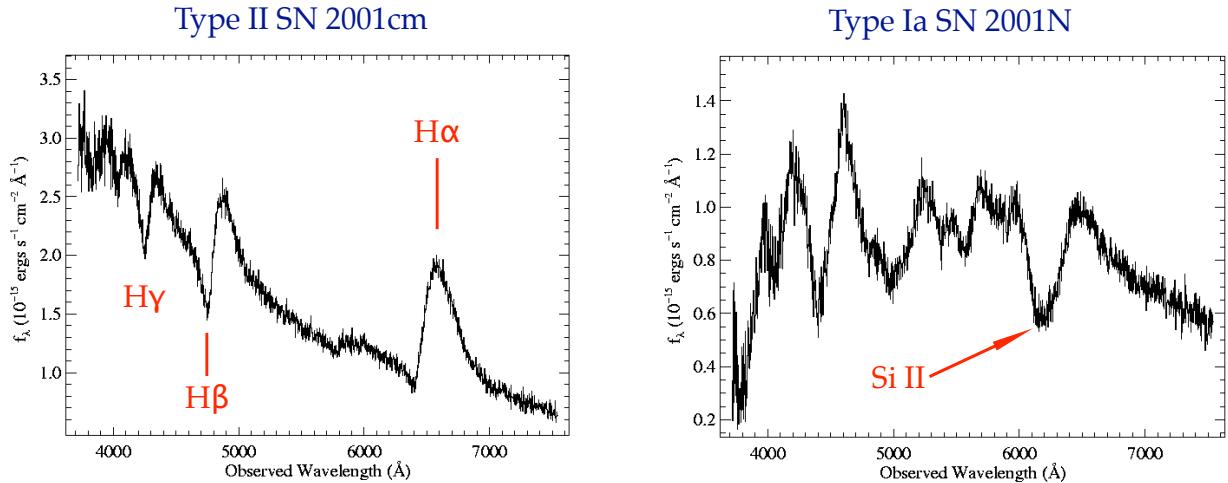


Figure 16.3: Typical optical spectra of Type II and Type Ia supernovae. (Spectra retrieved from www.supernovae.net).

SNe are divided between Type II and Type I depending on whether their optical spectra (usually at peak luminosity) exhibit hydrogen lines or not (see Figure 16.3). Supernovae of Type I are further subdivided into Type Ia if they show a prominent Si II $\lambda 6355$ absorption line, as in Figure 16.3, and Types Ib/Ic if they don't. Types Ib and Ic are differentiated on the basis of the presence or absence of He I lines. Type II SNe are labelled Type II-P (the more common type) and Type II-L depending on whether their light curves exhibit a plateau or a linear decay (see Figure 16.4).

A further distinction between Type Ia and other SNe is their location within a galaxy. Type Ia occur in galaxies of all types, and are the only SNe to occur in elliptical galaxies; since ellipticals have old stellar populations,

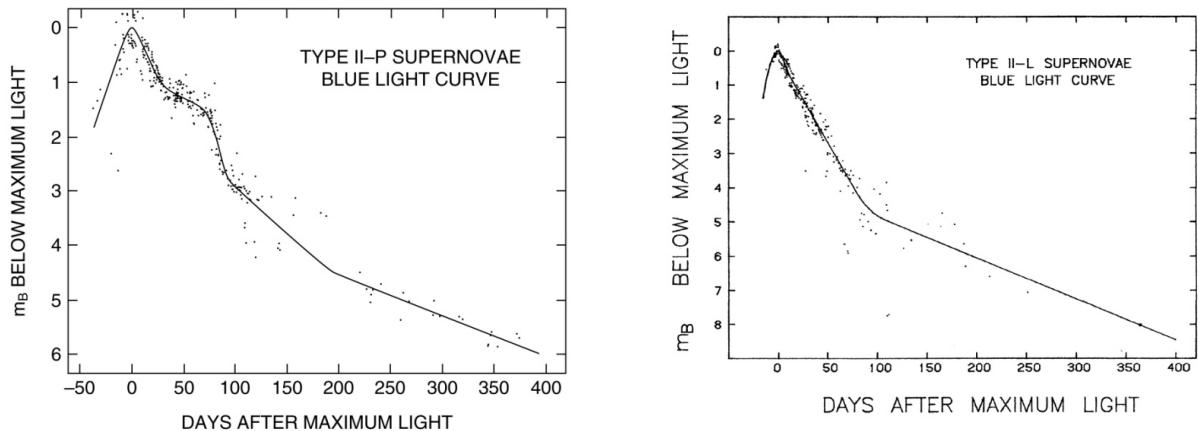


Figure 16.4: Typical light curves of Type II supernovae, from photometric m_B measurements of many Type II-P and Type II-L events. (Figures reproduced from Doggett & Branch 1985, AJ, 90, 2303).

this observation suggests that Type Ia SNe can have long-lived, low-mass progenitors. In contrast, Types II, Ib and Ic are associated with regions of star formation in spiral galaxies, and therefore are likely to have short-lived, high-mass progenitors. These observations have led to the current view that:

- (1) Type Ia are caused by the thermonuclear explosion of a CO white dwarf as a result of mass accretion in a binary system. About 25–30% of supernovae are of Type Ia. They are (on average) the most luminous of all supernova types and their light curves form a rather homogeneous group; the combination of these two characteristics makes them the best standard candles over cosmological distances. We shall return to Type Ia supernovae in a subsequent lecture.
- (2) Types II, Ib, and Ic SNe are collectively referred to as *core-collapse* supernovae, and are thought to be the end stage in the evolution of massive stars. While the progenitors of Type II SNe are still surrounded by a massive H-rich envelope when they explode, Type Ib progenitors have lost their H-rich envelope and Type Ic progenitors have also lost most of their He layer surrounding the core prior to the SN explosion. This sequence can be a consequence of mass loss from stars of increasing initial mass, as we saw in Lecture 15, but can also result from interaction with a binary companion. The remainder of this lecture is concerned with core-collapse supernovae.

16.3.1 The Progenitors of Core-Collapse Supernovae

Since the launch of the *Hubble Space Telescope* in 1990, a public-access archive of high resolution images of nearby galaxies has been accumulating. With high-precision astrometry, it has become possible to search for the progenitors of recent core-collapse supernovae in pre-supernova images of the host galaxies. This is fundamental work, providing empirical tests of our theories of stellar evolution for massive stars. Two teams, one based at Queen’s University, Belfast and the other at the University of California, Santa Cruz, have been particularly active in such endeavours; an excellent review of their results up to 2008 is given by Smartt, S. J., 2009, ARAA, 47, 63. Figure 16.5 shows three examples of nearby Type II-P SNe with unambiguous red supergiant progenitors.

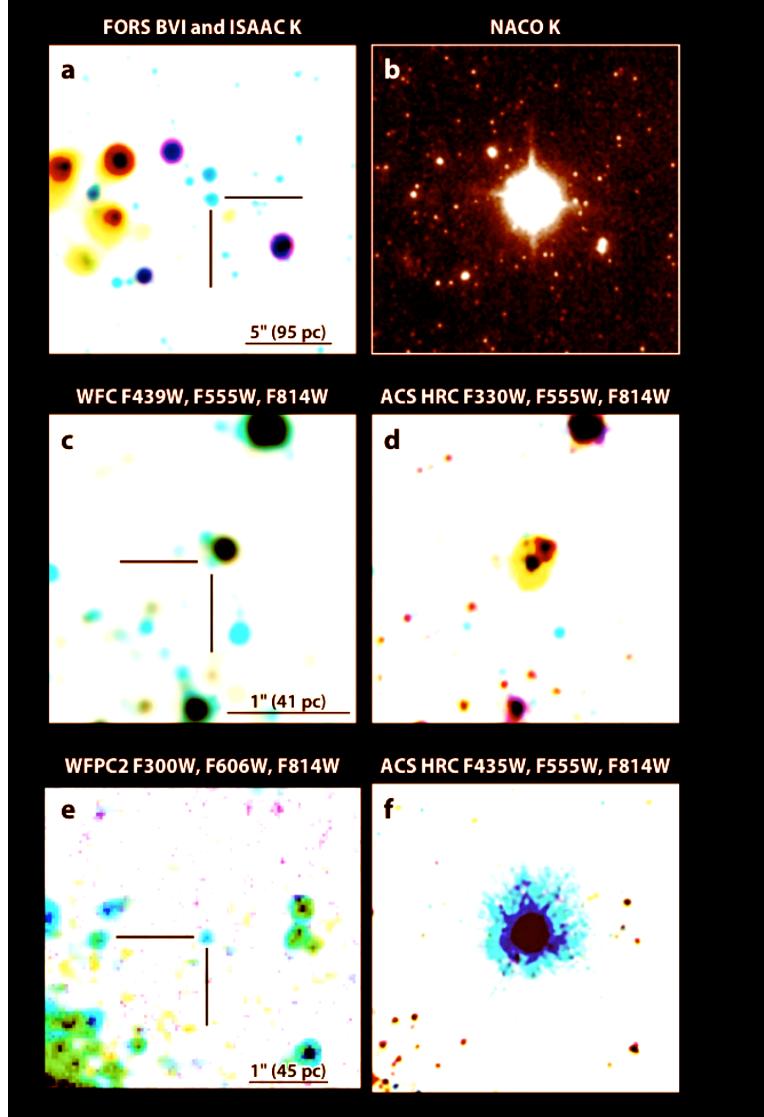


Figure 16.5: Hunting for SN progenitors. The left images show unambiguous red supergiants at the exact locations of subsequent Type II-P supernovae. *Top pair:* SN2008bk in the Scd spiral NGC7793 at a distance of ~ 4 Mpc. *Middle pair:* SN2005cs in M51 at a distance of ~ 7 Mpc. *Bottom pair:* SN2003gd in the face-on spiral galaxy M74 at a distance of ~ 9 Mpc. (Figure reproduced from Smartt 2009, ARAA, 47, 63).

Figure 16.6 summarises our current ‘best’ picture of the evolutionary paths to different core-collapse supernovae. The search for SN progenitors has produced a number of surprises, as well as confirming some aspects of the accepted scenario for the evolution of massive stars. As expected, red supergiants are the likely progenitors of most Type II-P SNe, but this channel seem to operate only for a relatively narrow mass range, for stars with initial masses between $M_{\min} = 8.5^{+1}_{-1.5} M_{\odot}$ and $M_{\max} = 16.5 \pm 1.5 M_{\odot}$. While the lower limit conforms to our ideas of the mass required to ignite carbon and heavier elements in the stellar core, as discussed in Lecture

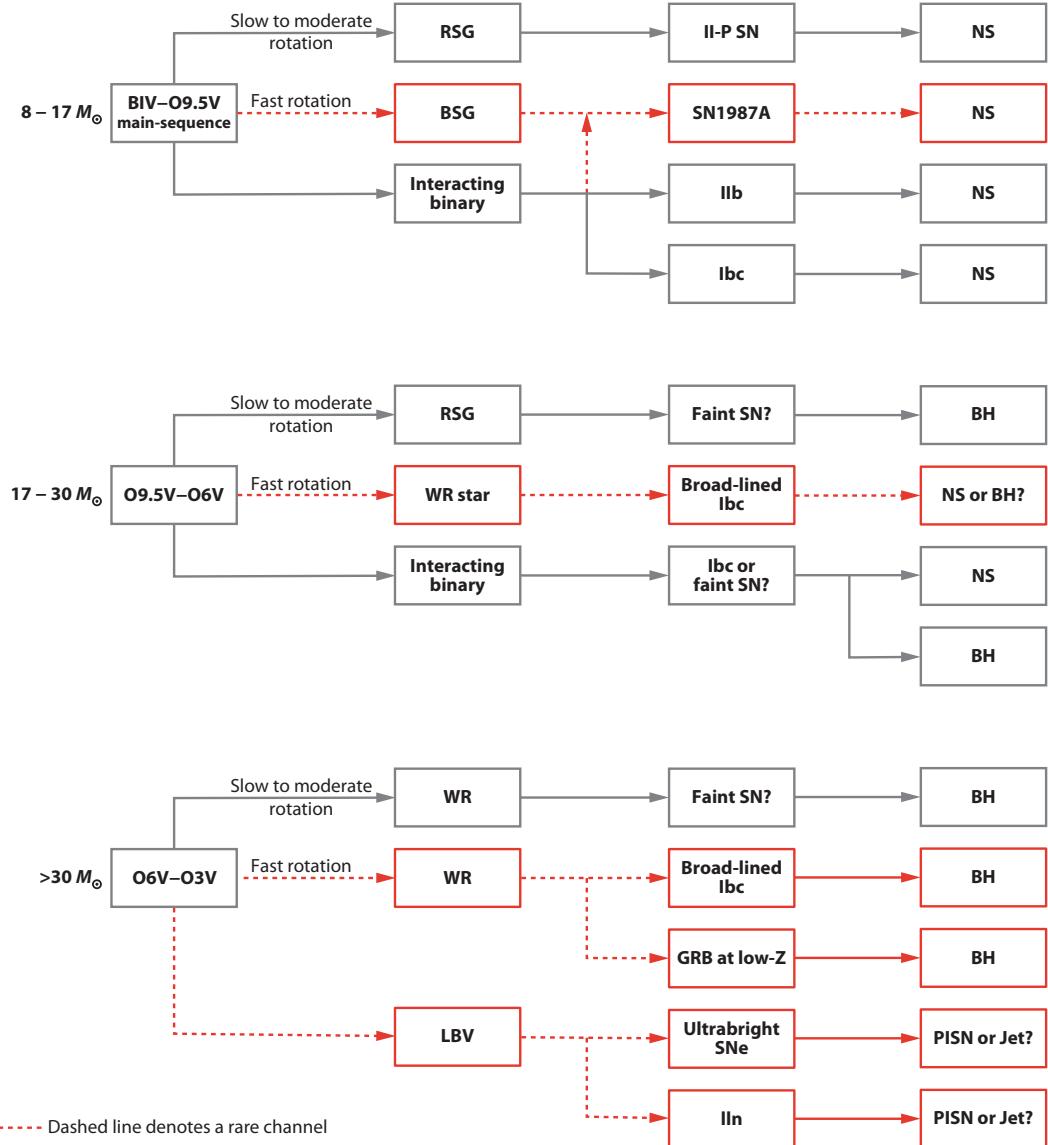


Figure 16.6: Summary diagram of possible evolutionary pathways and end states of massive stars—constructed on the basis of current empirical evidence and theoretical ideas—and illustrating the probable diversity in evolution and explosion. It is likely that metallicity, binarity and rotation all play important roles in determining the end states. The acronyms are for neutron star (NS), black hole (BH), and pair-instability supernova (PISN). Rare channels of evolution are shown with dashed lines. The ‘faint supernovae’ have been proposed but not yet detected. (Reproduced from Smartt 2009, ARAA, 47, 63).

15.1, the lack of SNe from RSG progenitors with initial masses between 17 and $30 M_{\odot}$ was unexpected. And yet, with ~ 20 Type II-P SN progenitors now identified, this ‘red supergiant problem’ seems fairly well established (see Figure 16.7).

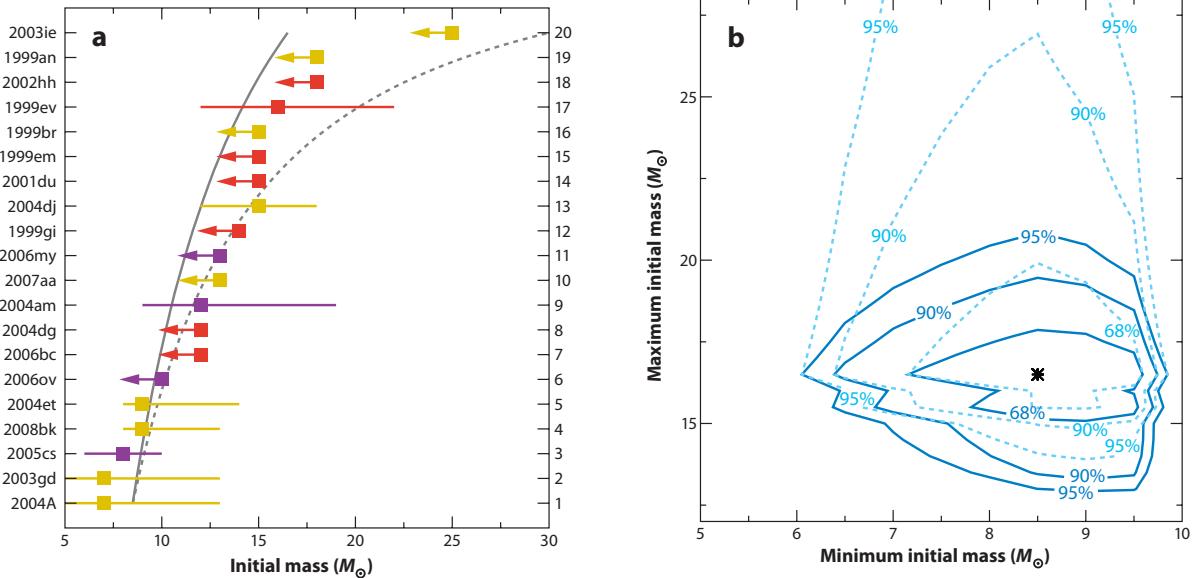


Figure 16.7: *Left:* Cumulative frequency plot of the masses of SN Type II-P progenitors, in order of increasing mass. The solid curve is a Salpeter Initial Mass Function with slope $\alpha = 2.35$ (Lecture 11.9) between $M_{\min} = 8.5M_{\odot}$ and $M_{\max} = 16.5M_{\odot}$, which gives the most likely fit to the data. The dotted line is also a Salpeter IMF, but with $M_{\max} = 30M_{\odot}$, showing that such massive progenitors should have been detected by now, if they gave rise to Type II SNe. The supernovae have been colour-coded according to the metallicity of the parent galaxy. *Right:* Maximum likelihood analysis of the Type II-P progenitor sample gives the most likely values for M_{\min} and M_{\max} (star symbol) and the likelihood contours, calculated for the detections only (dashed lines) and including the upper limits to the progenitor mass (solid lines). (Both figures reproduced from Smartt 2009, ARAA, 47, 63).

