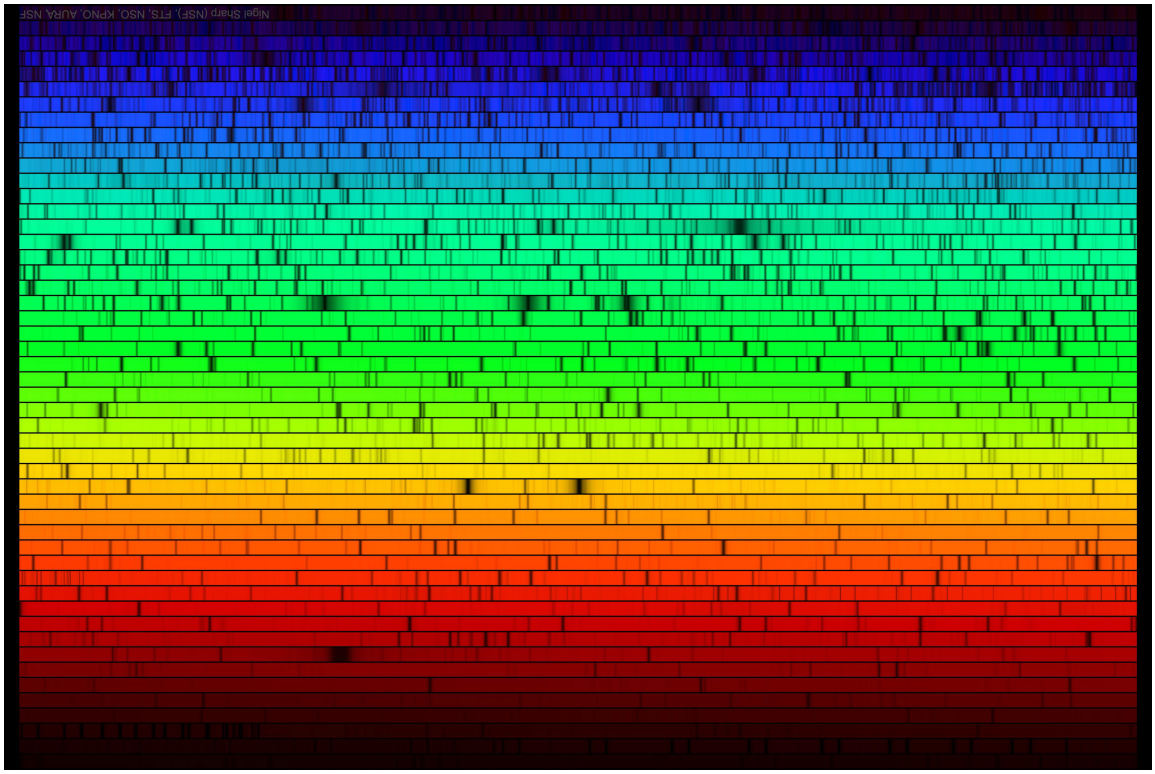


Figure 1: An echellogram of the solar spectrum.

The Physics of Stars

Notes for PHAS0036 [incorporating my legacy notes from PHAS2112 (now PHAS0018) where useful for completeness]

Preface

What follows are my notes for PHAS0036, 'The Physics of Stars'.

Necessary pre-knowledge for this course is provided by the second-year course 'Astrophysical Processes' (now PHAS0018), which I used to teach, and for which I developed substantial notes, including considerable additional material for my own interest. The PHAS0036 syllabus includes some of that additional material, which is incorporated into these notes. I've also freely included chunks of basic PHAS0018 material in the text where I felt it was useful for reference/revision purposes (notably Chapters 2–3 and 6), or to establish a consistent nomenclature (which may differ slightly from that currently adopted in PHAS0018). I've attempted to identify this material explicitly in section headings; it will not be covered in any depth in lectures (it's stuff you're *supposed* to know already), nor will it come up in the PHAS0036 exam (other than implicitly, as part of the broader context).

The two main themes of PHAS0036 (stellar atmospheres; structure & evolution) were formerly delivered as separate courses, and were originally delivered in PHAS0036 by separate lecturers. Rationalization of all this material into a coherent whole is now substantially complete; there may still be some duplication in the written notes, but I have tried to ensure that at least the nomenclature is homogeneous.

Nevertheless, do bear in mind that these are notes, written primarily for my own use. Boxed items, and indented or small-font sections, are 'extras', and not part of the syllabus (and therefore not examinable). You will find sections of incomplete text flagged up; some repetition; a numbering system that may depart from that given in lectures (as the written notes are under continual revision); and, probably, outright errors. If you think you've found any mistakes, or encounter anything that appears unclear (impenetrable, nonsensical…), please let me know.

Contents

Section 1

Setting the scene

1.1 Introduction

The foundations of stellar astrophysics are believed to be firm, allowing current research into more-detailed aspects – such as the roles of magnetic fields, rotation, or winds – to be conducted with confidence. The work of laying these foundations can be largely traced to a golden age for fundamental physics in the first part of the 20th century, in parallel with the development of quantum mechanics, nuclear physics, and relativity (although it is remarkable how far knowledge could be advanced even before Eddington's identification of nuclear fusion as the power source for stars).

With this background, the broad purposes of this course are twofold:

- first, we want to be able to understand the principles of how to model stellar spectra (such as that in Fig. 1.1) in order to infer photospheric properties;
- secondly, we aim to understand these stellar properties in the context of our theory of stellar structure and evolution. In effect, this means assembling a set of tools that allow us to interpret the distribution of stars in the Hertzsprung-Russell diagram (Fig. 1.2).

The content is therefore divided into two parts, delivered before and after Reading Week:

- we will first study how we can model stellar atmospheres and stellar spectra.
- then we will study the equations for the structure of the star's interior and its evolution.

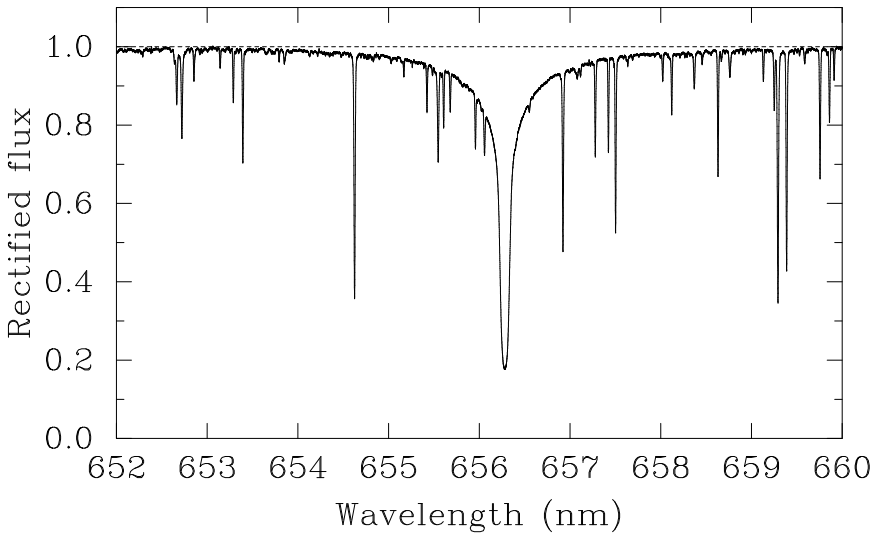


Figure 1.1: A short section of the solar flux spectrum around the strong Hα line (adapted from Wallace et al. 2011).

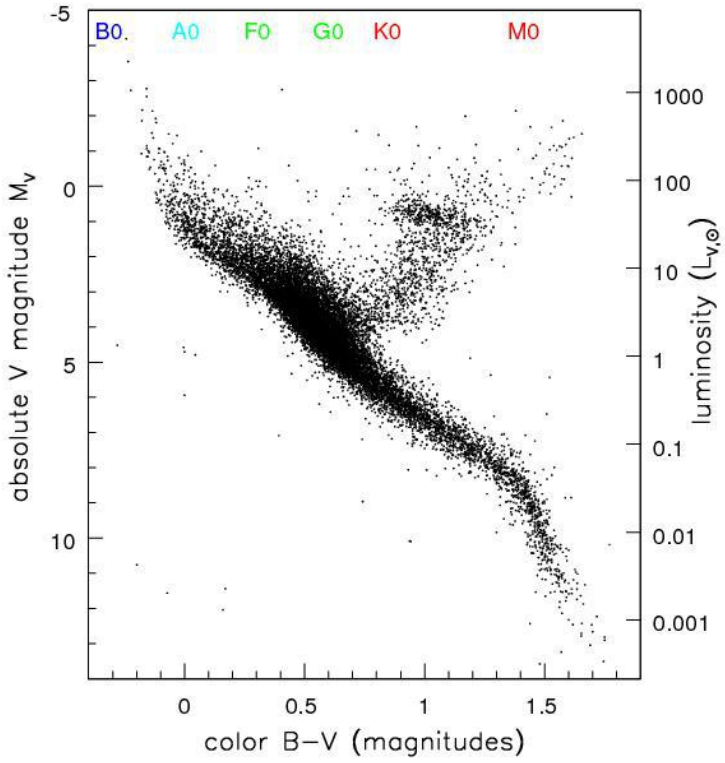


Figure 1.2: The Hertzsprung-Russell diagram for stars within 100 pc of the Sun (distances from *Hipparcos*).

1.1.1 Stellar Atmospheres

In the broadest terms, a stellar atmosphere can be considered as the transition region from the stellar interior to the interstellar medium. The behaviour of the atmosphere is controlled by the density of the gases in it and the energy escaping through it.[1] These quantities depend proximately on the temperature,[2] the pressure, and the atmospheric abundances (and potentially other factors, such as mass loss, or magnetic fields), which in turn are ultimately determined by the mass and age of the star (and potentially other factors, such as chemical composition and rotation rate at birth).

A *model* atmosphere is a numerical simulation of a real stellar atmosphere, typically presented as the run of physical parameters (such as temperature) as a function of depth; here 'depth' generally refers to *optical* depth (§4.4), measured inwards.

Observationally, the most easily accessible part of the atmosphere is the photosphere, because, by definition, the visible spectrum originates there. (We will generally use 'atmosphere' and 'photosphere' more or less interchangeably.) The photosphere is strongly affected by its characteristic temperature. Typically, the temperature drops by a factor ~ 2 from the bottom to the top of the photosphere, so the *local* temperature is not of itself a useful parameter to characterize the star. Instead, we can define (and measure) the *effective* temperature, $T_{\rm eff}$, in terms of the total radiative energy leaving the star:

$$\int_0^\infty F_\nu \, \mathrm{d}\nu \equiv \sigma T_{\rm eff}^4 \left(\int_0^\infty F_\lambda \, \mathrm{d}\lambda \right) \tag{1.1}$$

where F_ν is the flux[3] emitted by the star, per unit area, at frequency ν. The luminosity is then given by[4] $L = 4\pi R^2 \sigma T_{\rm eff}^4$. It is easy to recognize that $T_{\rm eff}$ corresponds to the temperature T of a black body[5] having the same power output as the star.

In the Sun, the photosphere is around 500 km thick[6] – that is, about half the distance from Land's End to John O'Groats, and less than about ~0.1% of the solar radius ($R_\odot \simeq 695{,}800$ km). Because of this, the atmosphere is well approximated as a plane-parallel medium. For other stars, the thickness scales with the surface gravity ($\sim M/R^2$) and effective temperature, and with the opacity (Section 4.3), as consequences of hydrostatic equilibrium (Section 2.1). Hydrostatic equilibrium also links surface gravity to pressure; and it is customary in stellar astrophysics to use surface gravity (which varies very little through a 'thin' stellar atmosphere), rather than pressure (which varies more more), to characterise

[1] If we were to attempt to identify a single paper as the starting point for modern research into radiative transfer in stellar atmospheres, it would probably be Sir Arthur Schuster's quaintly titled "Radiation Through a Foggy Atmosphere" (1905ApJ....21....1S).

[2] What is 'the' temperature (Appendix N)? Usually, the gas *kinetic* temperature is implied.

[3] A formal definition of flux is given in §3.2

[4] This simple formulation is true only for spherical stars. This is usually an excellent approximation, but can break down for very rapidly rotating stars, or stars in close binary systems, where $T_{\rm eff}$ is no longer a global constant, but can vary from point to point. An explicit integration over surface area is then required.

[5] Appendix C discusses black-body radiation.

[6] As determined from an analysis of limb darkening; cf. Section 8.4

the conditions giving rise to the observed spectrum. In principle, therefore, by studying the photosphere we can infer information on the gravity (a simple function of stellar mass and radius) and on the opacity (having a more complex dependence on temperature and abundances), as well as the effective temperature.

1.1.2 Stellar structure and evolution

The information we can infer from the spectrum is, literally, superficial: it tells us something about the surface conditions of the star. To connect these to global stellar parameters, such as mass, radius, luminosity, we need to know something about how these fundamental parameters relate to the photospheric conditions. This is what underpins the investigation of stellar structure; but the structure (and hence the photospheric properties) changes with time, as the nuclear fuel powering the star is consumed. Hence the studies of stellar structure and of stellar evolution are inextricably linked.

1.1.3 Linking to observations

Although the fundamental physical properties are parameters such as M, R, and L, these translate into observational parameters such as absolute magnitude and colour index (cf. Appendix E), or spectral type. The ultimate goal is to relate these observationally accessible quantities to the physical parameters, and thence to make inferences about stellar structure and evolution.

The spectrum is a particularly powerful tool in this endeavour, because, as already noted, the continuum and line spectra depend principally on the gravity, temperature, and composition (as well as rotation, magnetic field, mass loss, maculation, and other properties). The line spectra are used to classify stars, according to their temperature and photospheric pressure. With this criterion, we can assign a spectral type to a star. Originally, such classifications used photographic records (spectrograms) that centred in the blue region of the spectrum. Spectra toward the cool end of the sequence are called ‘late-type’ spectra, and those toward the hot end ‘early-type’ (for historical reasons; they bear no relationship to actual ages or evolutionary stages). There are ~60 categories of stars, from O2 to M8. From the hottest to coolest, the standard temperature sequence is: OBAFGKM. (Additionally, types C [and R, N] and S are used for spectroscopically distinct objects at M-star temperatures; types L, T, and Y are used for brown dwarfs.)

In the short introductory sections 1–4, we’ll briefly recap some basic material (which should be familiar from PHAS 0018) in order to ensure we’re fully equipped for subsequent topics. Section 5 then examines opacity sources in some detail, introducing the concept of the Rosseland Mean Opacity; while Section 6 reviews LTE. Section 7 introduces some basic concepts in radiative transfer.

We then look at ‘real world’ model atmospheres.

Section 2

Structure equations [review of PHAS0018 material]

For reference, we recall some basic equations of stellar (and atmospheric) structure:

2.1 Hydrostatic (pressure) equilibrium

Consider a volume element of density ρ, thickness dr, area dA. The gravitational force on the element (its weight) is the gravity times the mass:

$$g \times \rho \, \mathrm{d}A \, \mathrm{d}r.$$

In hydrostatic equilibrium (HSE), the downward force of gravity is balanced by the upwards pressure:[1]

$$\mathrm{d}P \, \mathrm{d}A = -g\rho \, \mathrm{d}A \, \mathrm{d}r; \quad \text{i.e.,}$$
$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho \, g \tag{2.1}$$

(the minus sign indicating that the two forces act in opposite directions). In spherical symmetry (e.g., for a star) the local gravity is $Gm(r)/r^2$, where $m(r)$ is the total mass contained within radius r

[1] Recall, pressure is force per unit area, so the force is pressure times area

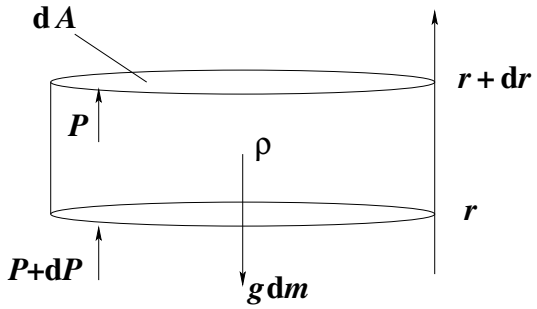


Figure 2.1: Hydrostatic equilibrium.

(measured from the centre), so in this case HSE can be expressed in the form

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = -\rho(r)\frac{Gm(r)}{r^2} \tag{2.2}$$

which is useful in discussing stellar atmospheres.

We can also express eqtn. (2.1) in terms of optical depth (anticipating eqtn. 4.6)[2]:

$$\frac{\mathrm{d}P(\tau)}{\mathrm{d}\tau_\nu} = \frac{g}{\kappa_\nu} \tag{2.3}$$

2.1.1 Equations of state

The principal sources of pressure throughout a 'normal' (non-degenerate) star are gas pressure, and radiation pressure.[3] We will take the corresponding equations of state to be, in general,

$$\begin{aligned} P_\mathrm{G} &= nkT; \\ &= (\rho kT)/(\mu m(\mathrm{H})) \end{aligned} \tag{2.4}$$

$$P_\mathrm{R} = \frac{1}{3}aT^4 \tag{3.21}$$

for number density n at temperature T, density ρ. Here μ is the mean molecular weight (§ 2.4), k is Boltzmann's constant, and $m(\mathrm{H})$ is the hydrogen mass; a is the radiation constant, $a = 4\sigma/c$, with σ the Stefan-Boltzmann constant.

2.2 Mass Continuity

The quantities $m(r)$ and $\rho(r)$ appearing in eqtn. (2.1) are not independent, but are related through the equation of mass continuity. For static configurations, the mass in a thin spherical shell of thickness $\mathrm{d}r$ at radius r is

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

that is,

$$\frac{\mathrm{d}m(r)}{\mathrm{d}r} = 4\pi r^2\rho(r). \tag{2.5}$$

In a dynamic medium, the mass passing through some unit area per unit time is just the mass density times the component of the velocity normal to the area. In particular, for a spherically symmetric flow, such as may apply in a stellar wind associated with mass-loss rate $\dot{M}$, we have

$$\dot{M} = 4\pi r^2 \rho(r) v(r) \tag{2.6}$$

where $\rho(r), v(r)$ are the density and (radial) flow velocity at radius r.

[2] $\mathrm{d}\tau_\nu = \pm k_\nu\,\mathrm{d}s = \pm\kappa_\nu\rho\,\mathrm{d}s$

[3] Electron degeneracy pressure is important in white dwarfs, and neutron degeneracy pressure in neutron stars. Magnetic pressure, $B^4/(4\pi)$, and turbulent pressure, $\rho v^2/2$, can also be significant under some circumstances.

2.3 Energy continuity

Our final equation of continuity is energy continuity; by inspection,

$$\frac{\mathrm{d}L}{\mathrm{d}r} = 4\pi r^2 \rho(r)\epsilon(r) \tag{2.7}$$

where

- r is radial distance measured from the centre of the star
- $P(r)$ is the total pressure at radius r
- $\rho(r)$ is the density at radius r
- $g(r)$ is the gravitational acceleration at radius r
- $m(r)$ is the mass contained with radius r
- $L(r)$ is the total energy transported through a spherical surface at radius r
- $\epsilon(r)$ is the energy generation rate per unit mass at radius r

The stellar radius is R, the stellar mass is $M \equiv m(R)$, and the emergent luminosity $L \equiv L(R)$ (dominated by radiation at the stellar surface).

2.4 Mean Molecular Weight

The 'mean molecular weight', μ, is[4] simply the average mass per particle in a gas, expressed in units of the hydrogen mass, $m(\mathrm{H})$. That is, the mean particle mass is $\mu m(\mathrm{H})$; since the number density of particles n is just the mass density ρ divided by the mean mass we have

$$n = \frac{\rho}{\mu m(\mathrm{H})} \qquad \text{and} \qquad P = nkT = \frac{\rho}{\mu m(\mathrm{H})} kT. \tag{2.8}$$

[For eqtn. 2.8 to be dimensionally correct, μ is necessarily dimensionless, consistent with the definition just given. Numerically, however, μ is equivalent to the molar mass expressed in g/mol.]

The mean molecular weight is trivially calculated if the fractional abundances *by number* are available for different nuclei and free electrons. In astrophysics, however, abundances are more often listed by

[4] Why 'molecular' weight for a potentially molecule-free gas or plasma? I don't know, though I suspect it may be because this nomenclature originated in other contexts (such as hydrostatic equilibrium in the Earth's atmosphere).

mass. For example, a medium having typical 'cosmic' or 'solar' abundances will have ~921 H nuclei, 78 He nuclei, and 1 heavier nucleus for every 1000 nuclei; but because the hydrogen nuclei weigh least, the respective *mass* fractions[5] are $X \simeq 0.74$ ($\sim 921/(1 \times 921 + 4 \times 78 + \sim 16 \times 1)$, $Y \simeq 0.25$, $Z \simeq 0.01$. In general, for an un-ionized gas

$$\mu = \left(X + \frac{Y}{4} + \sum_i (f_i / A_i) \right)^{-1}$$

where f_i, A_i are the mass fraction and atomic weight[6] of element i, which has atomic number z_i.[7]

For an ionized gas the total number density is the sum of the number densities of nuclei and free (i.e., unbound) electrons. For a *fully* ionized gas of mass density ρ these number densities are:

Element:	H	He	Metals
No. of nuclei	$\frac{X\rho}{m(\mathrm{H})}$	$\frac{Y\rho}{4m(\mathrm{H})}$	$\frac{Z\rho}{\overline{A}m(\mathrm{H})}$
No. of electrons	$\frac{X\rho}{m(\mathrm{H})}$	$\frac{2Y\rho}{4m(\mathrm{H})}$	$\frac{(\overline{A}/2)Z\rho}{\overline{A}m(\mathrm{H})}$

where $\overline{A}$ is the average atomic weight of metals (~16 for solar abundances), and for the final entry we make the approximation that $A_i \simeq 2z_i$. In this case

$$n \simeq \frac{\rho}{m(\mathrm{H})} (2X + {}^{3Y}/_{4} + {}^{Z}/_{2}) \equiv \frac{\rho}{\mu m(\mathrm{H})} \tag{2.9}$$

(where we have set ${}^{Z}/_{2} + {}^{Z}/_{\overline{A}} \simeq {}^{Z}/_{2}$, since $\overline{A} \gg 2$), and

$$\begin{aligned} \mu &\simeq (2X + {}^{3Y}/_{4} + {}^{Z}/_{2})^{-1} \\ &= 2\,(1 + 3X + Y/2)^{-1} \end{aligned} \tag{2.10}$$

(making use of the fact that $X + Y + Z \equiv 1$).