While several factors, including metallicity effects, variable IMF, dust obscuration, explosions as the rarer II-L, IIn, and Ib SNe could all be possible explanations, the most intriguing possibility is that stars more massive than $\sim 17M_{\odot}$ may collapse ‘quietly’ to form black holes and either very faint supernovae or no supernova at all.

Furthermore, it is almost certain that interacting binaries play an important role in determining the relative rates of types within SN populations. At the same time, there are indications that different channels can produce similar SNe, with examples of both an interacting binary (SN1993J) and a single WN star (SN2008ax) exploding as Types IIb—which are SNe whose initial spectra show some H, but then quickly evolve into Type Ib (Figure 16.2). However, so far there is no empirical confirmation that massive Wolf-Rayet stars produce the majority of Ib SNe in the local Universe. Undoubtedly, the schematic diagram in Figure 16.6 will become clearer as the work to identify SN progenitors continues.

16.4 Core Collapse

Despite the fact that different evolutionary paths can lead to a core-collapse supernova, the physical mechanism that produces the explosion is thought to be essentially similar for Type II, Ib and Ic SNe. As we saw in Lecture 15, by the end of its life a star with $M \gtrsim 11M_{\odot}$ has developed an iron core. When the iron core in the centre of a massive star grows, by silicon shell burning, to a mass around the Chandrasekhar mass limit of $\sim 1.44M_{\odot}$, electron degeneracy pressure can no longer stabilize the core. Unable to generate additional energy through nuclear fusion, the core collapses from an initial radius $R_{c,i} \sim 3000$ km to a final radius of only $R_{c,f} \sim 20$ km (from the size of the Earth to the size of Manchester!) on a very short timescale, of order of a fraction of a second. This starts what is called a core-collapse supernova, in the course of which the star explodes and parts of the star's heavy-element core and of its outer shells are ejected into the interstellar medium.

16.4.1 Energetics

Before considering further the details of core-collapse and the subsequent supernova explosion, let us calculate the energies involved. From the virial theorem, the gravitational energy released by the core collapse is:

$$U_{\text{gr}} = -\frac{3}{10} \frac{GM_c^2}{R_{c,i}} + \frac{3}{10} \frac{GM_c^2}{R_{c,f}} \simeq \frac{3}{10} \frac{GM_c^2}{R_{c,f}} \quad (16.1)$$

since $R_{c,i} \gg R_{c,f}$. With $R_{c,f} = 20$ km and $M_c = 1.4M_{\odot}$, we have:

$$U_{\text{gr}} = \frac{3}{10} \frac{6.7 \times 10^{-11} (1.4 \times 2 \times 10^{30})^2}{2 \times 10^4} \frac{\text{m}^3 \text{kg}^{-1} \text{s}^{-2} \text{kg}^2}{\text{m}} \quad (16.2)$$

$$U_{\text{gr}} = \frac{3}{10} \frac{6.7 \times 10^{-11} (2.8)^2 \times 10^{60}}{2 \times 10^4} \frac{\text{kg m}^2 \text{s}^{-2}}{\text{kg}} \quad (16.3)$$

$$U_{\text{gr}} = \frac{3}{10} \times 2.6 \times 10^{45} \sim 1 \times 10^{46} \text{ J} \sim 1 \times 10^{53} \text{ erg} \quad (16.4)$$

For comparison, the gravitational binding energy of the rest of the star

(the envelope) is:

$$U_{\text{env}} = -4\pi G \int_{R_{c,i}}^{R_{\text{env}}} M_r \rho r dr \quad (16.5)$$

where M_r is the mass internal to r , and the integration is carried out from the bottom of the envelope to the stellar radius (eq. 7.6). Integrating eq. 16.5 with an appropriate density distribution $\rho = f(r)$, yields $E_{\text{env}} \sim 10^{50}$ erg as the energy required to unbind the star.

As we have seen (section 16.2), a supernova explosion doesn't just unbind the star, but also ejects its outer layers into the surrounding interstellar medium at high speeds. Typical values for Type II SNe are $M_{\text{ej}} \sim 10M_{\odot}$ and $\langle v_{\text{ej}} \rangle \sim 3 \times 10^3 \text{ km s}^{-1}$. Thus,

$$E_{\text{ej}} \sim \frac{1}{2} 10 \times 2 \times 10^{30} \cdot (3 \times 10^6)^2 \text{ kg m}^2 \text{ s}^{-2} \quad (16.6)$$

or $E_{\text{ej}} \sim 10^{44} \text{ J} \equiv 10^{51} \text{ erg}$.

Finally, a Type II SN has a mean bolometric luminosity $\langle L_{\text{bol}} \rangle \sim 2 \times 10^8 L_{\odot}$ over a period of several months. Thus, the total energy lost as radiation is:

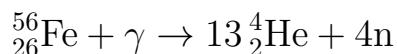
$$E_{\text{ph}} \sim 2 \times 10^8 \times 4 \times 10^{26} \cdot 100 \times 10^5 \text{ Js}^{-1} \text{ s} \quad (16.7)$$

or $E_{\text{ph}} \sim 10^{42} \text{ J} \equiv 10^{49} \text{ erg}$.

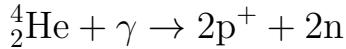
Summarising, $E_{\text{env}} + E_{\text{ej}} + E_{\text{ph}} \sim 10^{51} \simeq 0.01 U_{\text{gr}}$: only about 1% of the energy released in the final collapse of the iron core is used in the supernova event. However, the exact details of how this 1% of the gravitational energy is actually converted into the kinetic energy of the envelope—that is the actual mechanics of the explosion of a massive star—turns out to be a very difficult problem that has been the subject of intense theoretical activity in the modern computational era.

16.4.2 The Explosion Mechanism

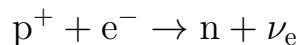
At the high temperatures required for Si burning, $T \gtrsim 3 \times 10^9 \text{ K}$, the photon energies are so high ($h\nu \sim 1 \text{ MeV}$) that photons can photodisintegrate iron nuclei, e.g.:



This process takes energy away from the radiation field and reduces the pressure supporting the core; the core contracts, increasing the local temperature further, until photons are sufficiently energetic to photodisintegrate He nuclei:



At this point, the core of the star is composed mainly of protons, neutrons and free electrons. As the core continues to collapse, the densities become so large that free electrons are captured by the protons produced by photodisintegration in an inverse β decay:



The removal of electrons from the plasma removes another source of pressure: the degeneracy pressure of free electrons. Without energy support the core collapses catastrophically on a timescale of milliseconds until the density in the inner core reaches values comparable to the density of matter inside atomic nuclei ($\rho_{\text{nuc}} \simeq 3 \times 10^{14} \text{ g cm}^{-3}$).

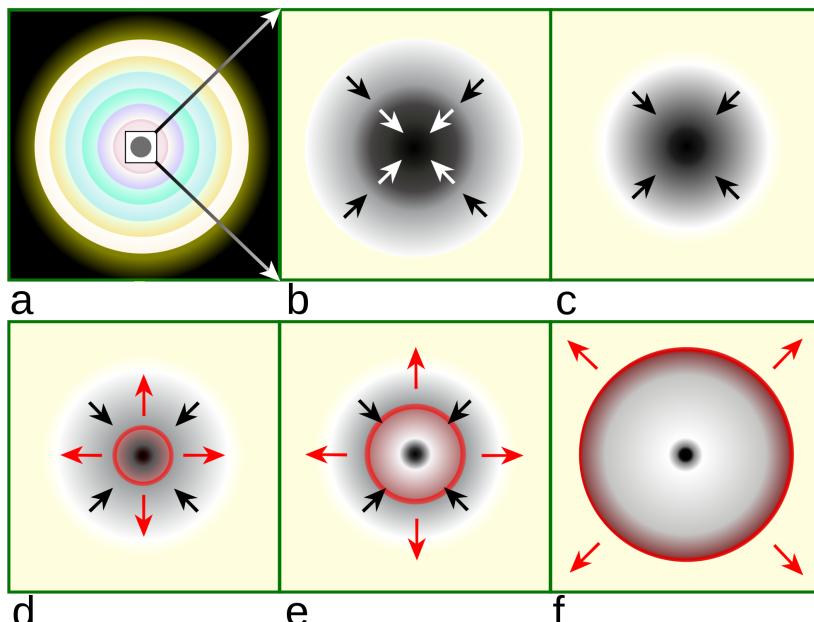


Figure 16.8: Within a massive, evolved star (a) the onion-layered shells of elements undergo fusion, forming an iron core. (b) Unable to generate energy by further fusion, the iron core starts to collapse. The inner part of the core is compressed into neutrons (c), causing infalling material to bounce (d) and form an outward-propagating shock front (red). The shock starts to stall (e), but it is re-invigorated by a process that may include neutrino interaction. The surrounding material is blasted away (f), leaving only a degenerate remnant.

The neutralisation of the plasma via the inverse β -decay is also thought to occur in stars with masses $8M_{\odot} \lesssim M \lesssim 11M_{\odot}$ which develop degenerate O-Ne cores, rather than Fe cores. The removal of the electron degeneracy pressure results in the core collapsing on the very short dynamical timescale, similar to the mechanism just described for stars with $M \gtrsim 11M_{\odot}$. This explains the lower limit $M_{\min} \simeq 8M_{\odot}$ for the progenitors of Type II SNe.

When the density in the inner core increases beyond nuclear densities, the neutrons (which are also fermions) become degenerate resulting in a strong increase in pressure which reverses the collapse. Almost instantaneously, the core becomes incompressible and 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 outer part of the core which is still in free-fall. This creates a shock wave that propagates into the outer layers of the star. 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, 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. Computer simulations indicate that, at this point, the shock stalls.

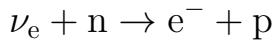
After the core bounce, a compact remnant begins to form at the centre of the collapsing star, rapidly growing by the accretion of infalling stellar material until the explosion sets in. This nascent remnant—the proto-neutron star—will evolve into a neutron star or may eventually collapse to

Stage	Core temperature (K)	Core density (kg/m^3)	Duration of stage
Hydrogen burning	4×10^7	5×10^3	7×10^6 years
Helium burning	2×10^8	7×10^5	7×10^5 years
Carbon burning	6×10^8	2×10^8	600 years
Neon burning	1.2×10^9	4×10^9	1 year
Oxygen burning	1.5×10^9	10^{10}	6 months
Silicon burning	2.7×10^9	3×10^{10}	1 day
Core collapse	5.4×10^9	3×10^{12}	$\frac{1}{4}$ second
Core bounce	2.3×10^{10}	4×10^{15}	milliseconds
Explosive (supernova)	about 10^9	varies	10 seconds

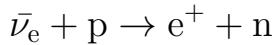
Figure 16.9: Evolutionary stages of a $25M_{\odot}$ star.

a black hole, depending on whether the progenitor star had a mass below or above $\sim 25M_\odot$.

Much computational work has focussed on the so-called ‘neutrino-driven’ explosion (see Figure 16.10). Neutrinos are trapped within the proto-neutron star because the density of matter is so high that their mean free path is significantly shorter than the radius of the neutron star and they take a fraction of a second to diffuse out. In the explosion scenario by the ‘delayed neutrino-heating mechanism’, the stalled shock wave can be revived by neutrinos depositing some of their energy in the layers between the nascent neutron star and the stalled shock front, mainly by charged-current ν_e and $\bar{\nu}_e$ captures on free nucleons:



and



This neutrino heating increases the pressure behind the shock and the heated layers begin to expand, creating between the shock front and the neutron star surface a region of low density but rather high temperature, the so-called ‘hot bubble’. The persistent energy input by neutrinos keeps the pressure high in this region and drives the shock outwards again, eventually leading to a supernova explosion (bottom panels of Figure 16.10). This may take a few 100 ms and requires that during this time interval a few percent of the neutrino energy is converted into thermal energy of nucleons, leptons and photons.

Other recent work has proposed a novel idea: the excitation of core pulsation modes in the deep interior of the proto-NS. The sound pulses radiated from the core steepen into shock waves that merge as they propagate into the outer mantle, depositing their energy and momentum. The ultimate source of the acoustic power is (as in other scenarios) the gravitational energy of infall—core oscillation acts like a transducer to convert this accretion energy into sound. An advantage of the acoustic mechanism is that acoustic power does not abate until accretion subsides, so that it is available as long as it may be needed to explode the star.

It may well be that some supernovae explode by the neutrino mechanism, while others, if the neutrino mechanism fizzles, explode by the acoustic mechanism. Or maybe neither works! Whatever the mechanism, it is clear

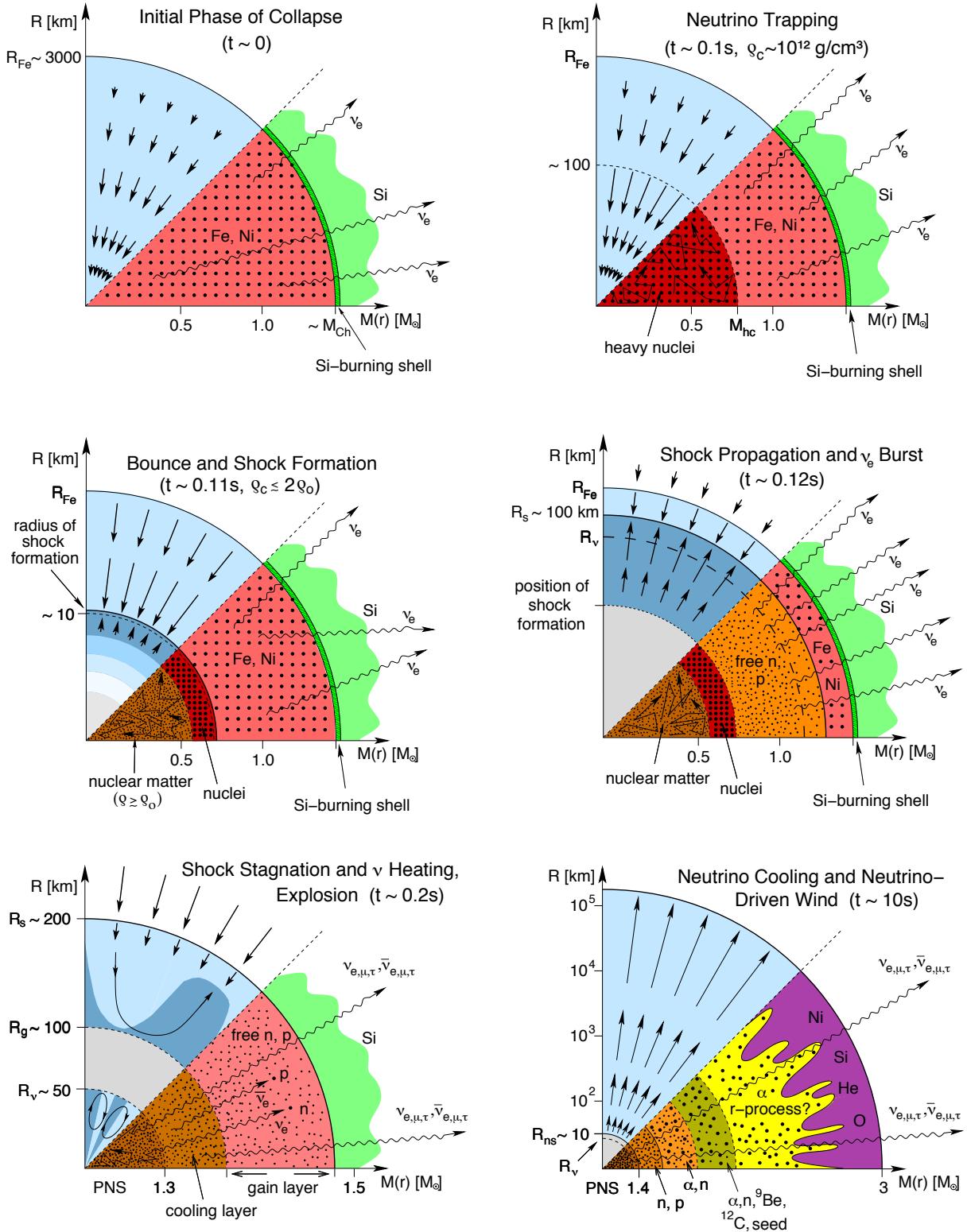


Figure 16.10: Computer simulations of the ‘delayed neutrino-heating mechanism’ driving Type II supernova explosions. The panels display the dynamical conditions in their upper half, with arrows representing velocity vectors. The nuclear composition, as well as the nuclear and weak processes, are indicated in the lower half of each panel. The horizontal axis gives mass information. (Figure reproduced from Janka et al. 2007, Physics Reports, 442, 38).

that massive stars do explode as SNe. In computer simulations of Type II SN explosions, our lack of knowledge of the explosion mechanism is circumvented by introducing by hand a ‘piston’—an injection of a specified amount of energy that drives the explosion—so that its consequences can be followed numerically. Both E_{kin} and the mass boundary between core and envelope (usually referred to as the ‘mass cut’) are uncertain and are usually treated as free parameters. While ultimately disappointing, this type of ‘fudge’ is not uncommon in numerical simulations of astrophysical processes.

16.5 Explosive Nucleosynthesis

The photodisintegration and inverse β -decay reactions which take place during core collapse and the shock expansion following core bounce all produce free neutrons. Neutrons do not experience a Coulomb barrier, so that they can be incorporated more easily than charged particles into Fe-peak nuclei to form heavier elements.

Neutron-capture nucleosynthesis was already discussed in Lecture 7.5. The process of interest here is the ‘rapid neutron capture’, or r-process, whereby a nucleus can absorb several neutrons quickly before decaying. As the

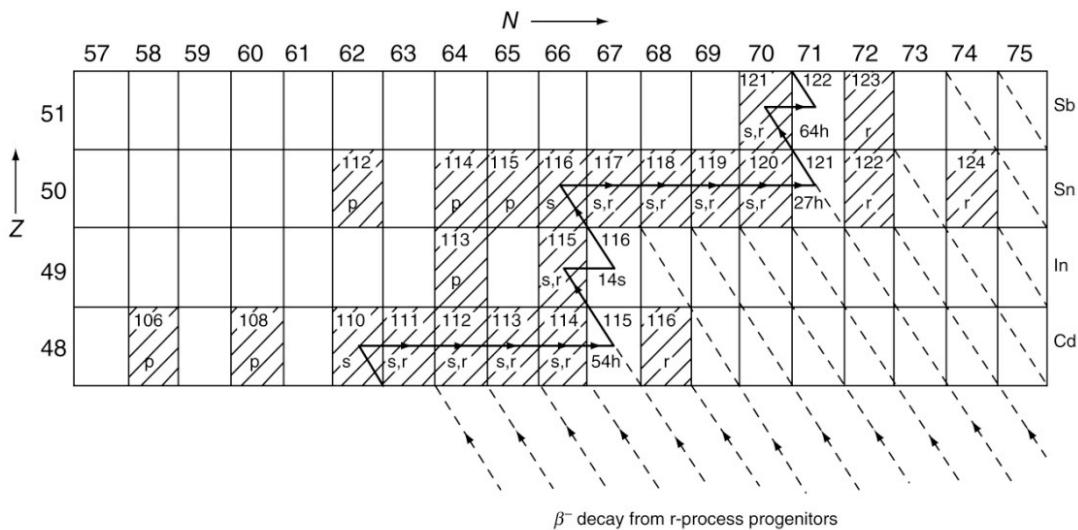


Figure 16.11: Nucleosynthesis of the elements Cd through to Sb. Stable isotopes are shown hatched. The solid line shows the path of the ‘slow neutron capture’ process. (Figure reproduced from Pearson, J.M. 1986, Nuclear Physics: Energy and Matter, Adam Hilger, Bristol).

number of neutrons increases, the half-life of the isotope decreases until it is shorter than the interval between successive neutron captures, at which point the neutron-rich isotope will undergo β -decay.

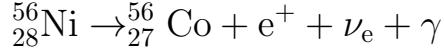
For example, referring to Figure 16.11, if $^{115}_{49}\text{In}$ nuclei absorb seven neutrons on timescales smaller than the half-lives of the various isotopes along the way, $^{122}_{49}\text{In}$ can be synthesised. If its half-life is short compared to the interval between successive neutron captures, then $^{122}_{49}\text{In}$ will decay into $^{122}_{50}\text{Sn}$ which, by definition, is then an r-process element. From Figure 16.11 it can be seen that some isotopes can be synthesised by both the s- and r-process.

16.5.1 Light Curves of Core Collapse Supernovae

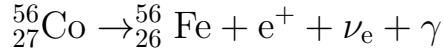
The visible supernova explosion starts when the shock wave reaches the stellar surface, giving rise to a short pulse (~ 30 minutes) of soft X-ray emission. The luminosity then declines rapidly as the stellar surface expands and cools. When a massive H-rich envelope is present, the gas which was ionised by the shock enters a stage of prolonged recombination, releasing energy at a nearly constant temperature of ~ 5000 K; the photons emitted as the recombining electrons cascade down the H energy levels produce the plateau in the light curves of Type II-P SNe. In cases where the H-rich envelope is not sufficiently massive, the plateau phase is absent (Type II-L light curves).

The overall light curve on which the plateau is superposed in Type II-P SNe is due to the radioactive decay of isotopes produced by the shock front as it travels outwards through the star. As the shock wave moves through the Si layer, it raises the temperature to $T \sim 5 \times 10^9$ K; nuclear statistical equilibrium is reached in a few seconds, and explosive nucleosynthesis of Si takes place. As was the case with the slower hydrostatic burning stages prior to core collapse, Fe-group elements are produced, but the main product is $^{56}_{28}\text{Ni}$ rather than $^{56}_{26}\text{Fe}$. The reason is that the timescale is too short for β -decays to take place and change the ratio of protons to neutrons. Since the ‘fuel’, $^{28}_{14}\text{Si}$, has $Z/A = 1/2$, (most of) the products also have $Z/A = 1/2$, hence $^{56}_{28}\text{Ni}$ instead of $^{56}_{26}\text{Fe}$. As the shock waves move out and the temperature drops below $\sim 10^9$ K (around the O, Ne layer) explosive nucleosynthesis stops.

The most abundant isotope produced by explosive nucleosynthesis, $^{56}_{28}\text{Ni}$, is not stable but decays to $^{56}_{27}\text{Co}$ via the β^+ decay reaction:



with a half-life $\tau_{1/2} = 6.1$ days.¹ $^{56}_{27}\text{Co}$ subsequently decays to stable $^{56}_{26}\text{Fe}$ by a similar reaction:



with the longer half-life $\tau_{1/2} = 77.7$ days. The high-energy photons produced by these decays power the emission from the expanding SN remnant.

Since radioactive decay is a statistical process, the rate of decay must be proportional to the number of atoms remaining in a sample (in this case the supernova remnant):

$$\frac{dN}{dt} = -\lambda N, \quad (16.8)$$

where λ is a constant. Integrating 16.8, we find:

$$N(t) = N_0 e^{-\lambda t}, \quad (16.9)$$

where N_0 is the original number of radioactive atoms in the sample and $\lambda = \ln 2/\tau_{1/2}$.

When radioactive decay is powering the SN light curve, the rate at which energy is deposited into the supernova remnant is proportional to the decay rate dN/dt . Thus, we have:

$$\frac{d \log_{10} L}{dt} = -\log_{10} e \cdot \lambda = -0.434\lambda \quad (16.10)$$

where L is the bolometric luminosity or, in terms of magnitudes:

$$\frac{d M_{\text{bol}}}{dt} = 1.086\lambda \quad (16.11)$$

In a typical Type II-P SN, the exponential decline of the light curve after 50–100 days reflects the decay of $^{56}_{27}\text{Co}$ to $^{56}_{26}\text{Fe}$ and implies that $\sim 0.07M_\odot$ of $^{56}_{28}\text{Ni}$ was ejected in the explosion. This puts constraints on the position of the ‘mass cut’ between the collapsing core and the ejected envelope (the

¹Recall the definition of half-life: if we take a sample of radioactive material initially consisting of X_0 atoms of isotope X , there is a 50% chance that any atom in the sample will decay to isotope Y over a time interval $\tau_{1/2}$. Thus, after a time $n \times \tau_{1/2}$ the fraction of isotope X remaining is $X/X_0 = 2^{-n}$.

remainder of the synthesised ^{56}Ni is locked up in the collapsed compact object). This mass cut is a key parameter in computer calculations of the nucleosynthetic yields of supernovae. The light curves of Type Ib and Ic supernovae are completely dominated by the $^{56}_{28}\text{Ni} \rightarrow ^{56}_{27}\text{Co} \rightarrow ^{56}_{26}\text{Fe}$ radioactive decay after the initial luminosity peak caused by shock breakout.

It is expected that the explosive nucleosynthesis of the supernova shock also produces significant amounts of other radioactive isotopes, such as $^{57}_{27}\text{Co}$, with a half-life $\tau_{1/2} = 271$ days; $^{22}_{11}\text{Na}$ ($\tau_{1/2} = 2.6$ years) and $^{44}_{22}\text{Ti}$ ($\tau_{1/2} \simeq 47$ years). If these isotopes are present in sufficient quantities, each in turn may contribute to the overall light curve, causing the slope of the curve to change with time (see Figure 16.12).

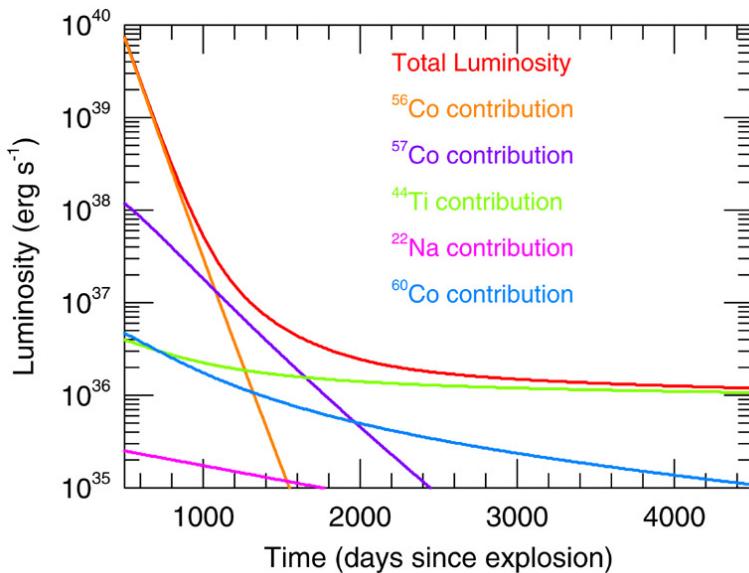


Figure 16.12: Radioactive decays contributing to supernova light curves. (Reproduced from <http://cococubed.asu.edu>, Frank Timmes's webpages at Arizona State University).

16.6 SN1987A in the Large Magellanic Cloud

The closest supernova in modern times was SN1987A in the Large Magellanic Cloud, our companion galaxy at a distance of 50 kpc. Being so close, this event was studied (and continues to be so) in more detail and with higher precision than any other supernova. The huge amount of data acquired on this object confirmed some of the theoretical ideas on core-collapse SNe but also improved our understanding of such events by presenting us with a number of surprises.

16.6.1 Neutrinos

Arguably the most exciting observation of SN1987A was the detection of its neutrinos—the first time that neutrinos had been detected from an astronomical source other than the Sun, validating the basic picture of the formation of a neutron star from a collapsed iron core.

A total of 20 electron anti-neutrinos were detected simultaneously in the span of 12.5 seconds by two underground Cerenkov experiments (designed to search for proton decay), one in Japan and the other in the USA. The typical energies of the $\bar{\nu}_e$ were ~ 20 MeV. Neutrinos hardly interact with ordinary matter. Correcting for the very low detection efficiency, implies that a time-integrated flux, or fluence, of $2 \times 10^9 \text{ cm}^{-2}$ $\bar{\nu}_e$ particles had reached the Earth. Further correcting for the other five types of neutrinos (ν_e , ν_μ , $\bar{\nu}_\mu$, ν_τ and $\bar{\nu}_\tau$) and for the surface of a sphere of radius 50 kpc, we find that the total energy of the neutrinos emitted by SN1987A was:

$$\begin{aligned} E_\nu &\simeq 2 \times 10^9 \text{ cm}^{-2} \times 6 \times (20 \text{ Mev} \times 1.6 \times 10^{-6} \text{ erg MeV}^{-1}) \\ &\quad \times 4\pi (5 \times 10^4 \text{ pc} \times 3.1 \times 10^{18} \text{ cm pc}^{-1})^2 \\ &\simeq 10^{53} \text{ erg} \end{aligned}$$

close to the total energy expected from the collapse of a stellar core with the parameters given in Section 16.4.1.

The neutrinos from SN1987A were detected on February 23.31 UT, three hours *before* the arrival of the first photons from the optical event. The neutrinos began their trip to Earth ahead of the photons, presumably because the exploding star became optically thin to neutrinos before the shock wave reached the surface. To keep ahead of the photons over the 163 000 light-years distance from the LMC to Earth, the neutrinos must have travelled close to the speed of light. This conclusion, together with lack of significant dispersion in the arrival times of neutrinos of different energies, has been used to place an upper limit to the mass of the electron neutrino $m_{\nu_e} \leq 16$ eV, consistent with the upper limit of 2.2 eV from the most sensitive laboratory experiments. Here is a good example of an interface between astronomy and particle physics.

16.6.2 The Progenitor of SN1987A

With so many images of the LMC obtained at many ground-based observatories over the years, it was relatively straightforward to identify the progenitor on pre-explosion photographic plates as a 12th-mag star previously catalogued as Sk –69 202 (see Figure 16.13).

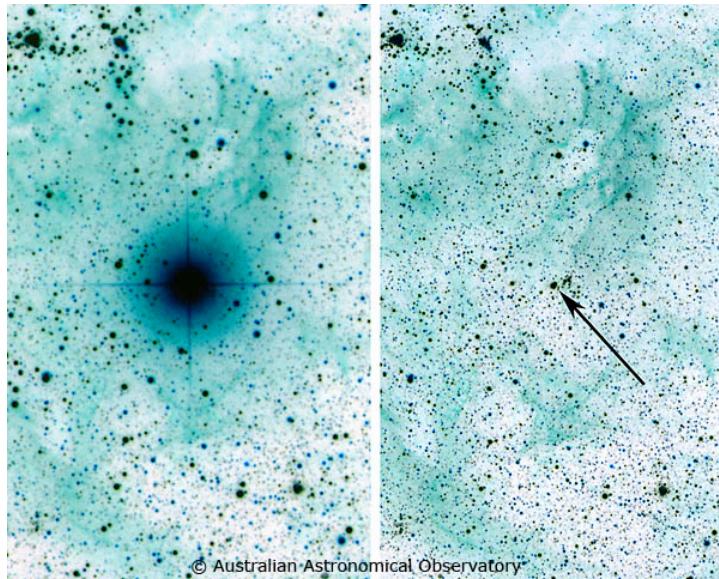


Figure 16.13: SN1987A in the Large Magellanic Cloud. (Photos by David Malin, Anglo-Australian Observatory).

Surprisingly, Sk –69 202 was a *blue*, rather than red, supergiant, classified as B3 I, with $L \approx 1.1 \times 10^5 L_\odot$, $T_{\text{eff}} \approx 16\,000\,\text{K}$ and probable initial mass $M_{\text{ZAMS}} \approx 18M_\odot$. This quickly led people to reconsider models of stellar evolution for massive stars and to realise the perhaps previously unappreciated importance of metallicity in shaping their evolution. The LMC has a present-day metallicity $Z_{\text{LMC}} \approx 1/2Z_\odot$. Even a modest factor of two drop in metallicity has several consequences: it results in an extended ‘blue loop’ in the path in the H-R diagram taken by an evolving massive star (see Figure 16.14); it reduces the amount of mass loss through the stellar wind (as we saw in Lecture 15); and the star has a hotter photosphere.

16.6.3 The Light Curve

The right panel of Figure 16.14 shows the light curve of SN1987A from photometric observations carried out during four years after the explosion.