We can drop the approximations to obtain a more general (but less commonly used) definition,

$$\mu^{-1} = \sum_i \frac{z_i + 1}{A_i} f_i$$

(where, of course, the summation now includes hydrogen and helium).

[5] The mass fractions of hydrogen, helium, and metals are conventionally labelled X, Y, and Z (cf. Appendix H), where, in astrophysics, 'metals' should be read to mean 'all elements heavier than helium'.

[6] Averaged over isotopes if necessary.

[7] Of course, we could include hydrogen and helium implicitly in this definition, but it's normal practice to identify them separately.

By way of illustration:

for a neutral, pure-hydrogen gas ($X = 1$, $Y = Z = 0$), $\mu = 1$;

for a fully ionized pure-hydrogen gas[8] ($X = 1$, $Y = Z = 0$), $\mu = 1/2$;

for a neutral, pure-helium gas ($Y = 1$, $X = Z = 0$), $\mu = 4$;

for a fully ionized pure-helium gas ($Y = 1$, $X = Z = 0$), $\mu = 4/3$ (mass 4 distributed between three particles: a nucleus and two free electrons).

These simple examples serve to demonstrate that the mean molecular weight depends on both the chemistry (abundances) and the state of ionization. For a neutral gas of solar abundances ($X = 0.7381$, $Y = 0.2485$, $Z = 0.0134$), $\mu \simeq 1.25$; if all elements are singly ionized, there are twice as many particles per unit mass, and $\mu \simeq 1.25/2 = 0.61$, while a fully fully ionized solar-abundance plasma has $\mu \simeq 0.60$. (Because of its abundance, it is primarily the ionization state of hydrogen that matters.)

It's sometimes convenient to make use of the mean molecular weight per free electron, $\mu_e = \rho/(n_e m(\mathrm{H})$, or the mean molecular weight per ion, $\mu_i = \rho/n_i m(\mathrm{H})$. If we allow the term 'ions' to include neutral atoms, then the total number density is $n = n_e + n_i$, and $1/\mu = 1/\mu_e + 1/\mu_i$; and for a perfect-gas equation of state we can write (e.g.) the electron pressure as

$$P_e = \frac{\rho}{\mu_e m(\mathrm{H})} kT.$$

[8]The astute reader may notice an 'issue' here; the mean molecular weight of a gas of pure atomic hydrogen is, by our definition, exactly 1. However, the mean molecular weight of a fully ionized hydrogen gas is not *exactly* 0.5, because of binding energy – the sum of the masses of a free proton and a free electron is marginally greater than the mass of a hydrogen atom, so the mean molecular weight of an ionized hydrogen gas is actually a shade greater than 0.5. Such niceties are rarely of any consequence in general astrophysical studies.

Section 3

Radiation [review of 0018 material]

Almost all the astrophysical information we can derive about distant sources results from the radiation that reaches us from them. Our starting point is, therefore, a review of the principal ways of describing radiation. (In principle, this could include polarization properties, but we neglect that for simplicity).

The fundamental definitions of interest are of (specific) intensity and (physical) flux.

3.1 Specific Intensity, I_ν

The specific intensity (or radiation intensity, or surface brightness, or monochromatic radiance) is defined as:

> the rate of energy flowing at a given point,
> per unit area,
> per unit time,
> per unit frequency interval,
> per unit solid angle Ω (in azimuth ϕ and at direction θ to the normal; refer to the geometry sketched in Fig. 3.1)

or, expressed algebraically,

$$I_\nu(\theta,\phi) = \frac{\mathrm{d}E_\nu}{\mathrm{d}S\ \mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega} = \frac{\mathrm{d}E_\nu}{\mathrm{d}A\cos\theta\ \mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega} \quad [\mathrm{J\ m^{-2}\ s^{-1}\ Hz^{-1}\ sr^{-1}}]. \tag{3.1}$$

We've given a 'per unit frequency definition', but we can always switch to 'per unit wavelength' by noting that, for some frequency-dependent physical quantity 'X_ν', we can write

$$X_\nu\,\mathrm{d}\nu = X_\lambda\,\mathrm{d}\lambda$$

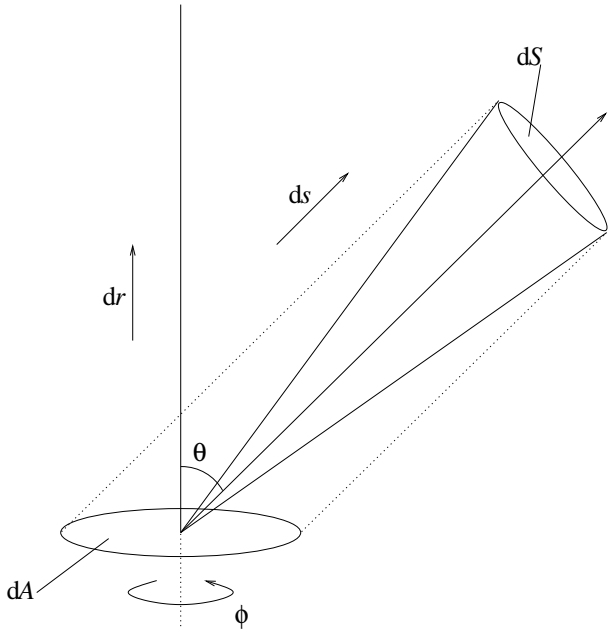


Figure 3.1: Geometry used to define radiation quantities. The element of area dA might, for example, be on the 'surface' of a star (whatever that means).

or

$$X_\lambda = X_\nu \left| \frac{\mathrm{d}\nu}{\mathrm{d}\lambda} \right|$$

(which has the same dimensionality on each side of the equation). Mathematically, $\mathrm{d}\nu/\,\mathrm{d}\lambda = -c/\lambda^2$, but physically the minus sign just reflects the fact that increasing frequency means decreasing wavelength (clearly, we require a positive physical quantity on either side of the equation). For example, specific intensity per unit wavelength is related to I_ν by

$$I_\lambda = I_\nu \left| \frac{\mathrm{d}\nu}{\mathrm{d}\lambda} \right| = I_\nu \frac{c}{\lambda^2} \quad [\mathrm{J\ m^{-2}\ s^{-1}\ m^{-1}\ sr^{-1}}]$$

where the θ, ϕ dependences are implicit (as will generally be the case in these notes). Equation (3.1) defines the monochromatic specific intensity ('monochromatic' will usually also be implicit); we can define a total intensity by integrating over frequency:

$$I = \int_0^\infty I_\nu \,\mathrm{d}\nu \quad [\mathrm{J\ m^{-2}\ s^{-1}\ sr^{-1}}].$$

3.1.1 Mean Intensity, J_ν

The mean intensity is, as the name suggests, the average of I_ν over solid angle Ω; it is of use when evaluating the rates of physical processes that are photon dominated but independent of the angular

distribution of the radiation (e.g., photoionization and photoexcitation rates). We define it as

$$J_\nu = \frac{\int_{4\pi} I_\nu \, \mathrm{d}\Omega}{\int_{4\pi} \mathrm{d}\Omega} = \frac{1}{4\pi} \int_{4\pi} I_\nu \, \mathrm{d}\Omega$$

or, since

$$\int_{4\pi} \mathrm{d}\Omega = \int_0^{2\pi} \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi,$$

$$J_\nu = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} I_\nu \, \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi \quad [\mathrm{J\ m^{-2}\ s^{-1}\ Hz^{-1}\ sr^{-1}}] \tag{3.2}$$

Introducing the standard astronomical nomenclature[1] $\mu = \cos\theta$ (whence $\mathrm{d}\mu = -\sin\theta \, \mathrm{d}\theta$), we have

$$\int_{4\pi} \mathrm{d}\Omega = \left(-\int_0^{2\pi} \int_{+1}^{-1} \mathrm{d}\mu \, \mathrm{d}\phi = \right) \int_0^{2\pi} \int_{-1}^{+1} \mathrm{d}\mu \, \mathrm{d}\phi \tag{3.3}$$

and eqtn. (3.2) becomes

$$J_\nu = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^{+1} I_\nu(\mu, \phi) \, \mathrm{d}\mu \, \mathrm{d}\phi \tag{3.4}$$

(where for clarity we show the μ, ϕ dependences of I_ν explicitly).

If the radiation field is independent of ϕ (i.e., has azimuthal symmetry, as may be the case in a stellar atmosphere without starspots, for example) then this simplifies to

$$J_\nu = \frac{1}{2} \int_{-1}^{+1} I_\nu(\mu) \, \mathrm{d}\mu. \tag{3.5}$$

(From this it is evident that if I_ν is completely isotropic – i.e., no $\theta[\equiv \mu]$ dependence, as well as no ϕ dependence – then $J_\nu = I_\nu$. This should be intuitively obvious; if the intensity is the same in all directions, then the mean intensity must equal the intensity [in any direction].)

3.2 Physical Flux, F_ν

The physical flux (or radiation flux density, or radiation flux, or just 'flux'[2]) is the net rate of energy flowing across unit area (e.g., at a detector), from all directions, per unit time, per unit frequency

[1] As we've seen, it's also completely standard to use μ for mean molecular weight; fortunately, the context rarely permits any ambiguity about which 'μ' is meant.

[2] A note for the pedantic, best ignored by others: few, if any, astronomers would have qualms about referring to F_ν as 'the flux'. However, more properly it is a flux *density;* formally, a flux $\mathcal{F}_\nu$ has (SI) units of J m^{-2} s^{-1}, and so $\mathcal{F}_\nu = \nu F_\nu$.

interval:

$$F_\nu = \frac{\int_{4\pi} \mathrm{d}E_\nu}{\mathrm{d}A\,\mathrm{d}t\,\mathrm{d}\nu} \quad [\mathrm{J\ m^{-2}\ s^{-1}\ Hz^{-1}}].$$

The absence of explicit directionality, and the concept of energy 'flow', are what distinguishes flux from intensity, but the two are clearly related. Using eqtn. (3.1) we see that

$$F_\nu = \int_{4\pi} I_\nu \cos\theta\,\mathrm{d}\Omega \tag{3.6}$$

$$= \int_0^{2\pi}\int_0^{\pi} I_\nu \cos\theta \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \quad [\mathrm{J\ m^{-2}\ Hz^{-1}}] \tag{3.7}$$

$$= \int_0^{2\pi}\int_{-1}^{+1} I_\nu(\mu,\phi)\mu\,\mathrm{d}\mu\,\mathrm{d}\phi$$

or, if there is no ϕ dependence,

$$F_\nu = 2\pi \int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu. \tag{3.8}$$

Because we're simply measuring the energy flowing across an area, there's no explicit directionality involved – other than if the energy impinges on the area from 'above' or 'below'.[3] For stellar-atmosphere considerations, it's therefore often convenient to divide the contributions to the flux into the 'upward' (emitted, or 'outward') radiation ($F_\nu^+; 0 \le \theta \le \pi/2$, Fig 3.1) and the 'downward' (incident, or 'inward') radiation ($F_\nu^-; \pi/2 \le \theta \le \pi$), with the net upward flux being $F_\nu = F_\nu^+ - F_\nu^-$:

$$F_\nu = \int_0^{2\pi}\int_0^{\pi/2} I_\nu \cos\theta \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi \qquad + \int_0^{2\pi}\int_{\pi/2}^{\pi} I_\nu \cos\theta \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$$

$$\equiv F_\nu^+ \qquad - F_\nu^-$$

As an important example, the surface flux emitted by a star is just F_ν^+ (assuming there is no incident external radiation field);

$$F_\nu = F_\nu^+ = \int_0^{2\pi}\int_0^{\pi/2} I_\nu \cos\theta \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi$$

[3] In principle, flux is a vector quantity, but the directionality is almost always implicit in astrophysical situations; e.g., from the centre of a star outwards, or from a source to an observer.

or, if there is no ϕ dependence,

$$= 2\pi \int_0^{\pi/2} I_\nu \cos\theta \sin\theta \, \mathrm{d}\theta.$$

$$= 2\pi \int_0^{+1} I_\nu(\mu) \mu \, \mathrm{d}\mu. \tag{3.9}$$

If, furthermore, I_ν has no θ dependence over the range 0–$\pi/2$ then

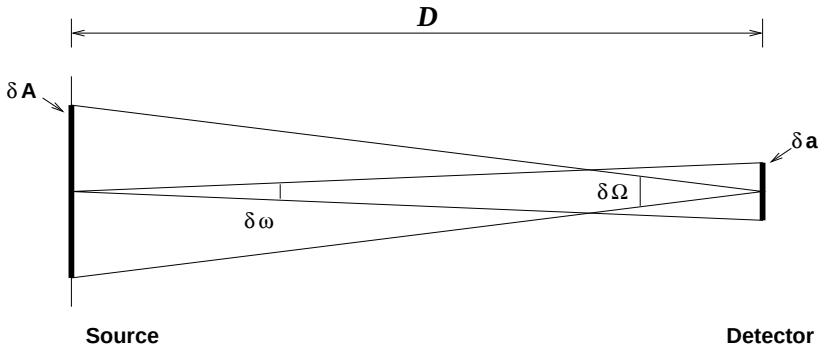
$$F_\nu = \pi I_\nu \tag{3.10}$$

(since $\int_0^{\pi/2} \cos\theta \sin\theta \, \mathrm{d}\theta = 1/2$). In reality, I_ν is never independent of θ; nevertheless, if for some reason we make the approximation that $I_\nu \simeq B_\nu$, the Planck function, then $F_\nu = \pi B_\nu$.

3.3 Flux vs. Intensity

If I_ν is completely isotropic, then $F_\nu^+ = F_\nu^-$, and $F_\nu = 0$. This highlights an important point: essentially, an intensity (particularly the mean intensity) describes how much radiation there is at some point, while flux measures the nett flow of radiation. One can have very large intensities associated with relatively small fluxes (cf., e.g., Box 12.2).

In the context of observations of astronomical sources, a related difference between F_ν and I_ν is that while the physical flux falls off as r^{-2}, the specific intensity is independent of distance to the source[4] (but requires the source to be resolved). This can be understood simply by noting that specific intensity is defined in terms of 'the rate of energy flow per unit area of surface. . . *per unit solid angle*'. The energy flow per unit area falls off as r^{-2}, but the area per unit solid angle increases as r^2, and so the two cancel.



[4] In each case, true only if intervening material doesn't attenuate the radiation.

To express this more formally: suppose some area δA on a source at distance D subtends a solid angle $\delta\Omega$ at a detector; while the detector, area δa, subtends a solid angle $\delta\omega$ at the source. The energy emitted towards (and received by) the detector is

$$\begin{aligned} E &= I_\nu \, \delta A \, \delta\omega; \text{ but} \\ \delta A &= D^2 \delta\Omega \text{ and } \delta\omega = \delta a / D^2, \text{ so} \\ \frac{E}{\delta\Omega} &= I_\nu D^2 \, \frac{\delta a}{D^2}; \end{aligned}$$

that is, the energy received per unit solid angle (i.e., the intensity) is distance independent. In astronomical parlance, we say that the surface brightness of source is distance independent (in the absence of additional processes, such as interstellar extinction).

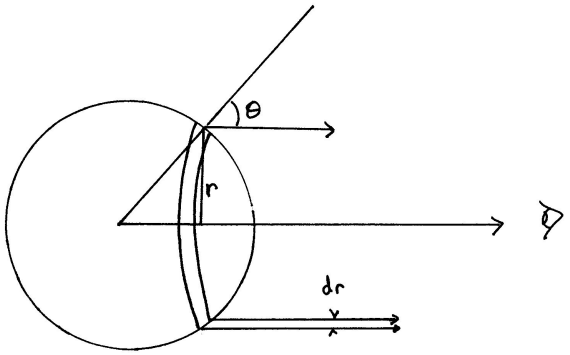
A source must be spatially resolved for us to be able to measure the intensity; otherwise, we can measure 'only' the flux – if the source is unresolved, we can't identify different directions towards it. Any spatially extended source will, at some large enough distance D, produce an image source at the focal plane of a telescope that will be smaller than the detector (pixel) size. For such an unresolved source, the detected energy is

$$\begin{aligned} E &= I_\nu \, \delta a \, \delta\Omega \\ &= I_\nu \, \delta a \, \frac{\delta A}{D^2} \end{aligned}$$

and we recover the expected inverse-square law for the detected flux.

3.3.1 Flux from a star

As an example, consider the flux from a spherical star at distance D.



The observer sees the projected area of the annulus as

$$\mathrm{d}A = 2\pi r\,\mathrm{d}r$$

and since

$$r = R\,\sin\theta \qquad (\text{and } \mathrm{d}r = R\cos\theta\,\mathrm{d}\theta)$$

we have

$$\begin{aligned}\mathrm{d}A &= 2\pi R\sin\theta\; R\cos\theta\,\mathrm{d}\theta \\ &= 2\pi R^2\sin\theta\cos\theta\,\mathrm{d}\theta \\ &= 2\pi R^2\mu\,\mathrm{d}\mu\end{aligned}$$

where as usual $\mu = \cos\theta$. The annulus therefore subtends a solid angle

$$\mathrm{d}\Omega = \frac{\mathrm{d}A}{D^2} = 2\pi\left(\frac{R}{D}\right)^2\mu\,\mathrm{d}\mu.$$

The flux received from this solid angle is

$$\mathrm{d}f_\nu = I_\nu(\mu)\,\mathrm{d}\Omega$$

so that the total observed flux is

$$f_\nu = 2\pi\left(\frac{R}{D}\right)^2\int_0^1 I_\nu\mu\,\mathrm{d}\mu$$

or, using eqtn. (3.9),

$$\begin{aligned}&= \left(\frac{R}{D}\right)^2 F_\nu \\ &= \theta_*^2 F_\nu \qquad [\mathrm{J\ m^{-2}\ s^{-1}\ Hz^{-1}}]\end{aligned}$$

where θ_* is the solid angle subtended by the star (measured in radians).

3.4 Moments of the radiation field

The moments of the radiation field are of great importance in the study of radiation transport in stellar atmospheres (e.g., §8). We define the n^{th} moment of the radiation field as

$$\begin{aligned}M_\nu(n) &\equiv \frac{\int_{4\pi} I_\nu(\mu,\phi)\mu^n\,\mathrm{d}\Omega}{\int_{4\pi}\mathrm{d}\Omega} \\ &= \frac{1}{4\pi}\int_{4\pi} I_\nu(\mu,\phi)\mu^n\,\mathrm{d}\Omega \\ &= \frac{1}{4\pi}\int_0^{2\pi}\int_{-1}^{+1} I_\nu(\mu,\phi)\mu^n\,\mathrm{d}\mu\,\mathrm{d}\phi, \qquad = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^n\,\mathrm{d}\mu \quad \text{for azimuthal symmetry.}\end{aligned}$$

We've therefore already encountered the zeroth-order moment, which is the mean intensity:

$$J_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\,\mathrm{d}\mu. \tag{3.5}$$

We have previously written the flux as

$$F_\nu = 2\pi\int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu; \tag{3.8}$$

to cast this in the same form as eqtns. (3.5) and (3.5), we define the 'Eddington flux' as $H_\nu = F_\nu/(4\pi)$, i.e.,

$$H_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu. \tag{3.11}$$

We see that H_ν is the first-order moment of the radiation field.

The second-order moment, the so-called 'K integral', is, from the definition of moments,

$$K_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^2\,\mathrm{d}\mu \tag{3.12}$$

We will see in Section 3.7 that the K integral is straightforwardly related to the radiation pressure, $P_\nu(= 4\pi K_\nu/c)$.

In the special case that I_ν is isotropic we can take it out of the integration over μ, and

$$\begin{aligned} K_\nu &= \frac{1}{2}\frac{\mu^3}{3} I_\nu\Big|_{-1}^{+1} \\ &= \frac{1}{3} I_\nu \qquad \left[\text{also} = \frac{1}{3}J_\nu \text{ for isotropy}\right] \end{aligned} \tag{3.13}$$

Higher-order moments are not used in general. So, to recap (and using the notation first introduced by Eddington himself), for $n = 0, 1, 2$:

$n = 0$	Mean Intensity	$J_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\,\mathrm{d}\mu$
$n = 1$	Eddington flux	$H_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu$
$n = 2$	K integral	$K_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^2\,\mathrm{d}\mu$

(all with units [J m^{-2} s^{-1} Hz^{-1} sr^{-1}]).

We can also define the integral quantities

$$J = \int_0^\infty J_\nu\,\mathrm{d}\nu$$
$$F = \int_0^\infty F_\nu\,\mathrm{d}\nu$$
$$K = \int_0^\infty K_\nu\,\mathrm{d}\nu$$

3.5 Other ‘Fluxes’, ‘Intensities’

Astronomers can be rather careless in their use of the terms ‘flux’ and ‘intensity’. The ‘fluxes’ and ‘intensities’ discussed so far can all be quantified in terms of physical (e.g., SI) units. Often, however, astronomical signals are measured in more arbitrary ways (such ‘integrated signal at the detector’, ‘counts’, or even ‘photographic density’); in such cases, it’s not unusual to see the signal referred to as a ‘flux’ (or ‘intensity’) in a spectrum, but this is just a loose shorthand, and doesn’t allude to the quantities defined in this section.

There are other physically-based quantities that one should be aware of. For example, discussions of model stellar atmospheres may refer to the ‘astrophysical flux’; this is given by F_ν/π (also called, rarely, the ‘radiative flux’), which is evidently similar to the Eddington flux, $H_\nu = F_\nu/(4\pi)$, which has itself occasionally been referred to as the ‘Harvard flux’. Different authors may refer to any of what we’ve called the ‘physical flux’, ‘astrophysical flux’, ‘Eddington flux’ etc. as simply ‘the flux’.

3.6 Radiation Energy Density, U_ν

Consider some volume of space containing a given number of photons; the photons have energy, so we can discuss the density of radiant energy. From eqtn. (3.1), and referring to Fig. 3.1,

$$\mathrm{d}E_\nu = I_\nu(\theta)\,\mathrm{d}S\;\mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega.$$

We can eliminate the time dependence[5] by noting that there is a single-valued correspondence[6] between time and distance for radiation. Defining a characteristic length $\ell = ct$, $\mathrm{d}t = \mathrm{d}\ell/c$, and

$$\begin{aligned}\mathrm{d}E_\nu &= I_\nu(\theta)\,\mathrm{d}S\,\frac{\mathrm{d}\ell}{c}\,\mathrm{d}\nu\,\mathrm{d}\Omega \\ &= \frac{I_\nu(\theta)}{c}\,\mathrm{d}V\,\mathrm{d}\nu\,\mathrm{d}\Omega \end{aligned} \tag{3.14}$$

where the volume element $\mathrm{d}V = \mathrm{d}S\;\mathrm{d}\ell$. We integrate over solid angle and volume to obtain the mean radiation energy density per unit frequency (implicitly, per unit volume):

$$\begin{aligned}U_\nu\,\mathrm{d}\nu &= \frac{1}{V}\int_V\int_\Omega \mathrm{d}E_\nu \\ &= \frac{1}{c}\int_{4\pi} I_\nu\,\mathrm{d}\nu\,\mathrm{d}\Omega\end{aligned}$$

[5] Assuming that no time dependence exists; that is, that for every photon leaving some volume of space, a compensating photon enters. This is an excellent approximation under many circumstances.

[6] Well, nearly single-valued; the speed at which radiation propagates actually depends on the refractive index of the medium through which it moves – e.g., the speed of light in water is only ¾c. The actual value doesn’t matter in the present context, as the implied volumes cancel out.

whence

$$U_\nu = \frac{1}{c}\int_{4\pi} I_\nu \, \mathrm{d}\Omega$$
$$= \frac{4\pi}{c} J_\nu \qquad [\mathrm{J\ m^{-3}\ Hz^{-1}}] \qquad \left[\text{ from eqtn. (3.2): } J_\nu = {}^1\!/_{4\pi}\textstyle\int I_\nu \, \mathrm{d}\Omega \right] \tag{3.15}$$

Again, this is explicitly frequency dependent; the total energy density is obtained by integrating over frequency:

$$U = \int_0^\infty U_\nu \, \mathrm{d}\nu.$$

For black-body radiation, $J_\nu (= I_\nu) = B_\nu$, and

$$U = \int_0^\infty \frac{4\pi}{c} B_\nu \, \mathrm{d}\nu$$

but $\int \pi B_\nu = \sigma T^4$ (eqtn. (C.7)) so

$$U = \frac{4\sigma}{c} T^4 \equiv aT^4$$
$$= 7.55 \times 10^{-16}\, T^4 \qquad \mathrm{J\ m^{-3}} \tag{3.16}$$

where T is in kelvin, σ is the Stefan-Boltzmann constant, and a is the 'radiation constant'. Note that the energy density of black-body radiation is a fixed quantity (for a given temperature).

For a given form of spectrum, the energy density in radiation must correspond to a specific number density of photons:

$$n_{\mathrm{phot}} = \int_0^\infty \frac{U_\nu}{h\nu} \, \mathrm{d}\nu.$$

For the particular case of a black-body spectrum, this evaluates to

$$n_{\mathrm{phot}} \simeq 2 \times 10^7\, T^3 \qquad \text{photons m}^{-3}. \tag{3.17}$$

Dividing eqtn. (3.16) by (3.17) gives the mean energy per photon for black-body radiation,

$$\overline{h\nu} = 3.78 \times 10^{-23} T = 2.74kT \tag{3.18}$$

(although there is, of course, a broad spread in the energies of individual photons).

3.7 Radiation pressure, and the equation of state for radiation

A photon carries momentum $E/c\,(= h\nu/c)$.[7] Because there is always a flux of photons, there is a corresponding momentum flux; and a momentum flux (momentum per unit time, per unit area) represents *pressure*.[8] If photons encounter a surface at some angle θ to the normal, the component of momentum perpendicular to the surface per unit time per unit area is that pressure,

$$\mathrm{d}P_\nu = \frac{\mathrm{d}E_\nu}{c} \times \cos\theta \, \frac{1}{\mathrm{d}t\,\mathrm{d}A\,\mathrm{d}\nu}$$

(where we have chosen to express the photon pressure 'per unit frequency', $\mathrm{d}\nu$); but the specific intensity is

$$I_\nu = \frac{\mathrm{d}E_\nu}{\mathrm{d}A\cos\theta\,\mathrm{d}\Omega\,\mathrm{d}\nu\,\mathrm{d}t}, \tag{3.1}$$

whence

$$\mathrm{d}P_\nu = \frac{I_\nu}{c}\cos^2\theta\,\mathrm{d}\Omega$$

i.e.,

$$P_\nu = \frac{1}{c}\int_{4\pi} I_\nu \mu^2\,\mathrm{d}\Omega \qquad [\mathrm{J\ m^{-3}\ Hz^{-1}} \equiv \mathrm{Pa\ Hz^{-1}}] \tag{3.19}$$

We know that

$$\int_{4\pi}\mathrm{d}\Omega = \int_0^{2\pi}\int_{-1}^{+1}\mathrm{d}\mu\,\mathrm{d}\phi \tag{3.3}$$

so

$$P_\nu = \frac{2\pi}{c}\int_{-1}^{+1} I_\nu\mu^2\,\mathrm{d}\mu.$$

We encountered the K integral (the second moment of the radiation field),

$$K_\nu = \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^2\,\mathrm{d}\mu, \tag{3.12}$$

in Section 3.4; we now see that

$$P_\nu = \frac{4\pi}{c}K_\nu. \tag{3.20}$$

[7] It may help to note that, classically, momentum is mass times velocity. From $E = mc^2 = h\nu$, the photon rest mass is $h\nu/c^2$, and its velocity is c, hence momentum is $h\nu/c$.

[8] Dimensional arguments validate this; in the SI system, momentum has units of kg m $\mathrm{s^{-1}}$, and momentum flux has units of kg m $\mathrm{s^{-1}}$, $\mathrm{m^{-2}}$, $\mathrm{s^{-1}}$; i.e., kg $\mathrm{m^{-1}\ s^{-2}}$, = N $\mathrm{m^{-2}}$ = Pa – the units of pressure. Pressure in turn is force per unit area (where force is measured in Newtons, = J $\mathrm{m^{-1}}$ = kg m $\mathrm{s^{-2}}$).

For an isotropic radiation field we know that $K_\nu = {}^1\!/_3 I_\nu = {}^1\!/_3 J_\nu$ (eqtn. 3.13), and so

$$P_\nu = \frac{4\pi}{3c} I_\nu = \frac{4\pi}{3c} J_\nu.$$

In this isotropic case we also have

$$U_\nu = \frac{4\pi}{c} J_\nu \quad = \frac{4\pi}{c} I_\nu$$

(eqtn. 3.15) so – for an isotropic radiation field – the radiation pressure is

$$P_\nu = \frac{1}{3} U_\nu.$$

For black-body radiation we can integrate the radiation energy density analytically over frequency ($\int U_\nu \, \mathrm{d}\nu = aT^4 = {}^{4\sigma}\!/_c T^4$; eqtn. 3.16), giving

$$P_\mathrm{R} = \frac{1}{3} aT^4 = \frac{4\sigma}{3c} T^4 \qquad [\mathrm{J\ m^{-3}} \equiv \mathrm{N\ m^{-2}} \equiv \mathrm{Pa}]. \tag{3.21}$$

In that it expresses the relationship between pressure and temperature, eqtn. (3.21) is the equation of state for (black-body) radiation.

3.8 The Eddington limit

In the isotropic case, P_R (or P_ν) is a scalar quantity – it has magnitude but no explicit direction (like air pressure, locally, on Earth). For an anisotropic radiation field, the radiation pressure has a direction (normally outwards from a star), and is a vector quantity. It will not be surprising that radiation pressure plays a particularly important role in luminous stars; and at their photospheres the radiation field is, evidently, no longer even approximately isotropic (since the radiation is escaping from the surface).

If we consider some spherical surface at distance r from the energy-generating centre of a star, where all photons are flowing outwards (i.e., the surface of the star), the total photon momentum flux, per unit area per unit time, is

$$\frac{L}{c} \Big/ (4\pi r^2) \qquad [\mathrm{J\ m^{-3}} = \mathrm{kg\ m^{-1}\ s^{-2}}]$$

(recalling that the momentum of a single photon is $E/c, = h\nu/c$, so the total momentum is L/c). A major opacity source in hot-star atmospheres is Thomson (electron) scattering, for which the cross-section is

$$\sigma_\mathrm{T} \left[= \frac{8\pi}{3} \left(\frac{e^2}{m_\mathrm{e} c^2} \right)^2 \right] = 6.7 \times 10^{-29}\ \mathrm{m^2}$$

(§5.5.1). The radiation force exerted on an electron is electron is

$$F_\mathrm{R} = \frac{\sigma_\mathrm{T} L}{4\pi r^2 c} \qquad [\mathrm{J\ m^{-1}} \equiv \mathrm{N}]$$

which is then transmitted to positive ions by electrostatic interactions. For stability, this outward force must be no greater than the inward gravitational force; for equality (and assuming that the atmosphere consists simply of fully ionized hydrogen)

$$\frac{\sigma_{\rm T} L}{4\pi r^2 c} = \frac{GM\,(m({\rm H}) + m_{\rm e})}{r^2} \simeq \frac{GMm({\rm H})}{r^2}.$$

This equality gives a limit on the maximum luminosity as a function of mass for a stable star – the Eddington Luminosity,

$$\begin{aligned} L_{\rm Edd} &= \frac{4\pi GMc\, m({\rm H})}{\sigma_{\rm T}} \\ &\simeq 1.3 \times 10^{31} \frac{M}{{\rm M}_\odot} \qquad [{\rm J\ s^{-1}}], \text{ or} \\ \frac{L_{\rm Edd}}{{\rm L}_\odot} &\simeq 3.4 \times 10^4 \frac{M}{{\rm M}_\odot} \end{aligned} \tag{3.22}$$

Since luminosity is proportional to mass to some power (roughly, $L \propto M^{3-4}$ on the upper main sequence; cf. Section 12.4), we infer that the Eddington Luminosity imposes an upper limit on the mass of stable stars.

(In practice, instabilities may cause a super-Eddington atmosphere to become clumpy, or 'porous'; radiation is then able to escape through paths of reduced optical depth between the clumps. Nevertheless, the Eddington limit represents a good approximation to the upper limit to stellar luminosity. We see this limit as an upper bound to the distribution of stars in the Hertzsprung-Russell diagram, the so-called 'Humphreys-Davidson limit'.)

Section 4

The interaction of radiation with matter

[review of 0018 material]

4.1 Emission: increasing intensity

A common astrophysical[1] definition of the (monochromatic) emissivity is the energy generated per unit volume,[2] per unit time, per unit frequency, per unit solid angle:

$$j_\nu = \frac{\mathrm{d}E_\nu}{\mathrm{d}V\,\mathrm{d}t\,\mathrm{d}\nu\,\mathrm{d}\Omega} \quad [\mathrm{J\ m^{-3}\ s^{-1}\ Hz^{-1}\ sr^{-1}}]. \tag{4.1}$$

The change (increase) in I_ν along an element of distance $\mathrm{d}s$ caused by the emissivity of a volume of material of unit cross-sectional area is just:

$$\mathrm{d}I_\nu = +j_\nu \mathrm{d}s. \tag{4.2}$$

4.2 Extinction: decreasing intensity

'Extinction' is a general term for the removal of light from a beam. Two different classes of process contribute to the extinction: absorption and scattering. Absorption (sometimes called 'true absorption') results in the destruction of photons; scattering merely involves redirecting photons in some new direction. For a beam directed towards the observer, scattering still has the effect of diminishing the recorded signal, so for the moment the two types of process can be treated together.

The amount of intensity removed from a beam by extinction in (say) a gas cloud must depend on

– The initial strength of the beam (the more light there is, the more you can remove)

[1] Other definitions of 'emissivity' occur in physics.

[2] The emissivity can also be defined per unit mass (j_ν/ρ); or, in principle, per particle.

– The number of particles (absorbers/scatterers)

– The microphysics of the particles – specifically, how likely they are to absorb (or scatter) an incident photon. This microphysics is characterized by an effective cross-section per particle presented to the radiation.

By analogy with eqtn. (4.2), we can therefore write the change (decrease) in intensity along length $\mathrm{d}s$ as

$$\mathrm{d}I_\nu = -a_\nu n I_\nu \,\mathrm{d}s \tag{4.3}$$

for a number density of n extinguishing particles per unit volume, with a_ν the 'extinction coefficient', or cross-section (in units of area) per particle.

4.3 Opacity

In astrophysical applications, it is customary to combine the cross-section per particle (having dimensions of area) with the number of particles (having dimensions of 'per volume'); i.e.,

$$a_\nu n \equiv k_\nu (= k_\nu^{\mathrm{a}} + k_\nu^{\mathrm{s}})$$

where k_ν, the (volume) opacity, has dimensions of area per unit volume (i.e., length), or SI units of m^{-1}; and we have acknowledged the separate contributions of absorption and scattering processes.

We can now write eqtn. (4.3) as

$$\mathrm{d}I_\nu = -k_\nu I_\nu \,\mathrm{d}s, \tag{4.4}$$

recalling that k_ν may have contributions from both true absorption *and* scattering (k_ν^{a}, k_ν^{s}).

Alternatively, expressed as 'per unit mass', the opacity is

$$a_\nu n \equiv \kappa_\nu \rho$$

whence

$$\mathrm{d}I_\nu = -\kappa_\nu \rho(s) I_\nu \,\mathrm{d}s$$

for mass density ρ, where κ_ν is the (monochromatic) mass extinction coefficient or, more usually, the opacity per unit mass (dimensions of area per unit mass; SI units of $\mathrm{m}^2\ \mathrm{kg}^{-1}$).

The opacity is related to the probability that, at a given point, a photon of frequency ν will interact with matter; that is, in zones with high opacity, the radiation field and matter are strongly coupled, while

zones with small opacity the photons flow freely. In this context, the volume opacity has a straightforward and useful physical interpretation – it is the reciprocal of the photon mean free path, ℓ_ν:

$$\ell_\nu \equiv 1/k_\nu.$$

(This will be demonstrated Section 7.2.)

[Be aware that the nomenclature we use here is not universally adopted; in the literature, κ is generally used to indicate opacity 'per unit mass', but some authors use it for opacity 'per unit volume'; other authors use, e.g., a_ν in place of k_ν; and so on.]

4.4 Optical depth

We can often calculate, but rarely measure, opacity as a function of position along a given path. Observationally, all that is usually accessible is the cumulative effect of the opacity integrated along the line of sight; this is quantified by the optical depth,

$$\tau_\nu = \int_0^D k_\nu(s)\,\mathrm{d}s \qquad = \int_0^D \kappa_\nu\,\rho(s)\,\mathrm{d}s \qquad = \int_0^D a_\nu\,n(s)\,\mathrm{d}s \tag{4.5}$$

over distance D.

Similarly, the parameter 's' that expresses the physical location along the photon propagation direction is often observationally inaccessible; consequently, it is useful (or essential) to change the independent variable and express the radiative and matter variables (I_ν, ρ, T, etc) in terms of τ_ν. The differential relation between $d\tau_\nu$ and ds is

$$\mathrm{d}\tau_\nu = \pm k_\nu\,\mathrm{d}s = \pm\kappa_\nu\rho\,\mathrm{d}s \tag{4.6}$$

where the '$\pm$' depends on the geometry applicable in a given situation – '+' if s and τ are chosen to increase in the same direction, '−' otherwise. (In stellar atmospheres, physical distance is often measured 'outwards', while optical depth is conveniently measured 'inwards'.)

A simple illustration of optical depth is afforded by the effect of interstellar extinction. A star at some fixed distance D undergoes different amounts of extinction at different wavelengths – over the optical regime, more blue light than red is removed. We understand this to reflect the wavelength dependence of the opacity.

Interstellar extinction is often measured and expressed photometrically; the ratio of the observed flux, f_λ, to the flux that would be observed in the absence of extinction, $f_\lambda(0)$, is expressed in magnitudes as

$$\begin{aligned} A_\lambda &= -2.5\log_{10}\left(\frac{f_\lambda}{f_\lambda(0)}\right) \\ &= -2.5\log_{10}\left(\exp\{-\tau_\lambda\}\right), \qquad = 1.086\tau_\lambda \end{aligned}$$

(using eqtn. 7.3 to relate flux ratio to optical depth).

Section 5

Opacity sources

5.1 Introduction

Recall the distinction between absorption and scattering opacity sources:

- Under most circumstances of interest to us, *scattering* effectively involves a photon emerging from the scattering event with virtually the same energy (in the rest frame of the scatterer), but in a new direction.[1]
- *Absorption* processes can be regarded as resulting in the destruction of a photon (through conversion into other forms of energy, whether radiative or kinetic [thermal]; the inverse processes are sources of emission).

Both general mechanisms can give rise to continuous opacities (across a wide range of frequencies/wavelengths) or to line opacities (across a narrow range of frequencies). Processes which can be important in the stellar-astrophysics context include:

bound–bound	line	absorption, or scattering	(scat.: photoexcitation followed promptly by radiative decay through the same channel)
ro-vibrational	line	absorption	changes in molecular quantum states
bound–free	cont.	absorption	photoionization
free–free	cont.	absorption	bremsstrahlung
electron scattering	cont.	scattering	of photons by free electrons
Rayleigh scattering	cont.	scattering	of photons by bound electrons

[1] In Compton scattering, energy is transferred from a high-energy photon to the scattering electron (or vice versa for inverse Compton scattering). These processes are important at X-ray and γ-ray energies; at lower energies, classical Thomson scattering dominates. For our purposes here, 'electron scattering' can be regarded as synonymous with Thomson scattering. Note, too that by 'electron scattering', astronomers almost always mean scattering *by* electrons – not *of* electrons.

[This list is not comprehensive; in particular, high-energy processes are largely excluded. For example, Compton scattering (the general form of electron scattering), pair production (decay of a photon to particle pair), and photodissociation of nuclei may need to be considered in extreme environments.]

5.2 Mean opacities

The calculation of opacities is a considerable undertaking, even for time-independent situations.[2] The analyst has to consider an arbitrarily large number of transitions, for an arbitrarily large range of physical conditions. The work is conceptually straightforward (for an atomic or molecular physicist), but represents a substantial computational challenge, requiring vast quantities of atomic and molecular data. The challenge has been largely addressed for stellar interiors (where a number of simplifications are possible), particularly through the OPAL and OPACITY PROJECT (OP) consortia; and while sophisticated numerical methods have allowed good progress for atomic and ionic transitions, the calculation of the accurate data required for molecular opacities is an ongoing industry.

To compute a detailed emergent spectrum, a full, frequency-dependent description of all important opacities is required. However, for some applications (particularly when computing the overall radiative energy transport – for example, in interior structures), we just want to know how much radiative energy gets through in total, without being too bothered about the detailed frequency dependence. In these cases we still need to *start* with some knowledge of the frequency dependences of both the opacities and the radiation field, since most energy is transported at those frequences where the flux is high and the opacity is low. However, for practical purposes we can construct a variety of frequency-independent *mean* opacities, choosing suitable frequency-dependent weighting factors related to the radiation field. Precomputing these mean opacities, and tabulating them as functions of temperature and density, is a one-time exercise that allows for considerable efficiencies in calculation of stellar-interior structures, for example.

5.2.1 Planck mean opacity [not for lectures]

Perhaps the most straightfoward 'mean opacity' is the flux-weighted, frequency-integrated Planck mean, defined as

$$\overline{k}_{\rm P}(\rho, T) = \frac{\int_0^\infty k_\nu^a(\rho, T)\, B_\nu(T)\, {\rm d}\nu}{\int_0^\infty B_\nu(T)\, {\rm d}\nu},$$

[2] Time-independent opacities are applicable when astrophysical conditions (e.g., pressure, temperature, abundances) vary on much longer timescales than the timescales for microscopic processes (photoionization, excitation, etc.). This is true in all circumstances considered in these notes.

which weights the opacities by the intensity of radiation. At frequencies where the intensity is very small, the opacities don't much matter in the context of energy transport, because there's just so little energy to worry about. Note that the Planck mean is obtained by averaging over only absorption opacities; scattering cancels out.

The Planck mean opacity has applications to energy transport in an optically thin plasma. The weighting function, $B_\nu(T)$, peaks at $h\nu \simeq 2.7kT$ (eqtn. 3.18); that is, opacities around that frequency have the greatest effect.