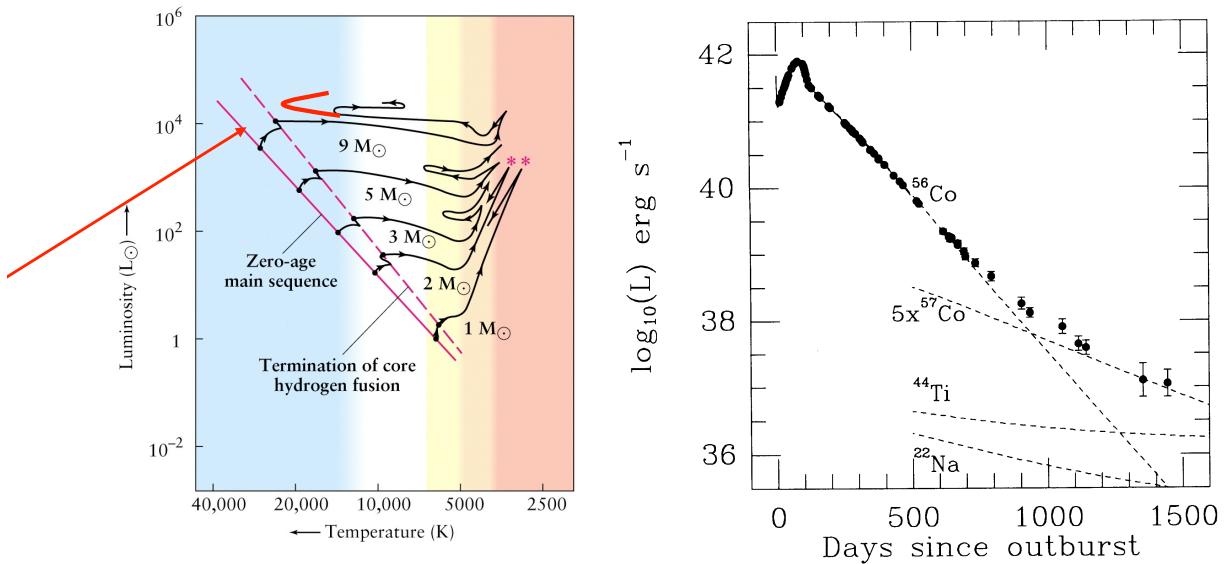


Figure 16.14: *Left:* At low metallicities, massive stars exhibit more extended ‘blue loops’ in the upper regions of the H-R diagram. This probably explains why the progenitor of SN1987A exploded when it was a blue, rather than red, supergiant. *Right:* Bolometric light curve of SN1987A through the first 1444 days after the explosion. The dashed lines show the contributions from radioactive isotopes produced by explosive nucleosynthesis. (Figure reproduced from Carroll & Ostlie, *Modern Astrophysics*; original data from Suntzeff et al. 1992, ApJ, 384, L33).

Not immediately evident from this figure is the subluminous nature of this supernova—whereas typical Type II SNe reach $M_{\text{bol}} = -18$, SN1987A peaked at $M_{\text{bol}} = -15.5$, ten times fainter. Also, its rise to maximum light was much slower than normal, the rising portion of the light curve lasting for 80 days.

These ‘anomalies’ go hand-in-hand with the fact that the star exploded as a blue supergiant, much smaller and denser than a red supergiant. Consequently, a larger fraction of the thermal energy released in the collapse was converted into the mechanical energy required to lift the envelope of the star out of the deeper potential well of a blue supergiant.

The proximity of SN1987A allowed us to follow its light curve for longer, and more accurately, than is normally possible for other supernovae. The change in slope as the decay of ^{57}Co takes over from ^{56}Co near day 1000 is evident from Figure 16.14. An independent confirmation of our ideas of radioactive powering of the light curve was the direct detection in late 1987 of the γ -ray emission lines at 847 keV and 1238 keV that accompany the decay of ^{56}Co to ^{56}Fe . These emission lines were recognised in accumulated spectra recorded by the NASA Solar Maximum Mission satellite (Matz et al. 1988, Nature, 331, 416). These observations showed that

about $0.075M_{\odot}$ of $^{56}_{28}\text{Ni}$, the parent nucleus of $^{56}_{27}\text{Co}$, was formed in the explosion. Later, the 122 keV emission line from $^{57}_{27}\text{Co}$ was observed in 1991 by the *Compton Gamma Ray Observatory*, and $0.003M_{\odot}$ of this nuclide was inferred to have been synthesised.

16.6.4 The Supernova Remnant

While the progenitor of SN1987A experienced a reduced amount of mass loss compared with a more metal-rich star of the same mass, we have direct confirmation of pre-explosion episodes of mass loss in the appearance of the supernova remnant, recorded with an exquisite series of images by the *Hubble Space Telescope* (Figure 16.15).

The radioactive debris of SN1987A, expanding with velocities of up to several thousand km s^{-1} , is at the centre of a triple ring system; this is circumstellar gas which was probably dark before the SN and was heated and ionized by the flash of radiation at the time of shock breakout. The inner ring exhibits an enhanced abundance of nitrogen (by a factor of 10 compared with the interstellar abundance of N in the LMC) implying that it has been through the CNO nuclear burning process, and is expanding with a velocity $v_{\text{exp}} = 10 \text{ km s}^{-1}$; both findings strongly point to an origin

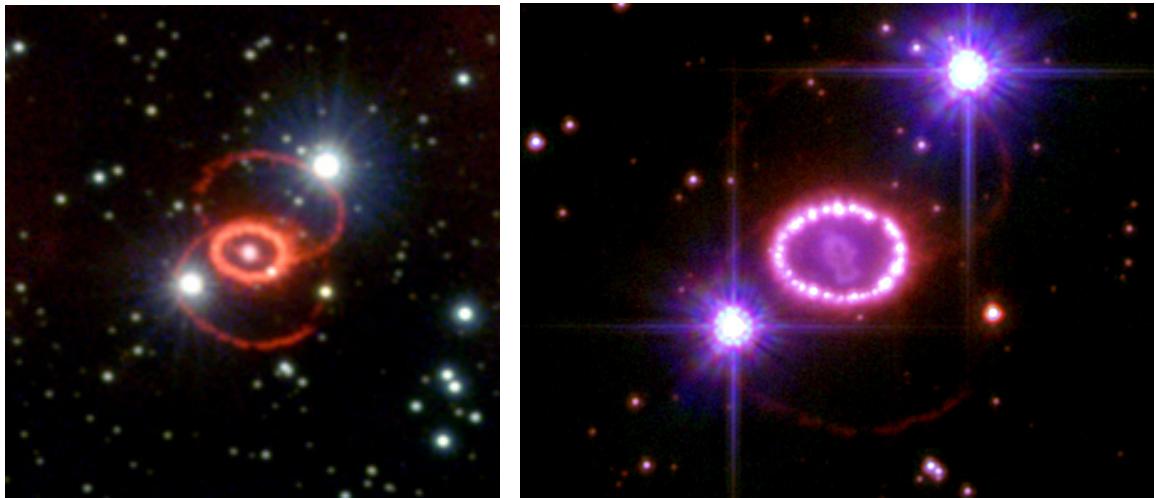


Figure 16.15: Images of SN1987A and its immediate surroundings recorded with the *Hubble Space Telescope* on 6 February 1996 (*Left*) and 6 December 2006 (*Right*). The inner ring has been brightening in clumps as the shock front from the supernova began colliding with the slower moving material ejected by the SN progenitor during its red supergiant phase.

in a mass loss event during an earlier red supergiant phase. The radius of the inner ring (which is circular but appears elliptical because it is inclined to our line of sight) is ~ 0.2 pc; combined with its velocity, it implies that this gas was ejected $\sim 20,000$ years before the SN explosion.

Models of the ring structure require the density of the slow wind in the equatorial plane be 20 times the density of the wind in the polar direction. The reason for such an aspherical, but axisymmetric, wind remains unclear, although similar structures are inferred to exist in planetary nebulae. One possibility is that a binary companion was involved in the equatorial mass loss. There is no direct evidence for a binary companion, although such a model may be attractive on other grounds, including mixing of the progenitor- star envelope and even the blueward evolution of the star before the explosion.

The shock front from the expanding supernova remnant began to collide with the slower moving stellar wind comprising the inner ring in 1996. The result was a brilliant display of bright clumps in the inner ring that developed over the next several years (right panel of Figure 16.15). The two larger rings are not in planes containing the central explosion, but lie in front of and behind the star. Their origin is still unclear. While several models have been put forward, involving a binary companion and/or a fast blue supergiant wind, they generally seem somewhat contrived and there is no general consensus as to the interpretation of these enigmatic features.

16.6.5 The Light Echo of SN1987A

In 1939, the French astronomer Paul Couderc developed the basic mathematical description of light echoes and applied it to earlier observations of Nova Persei 1901. Soon after the explosion, Roger Chevalier and others predicted that SN1987A should give rise to very conspicuous light echoes, first detected a year after the explosion and monitored over several years, as shown in Figure 16.16.

The formation of the light echoes is illustrated in Figure 16.17. At any one time after the explosion, there is an ellipsoidal surface of equal delay in light-travel time. If mirrors were placed on such a surface, we would see reflected light from the supernova. In an astronomical situation, the

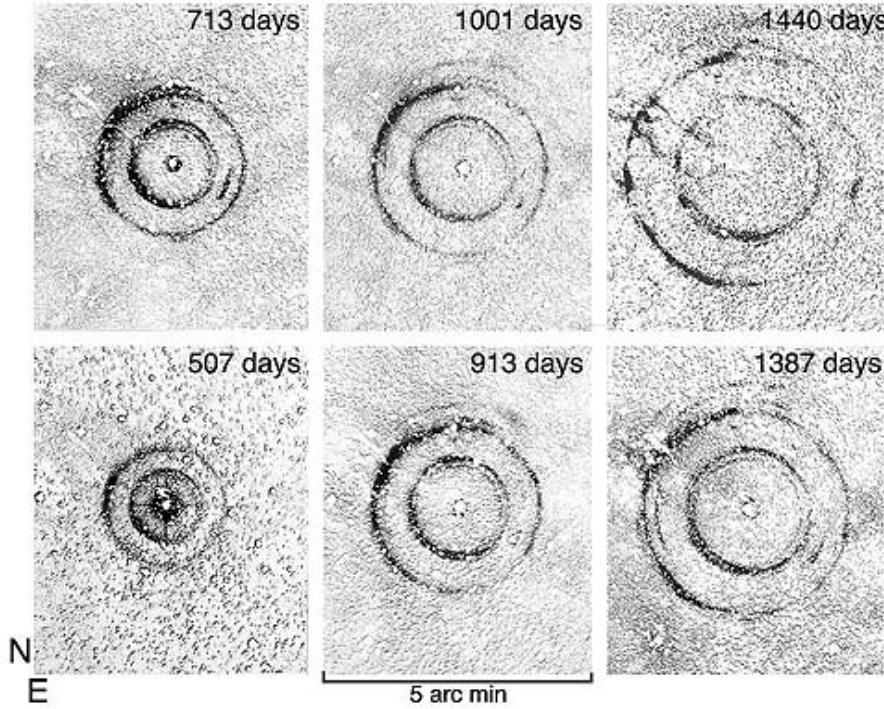


Figure 16.16: Scattered light echoes from SN1987A photographed with the Anglo-Australian Telescope.

'mirrors' are dust particles which scatter the incident starlight; if the dust is located in planar sheets, we see rings where the sheets intersect the light-travel time delay ellipse. By modelling the rings produced by SN1987A and their evolution, Chevalier and collaborators showed that they occur in sheets of dust (and gas) located ~ 120 and ~ 330 pc in front of the supernova, and that the dust producing the light echoes is comparable to the total column density of dust in front of the stars in this region

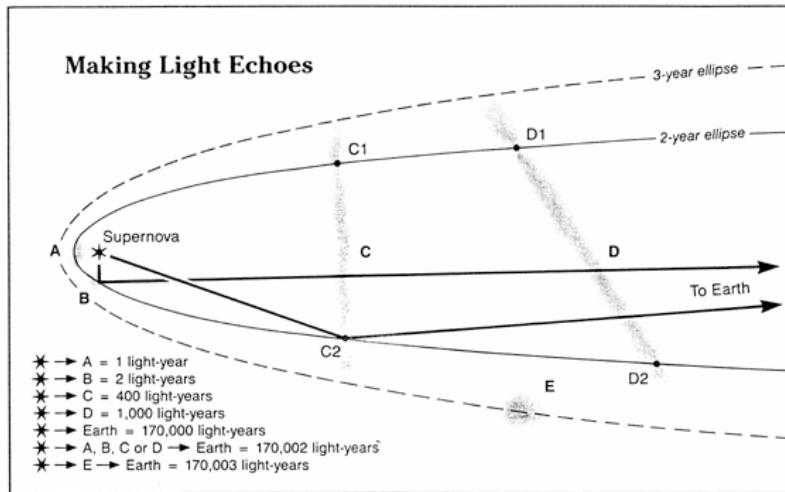


Figure 16.17: Schematic diagram of light echoes from SN1987A (not to scale).

of the LMC. The two main concentrations of interstellar material in this general direction are also seen in the absorption lines they produce in the spectra of background stars. The changing detail of the light echoes is a vivid demonstration of the small-scale structure if the interstellar medium, which had perhaps not been fully appreciated before.

16.7 Gamma-ray Bursts

We conclude this lecture with a brief mention of γ -ray bursts (GRBs) which, at least in some cases, appear to be associated with the core-collapse of massive stars.

About once a day, at some random location in the sky, a shower of γ -ray photons with energies between 1 keV and many GeV appears. The bursts last from 1/100 to 1000s of seconds, have rise times as fast as 10^{-4} seconds followed by an exponential decay. After 25 years of controversy, GRB events were finally identified as being extragalactic. This realisation also implied that GRBs are by far the most energetic astrophysical sources known; even allowing for the fact that the radiation is highly beamed, at peak GRBs can be three orders of magnitude more luminous than the most luminous QSOs and 10^7 times brighter than the most luminous SNe!

After thousands of catalogued events, it is clear that there are two basic classes of GRBs: events that last less than 2 s are short-hard events, while those that last longer than 2 s are long-soft GRBs (the ‘hard’ and ‘soft’ labels refer to the spectral energy distribution—a hard spectrum source has more energy at higher energies and shorter wavelengths). Short bursts are thought to be associated with mergers of compact objects—neutron star-neutron star or neutron star-black hole mergers. Long-duration GRBs are thought to accompany the core collapse of very massive stars under the ‘right’ circumstances.

In the ‘collapsar’ model developed by Stan Woosley and collaborators at the University of California, Santa Cruz, a progenitor star with sufficiently high mass (possibly a Wolf-Rayet star) collapses to form a black hole. If the star is rotating, matter would fall freely along the rotation axis onto the black hole on timescales of seconds; however, along the equatorial plane infalling matter would be held up in an accretion disk by a centripetal

force. Detailed numerical simulations indicate that after a few seconds the polar regions become sufficiently evacuated of matter to allow the deposited gravitational energy to begin to propagate outwards, forming a relativistic and highly collimated jet. The jet is thought to carry enough energy that, even if only a small fraction couples to the matter in the jet boundary, it will be enough to explode the star.

Empirical confirmation of this scenario was provided by the detection in 1998 of GRB 980425 phenomenologically connected to SN1998bw, a hyper-luminous Type Ib SN with very broad spectral lines (a Type Ib-BL SN), indicative of outflows with speeds $v_{\text{ej}} \gtrsim 0.1c$. As of July 2010, three more spectroscopically confirmed Type Ib SNe associated with GRBs have been reported; several other events have shown photometric evidence for a SN counterpart.

Most GRB events occur at redshifts which are too high to discern an associated supernova (which has to be separated from the light of the GRB afterglow and of the host galaxy). However, there are also a few cases, at lower redshifts, where a conventional SN should have been detected—had it occurred—but it wasn’t. Perhaps these are the ‘failed SNe’ of Woosley’s original (i.e. pre-GRB) collapsar model, although opinions are still divided.

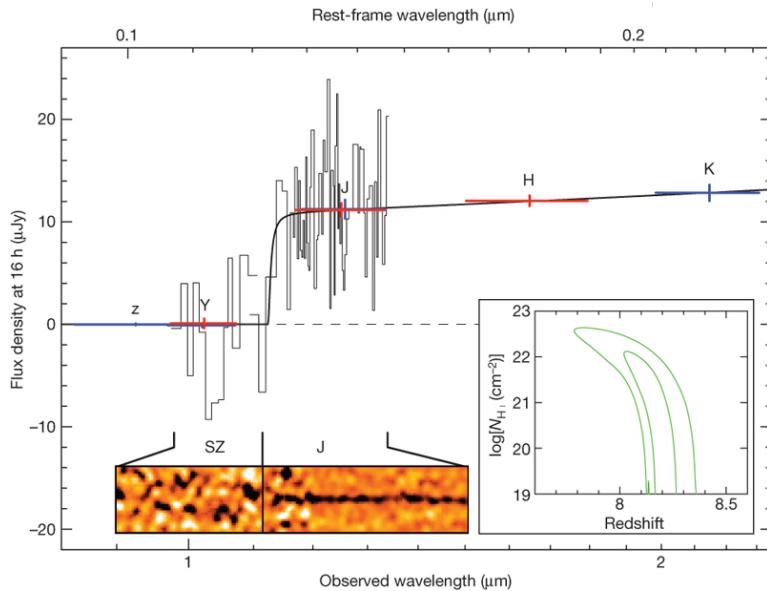


Figure 16.18: Near-IR spectrum and photometry of GRB 090423 at $z \approx 8.2$, reproduced from Tanvir et al. 2009, Nature, 461, 1254. The inset shows the 1σ and 2σ confidence contours on the correlated values of redshift and absorbing H I column density in the host galaxy.

The extreme brightness of GRBs, albeit for only brief periods, makes them visible out to cosmological distances. Indeed, one of the highest redshift objects known is the host galaxy of GRB 090423 at $z \simeq 8.2$, only 600 Myr ($\sim 5\%$ of the present age of the Universe) after the Big Bang (see Figure 16.18). Thus, GRBs can be some of the best probes of the earliest star formation episodes in the Universe. Furthermore, they act as bright background light sources in whose spectrum we can see the absorption produced by intervening material in galaxies and in the intergalactic medium over very long sight-lines (see Figure 16.19). With the technique of absorption-line spectroscopy, astronomers have used GRBs to study the properties of the Universe at different stages of its evolution, from its infancy to the present day.