5.2.2 The Rosseland mean

The Norwegian astrophysicist Sven Rosseland introduced another form of frequency-averaged mean opacity, now called the Rosseland mean opacity, $\overline{k}_\mathrm{R}$ ($= \overline{\kappa}_\mathrm{R}\rho$). It is defined by

$$\frac{1}{\overline{k}_\mathrm{R}(\rho,T)}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu \equiv \int_0^\infty \frac{1}{k_\nu(\rho,T)}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu; \tag{5.1}$$

i.e.,

$$\frac{1}{\overline{k}_\mathrm{R}} = \frac{\int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu}{\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu} \tag{5.2}$$

where $B_\nu(T)$ is the Planck function,

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\Big/\left(\exp\{h\nu/kT\}-1\right),$$

whence

$$\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T} = \frac{2h^2\nu^4}{c^2kT^2}\frac{\exp\{h\nu/kT\}}{\left(\exp\{h\nu/kT\}-1\right)^2}.$$

We see that the Rosseland mean is obtained from the weighted sum of the *inverse* of all (absorption + scattering) opacity sources – that is, it is the weighted harmonic mean of k_ν, where
(i) the calculation of $\overline{k}_\mathrm{R}(\rho,T)$ favours frequency regions of *low* opacity;
(ii) the weighting factor peaks at $h\nu_\mathrm{p} = 3.8kT$, and is small for very low and very high frequencies.

The Rosseland mean opacity isn't some arbitrary construct, but was derived as the appropriate description of the mean opacity required to yield the correct value for the frequency-integrated radiative energy flux in an *optically thick* medium; we shall investigate this in Section 12 (eqtn. 12.4). It therefore offers a number of advantages for analytical work in stellar structure and stellar atmospheres; and it remains (in only slightly modified form) an important and commonly used characterization of opacities in stellar interiors for modern computational studies.

5.3 Free–free absorption

We begin our discussion of individual astrophysical opacity sources with free–free absorption, partly because its Rosseland mean is straightforward to evaluate, and introduces us to 'Kramers' opacity', a useful approximate analytical representation of several opacity sources.

Free–free processes arise when an unbound electron interacts with an ion, but is not captured – i.e., it is a free electron before the interaction, and is a free electron afterwards. Conceptually, as the electron moves through the electric field[3] of the ion it is initially accelerated, gaining energy, and subsequently decelerated, losing energy; in general, the electron/ion pair emerges with a different kinetic energy to the initial state, and the nett effect is that a photon is emitted or absorbed.

Although our aim in this section is to examine free–free *absorption*, it is convenient to start by considering the inverse process of free–free *emission* (often referred to a 'bremsstrahlung', usually translated as 'braking radiation'). As an ansatz, we might reasonably anticipate that the strength of free–free emission will depend on the number of particles, and the frequency dependence on their velocity distribution of the particles. A particularly important case is when the electron velocity distribution is Maxwellian, giving rise to 'thermal bremsstrahlung', whose form we can infer from general principles.

• First, for a thermal (Maxwellian) velocity distribution, the number of electrons with kinetic energy E is proportional to the Boltzmann factor, $\exp(-E/kT)$; so we reasonably expect any corresponding free–free radiation to reflect this, and hence to have a frequency dependence proportional to $\exp(-h\nu/kT)$.

• Secondly, since free–free emission involves an interaction between ions and electrons, the rate of emission per unit volume must be proportional to the product of the densities, $n_\mathrm{i} n_\mathrm{e}$; but for a fully ionized plasma, $n_\mathrm{e} = \bar{Z} n_\mathrm{i}$ (where $\bar{Z}$ is the average atomic number), so the strength should depend on $n_\mathrm{i}^2 (\propto n_\mathrm{e}^2, \propto \rho^2)$.

• Finally, the mean thermal velocity of the electrons scales with temperature as $T^{1/2}$ (since $kT \propto mv^2/2$), so the time over which they are able to interact radiatively with (much more massive, hence relatively static) ions is $\propto T^{-1/2}$.

Hence, overall, we expect the volume *emissivity* of free–free radiation to scale as

$$j_\nu \propto T^{-1/2} n_\mathrm{i}^2 \exp(-h\nu/kT), \qquad \propto T^{-1/2} \rho^2 \exp(-h\nu/kT).$$

Returning to the question of free–free *absorption*, Kirchhoff's law tells us that, in thermodynamic

[3] A free electron may also be accelerated by a magnetic field, giving rise to cyclotron radiation or, for extremely relativistic particles, synchrotron radiation. (Cyclotrons and synchrotrons were particle accelerators developed in the 1930s and 1940s.)

equilibrium,

$$j_\nu = k_\nu B_\nu, \textit{ i.e.,}$$
$$\propto k_\nu \nu^3 \left(\exp\{h\nu/kT\} - 1\right)^{-1}$$

(using the Planck function), so our heuristic result for free–free opacity is

$$k_\nu \propto \rho^2 \nu^{-3} T^{-1/2} \left(1 - \exp\{-h\nu/kT\}\right).$$

A more rigorous analysis gives virtually the same form,

$$k_\nu = 3.7 \times 10^{10}\, Z^2 n_\mathrm{e} n_\mathrm{i} \nu^{-3} T^{-1/2} \left(1 - \exp\left\{-\frac{h\nu}{kT}\right\}\right) \overline{g_\mathrm{ff}} \quad \mathrm{m}^{-1} \tag{5.3}$$

where $\overline{g_\mathrm{ff}}$ is the 'Gaunt factor',[4] a quantum-mechanical correction factor which is of order unity at optical/UV wavelengths (but which may be an order of magnitude larger in the radio regime).

5.3.1 The Rosseland mean opacity for free–free absorption

Recall that the Rosseland mean opacity is defined by

$$\frac{1}{\bar{k}_\mathrm{R}(\rho, T)} \int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu \equiv \int_0^\infty \frac{1}{k_\nu(\rho, T)} \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu. \tag{5.1}$$

We can easily evaluate the integralon left-hand side of eqtn. (5.1), for given temperature T:

$$\begin{aligned}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu &\rightarrow \frac{\mathrm{d}}{\mathrm{d}T}\int_0^\infty B_\nu\,\mathrm{d}\nu \\ &= \frac{\mathrm{d}(\sigma T^4/\pi)}{\mathrm{d}T} \\ &= 4\sigma T^3/\pi\end{aligned}$$

(where we can take the '$\mathrm{d}T$' out of the integral because we're integrating over frequency, and temperature isn't a function of frequency).

Next, consider the right-hand side; substituting for $\mathrm{d}B_\nu(T)/\mathrm{d}T$ we obtain

$$\int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\,\mathrm{d}\nu \rightarrow \int_0^\infty \frac{1}{k_\nu}\frac{2h^2\nu^4}{c^2kT^2}\frac{\exp\{h\nu/kT\}}{\left(\exp\{h\nu/kT\} - 1\right)^2}\,\mathrm{d}\nu;$$

and using eqtn. (5.3) for k_ν,

$$\begin{aligned}&\propto \int_0^\infty \frac{1}{\rho^2\nu^{-3}T^{-1/2}\left(1 - \exp\{-h\nu/kT\}\right)}\ \frac{2h^2\nu^4}{c^2kT^2}\frac{\exp\{h\nu/kT\}}{\left(\exp\{h\nu/kT\} - 1\right)^2}\,\mathrm{d}\nu \\ &\propto \rho^{-2}T^{-3/2}\int_0^\infty \nu^7 \frac{\exp\{2h\nu/kT\}}{\left(\exp\{h\nu/kT\} - 1\right)^3}\,\mathrm{d}\nu.\end{aligned}$$

[4] Sometime called the Kramers-Gaunt factor

We simplify this by setting $h\nu/kT \equiv x$,

$$\propto \rho^{-2} T^{-3/2} \int_0^\infty T^8 x^7 \frac{\exp\{2x\}}{(\exp\{x\}-1)^3}\,\mathrm{d}x.$$

Note that the integral is still over *frequency,* so the T^8 term can come out of the integral. Combining our results so far, we obtain

$$\frac{1}{\bar{k}_\mathrm{R}} \propto \rho^{-2} T^{7/2} \int_0^\infty x^7 \frac{\exp\{2x\}}{(\exp\{x\}-1)^3}\,\mathrm{d}x.$$

The integral is called the 'Strömgren integral', and, by inspection, we can see that it evaluates to a constant (~ 5020). We therefore arrive at our final expression for the Rosseland mean opacity for free–free absorption,

$$\begin{aligned} \bar{\kappa}_\mathrm{R}^\mathrm{ff} = \bar{k}_\mathrm{R}^\mathrm{ff}/\rho &= \kappa_0^\mathrm{ff} \rho T^{-7/2}, \\ &\simeq 10^{22} Z^2 \rho T^{-7/2} \ \ \mathrm{m^2\ kg^{-1}} \end{aligned} \tag{5.4}$$

where κ_0 is a constant for a given chemical composition, and the final result is an approximation to numerical calculations.

5.3.2 Opacity 'laws' & Kramers' opacity

By construction, the Rosseland mean opacity is frequency independent, but of course it still depends on temperature and density (because k_ν is a function of these variables). In many circumstances (particularly for analytical work), this dependence can be reasonably approximated by power-law representations of the form

$$\kappa \simeq \kappa_0 \rho^p T^q. \tag{5.5}$$

Eqtn. 5.4 is evidently an example of a such an 'opacity law' (of course, 'law' here doesn't imply a basic law of nature, like Newton's laws of motion). For free–free, we've seen that

$$\kappa \propto \rho T^{-3.5};$$

this specific opacity law is called Kramers' opacity law (after Hendrik 'Hans' Kramers, who derived it in 1923). Kramers' law is particularly important, since as well as being an exact result for free–free absorption, it turns out also to be a reasonable approximation for some other major radiative processes, and is useful for computing models of stellar interiors. Other opacity 'laws' that you may come across include

Schwarzschild's opacity:	$p = 0.75$	$q = -3.5$
electron-scattering opacity (§5.5):	$p = 0$	$q = 0$

5.4 Bound–free, bound–bound

Of course, a thorough analysis of the opacities from these sources requires detailed numerical evaluation of cross-sections and the ionization and excitation conditions. This is now feasible computationally, but some insight can be gained from more general considerations.

5.4.1 Bound–free

To quote Schwarzschild (1958, *Structure and Evolution of the Stars*), "we shall now drop all considerations of accuracy and derive some very approximate formulae for the [bound–free] absorption coefficient".

For some element E in ionization state i ($i = 0$ for a neutral species, $i = 1$ for once ionized, etc.), with number density $n(E_i)$, the bound-free volume opacity is

$$k_\nu(E_i) = \sigma_\nu^{\rm bf}(E_i) \times n(E_i)$$

Although accurate photoionization cross-sections, σ_ν, require detailed quantum-mechanical calculations, but we can gain insight by adopting the result for a simple hydrogenic system.[5] Dropping the 'E_i' qualifier to simplify notation,

$$\sigma_\nu^{\rm bf} = \frac{64\pi^4 m_{\rm e} e^{10}}{3\sqrt{3} c h^6} S_{n\ell}^4 \frac{Z_i^4 g_\nu^{\rm bf}}{n^5 \nu^3}$$

where Z_i is the atomic number (nuclear charge), $S_{n\ell}$ is a 'screening factor' for quantum numbers n, ℓ (resulting from electrons in interior orbitals), and $g_\nu^{\rm bf}$ is the bound–free gaunt factor (typically of order unity). Of course, this holds only when the energy of the incident photon exceeds the ionization energy, given by

$$h\nu > \chi_i, = \frac{2\pi^2 m_{\rm e} e^4}{h^2} \frac{Z_i^2}{n^2}.$$

[That is, the bound-free cross-section for a given element in a given excitation state increases as $\sim \nu^{-3}$, up to this threshold frequency, at which point it drops to zero. This introduces a series of absorption 'edges' corresponding to different excitation states – for example, the Lyman, Balmer, Paschen... in the hydrogen spectrum.] Subsuming many constant (or fixed) terms into a constant of proportionality, we see that

$$\sigma_\nu^{\rm bf} \propto \frac{Z_i^2 Z_i^2}{\nu^3} \propto \frac{Z_i^2 \chi_i}{\nu^3}.$$

[5] A 'hydrogenic system' in this context means an ion with only one electron, or one electron in the valence shell. And be careful: in this discussion, an unqualified n represents the principal quantum number, not the density!

We next recall that the Saha equation, which gives us the ratio of two consecutive stages of ionization:

$$\frac{n_i}{n_{i+1}} = \frac{n_\mathrm{e}}{2}\frac{U_i(T)}{U_{i+1}(T)}\left(\frac{h^2}{2\pi m_\mathrm{e} kT}\right)^{3/2} \exp\{\chi_i/kT\}$$

(where $n_i \equiv n(E_i)$ and $U_i(T)$ is the partition function; cf. Appendix D). To make progress we make the simplifying (and not unreasonable) assumption that an element exists primarily in a single, dominant stage of ionization. The principal source of bound–free opacity for that element is then photoionization *to* that stage, from the next lower stage of ionization, in which case $n_{i+1} \simeq n_E$, the number density of the element, and

$$\begin{aligned} n_i &\propto n_E n_\mathrm{e} kT^{-3/2} \exp\{\chi_i/kT\}\,, \\ &\propto \rho^2 kT^{-3/2} \exp\{\chi_i/kT\} \end{aligned}$$

for given abundances.

Overall, therefore, for element E,

$$\begin{aligned} k_\nu &\propto \frac{Z_i^2\chi_i}{\nu^3}\,\rho^2(kT)^{-3/2}\exp\{\chi_i/kT\} \\ &\propto Z_i^2\frac{\chi_i}{kT}\nu^{-3}\,\rho^2(kT)^{-1/2}\exp\{\chi_i/kT\} \end{aligned}$$

In the spirit of dominant ions, we note that stages with ionization energies greatly different to $\sim kT$ won't contribute significantly to the opacity.[6] Consequently, we need only be concerned with the regime for which $\chi/kT \simeq 1$, leaving us with

$$k_\nu/\rho \propto \nu^{-3}\,\rho^2 T^{-1/2}$$

which we should recognize as similar to the functional form of free–free, eqtn. (5.3).

The total bound–free opacity will be the sum over all elements Z, in all ionization stages i, and all excitation states n. This summation, and the integration to construct the Rosseland mean, smooth out the effects of the discontinuites at ionization edges, leading to a Rosseland mean that has a Kramers'-type form,

$$\begin{aligned} \bar{\kappa}_\mathrm{R}^\mathrm{bf} &\simeq 10^{24} Z(1+X)\rho T^{-7/2}\ \mathrm{m^2\ kg^{-1}} \\ &= \bar{\kappa}_0^\mathrm{bf}\rho T^{-7/2}. \end{aligned}$$

[6] Those with $\chi \ll kT$ will nearly all be ionized to higher stages, and hence not numerous; for those with $\chi \gg kT$ there are few sufficiently energetic photons to be absorbed, so any opacity is unimportant.

5.4.2 Bound–bound

At high temperatures ($\gtrsim 10^6$ K) the radiation is energetic enough to ionize the plasma, but at cooler temperatures bound–bound transitions are a major source of opacity. Evaluating this opacity involves calculations for millions of lines (transitions). However, in stellar interiors, pressure and temperature broadening are substantial, resulting in a quasi-continuous opacity arising from numerous broad, overlapping lines. Moreover, the bound–bound cross-sections for hydrogenic species have the same form as the bound–free. We can therefore anticipate that, overall (though not within a particular line!), the broad frequency dependence may again vary like ν^{-3}. Moreover, we might also expect a ρ^2 density dependence (as the rate of bound-bound processes depends on the number density of targets *and* the number density of colliders), and therefore, plausibly, a Kramers-like opacity law. This expectation is born out by the results of detailed calculations (Fig. 5.1), which yield

$$\bar{\kappa}_{\rm R}^{\rm bb} \simeq 10^{24} Z \rho T^{-7/2} \ \ {\rm m}^2\ {\rm kg}^{-1}$$

[In practice, modern work takes advantage of a Kramers-*like* temperature dependence, tabulating $\kappa(T)$ as a function of $\mathcal{R} = \rho/T^3$, rather than simply ρ.]

5.4.3 H^- opacity

A free electron can attach to a neutral hydrogen atom, forming the negative hydrogen ion, H^- (chemists call this hydride, or the hydrogen anion, but astronomers just say 'H-minus'). Like other species, this ion can absorb photons through bound–free or free–free processes:[7]

$$\begin{aligned} {\rm H}^- + h\nu &\rightleftharpoons {\rm H} + e^- \\ {\rm H}^- + e^- + h\nu &\rightleftharpoons {\rm H}^- + e^- \end{aligned}$$

The H^- ion has an ionization potential[8] of only 0.754 eV (corresponding to a wavelength of ~1.65 μm), and so is easily destroyed; in the solar atmosphere, only about 1 in 10^7 hydrogen atoms is in the form of H^-.

However, for the neutral hydrogen, only electrons in $n = 3$ contribute to opacity in the optical (where most of the radiation emerges), and the Boltzmann equation tells us that as few as ~1 in 10^9 H^0 atoms has its electron in $n = 3$ at ~6000 K. The photoionization cross-sections for $n = 3$ H^0 and for H^- are

[7] Note that, while we refer to 'H^- opacity', it is *not* a separate mechanism – just a specific instance of bound–free & free-free processes that happens to be particularly important in the atmospheres of the Sun and solar-type stars. In this case, the bound–free process is usually refered to as photodetachment, or photodissociation, rather than photoionization.

[8] Or binding energy, or dissociation energy. The ion can only exist in the ground state – there are no other bound states with excitation energies less than the dissociation energy – which is why it can only form from ground-state H^0. The quoted numerical value is from Andersen, Haugen & Hotop, J. Phys. Chem. Ref. Data, 28, 1511, 1999.

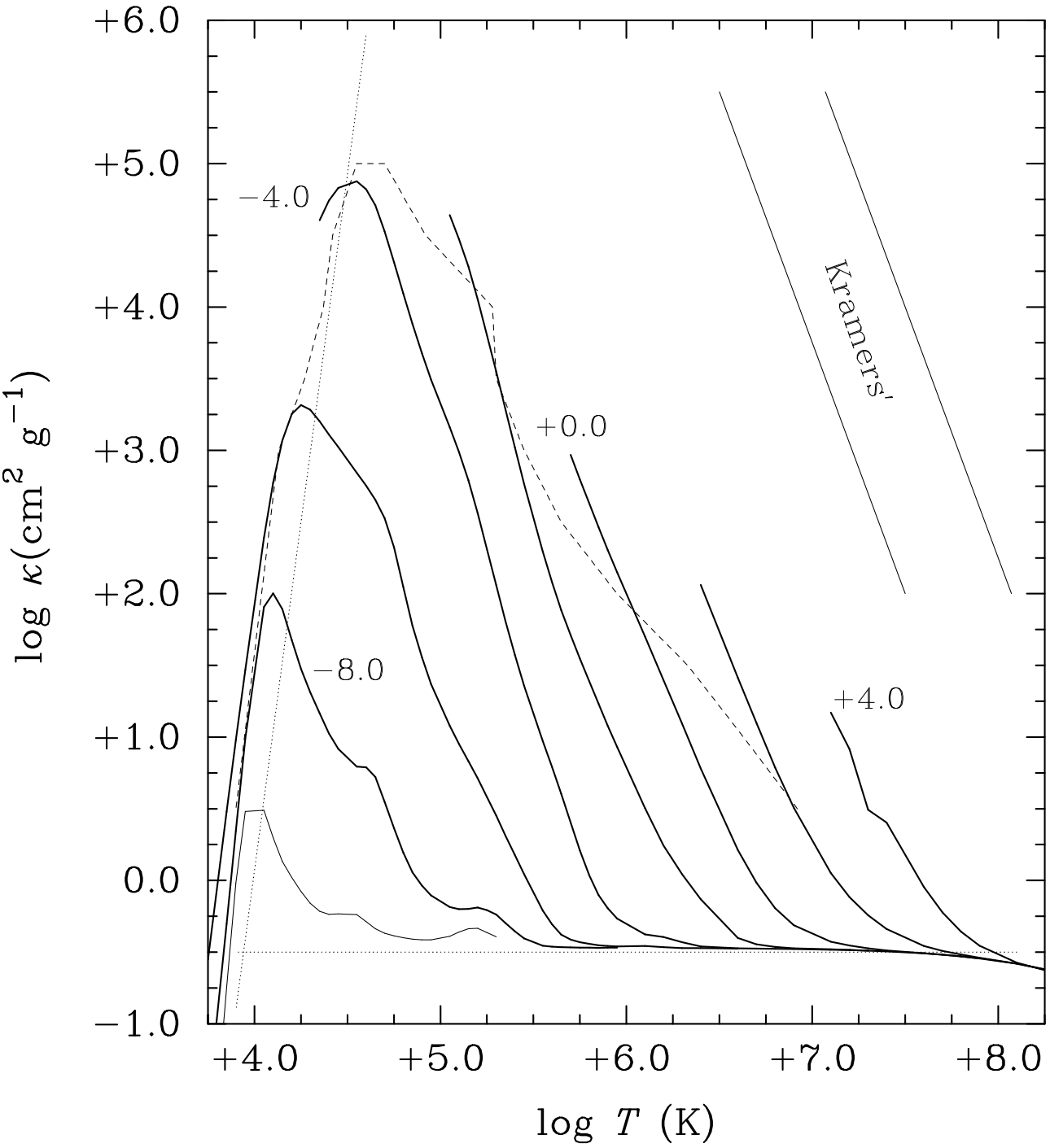
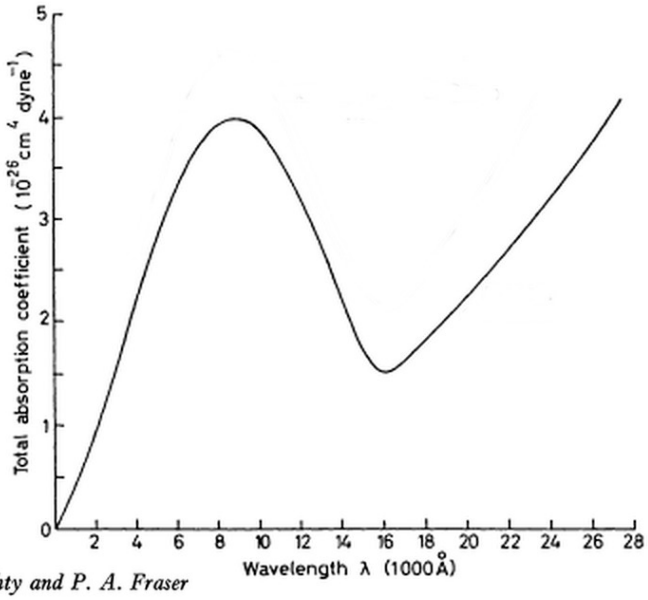


Figure 5.1: Radiative opacities from OPAL (Iglesias & Rogers 1996), for solar metallicity ($X = 0.70, Y = 0.28, Z = 0.02$) and densities $\log\rho = -10{:}{+4}$ (g cm^{-3}; 1 g cm^{-3} = 10^3 kg m^{-3}) at steps of two dex. Kramer's opacity law is shown for two (arbitrary) densities separated by $\Delta\log\rho = 2$.