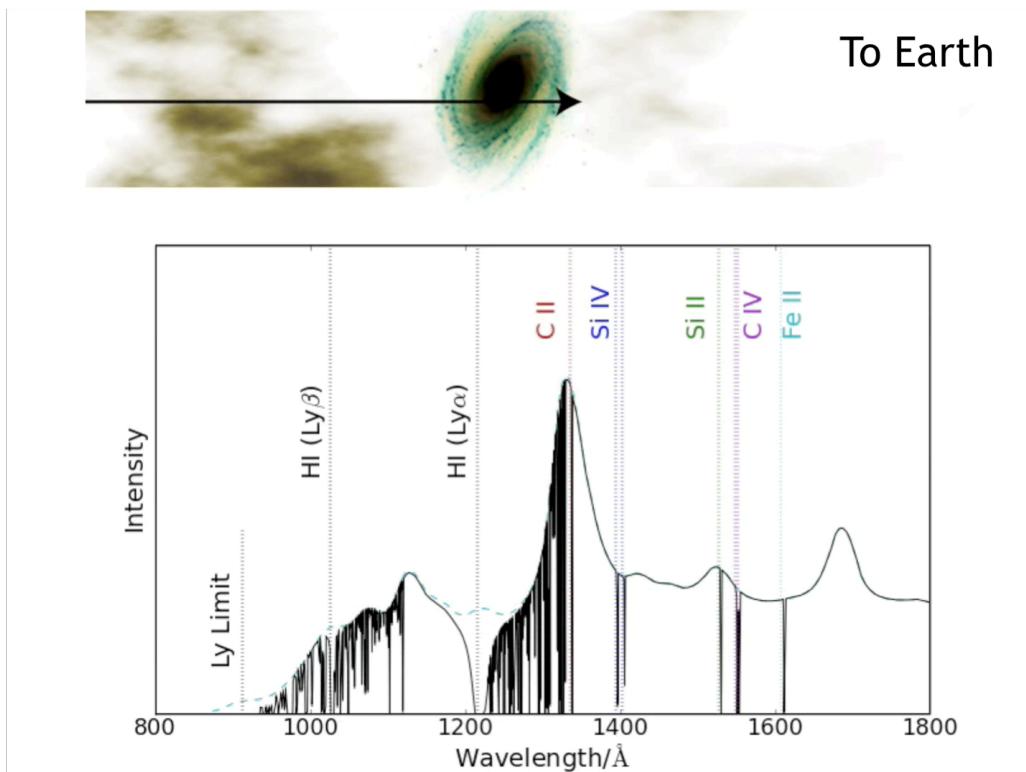


Figure 16.19: Absorption line spectroscopy of bright GRB afterglows can provide detailed physical information on intervening gas in and between galaxies fortuitously located along the line of sight. (Figure courtesy of Andrew Pontzen).

BINARY STARS and TYPE Ia SUPERNOVAE

18.1 Close Binary Star Systems

It is thought that about half of the stars in the sky are in multiple systems, consisting of two or more stars in orbit about the common centre of mass. In most of these systems, the stars are sufficiently far apart that they have little impact on one another, and evolve independently of one another, except for the fact that they are bound to each other by gravity.

In this lecture we will consider close binary systems, where the distance separating the stars is comparable to their size. In this situation, the outer layers of the stars can become deformed by the gravity of the companion. Under the right circumstances, matter can be transferred from one star to the other with far-reaching consequences for the evolution of the two stars. We begin by considering how gravity operates in a close binary system.

18.1.1 Lagrangian Points and Equipotential Surfaces

Consider two stars in circular orbits about their common centre of mass in the $x-y$ plane with angular velocity $\omega = v_1/r_1 = v_2/r_2$, where v is the orbital speed and r the distance from the centre of mass of the system. When considering such a system, it is convenient to work in a corotating coordinate system with the centre of mass at the origin (Figure 18.1), and the mutual gravitational attraction between the two stars balanced by the outward push of a centrifugal force.

In this frame of reference, the centrifugal force vector on a mass m at a distance r from the origin is:

$$\mathbf{F}_c = m\omega^2 r \hat{\mathbf{r}} \quad (18.1)$$

in the outward radial direction. When considering the potential energy of the system in a corotating coordinate system, we add to the gravitational

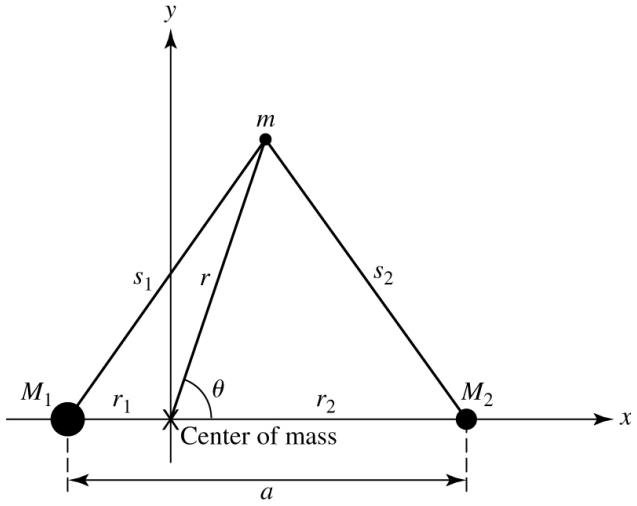


Figure 18.1: Corotating coordinates for a binary star system. Note that $a = r_1 + r_2$ and $M_1 r_1 = M_2 r_2$. In the examples considered in the text, $M_1 = 0.85M_\odot$, $M_2 = 0.17M_\odot$ and $a = 0.718R_\odot$.

potential energy:

$$U_g = -G \frac{Mm}{r} \quad (18.2)$$

a “centrifugal potential energy”

$$U_c = -\frac{1}{2}m\omega^2 r^2 \quad (18.3)$$

obtained by integrating eq. 18.1 with the boundary condition $U_c = 0$ at $r = 0$. Including the centrifugal term, the effective potential energy for a small test mass m located in the plane of the orbit is:

$$U = -G \left(\frac{M_1 m}{s_1} + \frac{M_2 m}{s_2} \right) - \frac{1}{2}m\omega^2 r^2. \quad (18.4)$$

Dividing by m , we obtain the effective potential energy per unit mass, or the **effective gravitational potential**, Φ :

$$\Phi = -G \left(\frac{M_1}{s_1} + \frac{M_2}{s_2} \right) - \frac{1}{2}\omega^2 r^2. \quad (18.5)$$

Returning to Figure 18.1, we have (from the law of cosines):

$$s_1^2 = r_1^2 + r^2 + 2r_1 r \cos \theta \quad (18.6)$$

$$s_2^2 = r_2^2 + r^2 - 2r_2 r \cos \theta. \quad (18.7)$$

The angular frequency of the orbit, ω , is given by Kepler’s third law:

$$\omega^2 = \left(\frac{2\pi}{P} \right)^2 = \frac{G(M_1 + M_2)}{a^3}. \quad (18.8)$$

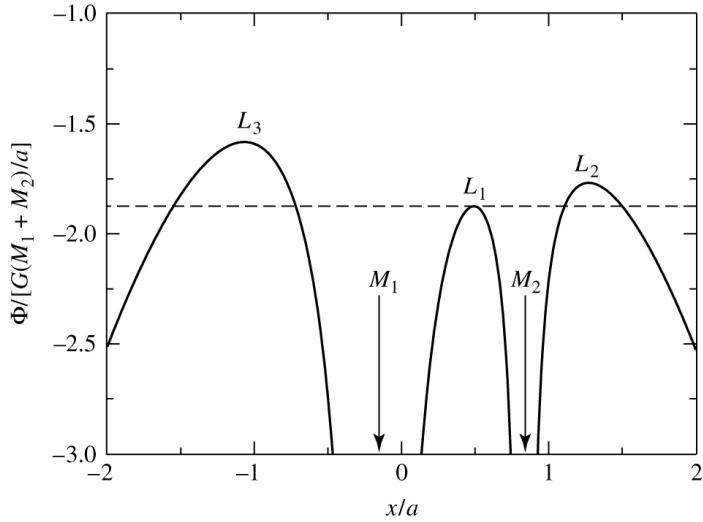


Figure 18.2: The effective gravitational potential Φ along the x -axis for two stars with masses $M_1 = 0.85M_\odot$ and $M_2 = 0.17M_\odot$. The stars are separated by a distance $a = 0.718R_\odot$ (see Figure 18.1). The dashed line is the value of Φ at the inner Lagrangian point. If the total energy per unit mass of a particle exceeds this value of Φ , it can flow through the inner Lagrangian point between the two stars.

The last four equations can be used to evaluate the effective gravitational potential Φ at every point in the orbital plane of a binary star system. For example, Figure 18.2 shows the value of Φ along the x -axis. The significance of this graph becomes clear when we consider the x -component of the force on a small test mass m , initially at rest on the x -axis:

$$F_x = -\frac{dU}{dx} = -m \frac{d\Phi}{dx} \quad (18.9)$$

At the values of x/a labelled L_n , $d\Phi/dx = 0$, and therefore there is no net force on the test mass: the gravitational pull exerted on m by M_1 and M_2 is just balanced by the centrifugal force of the rotating reference frame. These are the **Lagrangian points**. In a non-rotating reference frame, the Lagrangian points mark positions where the combined gravitational pull of the two masses on a test mass provides precisely the centripetal force required for the test mass to rotate with them. Thus, at these points a test mass maintains its position relative to the two stars. These equilibrium points are clearly unstable because they are local maxima of Φ : perturb the test mass slowly and it will accelerate down the potential.

As we shall see presently, the *inner Lagrangian point*, L_1 , is central to the evolution of close binary systems. Approximate expressions for the

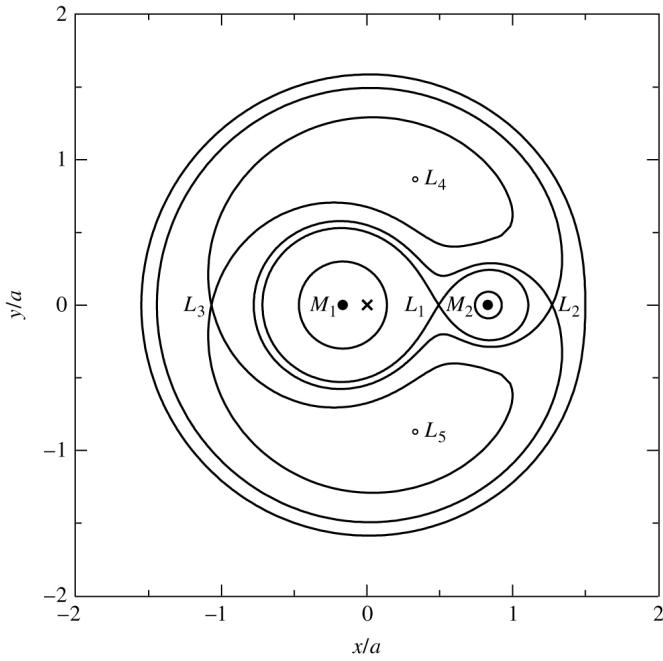


Figure 18.3: Intersections of equipotential surfaces with the plane of the orbit of a close binary system with masses $M_1 = 0.85M_\odot$, $M_2 = 0.17M_\odot$ and separation $a = 0.718R_\odot$ (see Figure 18.1). The centre of mass of the system is indicated with the ‘ \times ’ symbol.

distances from L_1 to M_1 and M_2 , denoted ℓ_1^1 and ℓ_1^2 respectively, are:

$$\ell_1^1 = a \left[0.500 - 0.227 \log_{10} \left(\frac{M_2}{M_1} \right) \right] \quad (18.10)$$

$$\ell_1^2 = a \left[0.500 + 0.227 \log_{10} \left(\frac{M_2}{M_1} \right) \right] \quad (18.11)$$

Points in space that share the same value of Φ define an equipotential surface. Figure 18.3 shows equipotential contours on the plane of the orbit for the binary system illustrated in Figure 18.1. Very close to each of the two masses, the equipotential surfaces are nearly spherical and centred on each mass. However, further away, the combined gravitational influence of M_1 and M_2 distorts the equipotential surfaces into teardrop shapes until they touch at the inner Lagrangian point. At greater distances, the equipotential surfaces assume a ‘dumbbell’ shape surrounding both masses.

These equipotential surfaces are *level surfaces* for binary stars. In such systems, as a star evolves it will expand to fill successively larger equipotential surfaces. This is easy to see when we consider that the effective gravity at each point is always perpendicular to the equipotential surface at that point. As there is no gravity parallel to an equipotential surface, a pressure difference in that direction cannot be maintained. It follows that

the density must also be the same along each equipotential surface in order to have a constant pressure there.

18.1.2 Classes of Binary Stars

Binary stars are classified on the basis of which equipotential surfaces are filled. In detached binaries, the distance separating the two stars is much greater than their radii. The stellar surfaces are close to spherical (see Figure 18.4) and the two stars evolve nearly independently of each other. These are the systems which, as we discussed in Lecture 4, provide us with measures of stellar masses from observations of their orbital periods.

However, if one of the two stars in the course of its evolution expands to fill its equipotential surface up to the inner Lagrangian point L_1 , its atmospheric gases can escape and be drawn towards its companion. The teardrop-shaped regions of space bounded by this particular equipotential surface are called *Roche lobes*; when one of the stars in a binary system has expanded beyond its Roche lobe, mass transfer to its companion can take place. Such a system is called a semidetached binary (middle panel of

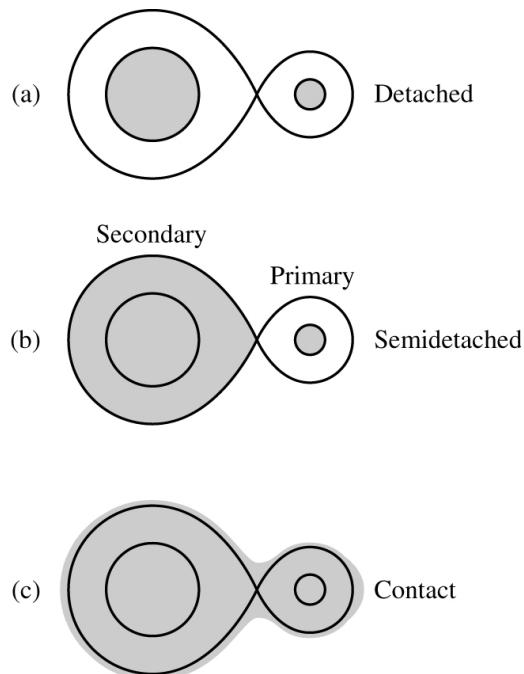


Figure 18.4: Different classes of binary systems. In semidetached binaries (b), the secondary has expanded to fill its Roche lobe. In contact binaries (c), the two stars share a common atmosphere.

Figure 18.4). The star that has filled its Roche lobe and is losing mass is usually referred to as the *secondary star* in the system with mass M_2 , and its accreting companion is the *primary star* with mass M_1 . Note that the primary star can be either more or less massive than the secondary star.

It may also be the case that *both* stars expand to (over)fill their Roche lobes. In this case, the two stars share a common atmosphere bounded by a dumbbell-shaped equipotential surface, such as the one passing through the Lagrangian point L_2 . Such systems are referred to as *contact binaries* (Figure 18.4).

The three cases illustrated in Figure 18.4, together with a range of stellar types, give rise to a rich nomenclature of different classes of interacting binary systems. Here we mention just two.

Cataclysmic Variables consist of a white dwarf primary and an M-type secondary filling its Roche lobe. They have short periods and irregularly increase in brightness by a large factor, then drop back to a quiescent state. Much attention is focused on CVs because they provide valuable information on final stages of stellar evolution and on accretion disks (see below).

X-ray Binaries have a neutron star or black hole component. The X-rays are generated by the accretion of gas onto the degenerate star from a non-degenerate companion. Observations of neutron star X-ray binaries complement the information on their physical properties (such as masses, radii, rotation and magnetic fields) obtained from pulsar studies.

Cygnus X-1 was the first X-ray source widely accepted to be a black hole candidate. Its mass is estimated to be $8.7M_\odot$, and it has been shown to be too compact to be any kind of object other than a black hole. Cygnus X-1 belongs to a high-mass X-ray binary system that includes the blue supergiant HDE 226868; the separation of the two objects is only ~ 0.2 AU. A stellar wind from the blue supergiant provides material for an accretion disk around the X-ray source. Matter in the inner disk is heated to millions of degrees, generating the observed X-rays. A pair of jets perpendicular to the disk are carrying part of the infalling material away into interstellar space.

18.1.3 Accretion Disks

The orbital motion of a semidetached binary can prevent the mass that escapes from the swollen secondary star from falling directly onto the primary. Instead, the mass stream goes into orbit around the primary to form a thin accretion disk of hot gas in the orbital plane, as shown in Figure 18.5.