The 'floor' at $\kappa \simeq -0.5$ dex cm^{-2} kg^{-1} arises from electron scattering opacity (corresponding to a fully ionized plasma); to the left of the near-vertical dotted line, hydrogen and helium are largely neutral.

The T, ρ locus of a solar-interior model is indicated by a dashed line. Note that it falls mainly in the regime for which Kramers' does a pretty good job (i.e., within the dotted lines).



N. A. Doughty and P. A. Fraser

FIG. 2. *The absorption coefficient of the negative hydrogen ion at a temperature of 6 300 °K (θ=0·8) due to bound–free and free–free transitions.*

Figure 5.2: H^- cross-section. The bound–free ionization (or dissociation) threshold is at ∼16 500Å; absorption at longer wavelengths arises from free-free opactity.

roughly similar, so H^- dominates in the Paschen continuum ($\lambda < 821$ nm; in the Balmer continuum, $\lambda < 365$ nm, H^0 and H^- have comparable overall opacity.)

H^- can only form when the temperature is low enough for hydrogen to be largely neutral *and* low enough that the ion isn't dissociated; this means photospheres with $T_{\rm eff} \lesssim 8000$ K. However, we also need a source of free electrons in order to create H^-; these can only come from metals with first ionization potentials less than that of hydrogen (i.e., < 13.6 eV). Relevant easily ionized, cosmically abundant sources include aluminium (first ionization potential 6.0 eV), magnesium (7.6 eV), iron (7.9 eV), and silicon (8.2 eV), requiring temperatures above ∼3000 K to produce free electrons. Thus H^- is important over only a limited temperature range – but the Sun's effective temperature happens to be right in the middle of that range.

The total H^- opacity is proportional to the electron density. Since the electrons come from the metals (not H, He), the opacity rises with metallicity, and with temperature – until the ion begins to be collisionally destroyed.

The cross-section is not a simple power law in frequency; the bound–free component peaks around 8000Å, with the free–free component rising to longer wavelengths (Fig. 5.2). It should not be a surprise, therefore, that, when integrated over frequency, H^- opacity does not follow a Kramers' law; rather, very roughly,

$$\bar{\kappa}_{\rm R}(H^-) \simeq 10^{-30} Z \rho^{1/2} T^9 \ {\rm m}^2 \ {\rm kg}^{-1}.$$

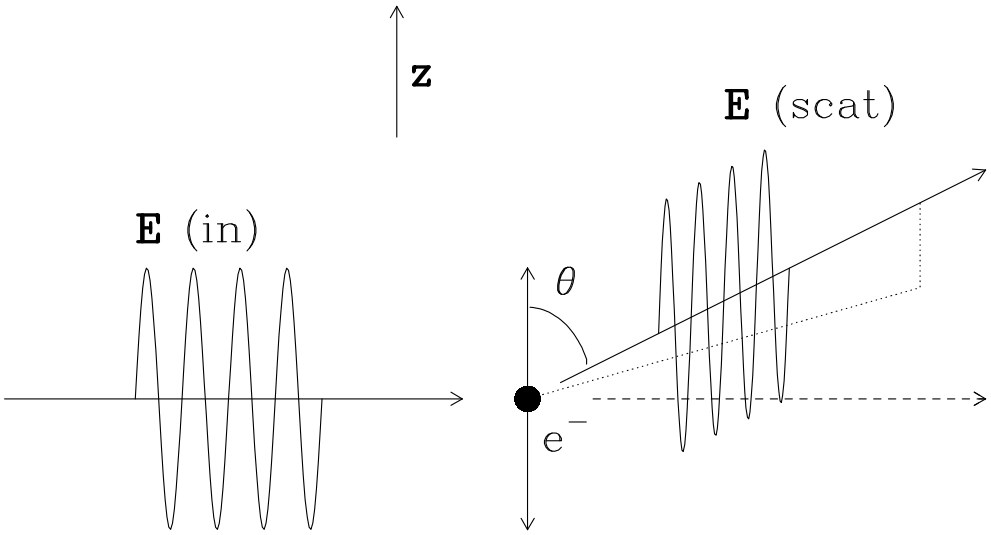


Figure 5.3: Electron scattering in the Thomson limit; the case of a linearly polarized incident wave. **E** indicates the electric field of the incident and scattered waves (with modulus E_0).

5.5 Electron scattering: Thomson scattering

In astrophysical applicatiosn, 'electron scattering' refers to the scattering of photons by free electrons in a plasma. A photon incident on an electron induces a force (because it's an electromagnetic wave interacting with an electrically charged particle). In response, the starts oscillating as a dipole; and an oscillating dipole emits radiation. This process was first studied by J.J. Thomson (1856–1940),[9] in the non-relativistic case. It can be considered the low(-photon)-energy limit of the more general case of Compton scattering.

In this low-energy limit we require the photon energy to be less than the electron rest-mass energy: $h\nu \ll m_\mathrm{e} c^2$. The simple physical interpretation is that the electron can't gain momentum from the photon, and can be considered static. Suppose that the incoming photons can be considered as continuous (and linearly polarized) plane wave. The electric field as a function of time t is described by

$$\mathbf{E} = E_0 \, \sin(\omega t) \, \hat{\mathbf{z}}$$

(where $\omega = 2\pi\nu$ is the angular frequency and $\hat{\mathbf{z}}$ is the unit vector in the direction of the electric-field variability). As long as the electron remains non-relativistic, the (Lorentz) force exerted on it by the electromagnetic wave comes predominantly from this electric field, and the electron's equation of motion is

$$m_\mathrm{e} \frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = -e \, E_0 \, \sin(\omega t) \tag{5.6}$$

[9] Thomson, seven of his research assistants, and his son, G.P. Thomson, all won Nobel Prizes in physics! As far as I know, Lord Kelvin (William Thomson) was no relation.

which has the solution

$$z = \frac{e\,E_0}{m_\mathrm{e}\,\omega^2}\sin(\omega t)$$

(where e is the electron charge). That is, the electron oscillates in the direction of the incoming wave's electric field (as one might intuitively expect). The electron can now be thought of as an oscillating electric dipole – which emits electromagnetic radiation. The emitted radiation will have the same frequency as the incoming radiation, but, in general, a different direction; that is, an incoming photon is scattered.

5.5.1 Cross-section

We first quote two results without proof:
(*i*) The Poynting flux – the energy flux of the incoming electromagnetic wave (the time-averaged magnitude of the Poynting vector) – is given by

$$\langle S \rangle = \frac{\epsilon_0 c E_0^2}{2};$$

(*ii*) the time-averaged power emitted by the dipole (electron) into solid angle $\mathrm{d}\Omega$ at angle θ to the dipole is (from Larmor's formula; cf. Fig. 5.3)

$$\left\langle \frac{\mathrm{d}W}{\mathrm{d}\Omega} \right\rangle = \left(\frac{e^2}{4\pi\epsilon_0 m_\mathrm{e} c^2} \right)^2 \frac{\epsilon_0 c E_0^2}{2} \sin^2\theta$$

where ϵ_0 is the permittivity of free space (vacuum permittivity), and SI forms are used.

From these results we can determine the differential scattering cross-section, defined as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \left\langle \frac{\mathrm{d}W}{\mathrm{d}\Omega} \right\rangle \Big/ \langle S \rangle$$

(which has dimensions of power per solid angle divided by flux; i.e., area per unit solid angle[10])

$$\begin{aligned} &= \left(\frac{e^2}{4\pi\epsilon_0 m_\mathrm{e} c^2} \right)^2 \sin^2\theta \\ &\equiv r_0^2 \sin^2\theta \end{aligned} \tag{5.7}$$

where the value of r_0, the 'classical electron radius', is 2.817×10^{-15} m. (Note that the cross-section is inversely proportional to mass squared; thus scattering by protons, for example, is generally negligible.)

The total (Thomson) cross-section is defined as

$$\sigma_\mathrm{T} = \int \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\,\mathrm{d}\Omega;$$

as usual, we set $\mu = \cos\theta$ and recall that $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = -\mathrm{d}\mu\mathrm{d}\phi$, giving

$$\begin{aligned}\sigma_{\mathrm{T}} &= \int_0^{2\pi}\int_{-1}^{+1} r_0^2(1-\mu^2)\,\mathrm{d}\mu\,\mathrm{d}\phi \\ &= 2\pi\, r_0^2 \int_{-1}^{+1} (1-\mu^2)\,\mathrm{d}\mu \\ &= \frac{8\pi}{3} r_0^2 \end{aligned} \tag{5.8}$$

Numerically, the Thomson cross-section is $\sigma_T = 6.65 \times 10^{-29}$ m^2.

[Formally, the discussion so far has focussed on a linearly polarized beam. We can compute the differential cross-section for unpolarized light by noting that it can be treated as the superposition of two orthogonally polarized waves. If we define ϑ as the angle between the scattered radiation and the original radiation (so that $\vartheta = \pi/2 - \theta$), the result is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{2}\left[r_0^2(1+\cos^2\vartheta)\right]$$

and the total cross-section is the same as before (as it must be, since an electron at rest has no intrinsic polarization, so it has to react to all linear polarizations in the same way [Miller])].

We can now recognise some of the more important properties of the cross-sections for Thomson scattering:

1. Eqtn. (5.7) has a symmetry: the scattering cross-section is unchanged under reflection ($\theta \to -\theta$) – but is not isotropic. That is. . .
2. Eqtn. (5.7) depends on the angle θ between incident and scattered photons. This can lead to electron scattering being a polarigenic process in asymmetric geometries.
3. The total cross-section is the same for polarized and unpolarized radiation (since an electron at rest has no intrinsic directionality)
4. Both the total and differential cross-sections are frequency independent. This means that (in the non-relativistic limit) scattering is equally effective at all frequencies – i.e., it is 'grey'.
5. Furthermore, the process does not alter the energy of the photon: the incoming and the scattered photon have the same frequency (although their direction of propagation is different) – that is, Thomson scattering is coherent.[11]

The Cosmic Microwave Background, and the solar K-corona, are familiar astrophysical phenomena in which Thomson scattering is of importance.

[11] If relativistic velocities pertain (photon energies comparable to electron rest mass), it's necessary to use the cross-section calculated in the Compton limit, which is energy dependent (i.e., involves a nett exchange of energy between photons and electrons).

5.6 Rayleigh scattering

If a photon scatters off an electron which is bound in an atom the process is called 'Rayleigh scattering'. It can occur only if the photon energy $h\nu$ is much less than the energy spacings of the atom (otherwise a bound–bound transition occurs).

Electrons are set into a resonant motion along their orbits by the photon. The electron equation of motion is

$$m_\mathrm{e}\frac{\mathrm{d}^2 z}{\mathrm{d}t^2} = -m_\mathrm{e}\omega_0^2 z - e\,E_0\,\sin(\omega t) \tag{5.9}$$

(cp. eqtn. 5.6, noting the extra term on the rhs from orbital motion), where we are, in effect, modelling the atom as a simple harmonic oscillator of natural frequency ω_0. Eqtn. (5.9) has the solution

$$z = \frac{e\,E_0}{m_\mathrm{e}\,(\omega^2 - \omega_0^2)}\,\sin(\omega t). \tag{5.10}$$

An analysis analogous to that for the Thomson cross-section (eqtn. 5.7) leads to a differential cross-section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{\omega^4}{(\omega^2 - \omega_0^2)^2} r_0^2 \sin^2\theta \tag{5.11}$$

(where the '$\sin^2\theta$' term indicates that Rayleigh scattering, like Thomson scattering, is a potentially polarigenic process) and total cross-section

$$\sigma_\mathrm{R} = \frac{\omega^4}{(\omega^2 - \omega_0^2)^2}\sigma_\mathrm{T} \tag{5.12}$$

In the limit in which the frequency of the incident radiation is much greater than the natural frequency of the atom ($\omega \gg \omega_0$, or $\nu \gg \nu_0$), eqtns. (5.11) and (5.12) reduce to the previously obtained expressions for scattering by a free electron (eqtns. 5.7, 5.8); in other words, an electron in an atom acts very much like a free electron as far as scattering of high frequency radiation is concerned (as long as bound–bound or bound–free interactions don't come into play).

In the opposite limit, in which the frequency of the incident radiation is much less than the natural frequency of the atom (i.e., at long wavelengths), eqtn. (5.11) yields

$$\begin{aligned}\sigma_\mathrm{R} &= \sigma_\mathrm{T}\left(\frac{\nu}{\nu_0}\right)^4, \\ &\propto \lambda^{-4};\end{aligned} \tag{5.13}$$

That is, the cross-section decreases with increasing wavelength.

Rayleigh scattering is important in cool stellar atmospheres (late K and M stars). However, the most familiar example of Rayleigh scattering is the scattering of visible radiation from the Sun by neutral atoms in the Earth's upper atmosphere.

(Nitrogen and oxygen are most important; the frequency of visible radiation is much less than the typical emission frequencies of these atoms, which lie in the ultraviolet band, so it is certainly the case that $\omega \ll \omega_0$.)

Because blue light is scattered more effectively than red light (eqtn. (5.13), the sky appears blue (mostly blue light is scattered into the line of sight). When the Sun is very low, the path length is longer, and the light undergoes more scatterings; most blue light is scattered out of the line of sight, leaving a red sky.

Furthermore, inspection of the sky using polarized sunglasses on a clear day should confirm that the daytime sky is linearly polarized.

5.7 Extinction by larger bodies [not for lectures]

For larger bodies (dust) we generally use 'Mie theory', developed by the German physicist Gustav Mie[12] in the early 20th century. His original classical theory of scattering applies to spheres of specified refractive index.

Very roughly speaking,
for wavelengths short compared to the size of the particles, the scattering cross-section is about twice as large as the geometric cross-section
for wavelengths comparable to or larger than the particle size, the scattering cross-section declines as 1/lambda where lambda is the wavelength of the light being scattered.

Interstellar dust has a typical size of $\sim 10^3$ Å (0.1μm), which matches the wavelength of UV radiation. 1,000 angstroms. In these circumstances, we can adopt 'Mie scattering'.

5.8 Summary

Overall, the cross-sections for free–free, bound–free, and bound–bound processes can all be represented, more or less accurately, as having ν^{-3} frequency dependencies; and their frequency-averaged Rosseland mean opacities are, individually and collectively, often well approximated by the Kramers' opacity law (Fig. 5.1).

H^- opacity (arising through bound–free and free–free processes) is particularly important in the atmospheres of solar-type stars.

[12] Rhymes with 'pea', not 'pie'.

Electron scattering and Rayleigh scattering are further important sources of opacity in stars.

Other sources of opacity play a role, under various circumstances – for example, molecular opacities in cool stellar atmospheres.

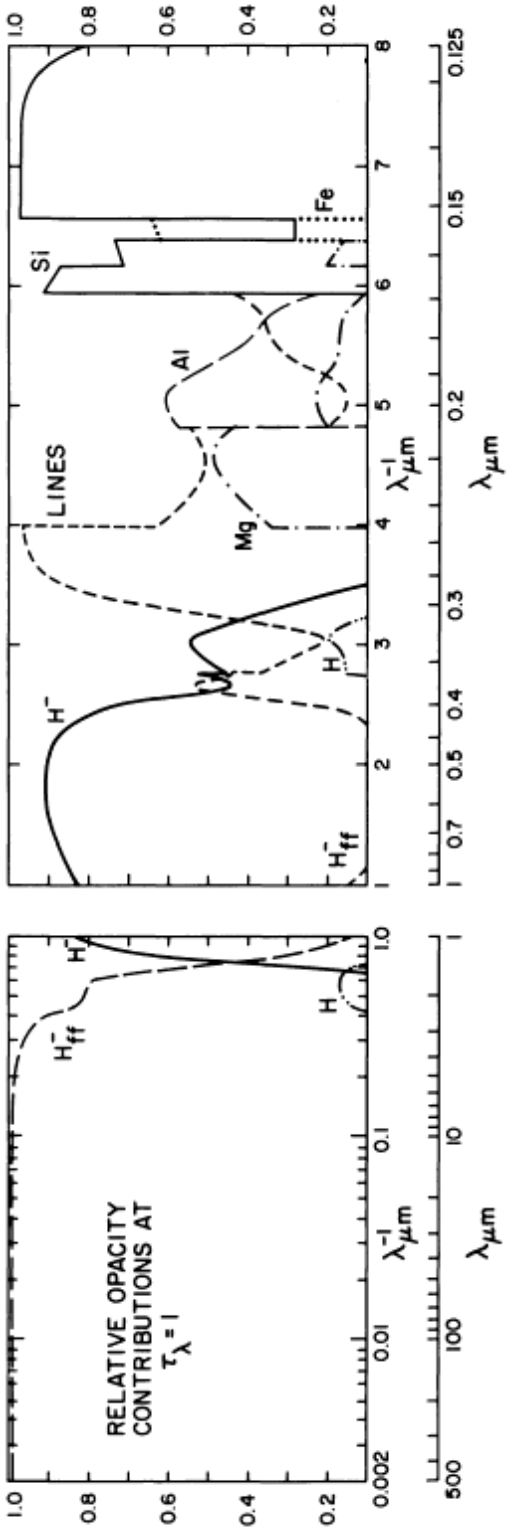


Figure 5.4: Principal opacity sources in the solar photosphere (from Vernazza, Avrett & Loeser, 1976, ApJS, 30, 1). The symbols H, Si, Al, Mg, and Fe refer to neutral-atom bound-free absorption, and LINES to a smoothly varying approximation to the effects of line opacity. Note the dominance of H^- (bound-free) opacity in the region of peak flux, around 0.5μm (5000Å).

Section 6

Thermodynamic Equilibrium and LTE

[review of 0018 material]

In stellar atmospheres we find:

- moderate (not 'extremely high') densities; the equation of state is well approximated by a perfect gas
- high temperatures; the gas is made of single atoms, ions, free electrons (and molecules in the cooler stars)
- the state of the gas is known when we know the particle distributions over all available bound and free energy levels; i.e., when we know the occupation numbers of these levels.
- from the occupation numbers we can compute gas pressure, opacity, emissivity of the gas. . .(in principle) everything!

Here we consider some of the circumstances that determine the occupation numbers.

6.1 Thermal equilibrium

In thermodynamic equilibrium, the state of the gas (the occupation numbers) is specified uniquely by the temperature T, through the equilibrium relations of statistical mechanics.

- The velocity distribution of free particles is given by a Maxwellian
- the excitation and ionization states of the gas are given by the Boltzmann+Saha equations
- the ratio between pure absorption/emission coefficients is given by the Planck function (Kirchhoff's law)

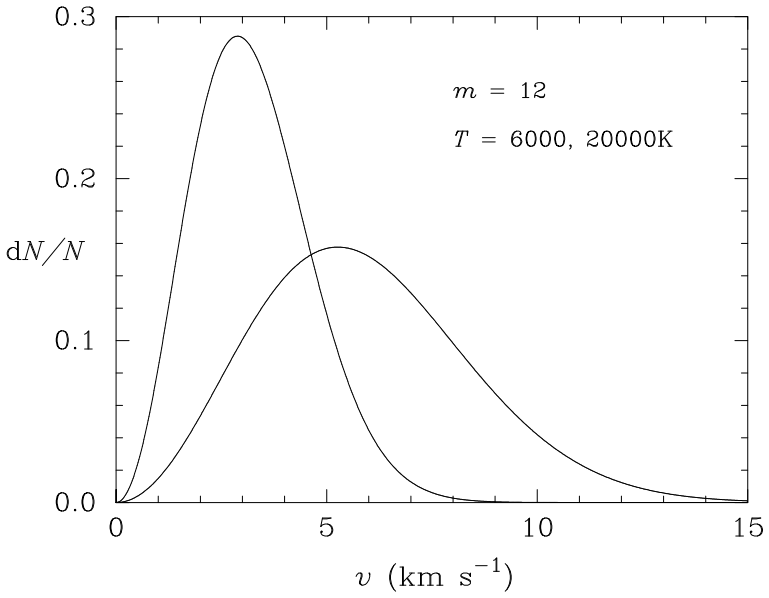


Figure 6.1: The Maxwell-Boltzmann velocity distribution

- The specific intensity I_ν is isotropic and independent on the geometry of the medium and on the chemical composition of the matter; hence for any given frequency, I_ν must be only a function of T, and is given by the Planck function,

$$I_\nu = B_\nu(T) = \frac{2h\nu^3}{c^2}\Big/(\exp\{h\nu/kT\} - 1)$$

All these distributions are functions only of temperature T.