A key component of accretion disk physics is *viscosity*, an internal friction that converts kinetic energy of bulk mass motion into random thermal motion. If matter is to fall inwards it must lose not only gravitational energy but also angular momentum. Since the total angular momentum of the disc is conserved, the angular momentum loss of the mass falling into the centre has to be compensated by an angular momentum gain of the mass far from the centre. In other words, angular momentum should be *transported* outwards for matter to accrete. Turbulence enhanced viscosity is the mechanism thought to be responsible for such angular momentum redistribution, although the origin of the turbulence itself is not fully understood. It is likely that magnetic fields play a role.

Accretion disks continue to be an active area of astrophysical research, both theoretical and observational. One of the reasons for the continued interest in this phenomenon is the ubiquity of accretion disks, from protostars and protoplanetary disks to binary stars, gamma-ray bursts, and active galactic nuclei.

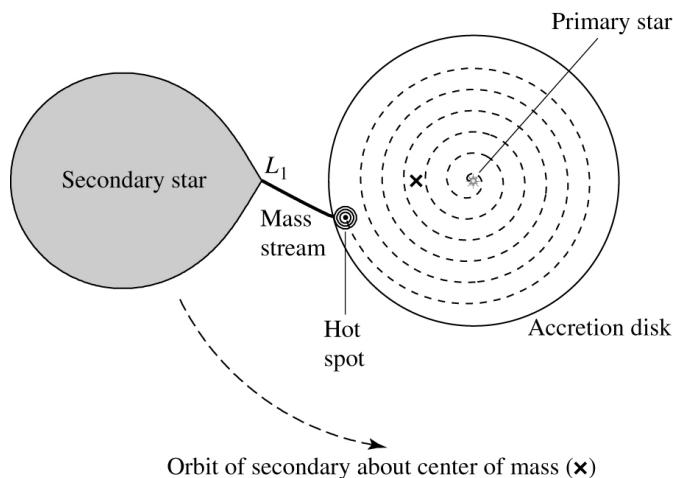


Figure 18.5: A semidetached binary showing the accretion disk around the primary star and the hot spot where the mass streaming through the inner Lagrangian point impacts the disk.

18.1.4 The Effects of Mass Transfer

The life history of a close binary system is quite complicated, with many possible variations depending on the initial masses and separations of the two stars involved. As mass is transferred from one star to the other, the mass ratio M_2/M_1 will change. The resulting redistribution of angular momentum affects the orbital period of the system, $P = \sqrt{2\pi}/\omega$, as well as the separation of the two stars, a . The extent of the Roche lobes, given by eqs. 18.10 and 18.11, depends on both a and M_2/M_1 , so it too will vary accordingly.

The effects of mass transfer can be illustrated with a simple analytical treatment that considers the total angular momentum of the binary system. Assuming circular orbits, the orbital angular momentum of the binary system is given by the expression:

$$L = \mu\sqrt{GMa} \quad (18.12)$$

where μ is the reduced mass:

$$\mu = \frac{M_1 M_2}{M_1 + M_2} \quad (18.13)$$

and $M = M_1 + M_2$ is the total mass of the two stars.

Assuming to a first approximation that no mass or angular momentum is removed from the system via stellar winds or gravitational radiation, both the total mass and the angular momentum of the system are conserved as mass is transferred between the two stars. That is, $dM/dt = 0$ and $dL/dt = 0$.

Taking the time derivative of the angular momentum, we have:

$$\begin{aligned} \frac{dL}{dt} &= \frac{d}{dt} \left(\mu\sqrt{GMa} \right) \\ 0 &= \sqrt{GM} \left(\frac{d\mu}{dt}\sqrt{a} + \frac{\mu}{2\sqrt{a}} \frac{da}{dt} \right) \\ \frac{1}{a} \frac{da}{dt} &= -\frac{2}{\mu} \frac{d\mu}{dt}. \end{aligned} \quad (18.14)$$

Differentiating (18.13), we have:

$$\frac{d\mu}{dt} = \frac{1}{M} \left(\dot{M}_1 M_2 + \dot{M}_2 M_1 \right) \quad (18.15)$$

since $M = M_1 + M_2$ is constant. Furthermore, from the condition that $dM/dt = 0$, it follows that $\dot{M}_1 = -\dot{M}_2$, and therefore,

$$\frac{d\mu}{dt} = \frac{\dot{M}_1}{M} (M_2 - M_1) . \quad (18.16)$$

Substituting 18.16 into 18.14, we arrive at our result:

$$\frac{1}{a} \frac{da}{dt} = 2\dot{M}_1 \frac{M_1 - M_2}{M_1 M_2} \quad (18.17)$$

which describes how the binary separation a varies as a result of mass transfer from M_2 to M_1 . Note that in the cases mentioned above, where the primary is a compact object, $M_1 < M_2$, so that da/dt is $-ve$: the stars get closer together.

The angular frequency ω of the orbit will also be affected, according to eq. 18.8. Since $M_1 + M_2$ is constant, Kepler's third law states that $\omega \propto a^{-3/2}$, and:

$$\frac{1}{\omega} \frac{d\omega}{dt} = -\frac{3}{2} \frac{1}{a} \frac{da}{dt} . \quad (18.18)$$

As the orbit shrinks, the angular frequency increases.

18.2 Type Ia Supernovae

As we discussed in Lecture 14, there is a limit to the mass of a white dwarf that can be supported by electron degeneracy pressure. This limit, known as the Chandrasekhar limit, is estimated to be $1.44M_{\odot}$ in the absence of significant rotation.

In close binary systems, mass transfer between the two stars may cause a white dwarf to approach the Chandrasekhar mass, leading to a catastrophic stellar explosion which we witness as a Type Ia supernova (Lecture 16.3). However, the details of the mechanism(s) that trigger the explosion are still unclear and the subject of much ongoing research. One of the problems is to understand why, when the Chandrasekhar limit is exceeded, the white dwarf does not 'just' collapse to form a neutron star.

18.2.1 Single Degenerate Scenario

The single degenerate scenario involves an evolving star transferring mass onto the surface of a white dwarf companion as in Figure 18.5. The current view among astronomers who model Type Ia supernova explosions is that in such systems the Chandrasekhar limit is never actually attained, so that collapse is never initiated. Instead, the increase in pressure and density due to the increasing mass of the white dwarf raises the temperature of the core and, as the white dwarf approaches to within $\sim 1\%$ of the Chandrasekhar limit, a period of convection ensues, lasting approximately 1,000 years. At some point in this simmering phase, carbon fusion is ignited. The details of the ignition are still unknown, including the location and number of points where the flame begins. Oxygen fusion is initiated shortly thereafter, but this fuel is not consumed as completely as carbon.

Once fusion has begun, the temperature of the white dwarf starts to rise. A main sequence star supported by thermal pressure would expand and cool in order to counterbalance an increase in thermal energy. However, as we have already discussed in previous lectures, degeneracy pressure is independent of temperature. Thus, the white dwarf is unable to regulate the burning process in the manner of normal stars, leading to a thermonuclear runaway. What happens next is also a matter of debate. It is unclear if the CO burning front occurs at subsonic speeds (normally referred to as a ‘*deflagration*’ event), or if the front accelerates and steepens to become a supersonic burning front, known as a ‘*detonation*’, or true explosion.

Regardless of the exact details of nuclear burning, it is generally accepted that a substantial fraction of the carbon and oxygen in the white dwarf is burned into heavier elements within a period of only a few seconds, raising the internal temperature to billions of degrees. The energy release from thermonuclear burning, $E \sim 1\text{--}2 \times 10^{51}$ erg, exceeds the binding energy of the star; the star explodes violently and releases a shock wave in which matter is typically ejected at speeds of $\sim 5\,000\text{--}20,000$ km s $^{-1}$. Whether or not the supernova remnant remains bound to its companion depends on the amount of mass ejected.

This scenario is similar to that of novae, in which a WD accretes matter more slowly and does not approach the Chandrasekhar limit. The infalling matter causes a H fusion surface explosion that does not disrupt the star.

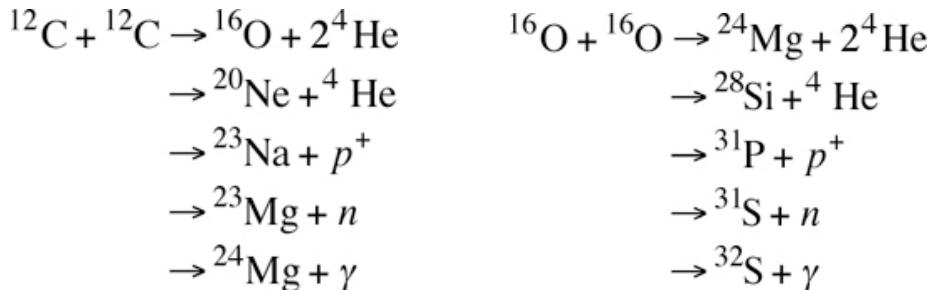
18.2.2 Double Degenerate Models

As the name implies, double degenerate models involve two white dwarfs in a binary orbit. Such systems are known to exist and it is a seemingly inevitable consequence that the two stars will spiral together as the system loses energy and angular momentum by gravitational radiation. On the other hand, computer simulations suggest that, as the two stars get very close together, the lighter of the two white dwarfs is torn apart and forms a thick disk around the other. This leads to an off-centre carbon ignition, resulting in ultimate collapse to a neutron star, rather than complete disruption of the white dwarf as a supernova.

Both the single- and the double-degenerate models have their strengths and their weaknesses. Of course, both mechanisms may be at work in nature, but we still do not know which is the dominant one. Furthermore, fundamental questions remain as to how the accretion of matter leads to the explosion in each progenitor model.

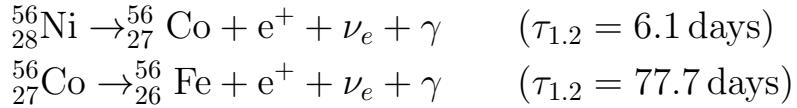
18.2.3 Nucleosynthesis in Type Ia Supernovae

The spectra of Type Ia supernovae taken near maximum light show absorption lines of intermediate mass elements, primarily O, Mg, Si, S, and Ca. These elements are produced by the rapid fusion of C and O via the channels already considered in Lecture 7.4.4:



At later epochs, the spectra become dominated by Fe and other heavy elements produced by explosive nucleosynthesis. Evidently, the outer ejecta show the products of C and O burning, while the inner, denser regions of the exploding star burn all the way to the Fe-group. The presence of high-velocity C and O in early-time spectra suggests that the explosion left behind some unburnt material, possibly in pockets.

The primary iron-peak element produced by the explosion is $^{56}_{28}\text{Ni}$, because the timescale of explosive nucleosynthesis is too short for β -decay to change the original proton to neutron ratio from $Z/A = 1/2$ of the fuel $^{28}_{14}\text{Si}$. $^{56}_{28}\text{Ni}$ decays to $^{56}_{26}\text{Fe}$ via the reactions:



powering the light curve of Type Ia supernovae which shows a two-step decline (see Figure 18.6). Analysis of the light curve indicates that, typically, the ejected mass of $^{56}_{28}\text{Ni}$ is $\sim 0.7\text{--}1 M_\odot$.

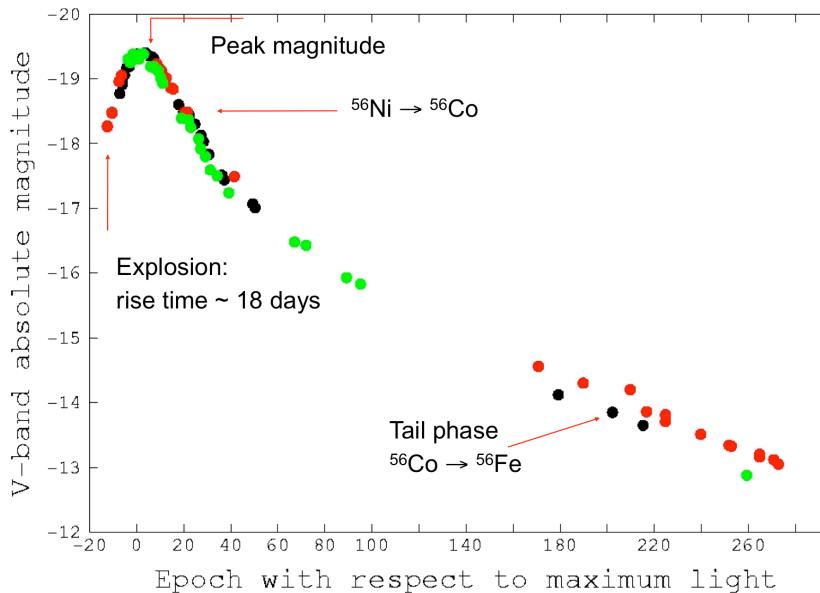


Figure 18.6: Typical light curve of a Type Ia supernova, constructed by combining the light curves of SN1990N, SN1996X and SN2002er.

Much effort has been devoted to computational modelling of nucleosynthesis in exploding CO white dwarfs, with a good degree of success in matching the temporal evolution of their spectra. SN Ia are thought to be the source of $\sim 2/3$ of the Fe which has accumulated in the interstellar medium up to the present-day. An origin in low and intermediate mass stars introduces a time-lag in the release of Fe compared to the prompt release of O and other alpha-capture elements from Type II supernovae which, as you will remember, have massive star progenitors. This fact has been much exploited by models that attempt to reconstruct the past history of star formation in galaxies from measurements of the relative abundances of different elements as a function of time.

18.2.4 SN 2011fe

Our ideas about the progenitors of Type Ia supernovae are largely driven by theoretical considerations based on circumstantial evidence, such as the lack of H and He lines in their spectra (few astrophysical objects lack these elements); their occurrence in galaxies of all types, rather than being associated exclusively with regions of star formation, as is the case for Type II SNe; and the consistency between the energy generated by burning a CO white dwarf and that associated with Type Ia SN events.

The reason for this rather unsatisfactory state of affairs is that we do not yet have any direct observation of a star prior to its explosion as Type Ia supernova. We recently came a little closer to achieving this ‘holy grail’ of supernova research, with the early detection of SN 2011fe in the nearby galaxy M101. At a distance of 6.4 Mpc, this is the closest SN Ia in the past 25 years. As Figure 18.7 shows, no star is detected in the best pre-supernova images of this frequently observed large spiral galaxy. The non-detection improves very significantly previous empirical limits on the luminosity of the secondary (which were not very constraining; see Figure 18.8), although they are not sufficiently stringent to distinguish between single- and double-degenerate scenarios.

Thanks to regular and frequent sky monitoring, many recent supernovae

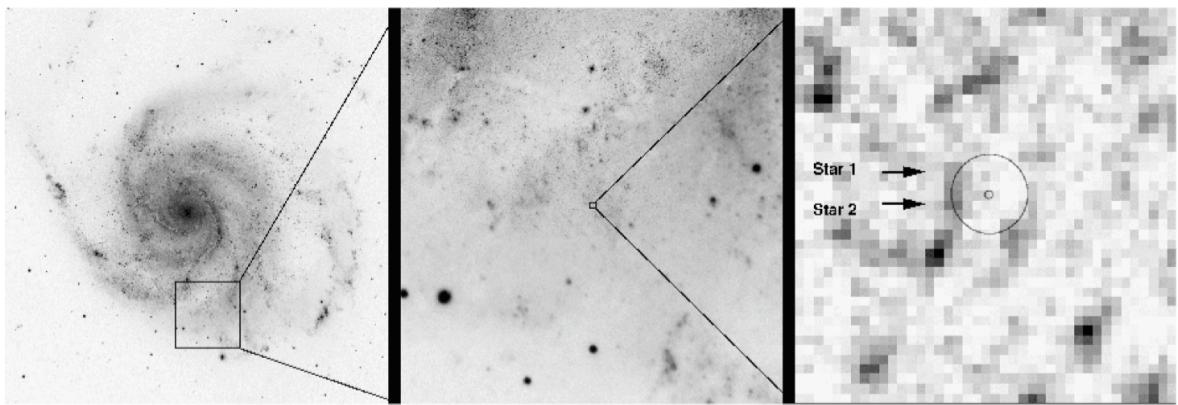


Figure 18.7: The recent bright ($V_{\max} = 10$) Type Ia supernova in the nearby ($d = 6.4$ Mpc) spiral galaxy M101 shows no obvious counterpart in archival *HST* images. The two concentric circles in the right-hand panel have radii corresponding to the 1σ (21 mas) and 9σ astrometric uncertainty in the position of the SN. (Figure reproduced from Li et al. 2011, arXiv:1109.1593).