6.1.1 The Maxwellian velocity distribution

In TE, the probability that a particle of mass m at a temperature[1] T has a velocity between v and $v + \mathrm{d}v$ is given by

$$f(v)\,\mathrm{d}v = \left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left\{\frac{-mv^2}{2kT}\right\} 4\pi v^2\,\mathrm{d}v$$

This distribution may be characterized in terms of the most probable speed,

$$v_0 = (2kT/m)^{1/2} = 12.85(T/10^4 A)^{1/2}\ \ \mathrm{km\ s^{-1}}$$

where A is the atomic weight of the particle.

[1] In this context, the kinetic temperature.

6.1.2 Boltzmann distribution

In TE at the temperature T, the excitation states of atoms are distributed over the bound levels according to the Boltzmann excitation equation:

$$\frac{n_n}{n_0} = g_n \frac{\exp\{-\chi_n/kT\}}{U_0}$$

where
$n_0 (= n_{jk})$ is the number density of atoms in ionization state j for element k;
$n_n (= n_{ijk})$ is the number density of atoms in excited state n of ionization state j for element k;
g_n is the statistical weight associated with level n;
χ_n is the excitation energy of level n; and
U_n is the partition function (cf. Appendix D.2).

6.1.3 Saha equation

Above the discrete bound eigenstates of an atom there exists a continuum of levels in which the electron is unbound and has a non-zero kinetic energy The relative number of atoms and ions in successive states of ionization is given by the Saha ionization equation,

$$\frac{n_{j+1} n_\mathrm{e}}{n_j} = \frac{(2\pi mkT)^{3/2}}{h^3} 2 \frac{U_{j+1}}{U_j} \exp\{-\chi_i/kT\}$$

where:
U_{j+1}, U_j are partition functions, and
χ_i is the ionization potential.

Essentially, the Saha equation is just an extension of the Boltzmann equation to continuum states. Again, the function on the right-hand side depends only on temperature.

Consider the process in which an atom is ionized from its ground level, resulting in an ion and a free electron moving with velocity v. The energy required for this process is $\chi_i + {}^1/_2 m_\mathrm{e} v^2$.

6.1.4 Kirchhoff's law

If the matter component (emitters and absorbers) is in TE, then the ratio between the true absorption and emission coefficients is a Planck function:

$$\frac{j_\nu}{k_\nu} = B_\nu$$

Note that this is a property of the *matter* – not a property of the radiation field. It's therefore valid only for those processes that involve matter–matter interactions (even if a photon results; i.e., collisional

processes, such as collisional excitation, free-free, etc.), but in principle, it may be applicable even if radiation is decoupled from the gas, as long as the matter component has reached its statistical equilibrium.

6.1.5 The specific intensity in TE

As radiation travels through a medium, the specific intensity varies due to emission, absorption, or scattering of photons. The simplest situation occurs when the interactions are so strong that matter and radiation are kept in mutual thermodynamical equilibrium *in all points of the system*.

The condition of perfect matter/radiation equilibrium can be obtained only in an ideal way. We need to imagine an adiabatic cavity, with walls acting as perfect absorbers, containing matter in equilibrium at a given temperature T. Since photons are continuously absorbed/emitted from the walls, not only the energy distribution but also the number of photons inside the cavity will adjust toward their equilibrium distributions. The number distribution of photons (bosons) is then given by the Planck distribution

$$\frac{\mathrm{d}n}{\mathrm{d}\nu} = \frac{8\pi}{c^3}\frac{\nu^2}{\exp\left\{\frac{h\nu}{kT}\right\} - 1} \tag{6.1}$$

while the energy distribution is given by the black-body law:

$$u_\nu = h\nu\frac{\mathrm{d}n}{\mathrm{d}\nu} = \frac{8\pi h\nu^3}{c^3}\frac{1}{\exp\left\{\frac{h\nu}{kT}\right\} - 1}. \tag{6.2}$$

Now, with simple thermodynamical considerations it is possible to infer some general properties of the radiation contained in the cavity. Imagine we will open a small hole in the cavity, so small that is not going to alter the equilibrium inside, and then we put a second cavity in contact with the hole. The second cavity has a different shape, different chemical composition, etc., *but it has the same temperature T of the first cavity* (Fig.6.2). Between the two cavities we set a filter (made of optics, lenses, screens, whatever) such that we allow the flow only for radiation with a certain frequency and a certain direction but, at the same time, we avoid any alteration of thermal or mechanical nature to the system.

Because T is the same, the two cavities are in mutual equilibrium so that the net flow of energy between the two cavities must be zero. Therefore, for definition of specific intensity, it follows that the value of I_ν flowing toward the other cavity must be the same on the two sides of the hole.

On the other hand, since we have chosen arbitrarily the position of the hole, we can conclude that

- under thermal equilibrium conditions, I_ν is isotropic, independent on the shape of the container and on the chemical composition of the matter contained inside the container.

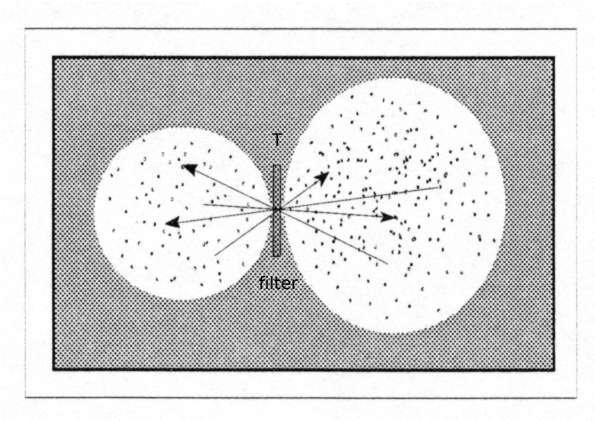


Figure 6.2: Thermodynamical Equilibrium between two cavities:

Therefore, for each given frequency the specific intensity at equilibrium *must* be a function of T only

$$I_\nu^{\rm eq} = f(T). \tag{6.3}$$

Now, due to isotropy, $I_\nu = J_\nu$. Since the mean intensity is related to the energy density through the relation

$$J_\nu = \frac{c}{4\pi} u_\nu \,, \tag{6.4}$$

by substituting the equilibrium distribution eqtn. (6.1), eqtn. (6.2) gives

$$I_\nu^{\rm eq} = J_\nu^{\rm eq} = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left\{\frac{h\nu}{kT}\right\} - 1} = B_\nu(T) \,. \tag{6.5}$$

- Therefore, if matter and radiation are in perfect thermal equilibrium then the specific intensity is given by a Planck function: $I_\nu = B_\nu$.

The energy density contained in the adiabatic cavity is then

$$u = \frac{1}{c} \int_0^\infty {\rm d}\nu \int_{4\pi} I_\nu^{\rm eq} \, {\rm d}\Omega = \frac{4\pi}{c} \int B_\nu d\nu = \frac{4\pi}{c} B(T) \tag{6.6}$$

where

$$B(T) = \int B_\nu d\nu = \frac{2\pi^4 k^4}{15 c^2 h^3} T^4 = \frac{ac}{4\pi} T^4 \tag{6.7}$$

(where a is the radiation constant).

Since radiation is isotropic, the net flux is zero at each point. Therefore, when dealing with physical situations close to the complete equilibrium it is more convenient to define the flux as integrated over one half only of the solid angle:

$$F_+ = \int_0^\infty B_\nu \, \mathrm{d}\nu \int_{2\pi} \cos\theta \, \mathrm{d}\Omega = \pi B(T) = \sigma T^4 \tag{6.8}$$

where $\sigma = ac/4 = 5.67 \times 10^{-5}$ erg cm^{-3}K^{-4}. $F_+ = \sigma T^4$ is the flux that flows (in one direction only) through an arbitrary imaginary surface contained inside a Planckian radiation field.

6.2 Local Thermodynamic Equilibrium

In TE the solution of the transfer problem is trivial: $I_\nu = B_\nu$. However, from the thermodynamic point of view the matter and radiation are two distinct systems (although they interact).

In general, the characteristic length scale to reach equilibrium is smaller for the matter component; even in presence of strong temperature and density gradients, particles can maintain a distribution very close to equilibrium locally (e.g., particle velocity distributions are Maxwellian; energy levels are distributed following Saha-Boltzmann, etc.)

The radiation field is generally more 'far-reaching'. However, if Kirchhoff's law is valid point by point at the local value of temperature, then the system is in Local Thermal Equilibrium (LTE). In LTE:

1. the excitation and ionization states of the gas are given by Saha-Boltzmann,
2. the velocity distributions are Maxwellian, and
3. and the Kirchhoff law is valid,

all at the local temperature.

LTE provides a convenient simplification to compute the state of the gas; it implies that we can apply the equilibrium relations at the local values of $T(r)$ in spite of gradients that exist in the atmosphere. The assumption of LTE should be applied with caution; it does not allow for coupling of the state of an element of gas with that of an element nearby (except from what we can impose by assuming global conservation equations, such as mass conservation or hydrostatic equilibrium).

Furthermore, be aware that, although $I_\nu = B_\nu$, this does *not* mean that an LTE model atmosphere emits a Planck spectrum. This is because we should properly write

$$I_\nu(\tau_\nu) = B_\nu(T(\tau_\nu)).$$

We 'see' into a model (or real) atmosphere down to optical depth $\tau_\nu \simeq 1$, but that corresponds to different linear depths – hence different temperatures, and different radiation fields – depending on the (frequency-dependent) opacity.

6.2.1 LTE and detailed balance

At the microscopic level, LTE requires that processes involving matter are in detailed balance – i.e., the rate at which any process occurs is exactly balanced by the rate at which its inverse occurs.

In general, processes that produce transitions fall in two categories: radiative and collisional.
–Collisional processes are in detailed balance when the velocity distributions are Maxwellian (as it is in LTE).
–Processes that involve the radiation field (photoexcitation and photoionization) depend directly on the radiation field, and will be in detailed balance only if the radiation field is Planckian.

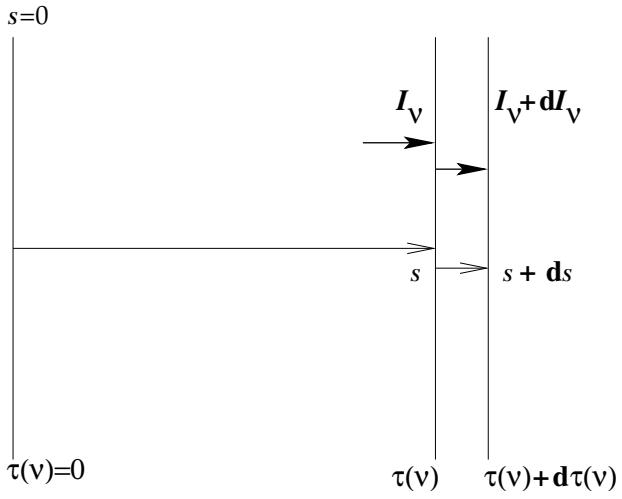
Section 7

Radiative Transfer

The major mode of energy transport through the surface layers of a star is via radiation; convective transport rarely carries a large fraction of outward flux in a stellar photosphere (although it can dominate in the interior, including the immediately subphotospheric layers).

7.1 The Equation of Radiative Transfer: First Principles

[reprises PHAS0018 material]



Consider a beam of radiation from a distant point source (e.g., an unresolved star), passing through some intervening material (e.g., interstellar gas). The intensity change as the radiation traverses an element of gas of thickness ds is just the intensity added, less the intensity taken away (per unit volume,

per unit frequency, per unit solid angle, per unit time):

$$\mathrm{d}I_\nu \cancel{(\mathrm{d}A\,\mathrm{d}\nu\,\mathrm{d}\omega\,\mathrm{d}t)} = +\, j_\nu \quad \mathrm{d}s \cancel{(\mathrm{d}A\,\mathrm{d}\nu\,\mathrm{d}\omega\,\mathrm{d}t)} \\ - k_\nu I_\nu\,\mathrm{d}s \cancel{(\mathrm{d}A\,\mathrm{d}\nu\,\mathrm{d}\omega\,\mathrm{d}t)}$$

i.e.,

$$\mathrm{d}I_\nu = (j_\nu - k_\nu I_\nu)\,\mathrm{d}s,$$

or

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = j_\nu - k_\nu I_\nu, \tag{7.1}$$

which is the simplest form of the Equation of Radiative Transfer (we will encounter other formulations shortly).

Equation (7.1) expresses the intensity of radiation as a function of position. In astrophysics, we often can't establish exactly where the absorbers are; for example, in the case of an absorbing interstellar gas cloud of given physical properties, the same absorption lines will appear in the spectrum of some background star, regardless of where the cloud is along the line of sight. It's therefore convenient to divide both sides of eqtn. (7.1) by k_ν; then using our definition of optical depth ($\mathrm{d}\tau_\nu = k_\nu\,\mathrm{d}s$; eqtn. 4.6) gives a more useful formulation,

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = \frac{j_\nu}{k_\nu} - I_\nu \\ \equiv S_\nu - I_\nu. \tag{7.2}$$

where S_ν is called the source function. Evaluating the source function is the central problem in constructing model stellar atmospheres – in principle, it requires evaluation of all sources of emission and extinction, frequency by frequency, depth by depth in the atmosphere.

We will consider the source function further in §7.3 (and will obtain a formal solution for I_ν in §7.4), but first we take a small detour...

7.2 Photon mean free path

If we simply want to establish how far photons in a beam will travel before being extinguished, we can ignore any 'new' photons – i.e., the emissivity. Setting $j_\nu = 0$ (which is appropriate for interstellar absorption lines, or headlights in fog) eqtn. (7.2) becomes

$$\mathrm{d}I_\nu/\mathrm{d}\tau_\nu = -I_\nu$$

Just by inspection, the solution is simply

$$I_\nu = I_\nu(0) \exp\{-\tau_\nu\}\,. \tag{7.3}$$

for incident intensity $I_\nu(0)$.[1] We can use this to establish the mean free path of a photon in a medium.

Evidently, as radiation travels through the absorbing medium, it is attenuated. The *average* distance over which photons will travel before being absorbed (or scattered), expressed in terms of optical depth, is

$$\langle \tau_\nu \rangle \equiv \int_0^\infty \tau_\nu \, p_a(\tau_\nu)\, \mathrm{d}\tau_\nu \Big/ \int_0^\infty p_a(\tau_\nu)\, \mathrm{d}\tau_\nu \tag{7.4}$$

where $p_a(\tau_\nu)\,\mathrm{d}\tau_\nu$ is the probability that a given photon will be absorbed in the interval $[\tau_\nu,\, \tau_\nu + \mathrm{d}\tau_\nu]$. The denominator of eqtn. (7.4) is the integral of this probability over all possible τ_ν, and so must evaluate to unity.[2]

If a photon is to be absorbed in the range $[\tau_\nu,\, \tau_\nu + \mathrm{d}\tau_\nu]$, then necessarily it has *not* already been absorbed in the interval $[0, \tau_\nu]$. The probability that a photon in a beam is absorbed over the latter interval is just the fractional change in intensity,[3]

$$\begin{aligned} p_a(0, \tau_\nu) &= \frac{I(0) - I(\tau_\nu)}{I(0)} \\ &= 1 - \frac{I(\tau_\nu)}{I(0)} \end{aligned}$$

so the probability that a photon has *not* been absorbed in the interval $[0, \tau_\nu]$ is

$$\begin{aligned} 1 - p_a(0, \tau_\nu) &= \frac{I(\tau_\nu)}{I(0)}, \\ &= \exp\{-\tau_\nu\} \end{aligned}$$

(from eqtn. 7.3).

The probability that a photon is absorbed over the interval $[\tau_\nu,\, \tau_\nu + \mathrm{d}\tau_\nu]$ is again the fractional change in intensity,

$$\begin{aligned} p_a(\tau_\nu, \tau_\nu + \mathrm{d}\tau_\nu) &= \left| \frac{\mathrm{d}I_\nu}{I(\tau_\nu)} \right|, \\ &= \mathrm{d}\tau_\nu \end{aligned}$$

(from eqtn. 7.2).

[1] In other contexts, versions of this result are known as Beer's Law, or the Beer–Lambert Law, or the Lambert–Beer Law, or the Beer–Lambert–Bouguer Law.

[2] The probability of anything, over all possible values of 'anything', is one.

[3] If the intensity drops by, say, 90%, then evidently there's a 90% probability that some photon initially in the beam will have been absorbed.

The combined probability of 'not absorbed in $[0, \tau_\nu]$' and 'absorbed in $[\tau_\nu, \tau_\nu + \mathrm{d}\tau_\nu]$' is the product of the two separate probabilities; that is,

$$p_a(\tau_\nu)\,\mathrm{d}\tau_\nu = \exp\{-\tau_\nu\} \times \mathrm{d}\tau_\nu.$$

Thus, from eqtn. (7.4),

$$\begin{aligned}\langle\tau_\nu\rangle &= \int_0^\infty \tau_\nu \exp\{-\tau_\nu\}\,\mathrm{d}\tau_\nu \\ &= 1\end{aligned}$$

[noting that $\int xe^{-x}\,\mathrm{d}x = -(1+x)e^{-x}$]; that is, the photon mean free path corresponds an optical depth of 1.

A medium is said to be 'optically thin' if $\tau_\nu < 1$, and 'optically thick' otherwise. Note that a medium can be optically thin at one wavelength, but optically thick at another; e.g., brick walls are optically thin to radio waves (we can listen to the radio indoors), but optically thick to visible light. In astrophysics, the change from optically thin to optically thick can happen over quite small wavelength changes (e.g., a nebula might be optically thin in the visible continuum, but optically thick in a line).

Although the mean free photon path corresponds to an optical depth of one, fewer than half the photons in a beam will travel that distance.[4] As eqtn. (7.3) shows, unit optical depth results in a reduction in intensity of a factor e^{-1}, or ~ 0.37; that is, rather *less* than half the photons survive passage through a medium of unit optical depth unscathed.

7.3 Source function revisited

Since $\mathrm{d}\tau_\nu = k_\nu\,\mathrm{d}s$ (eqtn. 4.6), the mean free path can be converted to (frequency-dependent) linear distance ℓ, if $k_\nu(s)$ is constant:

$$\ell\left(=\frac{\Delta\tau_\nu}{k_\nu}\right) = \frac{1}{k_\nu} \tag{7.5}$$

Eqtn. (7.5) offers insight into the meaning of the source function, which is defined as the ratio of emissivity to opacity (eqtn. 7.2). We see that

$$S_\nu\left(=\frac{j_\nu}{k_\nu}\right) = j_\nu\,\ell,$$

and hence that the source function has a straightforward physical interpretation: it is the the emissivity generated over a photon mean free path (for a column of gas of unit cross-sectional area).

[4] The mean is biassed by the largest values, i.e., by the few photons that travel much further than the m.f.p. This is analogous to a society in which a few people are paid much more than the average; most of the population then earn less than average.

Source function in LTE

Recall that, for systems in (local) thermodynamic equilibrium, j_ν and k_ν are related through the Kirchhoff relation for thermal emission/absorption,

$$j_\nu = k_\nu^{\rm a} B_\nu(T)$$

(where $k_\nu^{\rm a}$ explicitly identifies the opacity as arising through absorption). In this special (but important) case we see that the source function is given by the Planck function:

$$S_\nu = B_\nu. \tag{7.6}$$

In this case we can write eqtn. (7.1) as

$$\begin{aligned} \frac{{\rm d}I_\nu}{{\rm d}s} &= j_\nu - k_\nu I_\nu \\ &= k_\nu^{\rm a} B_\nu - k_\nu^{\rm a} I_\nu \\ &= k_\nu^{\rm a}(B_\nu - I_\nu) \end{aligned} \tag{7.7}$$

Formally, $S_\nu \to B_\nu$ only when collisions dominate over radiative processes; however, in practice, in any situation in which the opacity is very large, the source function can be well approximated by a Planckian. When the opacity is high, photons can only travel short distances, and even if radiative processes dominate over collisions, the *local* conditions have a dominant effect, and the *local* kinetic temperature will control the level populations via the occasional collisional process.

Thus, generally, LTE (with $S_\nu = B_\nu$) may be a reasonable approximation if
(1) the density is high (so that collisions dominate and drive excitation, de-excitation, and ionization levels towards Saha-Boltzmann), and
(2) if the opacity is large (so that photons can only travel short distances).

Scattering source function

Consider conditions under which the opacity arises solely through scattering; then the 'emitted' energy equals the 'absorbed' energy, at every frequency:

$$\begin{aligned} {\rm d}E_\nu({\rm e}) &= {\rm d}E_\nu({\rm a}); \\ \text{i.e.,} \quad \int_{4\pi} j_\nu \, {\rm d}\Omega &= \int_{4\pi} k_\nu^{\rm s} I_\nu \, {\rm d}\Omega, \end{aligned}$$

or, if (as is commonly the case) the emission/absorption coefficients have no overall directionality,

$$j_\nu \int_{4\pi} {\rm d}\Omega = k_\nu^{\rm s} \int_{4\pi} I_\nu \, {\rm d}\Omega. \tag{7.8}$$

Rearranging gives

$$\begin{aligned}\frac{j_\nu}{k_\nu^{\mathrm{s}}} &= \frac{\int_{4\pi} I_\nu \,\mathrm{d}\Omega}{\int_{4\pi} \mathrm{d}\Omega} \\ &= \frac{1}{4\pi}\int_{4\pi} I_\nu \,\mathrm{d}\Omega;\end{aligned}$$

i.e.,

$$S_\nu = J_\nu \tag{7.9}$$

(using eqtns. 7.2 and 3.2 for the left- and right-hand terms, respectively).