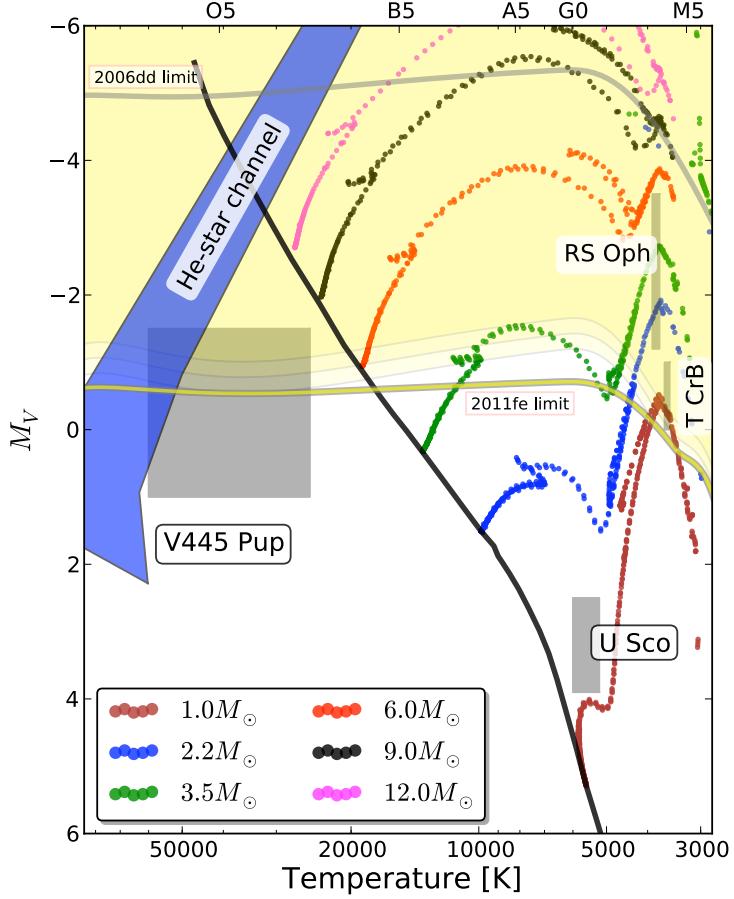


Figure 18.8: Progenitor system constraints for SN2011fe on the H-R diagram. The yellow area of the H-R diagram is excluded by the non-detection of a stellar counterpart at the position of SN2011fe in archival *HST* images of M101. The proximity of this galaxy has reduced significantly the allowed combination of effective temperature and luminosity of the progenitor, compared to the best previous limit, indicated by the grey line labelled ‘2006dd limit’ near the top of the diagram. In the single-degenerate model, the secondary companion to the star that exploded as SN2011fe must have been of relatively low luminosity, either a red giant star evolved from $M_{\text{ZAMS}} \lesssim 3.5 M_{\odot}$, or a main sequence star with $M \lesssim 5 M_{\odot}$. (Figure reproduced from Li et al. 2011, arXiv:1109.1593).

have been detected at very early times. It is estimated that the first photometry of SN2011fe was obtained only 11 hours after the explosion. At such early times, the luminosity can be related (with a few assumptions) to the initial radius of the star, R_0 . From such considerations, Bloom et al. (2011, arXiv:1111.0966) were able to place the limit $R_0 \lesssim 0.02 R_{\odot}$, consistent with a white dwarf progenitor. Early spectral observations also showed high velocity (up to $20,000 \text{ km s}^{-1}$) C and O features, as well heavier elements synthesised in the explosion. The presence of C and O in the spectra favour the progenitor being a CO white dwarf.

18.3 The Hubble Diagram of Type Ia Supernovae: Evidence for a Cosmological Constant

Suppose we know the absolute luminosity of an astronomical source, then we could use its observed flux to deduce its luminosity distance from the relationship:

$$F_{\text{obs}} = \frac{L}{4\pi d_L^2}. \quad (18.19)$$

which defines the cosmological luminosity distance d_L .

If we could be confident that the absolute luminosity is a constant in time and space, so that the object in question constitutes a *standard candle*, and if the source luminosity is sufficiently high that it can be detected over cosmological distances, then we could test for the cosmological parameters $\Omega_{m,0}$, $\Omega_{\Lambda,0}$, and $\Omega_{k,0}$ which determine the form of $d_L = f(z)$ according to the equations:

$$d_L(z) = \frac{c(1+z)}{\sqrt{|\Omega_{k,0}|}H_0} S_k \left(H_0 \sqrt{|\Omega_{k,0}|} \int_0^z \frac{dz}{H(z)} \right) \quad (18.20)$$

and

$$H(z) = H_0 \left[\Omega_{m,0} \cdot (1+z)^3 + \Omega_{k,0} \cdot (1+z)^2 + \Omega_{\Lambda,0} \right]^{1/2} = H_0 \cdot E(z)^{1/2} \quad (18.21)$$

where

$$S_k(x) = \begin{cases} \sin(x) & \text{for } k > 0 \\ x & \text{for } k = 0 \\ \sinh(x) & \text{for } k < 0 \end{cases} \quad (18.22)$$

In the above equations, z is the redshift, H is the Hubble parameter that measures the expansion rate of the Universe, the subscript 0 denotes present time and the three Ω give the ratios between the present-day densities of, respectively, matter m , curvature k , and a cosmological constant Λ to the *critical density* today:

$$\rho_{\text{crit}} = \frac{3H_0^2}{8\pi G} \simeq \frac{5 \text{ H atoms}}{\text{m}^3} \simeq \frac{1.35 \times 10^{11} M_\odot}{\text{Mpc}^3} \quad (18.23)$$

You should have encountered these equations already in the *Physical Cosmology* lectures of the Part II Astrophysics course.

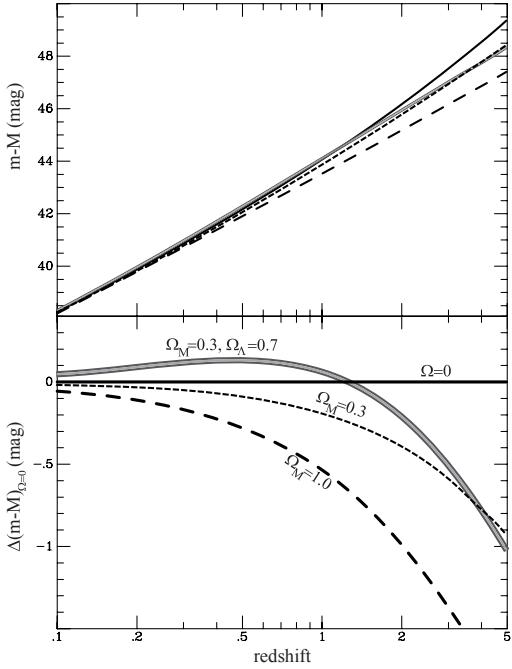


Figure 18.9: The distance modulus as a function of redshift for four relevant cosmological models, as indicated. In the lower panel the empty universe ($\Omega_{m,0} = \Omega_{\Lambda,0} = 0$) has been subtracted from the other models to highlight the differences.

Expressing the luminosity distance in terms of the distance modulus:

$$M - m = 2.5 \log \left(\frac{d_{L,0}}{d_L} \right)^2 = 5 \log \left(\frac{d_{L,0}}{d_L} \right) \quad (18.24)$$

where M is the magnitude of the standard candle at some nearby distance $d_{L,0}$. In the conventional definition of the distance modulus, $d_{L,0} = 10$ pc and M at this distance is usually referred to as the absolute magnitude. However, in cosmological situations this is a rather small distance and a more natural unit is 1 Mpc. If we measure the distance in this unit, the apparent magnitude is given by

$$m = M + 5 \log d_L + 25 \quad (18.25)$$

Figure 18.9 illustrates the dependence of the distance modulus on redshift for four different sets of cosmological parameters. It can be seen that if we could measure the distance modulus of a standard candle with a precision of about 10%, or 0.1 magnitudes, out to redshifts $z > 0.5$, we may be able to distinguish a Λ -dominated universe from a matter-dominated one.

A well-known example of standard candles are Cepheids, a class of variable stars which exhibit a period-luminosity relation which has allowed

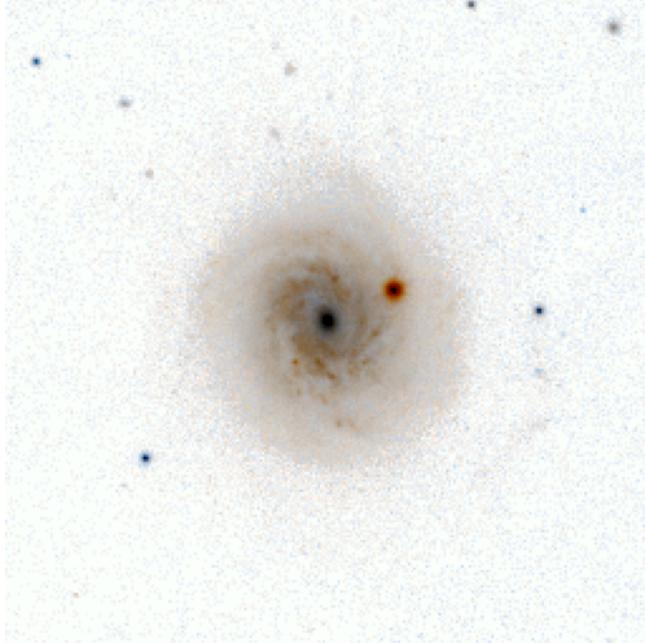


Figure 18.10: SN 1998aq in NGC 3982. This prototypical Type Ia supernova was discovered on 1998 April 13 by Mark Armstrong as part of the UK Nova/Supernova Patrol approximately two weeks before it reached its peak luminosity in the B -band. Its host galaxy, NGC 3982, is a nearly face-on spiral with a Seyfert 2 active nucleus. At a distance of 20.5 Mpc, NGC 3982 is a possible member of the Ursa Major cluster of galaxies.

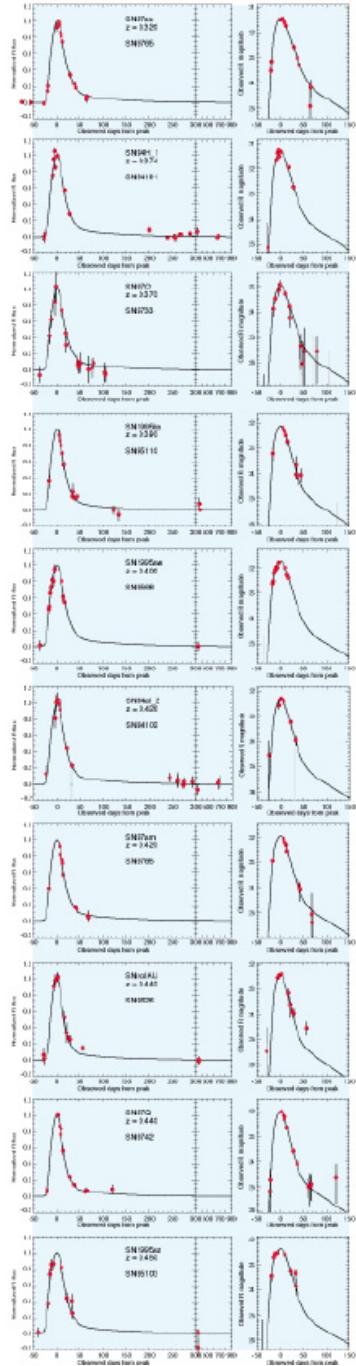
the determination of H_0 . However, Cepheids are intrinsically too faint to be followed beyond the local universe. The class of astronomical objects which has so far turned out to be closest approximation to a cosmological standard candle are Type Ia supernovae.

18.3.1 Type Ia Supernovae as Standard Candles

As early as 1938, Baade and Zwicky pointed out that supernovae were promising candidates for measuring the cosmic expansion. Their peak brightness seemed quite uniform, and they were bright enough to be seen at extremely large distances. In fact a supernova can, for a few weeks, be as bright as an entire galaxy [see Figure 18.10; SN 1998aq in NGC 3982 at a distance of ~ 20 Mpc reached peak magnitude $m_V = 11.4$, brighter than the whole galaxy which has $m_V = 11.8$]. Over the years, however, as more and more supernovae were measured, it became clear that they are in fact a heterogeneous group with a wide range of spectral characteristics and intrinsic peak brightnesses. Eventually, this led to the modern classification of supernovae summarised in Lecture 16.3.

Type Ia Supernovae

from $z = 0.32 \dots$
observed from the ground

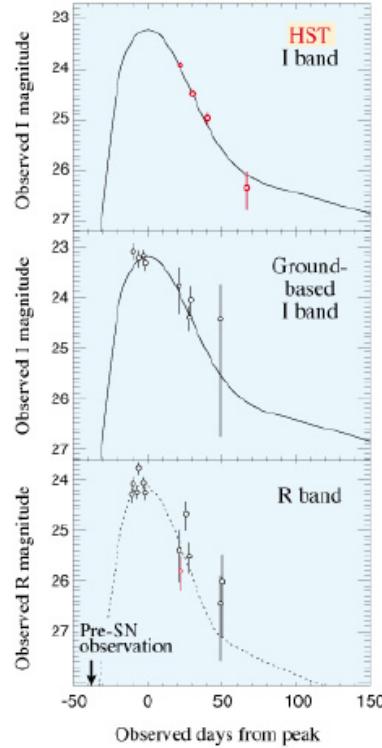


Light Curves

... to $z = 0.75$
observed from the ground

SN 1997ap at $z = 0.83$
observed from the
ground and with the HST

Perlmutter, et al., Nature (1996)



We observe most of the supernovae for approximately two months in both the R and I bands (corresponding approximately to the restframe B and V bands for the median redshift). At high redshifts, a significant fraction of this host galaxy light is within the seeing disk of the supernova, so final observations about one year later are usually necessary to observe (and subtract) the host galaxy light after the supernova has faded. The plots to the left and the right show just the R band light curves for about half of the 40 supernovae that have been completely observed and analyzed so far. The plots above show the highest redshift spectroscopically confirmed supernova, which was observed with the Hubble Space Telescope.

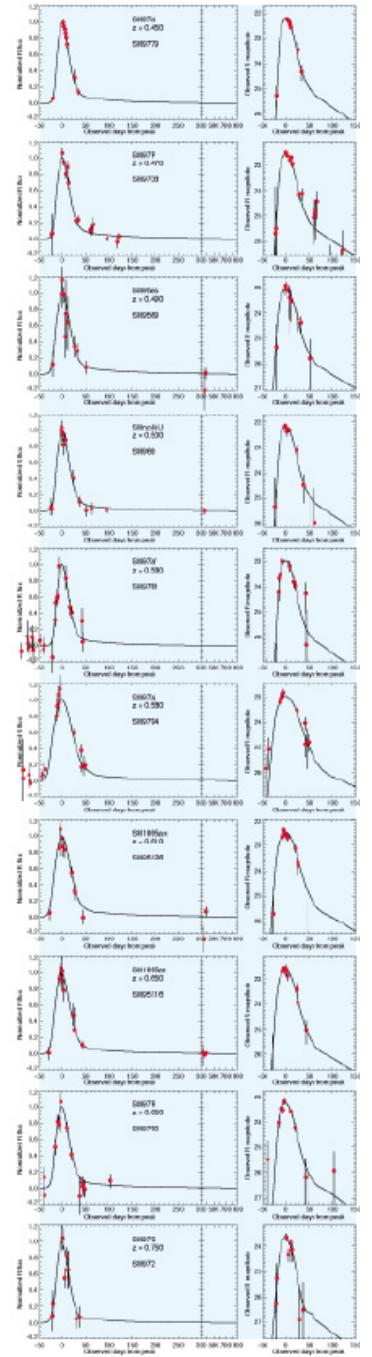


Figure 18.11: Light curves of Type Ia supernovae

The value of SNIa as cosmological probes arises from the high peak luminosity as well as the observational evidence (locally) that this peak luminosity is the sought-after standard candle. In fact, the absolute magnitude, at peak, varies by about 0.5 magnitudes which corresponds to a 50%-60% variation in luminosity; this, on the face of it, would make them fairly useless as standard candles. However, the peak luminosity appears to be well correlated with decay time: the larger L_{peak} , the slower the decay (see Figure 18.12). There are various ways of quantifying this effect, such as

$$M_B \approx 0.8(\Delta m_{15} - 1.1) - 19.5$$

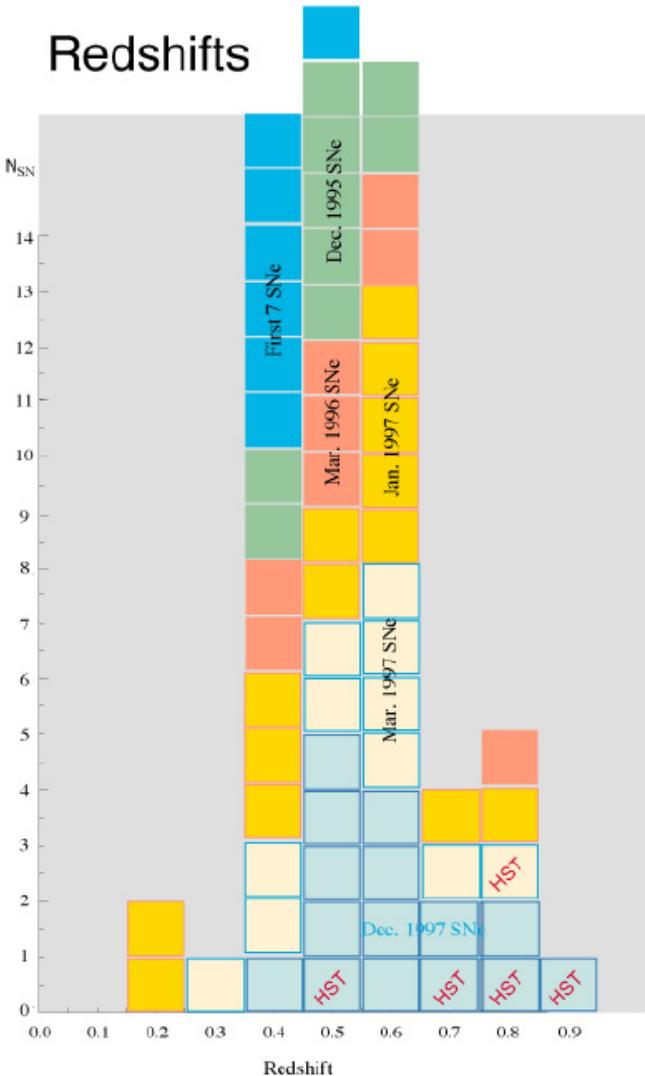
where M_B is the peak absolute magnitude in the B -band and Δm_{15} is the observed change in apparent magnitude 15 days after the peak. This is an empirical relationship, and there is no consensus about the theoretical explanation.¹ However, when this correction is applied, it appears that $\Delta L_{\text{peak}} < 20\%$. If true, this means that SNIa are candles that are standard enough to distinguish between cosmological models at $z \approx 0.5$ (see Fig. 18.9).