Taking this limiting case, eqtn. (7.1) becomes

$$\begin{aligned}\frac{\mathrm{d}I_\nu}{\mathrm{d}s} &= j_\nu - k_\nu I_\nu \\ &= k_\nu^{\mathrm{s}} J_\nu - k_\nu^{\mathrm{s}} I_\nu \\ &= k_\nu^{\mathrm{s}}(J_\nu - I_\nu)\end{aligned} \tag{7.10}$$

Generalized source function

If both scattering and true ('thermal') absorption contribute to the extinction, the source function is a simple linear combination of eqtns. (7.6) and (7.9), weighted by the relative contributions of scattering and absorption coefficients:

$$\begin{aligned}S_\nu &= \frac{k_\nu^{\mathrm{a}}}{k_\nu^{\mathrm{a}} + k_\nu^{\mathrm{s}}} B_\nu + \frac{k_\nu^{\mathrm{s}}}{k_\nu^{\mathrm{a}} + k_\nu^{\mathrm{s}}} J_\nu \\ &= \frac{k_\nu^{\mathrm{a}} B_\nu + k_\nu^{\mathrm{s}} J_\nu}{k_\nu^{\mathrm{a}} + k_\nu^{\mathrm{s}}}\end{aligned} \tag{7.11}$$

Combining eqtns. (7.7) and (7.10)

$$\begin{aligned}\frac{\mathrm{d}I_\nu}{\mathrm{d}s} &= k_\nu^{\mathrm{a}}(B_\nu - I_\nu) + k_\nu^{\mathrm{s}}(J_\nu - I_\nu) \\ &= (k_\nu^{\mathrm{a}} + k_\nu^{\mathrm{s}})(S_\nu - I_\nu), \qquad = k_\nu(S_\nu - I_\nu)\end{aligned}$$

(cp. eqtn. 7.2), where the source function is given by eqtn. (7.11).

Figure 7.1: Radiative transfer along a ray in the direction τ (and increasing s).

7.4 The 'formal solution' along a ray

We now return to consideration of the equation of radiative transfer, eqtn. (7.2), which is a simple differential equation. We seek to obtain a more general solution than the trivial one given by eqtn. (7.3), allowing for $j_\nu \neq 0$.

Starting with

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = S_\nu - I_\nu, \tag{7.2}$$

we multiply both sides by $\exp\{\tau_\nu\}$ to obtain

$$\begin{aligned}
\exp\{\tau_\nu\}\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} + \exp\{\tau_\nu\}I_\nu &= \exp\{\tau_\nu\}S_\nu; \quad \text{that is,} \\
\frac{\mathrm{d}}{\mathrm{d}\tau_\nu}\left(I_\nu \exp\{\tau_\nu\}\right) &= S_\nu \exp\{\tau_\nu\}, \quad \text{or} \\
I_\nu \exp\{\tau_\nu\} &= \int_0^{\tau_\nu} S_\nu(t_\nu)\, \exp\{t_\nu\}\, \mathrm{d}t_\nu + I_\nu(0)
\end{aligned}$$

(where t_ν is a dummy variable of integration and $I_\nu(0)$ is a constant of integration whose value is evident by considering $\tau_\nu = 0$); then, dividing by $\exp\{\tau_\nu\}$,

$$\begin{aligned}
I_\nu(\tau_\nu) &= \int_0^{\tau_\nu} S_\nu(t_\nu)\, \exp\{-\tau_\nu\} \exp\{t_\nu\}\, \mathrm{d}t_\nu \qquad + \quad I_\nu(0)\exp\{-\tau_\nu\} \\
&= \int_0^{\tau_\nu} S_\nu(t_\nu)\, \exp\{-(\tau_\nu - t_\nu)\}\, \mathrm{d}t_\nu \qquad + \quad I_\nu(0)\exp\{-\tau_\nu\}
\end{aligned} \tag{7.12}$$

7.4.1 Interpreting the formal solution

Equation (7.12) is called the *formal solution* for the specific intensity; it shows that
(*i*) the intensities generated at points t_ν suffer extinctions $\exp\{-(\tau_\nu - t_\nu)\}$ before being added to the specific intensity I_ν at the point τ_ν; and
(*ii*) at the point τ_ν along the ray s, the original intensity $I_\nu(0)$ has suffered an extinction $\exp\{-\tau_\nu\}$ (cf. Fig. 7.1). In the special case that S_ν is independent of τ_ν the formal solution simplifies to

$$I_\nu = I_\nu(0)\exp\{-\tau_\nu\} + S_\nu\left(1 - \exp\{-\tau_\nu\}\right).$$

In practice, eqtn. (7.12) is only a formal solution, as it still requires that the source function be known. Nonetheless, there are some simple cases in which it can be applied more or less directly; in particular, we have seen that $S_\nu = B_\nu(T)$ in LTE, so in this case we 'only' require knowledge of the temperature structure.

7.5 Radiative Transfer in Stellar Atmospheres

Having reviewed the principles in the simple case of radiative transfer along a ray, we turn to more general circumstances, where we have to consider radiation coming not just from one direction, but from arbitrary directions. The problem is now three-dimensional in principle; we could treat it in cartesian (xyz) coördinates,[5] but because a major application is in spherical objects (stars!), it's customary to use spherical polar coördinates.

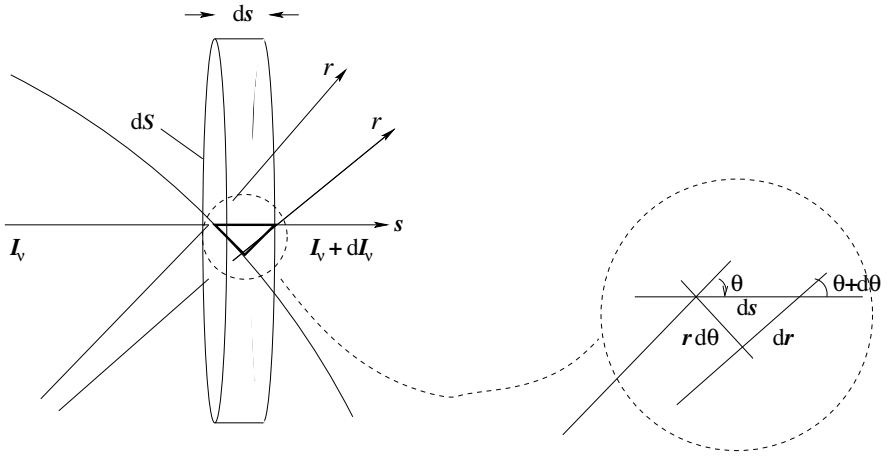


Figure 7.2: Geometry used in radiative-transfer discussion, section 7.5 (cp. Fig. 3.1). The curved line is intended to represent a spherical stellar surface; in the plane-parallel limit, this surface becomes a straight line ($d\theta \to 0$, and the two radii shown become parallel; cp. Fig. 3.1).
The 'outbound' radiation has $0 < \theta < \pi/2$, with $\pi/2 < \theta < \pi$ corresponding to 'inbound' rays.

Again consider a beam of radiation travelling in direction s, at some angle θ to the radial direction in a stellar atmosphere (Fig. 7.2). If we neglect the curvature of the atmosphere (the 'plane-parallel approximation'; $d\theta = 0$) and any azimuthal dependence of the radiation field, then the intensity change along this particular ray is

$$\frac{dI_\nu(\mu)}{ds} = j_\nu - k_\nu I_\nu(\mu), \tag{7.1}$$

[5] We could also treat the problem as time-dependent; but we won't ... A further complication that we won't consider is motion in the absorbing medium (which introduces a directional dependence in k_ν and j_ν); this directionality is important in stellar winds, for example.

as before (where we explicitly identify the μ [$= \cos\theta$] dependence of the specific intensity; in principle, other quantities may also have directional dependencies, but we will assume isotropy).

Overall, we're most often interested in the *radial* transfer of radiative energy; we see from Fig. 7.2 that

$$\mathrm{d}r = \cos\theta\,\mathrm{d}s \equiv \mu \mathrm{d}s$$

so the transfer in the radial direction is described by

$$\frac{\mu}{k_\nu}\frac{\mathrm{d}I_\nu(\mu)}{\mathrm{d}r} = \frac{j_\nu}{k_\nu} - I_\nu(\mu), \quad = S_\nu(\tau_\nu) - I_\nu(\mu) \tag{7.13}$$

(where we've divided through by k_ν); and since $\mathrm{d}\tau_\nu = -k_\nu\,\mathrm{d}r$ (eqtn. 4.6, introducing a minus because the sign convention in stellar-atmosphere work is such that optical depth *increases* with *decreasing* r) we have

$$\mu\frac{\mathrm{d}I_\nu(\mu)}{\mathrm{d}\tau_\nu} = I_\nu(\mu) - S_\nu(\tau_\nu). \tag{7.14}$$

This is the standard formulation of the equation of transfer in plane-parallel stellar atmospheres. It differs from the previous formulation, eqtn. (7.2), simply by the projection factor $\mu = \cos\theta$, allowing for the optical depth to be measured at some angle θ to the radial direction (and by the change of sign, which arises simply because of the change in the direction in which τ increases).

7.6 Generalized geometry, and spherical symmetry

[Omitted from lectures]

For arbitrary geometry we have to consider the full three-dimensional characterization of the radiation field; that is

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = \frac{\partial I_\nu}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}s} + \frac{\partial I_\nu}{\partial\theta}\frac{\mathrm{d}\theta}{\mathrm{d}s} + \frac{\partial I_\nu}{\partial\phi}\frac{\mathrm{d}\phi}{\mathrm{d}s}, \tag{7.15}$$

where r, θ, ϕ are our spherical polar coördinates. This is our most general formulation, but in the case of stellar atmospheres we can often neglect the ϕ dependence; and we rewrite the θ term by noting not only that

$$\mathrm{d}r = \cos\theta\,\mathrm{d}s \equiv \mu \mathrm{d}s$$

but also that

$$-r\,\mathrm{d}\theta = \sin\theta\,\mathrm{d}s.$$

(The origin of the minus sign may be clarified by reference to Fig. 7.2; for increasing s we have increasing r, but decreasing θ, so $r\,\mathrm{d}\theta$ is negative for positive $\mathrm{d}s$.)

Using these expressions in eqtn. (7.15) gives a two-dimensional form,

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}s} = \frac{\partial I_\nu}{\partial r}\cos\theta - \frac{\partial I_\nu}{\partial \theta}\frac{\sin\theta}{r}$$

but this is also

$$= j_\nu - k_\nu I_\nu \tag{7.1}$$

so, dividing through by k_ν as usual,

$$\begin{aligned}\frac{\cos\theta}{k_\nu}\frac{\partial I_\nu}{\partial r} - \frac{\sin\theta}{k_\nu r}\frac{\partial I_\nu}{\partial \theta} &= \frac{j_\nu}{k_\nu} - I_\nu \\ &= S_\nu - I_\nu\end{aligned}$$

Once again, it's now useful to think of a point in the atmosphere in terms of the optical depth measured radially inwards:

$$\mathrm{d}\tau = -k_\nu\,\mathrm{d}r,$$

which gives us the customary form of the equation of radiative transfer for use in extended stellar atmospheres, for which the plane-parallel approximation fails:

$$\frac{\sin\theta}{\tau_\nu}\frac{\partial I_\nu}{\partial \theta} - \mu\frac{\partial I_\nu}{\partial \tau_\nu} = S_\nu - I_\nu. \tag{7.16}$$

We recover our previous, plane-parallel, result if the atmosphere is very thin compared to the stellar radius. In this case, the surface curvature shown in Fig. 7.2 becomes negligible, and the first term tends to zero, recovering

$$\mu\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = I_\nu - S_\nu, \tag{7.14}$$

which is our previous formulation of the equation of radiative transfer in plane-parallel stellar atmospheres.

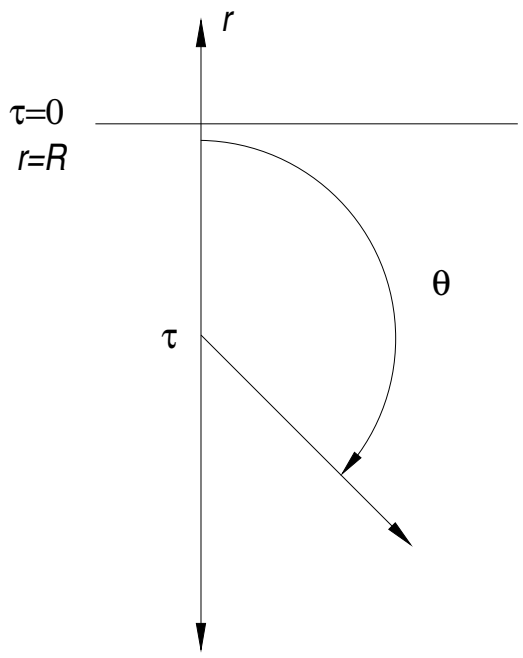
7.7 Plane-parallel atmospheres: formal solution

The difference between eqtns. (7.2) and (7.14) is simply the projection factor $\mu = \cos\theta$ (and a sign change arising because of a change in the convention adopted for the direction of increasing τ).[6] The formal solution to eqtn. (7.14) follows exactly as before (Section 7.4, eqtn. 7.12), but with τ_ν replaced by $-\tau_\nu/\cos\theta, = -\tau_\nu/\mu$ (where again the minus sign arises because we're now using the convention that τ_ν increases inwards, not outwards), yielding

$$I_\nu(\theta, \tau_\nu) = -\int_{\mathsf{c}}^{\tau_\nu} S_\nu(t_\nu)\exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\}\frac{\mathrm{d}t_\nu}{\mu}. \tag{7.17}$$

[6] We recover eqtn. (7.2) simply by setting $\mu = -1$ in eqtn. (7.14)

Here the lower limit of integration c depends on the boundary conditions (and substitutes for the 'constant of integration' term in eqtn. (7.12). The obvious boundary to consider is the outer limit of the atmosphere, which we define as being at $\tau_\nu = 0$; at this point there is no radiation coming from 'above'.[7]



Then at some depth point τ_ν, for the inwards radiation ($\pi > \theta > \pi/2$; see figure above) we have

$$I_\nu^{\text{in}}(\theta, \tau_\nu) = -\int_0^{\tau_\nu} S_\nu(t_\nu) \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{\mathrm{d}t_\nu}{\mu}. \tag{7.18}$$

For 'outbound' radiation ($0 < \theta < \pi/2$) we consider a t_ν-range up to ∞ (that is, as deep into the atmosphere as it's possible to see incoming radiation reaching the point τ_ν):

$$I_\nu^{\text{out}}(\theta, \tau_\nu) = -\int_\infty^{\tau_\nu} S_\nu(t_\nu) \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{\mathrm{d}t_\nu}{\mu} \tag{7.19}$$

The total intensity at the depth point characterized by τ_ν is just the sum of the inward and outward terms:

$$\begin{aligned} I_\nu(\theta, \tau_\nu) = &\int_{\tau_\nu}^{\infty} S_\nu(t_\nu) \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{\mathrm{d}t_\nu}{\mu} \\ &- \int_0^{\tau_\nu} S_\nu(t_\nu) \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{\mathrm{d}t_\nu}{\mu} \end{aligned} \tag{7.20}$$

(where the limits of integration of the first, 'outwards', term have been reversed, with an accompanying sign change). This is our formal solution for radiative transfer in a plane-parallel medium, analogous to eqtn. (7.12).

[7] Other than from distant stars, etc. – which is negligible.

Of special interest is the outwards (i.e., observable) radiation at $\tau_\nu = 0$:

$$I_\nu^{\rm out}(\theta, 0) = \int_0^\infty S_\nu(t_\nu) \exp\left\{\frac{-t_\nu}{\mu}\right\} \frac{{\rm d}t_\nu}{\mu} \tag{7.21}$$

For most stars, we can't measure the (angle-dependent) intensities, but in a few selected cases – most obviously the Sun – we can measure the specific intensity as a function of θ. More usually, we can measure only the *flux* from the distant stars – that is, the first moment of the radiation field.

7.7.1 Formal solutions for J_ν, F_ν, and K_ν (the Schwarzschild-Milne Equations)

By using the formal solution for the specific intensity, it is now straightforward to derive the formal solutions for the 0th, 1st and 2nd moments of the radiation field – the mean intensity J_ν, the Eddington flux H_ν (or, almost equivalently, the physical flux F_ν), and the K integral K_ν. These are of importance in calculating atmospheric structures and emergent spectra.

Recall that the mean intensity is defined as

$$J_\nu \equiv \frac{1}{4\pi} \int_{4\pi} I_\nu(\theta, \phi)\, {\rm d}\Omega = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^{+1} I_\nu(\mu, \phi)\, {\rm d}\mu\, {\rm d}\phi \tag{3.2}$$

(where $\mu = \cos\theta$, and the integral is the zeroth moment of the radiation field). If we again assume no azimuthal dependence[8] of I_ν then

$$J_\nu = \frac{1}{2} \int_{-1}^{+1} I_\nu(\mu)\, {\rm d}\mu, \quad = \frac{1}{2}\left[\int_0^{+1} I_\nu^{\rm out}(\mu)\, {\rm d}\mu + \int_{-1}^{0} I_\nu^{\rm in}(\mu)\, {\rm d}\mu\right].$$

Using eqtns. (7.18) and (7.19), and assuming that S_ν is isotropic[9] (so we can take it out of the μ integral), gives

$$\begin{aligned} J_\nu(\tau_\nu) = \frac{1}{2}\Bigg[&\int_{\tau_\nu}^\infty S_\nu(t_\nu) \int_0^{+1} \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{{\rm d}\mu}{\mu}\, {\rm d}t_\nu \\ &- \int_0^{\tau_\nu} S_\nu(t_\nu) \int_{-1}^{0} \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{{\rm d}\mu}{\mu}\, {\rm d}t_\nu\Bigg] \end{aligned} \tag{7.22}$$

Setting $w = 1/\mu$ (so that ${\rm d}\mu/\mu = -\,{\rm d}w/w$) and $x = t_\nu - \tau_\nu$, the integral over ${\rm d}\mu$ in the first term on the right-hand side of eqtn. (7.22) can be written as

$$\int_0^{+1} \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{{\rm d}\mu}{\mu} \rightarrow \int_1^\infty \exp\{-xw\} \frac{{\rm d}w}{w}, \equiv \int_1^\infty \frac{\exp\{-xw\}}{w^1}\, {\rm d}w.$$

The this has the form of an exponential integral, a standard form with well-studied properties:

$$E_n(x) = \int_1^\infty \exp\{(-xw)\}/w^n\, {\rm d}w$$

[8] That is, we set $\int_0^{2\pi} I_\nu\, {\rm d}\phi = I_\nu \int_0^{2\pi} {\rm d}\phi = 2\pi I_\nu$.

[9] This is a good assumption in static media.

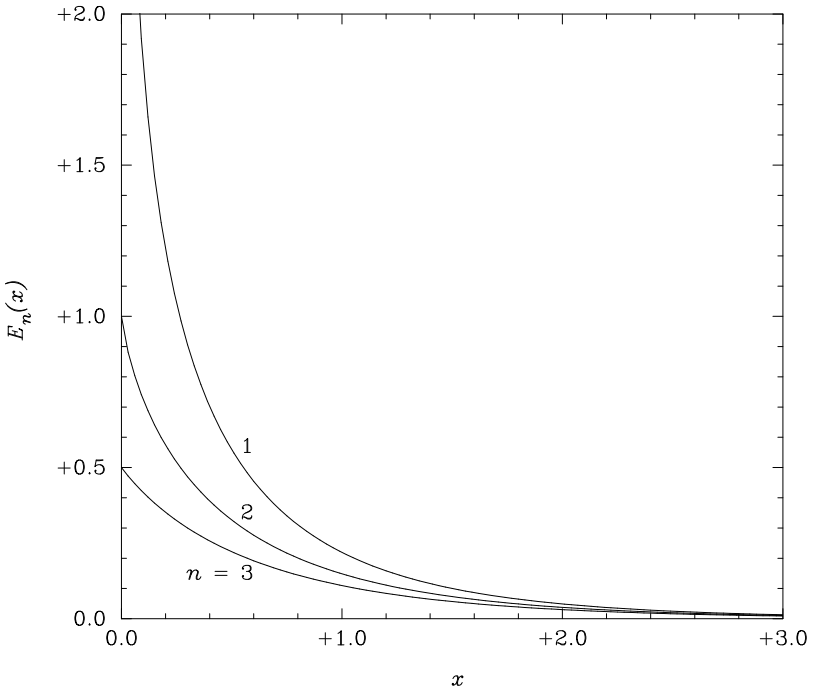


Figure 7.3: The exponential integral for $n = 1, 2, 3$, as a function of $x(= t_\nu - \tau_\nu)$.