In a given galaxy, supernovae are rare events (on a human time scale, that is), with one or two such explosions per century. But if thousands of galaxies can be surveyed on a regular and frequent basis, then it is possible to observe several events per year over a range of redshift. About 15 years ago two large international collaborations, the ‘*Supernova Cosmology Project*’, based at Berkeley, California, and the ‘*High-Z Supernova Search*’ based mostly in Baltimore, Maryland, began such ambitious programs. Observations with the *Hubble Space Telescope* have proved crucial for following SNe beyond $z \sim 0.5$ (see Figure 18.13). These efforts have been fantastically fruitful and the results have led to a major paradigm shift in our ‘consensus’ cosmology.

Fig. 18.14 shows the Hubble diagram for SNe of type Ia observed by the Supernova Cosmology Project up to 2003—the highest redshift supernova observed at that time was at $z = 0.86$. The conclusion seems to be that SNIa are 10% to 20% fainter at $z \approx 0.5$ than would be expected in an empty universe ($\Omega_{m,0} = \Omega_{\Lambda,0} = 0$ and $\Omega_{k,0} = 1$) and, more significantly, about 30% to 40% fainter than a model with $\Omega_{m,0} = 0.25$ (indicated by

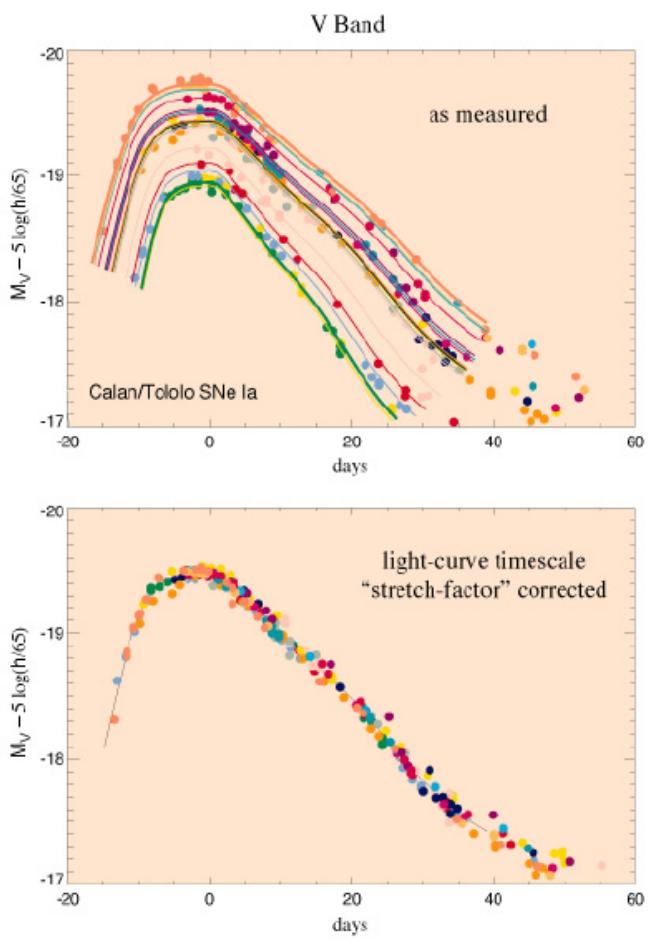
¹The existence of a well-defined mass threshold, $1.44 M_{\odot}$ for an accreting white dwarf to explode as a type Ia supernova is presumably at the root of this remarkable uniformity in their spectra and light curves, and the small residual degree of variation may reflect differences in accretion rates, rotational velocities and C/O ratios.

C. Pennypacker	M. DellaValle Univ. of Padova	R. Ellis, R. McMahon IoA, Cambridge
B. Schaefer Yale University	P. Ruiz-Lapuente Univ. of Barcelona	H. Newberg Fermilab



We have discovered well over 50 high redshift Type Ia supernovae so far. Of these, approximately 50 have been followed with spectroscopy and photometry over two months of the light curve. The redshifts shown in this histogram are color coded to show the increasing depth of the search with each new "batch" of supernova discoveries. The most recent supernovae, discovered the last week of 1997, are now being followed over their lightcurves with ground-based and (for those labeled "HST") with the Hubble Space Telescope.

Low Redshift Type Ia Template Lightcurves



Type Ia supernovae observed "nearby" show a relationship between their peak absolute luminosity and the timescale of their light curve: the brighter supernovae are slower and the fainter supernovae are faster (see Phillips, Ap.J.Lett., 1993 and Riess, Press, & Kirshner, Ap.J.Lett., 1995). We have found that a simple linear relation between the absolute magnitude and a "stretch factor" multiplying the lightcurve timescale fits the data quite well until over 45 restframe days past peak. The lower plot shows the "nearby" supernovae from the upper plot, after fitting and removing the stretch factor, and "correcting" peak magnitude with this simple calibration relation.

Figure 18.12:

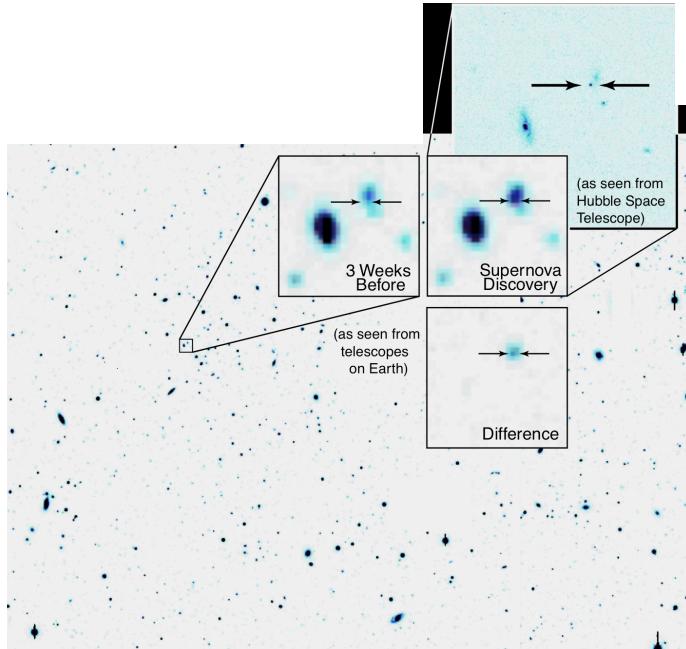


Figure 18.13: The superb resolution of the *Hubble Space Telescope* allows a more accurate measurement of the light curves of high redshift supernovae than is possible from the ground.

other considerations) and $\Omega_{\Lambda,0} = 0$. The introduction of a cosmological constant at the level $\Omega_{\Lambda,0} \simeq 0.7$ improves the fit to the SN magnitude vs. redshift relation significantly. The two teams concluded that we live in an accelerating universe (this is the effect of a cosmological constant), a discovery which *Science* magazine considered “*The Breakthrough of the Year*”. In 2011 Saul Perlmutter (The Supernova Cosmology Project), and Adam Riess and Brian Schmidt (The High- z Supernova Search Team) were awarded the Nobel Prize in Physics “*for the discovery of the accelerating expansion of the Universe through observations of distant supernovae*”.

18.3.2 Parameter Estimation

In this section we consider more closely the methods employed to determine the values of Ω_i which best fit the SN data shown in Figure 18.14. The approach to this ‘parameter estimation problem’ has many applications in the analysis of scientific measurements.

Let us assume that we have a sample of n SN measurements consisting of magnitude m_i , typical magnitude error $\pm\sigma_{m,i}$, and redshift z_i (there is also an error associated with z_i , but it can be neglected, for our purposes, compared with $\sigma_{m,i}$). We wish to compare quantitatively this data set

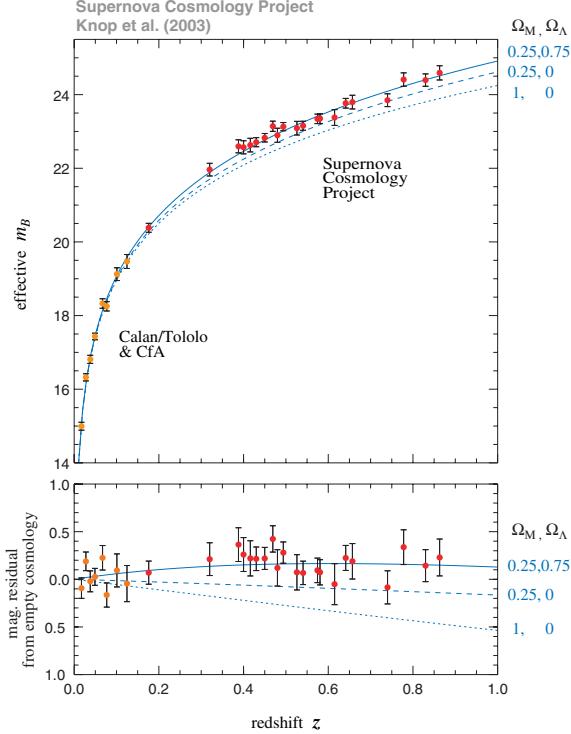


Figure 18.14: Hubble diagram for SNe Ia up to $z = 0.86$, reproduced from Knop et al. 2003, ApJ, 598, 102. The observed B -band magnitudes of the SNe Ia at maximum light are compared with the predictions for three cosmological models, as indicated. The lower panel shows the difference relative to an empty universe with $\Omega_{m,0} = \Omega_{\Lambda,0} = 0$ and $\Omega_{k,0} = 1$.

with theoretical expectations from Eqs. 18.25, 18.20 and 18.21 for different combinations of the parameters $(\Omega_{m,0}, \Omega_{\Lambda,0}, M)$.

There are two ways to tackle the absolute magnitude M . We could assume that we know M with sufficient precision from measurements of nearby SNe Ia. The absolute magnitudes of Type Ia supernovae in nearby galaxies do not depend on any value of Ω , but only on the Hubble parameter H_0 (and the assumption of negligible peculiar velocities relative to the Hubble flow). Alternatively, we could consider M to be a free parameter alongside $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$, and fit simultaneously for all three.

We'll consider the second approach. In order to get a compact notation, we define the parameter vector

$$\theta \equiv (\Omega_{m,0}, \Omega_{\Lambda,0}, M) \quad (18.26)$$

If we assume that the errors in the magnitude, σ_m, i are purely of a random nature and are drawn from a Gaussian distribution², then we can obtain the

²In scientific analysis this is often a crucial assumption, in the sense that generally we do not know all the sources of error in a measurement, random and systematic.

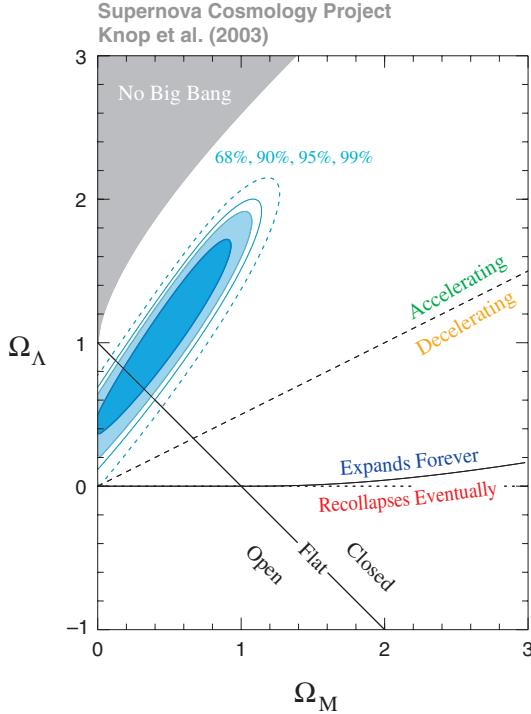


Figure 18.15: Joint likelihood contours in the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane for the SN data in Figure 18.14.

best fit parameters by maximising the posterior probability (likelihood):

$$L(\theta) \propto \exp \left[-\frac{1}{2} \chi^2 \right] \quad (18.27)$$

with

$$\chi^2 = \sum_{i=1}^n \left(\frac{m(z_i; \theta) - m_i}{\sigma_{m,i}} \right)^2 \quad (18.28)$$

It is then relatively straightforward to minimise eq. 18.28 to obtain the best-fit value of θ . More importantly, by calculating the value of $L(\theta)$ over a whole region in parameter space—which is relatively straightforward to do with numerical techniques—we can generate the full distribution of probabilities for the set of parameters considered.

If we are most interested in the cosmological parameters $\Omega_{m,0}$ and $\Omega_{\Lambda,0}$, and less concerned with the value of M , we can marginalize over the absolute magnitude and restrict ourselves to the two-dimensional probability distribution

$$L(\Omega_{m,0}, \Omega_{\Lambda,0}) = \int dM L(\Omega_{m,0}, \Omega_{\Lambda,0}, M) \quad (18.29)$$

Figure 18.15 shows contours of $L(\Omega_{m,0}, \Omega_{\Lambda,0})$ at the 68%, 90%, 95%, and 99% levels on the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane for the SN data in Figure 18.14. Clearly a range of $\Omega_{m,0}, \Omega_{\Lambda,0}$ combinations can reproduce the SNIa peak magnitudes, but it is noteworthy that at the 95% confidence level we do require $\Omega_{\Lambda,0} > 0$.

The confidence contours on the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane are stretched along a line $\Omega_{\Lambda,0} = 1.4\Omega_{m,0} + 0.4$. Some other cosmological test, which depends on Ω_i in a different way from the luminosity distance, is thus required to narrow down the allowed region. The angular scale of the temperature fluctuations of the Cosmic Microwave Background provides the most stringent of such constraints. The position of the first peak in the angular power spectrum, together with the amplitudes of the first two peaks, define a line on the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane which is nearly perpendicular to that of the SNIa measurements, at $\Omega_{m,0} + \Omega_{\Lambda,0} \simeq 1$, indicating that we live in a near-flat universe with $\Omega_{k,0} \simeq 0$. When we combine the CMB, SNIa and other measurements we arrive at today's consensus cosmology with $\Omega_{m,0} \simeq 0.3$, $\Omega_{\Lambda,0} \simeq 0.7$, $\Omega_{k,0} \simeq 0$ (see Figure 18.16).

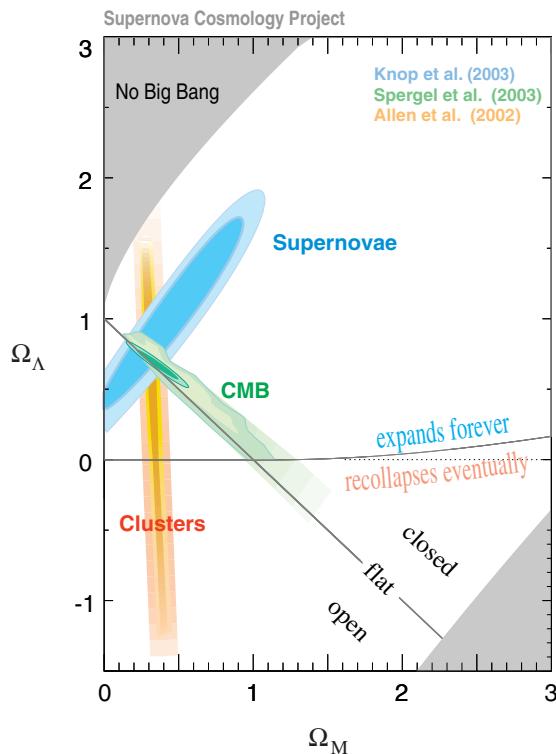


Figure 18.16: Joint likelihood contours in the $\Omega_{m,0} - \Omega_{\Lambda,0}$ plane from type Ia supernovae, the angular power spectrum of the cosmic background radiation, and massive galaxy clusters.