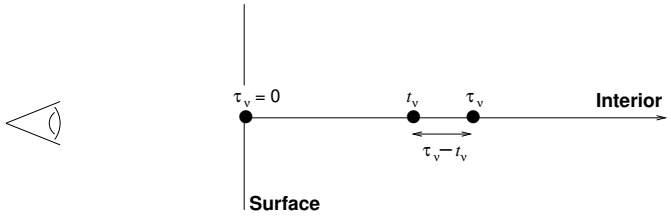
with $n = 1$ (cf. Fig. 7.3). Similarly, setting $x = \tau_\nu - t_\nu$ and $w = -1/\mu$, then the $\mathrm{d}\mu$ integral in the second term may also be written as

$$\int_{-1}^{0} \exp\left\{\frac{-(t_\nu - \tau_\nu)}{\mu}\right\} \frac{\mathrm{d}\mu}{\mu} \rightarrow + \int_{1}^{\infty} \exp\{-xw\} w \frac{\mathrm{d}w}{w^2}, = \int_{1}^{\infty} \frac{\exp\{-xw\}}{w^1}\, \mathrm{d}w.$$

We can therefore write eqtn. (7.22) as

$$J_\nu(\tau_\nu) = \frac{1}{2}\left[\int_{\tau_\nu}^{\infty} S_\nu(t_\nu)\, E_1(t_\nu - \tau_\nu)\, \mathrm{d}t_\nu + \int_{0}^{\tau_\nu} S_\nu(t_\nu)\, E_1(\tau_\nu - t_\nu)\, \mathrm{d}t_\nu\right]. \tag{7.23}$$

This is the formal solution we seek, and is analogous to eqtn. (7.20) for the intensity. It was first derived by Karl Schwarzschild, and is therefore sometimes called 'Schwarzschild's equation' in the context of stellar atmospheres.[10]



Over the given limits of integration in t_ν, $t_\nu - \tau_\nu$ is always positive in the first term of eqtn. (7.23), and $\tau_\nu - t_\nu$ is always positive in the second term (see figure above). We can therefore write Schwarzschild's

[10] 'Schwarzschild's equation' refers to a different form in discussions of planetary atmospheres.

equation in slightly more compact form,

$$J_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu)\, E_1(|t_\nu - \tau_\nu|)\, \mathrm{d}t_\nu.$$

This form is sufficiently common in stellar-atmosphere studies that it is further abbreviated using operator notation,

$$\begin{aligned} J_\nu(\tau_\nu) &= \Lambda_{\tau_\nu}[S_\nu(t_\nu)], \text{ where} \\ \Lambda_{\tau_\nu}[f(t_\nu)] &\equiv \frac{1}{2}\int_0^\infty f(t_\nu)\, E_1(|t_\nu - \tau_\nu|)\, \mathrm{d}t_\nu. \end{aligned}$$

— Completely analogous arguments show that the flux,

$$F_\nu \equiv \int I_\nu(\theta,\phi)\mu\, \mathrm{d}\Omega = \int_0^{2\pi}\int_{-1}^{+1} I_\nu(\mu)\,\mu\, \mathrm{d}\mu\, \mathrm{d}\phi$$

(the first moment of the radiation field) may be written as[11]

$$F_\nu(\tau_\nu) = 2\pi\left[\int_{\tau_\nu}^\infty S_\nu(t_\nu)\, E_2(t_\nu - \tau_\nu)\, \mathrm{d}t_\nu - \int_0^{\tau_\nu} S_\nu(t_\nu)\, E_2(\tau_\nu - t_\nu)\, \mathrm{d}t_\nu,\right] \tag{7.24}$$

($= \Phi_{\tau_\nu}[S_\nu(t_\nu)]$ in operator notation). The surface flux (at $\tau_\nu = 0$) is therefore

$$F_\nu(0) = 2\pi\int_0^\infty S_\nu(t_\nu) E_2(t_\nu)\, \mathrm{d}t_\nu; \tag{7.25}$$

We see that it is the integral of the source function at every depth, multiplied by an extinction factor $E_2(t_\nu)$ relevant to that depth.

— Finally, the K integral (the second-order moment) is given by

$$\begin{aligned} K_\nu(\tau_\nu) &= \frac{1}{2}\int_{-1}^{+1} I_\nu(\mu)\mu^2\, \mathrm{d}\mu \\ &= \frac{1}{2}\left[\int_{\tau_\nu}^\infty S_\nu(t_\nu)\, E_3(t_\nu - \tau_\nu)\, \mathrm{d}t_\nu + \int_0^{\tau_\nu} S_\nu(t_\nu)\, E_3(\tau_\nu - t_\nu)\, \mathrm{d}t_\nu\right] \\ &= \frac{1}{2}\int_0^\infty S_\nu(t_\nu)\, E_3(|t_\nu - \tau_\nu|)\, \mathrm{d}t_\nu, \qquad (\equiv X_{\tau_\nu}[S_\nu(t_\nu)] \text{ in operator notation}) \end{aligned} \tag{7.26}$$

Equations (7.23), (7.24), and (7.26) – collectively, the Schwarzschild–Milne equations – are the formal solutions for the first three moments of the specific intensity, i.e. they are the integral expression for mean intensity, flux, and K integral (J_ν, F_ν, K_ν). They are all depth-weighted averages of the source function S_ν, with the strongest contribution coming from the depth at which the argument of the exponential integrals is zero (i.e., at $t = \tau_\nu$).

[11] Note that event moments are represented by the sum of two integrals, and odd moments by the difference.

Section 8

A simple model atmosphere
Radiative equilibrium, the grey atmosphere, & the Eddington two-stream approximation

Classically, the model-atmosphere problem refers to the construction of mathematical models that provide a description of the physical structure of a stellar atmosphere and of its emerging spectrum. In principle, the formal solution of the equation of radiative transfer allows us to solve for the surface intensity (and hence a model spectrum). However, recalling the transfer equation in the form

$$\mu \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = \frac{\mu}{k_\nu}\frac{\mathrm{d}I_\nu}{\mathrm{d}r} = S_\nu - I_\nu \tag{7.13}$$

(appropriate to plane-parallel atmospheres) we see that, in order to solve for I_ν in practice, we need to know something about S_ν and k_ν (and both are functions of temperature, so we need to know the run of temperature through the atmosphere).

This is a problem of, potentially, enormous complexity: the occupation numbers depend on the radiation field, which depends on the occupation numbers; energy transport may be by radiation and convection; there may be a dynamical or time-dependent component of the atmosphere; the local effective temperature and gravity may vary over the surface (due to rotation, or a binary companion). In its most general formulation, the problem remains beyond practical solution at present (although increasing computational power is gradually allowing more and more approximations to be relaxed, and more and more physics to be incorporated into the models), and it is necessary to make several simplifications, creating models that are idealized to greater or lesser extent. In this section, we'll see how the condition of *radiative equilibrium* allows us to make progress with the temperature structure, and we'll use the *grey atmosphere* as the simplest approach to the opacity. Together, they allow a simple solution for the source function.

8.1 Basic assumptions

When using idealized models (as we shall do), it is crucial to understand in depth the restrictions that we are imposing: in this way can define the problem, understand the results in the context of the assumptions, and appreciate issues that remain to be explored.

The main simplifying assumptions that we will adopt in order to build a 'classical' model atmosphere relate to:

8.1.1 Geometry

We assume that the atmosphere is made of plane-parallel, homogeneous layers.
– The plane-parallel ('semi-infinite') atmosphere is a good approximation when the thickness of the atmosphere is small compared to the stellar radius.
– Homogeneity makes the problem one-dimensional (in the sense that all variables are functions only of depth). While this is a great simplification, we lose the ability to model many interesting phenomena, but a homogeneous, 1-D model should give us some kind of spatially averaged information.

We should note that for cool stars (F-type and later), many inhomogeneities arise through hydrodynamic phenomena in the convection zone; therefore we expect homogeneous models to better approximate stars with radiative envelopes. On the other hand, Ap stars (for example) show large gradients in the physical quantities, possibly because of strong magnetic fields, so in this case a plane-parallel homogeneous model is probably not a good approximation.

8.1.2 Steady state

We assume that the atmosphere is
– In steady state (i.e., we neglect all time-dependent phenomena such as pulsations, shocks etc.), and
– Static (we assume that level populations are constant in time and given by statistical equilibrium equations, and that hydrostatic equilibrium applies [no stellar winds]).

8.1.3 Momentum and energy balance

We assume hydrostatic equilibrium (momentum balance) and radiative equilibrium (energy balance). The former assumption implies that the pressure gradient balances the gravitational forces and is consistent with a static atmosphere; the latter implies that the energy-integrated flux is conserved along the x-coördinate which defines our 'one dimensional' directionality.

In summary, we will construct 1-D (plane-parallel), static models in radiative and hydrostatic equilibrium.

8.2 Flux constancy and radiative equilibrium

There are no important nett sources or sinks of energy in the envelope and atmosphere; the total (radiative + convective + conductive) energy flux generated in the core is simply transported outward, giving rise to the condition of *flux constancy.* Formally, the divergence of the total energy flux is zero, or more simply

$$\frac{\mathrm{d}\mathcal{F}(r)}{\mathrm{d}r} = 0; \tag{8.1}$$

that is, $\mathcal{F}(r)$, the total energy flux through some sphere of radius r, is a constant – say, $\mathcal{F}_0$.

At the photosphere, more or less by definition, radiation escapes readily, suggesting at once that this it is the principal means of energy transport (if only because it can carry away energy much faster than any other energy-loss process). K. Schwarzschild formally demonstrated this to be the case in practice (even for a star like the Sun, in which convective energy transport in the immediately sub-photospheric layers is important).

If we consider those regions in which the energy transport is dominated by radiation (such as the stellar atmosphere), then

$$\frac{\mathcal{F}(r)}{4\pi r^2} \simeq \int_0^\infty F_\nu(r)d\nu = \text{ a constant, } F_0 \left(= \sigma T_{\rm eff}^4\right) \tag{8.2}$$

where the formal solution for the radiative flux per unit area, F_ν, is given by eqtn. (7.24). If eqtn. (8.2) is valid throughout the atmosphere, then the atmosphere is said to be in radiative equilibrium. Radiative equilibrium is thus a special case of flux constancy (where the 'flux' in 'flux constancy' means the total energy flux, not just the radiative energy flux).

The (radiative) flux constant is related to the effective temperature through $F_0 = \sigma T_{\rm eff}^4$ (eqtn. C.7). When we construct a model atmosphere under the condition of radiative equilibrium, $T_{\rm eff}$ may therefore be selected as one of the fundamental parameters characterizing the model. However, we need to be able to express radiative equilibrium in terms of quantitities that enter directly into the radiative transfer we've previously discussed. This can be formulated in a number of ways, as pioneered by E.A. Milne.

8.2.1 Radiative equilibrium: Milne's formulation – 1

We start by taking the zeroth moment of the transfer equation in the form

$$\mu\frac{\mathrm{d}I_\nu}{\mathrm{d}r} = k_\nu S_\nu - k_\nu I_\nu; \tag{7.13}$$

that is, multiply both sides by μ^0 (= 1) and integrate over solid angle, assuming that k_ν and S_ν are isotropic (i.e., locally independent of direction):

$$\frac{\mathrm{d}}{\mathrm{d}r}\int_{4\pi} I_\nu(\mu,\phi)\mu\,\mathrm{d}\Omega = k_\nu S_\nu \int_{4\pi}\mathrm{d}\Omega - k_\nu\int_{4\pi} I_\nu(\mu,\phi)\,\mathrm{d}\Omega, \quad \text{i.e.,}$$

$$\frac{\mathrm{d}}{\mathrm{d}r}\int_0^{2\pi}\int_{-1}^{+1} I_\nu(\mu,\phi)\mu\,\mathrm{d}\mu\mathrm{d}\phi = k_\nu S_\nu \int_0^{2\pi}\int_{-1}^{+1}\mathrm{d}\mu\mathrm{d}\phi - k_\nu\int_0^{2\pi}\int_{-1}^{+1} I_\nu(\mu,\phi)\,\mathrm{d}\mu\mathrm{d}\phi$$

or, if there is no ϕ dependence of I_ν,

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[2\pi\int_{-1}^{+1} I_\nu(\mu)\mu\,\mathrm{d}\mu\right] = 4\pi k_\nu S_\nu - 2\pi k_\nu\int_{-1}^{+1} I_\nu(\mu)\,\mathrm{d}\mu. \tag{8.3}$$

The term in square brackets is 4π times the first moment of the radiation field – i.e., it is the flux F_ν (4π times H_ν; §3.4). Similarly, the second integral is twice J_ν (the mean intensity, or zeroth moment of the radiation field); that is,

$$\frac{\mathrm{d}F_\nu}{\mathrm{d}r} = 4\pi k_\nu S_\nu - 4\pi k_\nu J_\nu. \tag{8.4}$$

We're concerned with the total radiative energy flux (and not with the frequency dependence), so we should integrate over frequency:

$$\frac{\mathrm{d}}{\mathrm{d}r}\int_0^\infty F_\nu\,\mathrm{d}\nu = 4\pi\int_0^\infty k_\nu S_\nu\,\mathrm{d}\nu - 4\pi\int_0^\infty k_\nu J_\nu\,\mathrm{d}\nu.$$

Applying the condition of radiative equilibrium, the left-hand side is equal to zero (eqtns. 8.1, 8.2); thus

$$\int_0^\infty k_\nu J_\nu\,\mathrm{d}\nu = \int_0^\infty k_\nu S_\nu\,\mathrm{d}\nu. \tag{8.5}$$

This is our first way of formally expressing the condition of radiative equilibrium. It's important to be clear that it does *not* imply that $J_\nu = S_\nu$, frequency by frequency; only the integrated quantities are equal.

8.2.2 Radiative equilibrium: Milne's formulation – 2

To obtain a second formulation of interest, we again start from the equation of radiative transfer in the form of eqtn. (7.13), but now multiply both sides by $\mu^1(=\cos\theta)$. As in the derivation of eqtn. (8.5), we then perform an integral over solid angle (assuming that k_ν and S_ν are isotropic and that there is no ϕ dependence of I_ν; cp. eqtn. 8.3):

$$\frac{\mathrm{d}}{\mathrm{d}r}\int_{4\pi} I_\nu(\mu)\mu^2\,\mathrm{d}\Omega = k_\nu\int_{4\pi} S_\nu\mu\mathrm{d}\Omega - k_\nu\int_{4\pi} I_\nu(\mu)\mu\,\mathrm{d}\Omega \tag{8.6}$$

$$\Rightarrow \quad 2\frac{\mathrm{d}K_\nu}{\mathrm{d}r} = k_\nu S_\nu\int_{-1}^{+1}\mu\,\mathrm{d}\mu - 2k_\nu H_\nu$$

$$\Rightarrow \quad \frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu} = \frac{F_\nu}{4\pi}\ (\text{using } \mathrm{d}\tau_\nu = -k_\nu\mathrm{d}r) \tag{8.7}$$

where K_ν is the second moment of the radiation field. As before, we integrate over frequency to get the total radiative flux:

$$\int_0^\infty \frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu}\,\mathrm{d}\nu = \int_0^\infty \frac{F_\nu}{4\pi}\,\mathrm{d}\nu$$

or, in radiative equilibrium (eqtn. 8.2),

$$\int_0^\infty \frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu}\,\mathrm{d}\nu = \frac{F_0}{4\pi}, \quad \equiv H_0 \tag{8.8}$$

where H_0 is the (frequency-integrated) Eddington flux.

This is another useful way of formulating radiative equilibrium. Of course, eqtns. (8.5) and (8.8) are not independent; they are just different ways to express the same condition (radiative equilibrium, ≃flux conservation), which are useful under different circumstances.

8.2.3 Radiative equilibrium: Milne's formulation – 3 [Not for lectures]

We obtained the first and second Milne equations essentially by taking the zeroth and first moments of the equation of radiative transfer in a plane-parallel medium. A third equation can be obtained by taking the second moment – multiply both sides eqtn. (7.13) by μ^2, and integrate over solid angle. By using the formal solution for K_ν, eqtn. (7.26), we obtain

$$\int_0^\infty \frac{\mathrm{d}}{\mathrm{d}\tau_\nu}\left[\frac{1}{2}\int_{\tau_\nu}^\infty S_\nu(t_\nu)\,E_3(t_\nu-\tau_\nu)\,\mathrm{d}t_\nu + \frac{1}{2}\int_0^{\tau_\nu} S_\nu(t_\nu)\,E_3(\tau_\nu-t_\nu)\,\mathrm{d}t_\nu\right]\mathrm{d}\nu = \frac{F_0}{4\pi}$$

This relatively unwieldy form is Milne's third equation expressing radiative equilibrium.

8.3 The grey atmosphere

The case in which the opacity k_ν is independent of frequency – the 'grey case' – merits special consideration because of its amenability to semi-analytical solutions. However, the only important source of opacity that is truly grey is electron scattering, so the grey atmosphere is generally only useful for approximate or qualitative results; but it can nonetheless give useful physical insights that may be lacking in entirely numerical approaches.

Under the grey assumption, then integrating the transfer equation over frequency, and setting $k_\nu = k(\neq f(\nu))$, we obtain:

$$\mu\frac{\mathrm{d}I}{\mathrm{d}r} = kS - kI, \quad \text{or} \quad \mu\frac{\mathrm{d}I}{\mathrm{d}\tau} = I - S \tag{8.9}$$

where $I = \int_0^\infty I_\nu(\mu)\,\mathrm{d}\nu$ and $S = \int_0^\infty S_\nu\,\mathrm{d}\nu$ are the frequency-integrated specific intensity and the frequency-integrated source term, respectively, and τ is the optical depth (now independent of frequency).

This is an enormous simplication: the total specific intensity I will be known for each ray if we solve just one equation, instead of an effectively infinite set (one for each frequency). This simplification extends to the equations of radiative equilibrium; the simplified forms of eqtns. (8.2), (8.5) and (8.8) are

$$F = F_0 \qquad (8.10 \Leftarrow 8.2)$$

$$J = S \qquad (8.11 \Leftarrow 8.5)$$

$$\frac{\mathrm{d}K}{\mathrm{d}\tau} = \frac{F_0}{4\pi} \qquad (8.12 \Leftarrow 8.8)$$

(where $F = \int_0^\infty F_\nu(\mu)\,\mathrm{d}\nu$; $K = \int_0^\infty K_\nu\,\mathrm{d}\nu$; and the subscript ν has been dropped from the optical depth – i.e., $\tau_\nu \to \tau$, with no frequency dependence).

8.3.1 Source function: the Eddington two-stream approximation

In order to find a solution for the transfer equation in the grey case, note that, under the constraint of radiative equilibrium, the integral quantities J and S are the same (eqtn. $8.11 \Leftarrow 8.5$). In general, solving the transfer equation means having to find a solution for S; but in this simplified case we need 'only' to find a solution for J.

To do this, we begin by integrating eqtn. ($8.12 \Leftarrow 8.8$),

$$K(\tau) = \frac{F_0}{4\pi}\tau + \text{const.} \qquad (8.13)$$

A solution can be obtained if both K and F_0 can be expressed in terms of the intensity. In order to progress, we make use of the Eddington (1926) 'two-stream' approximation. Under this approximation, the specific intensity is characterized by only two angle-independent terms, one propagating inward and one outward:

$$I(\theta,\tau) = I^{\mathrm{out}}(\tau) \qquad \text{for } \theta \le \pi/2 \qquad (+1 \le \mu \le 0)$$

$$I(\theta,\tau) = I^{\mathrm{in}}(\tau) \qquad \text{for } \theta > \pi/2 \qquad (0 < \mu \le -1);$$

that is, under the two-stream approximation, at some given value of τ the specific intensity is constant in each hemisphere.

We can now evaluate the zeroth, first, and second moments of the radiation field. The mean intensity at

depth τ may be written as

$$\begin{aligned}
J(\tau) \equiv \int I(\theta,\phi,\tau)\frac{\mathrm{d}\Omega}{4\pi} \qquad &= \frac{1}{4\pi}\int_0^{2\pi}\int_{-1}^{+1} I(\mu,\phi,\tau)\,\mathrm{d}\mu\,\mathrm{d}\phi \\
&= \frac{1}{2}\int_0^{+1} I(\mu,\tau)\,\mathrm{d}\mu + \frac{1}{2}\int_{-1}^{0} I(\mu,\tau)\,\mathrm{d}\mu \qquad \text{(zeroth-order moment, assuming azimuthal symmetry)} \\
&= \frac{1}{2}I^{\mathrm{out}}(\tau)\int_0^{+1}\mathrm{d}\mu + \frac{1}{2}I^{\mathrm{in}}(\tau)\int_{-1}^{0}\mathrm{d}\mu \\
&= \frac{1}{2}\left[I^{\mathrm{out}}(\tau) + I^{\mathrm{in}}(\tau)\right] \qquad (8.14)
\end{aligned}$$

(that is, the mean intensity is just the average of intensities in the two hemispheres – unsurprisingly!). Similarly, the physical flux is given by

$$\begin{aligned}
F(\tau) = 4\pi H \qquad &= \int I(\theta,\phi,\tau)\mu\,\mathrm{d}\Omega \\
&= 2\pi\int_0^{+1} I(\mu,\tau)\mu\,\mathrm{d}\mu + 2\pi\int_{-1}^{0} I(\mu,\tau)\mu\,\mathrm{d}\mu \qquad \text{(first-order moment } \times 4\pi) \\
&= \pi\left[I^{\mathrm{out}}(\tau) - I^{\mathrm{in}}(\tau)\right] \qquad (8.15)
\end{aligned}$$

(which again matches common sense: the outwards flux of radiation should have something to do with the difference between the 'upwards' and 'downwards' intensities). Finally, the K integral (which we need in order to tackle eqtn. 8.13) is

$$\begin{aligned}
K(\tau) &= \int I(\theta,\phi,\tau)\mu^2\frac{\mathrm{d}\Omega}{4\pi} \qquad \text{(second-order moment)} \\
&= \frac{1}{6}\left[I^{\mathrm{out}}(\tau) + I^{\mathrm{in}}(\tau)\right] \\
&= \frac{J(\tau)}{3} \quad \text{(from eq. 8.14).} \qquad (8.16)
\end{aligned}$$

At the surface, $\tau \to 0$ and $I^{\mathrm{in}} \to 0$, so

$$\begin{aligned}
&J(0) = \frac{1}{2}I^{\mathrm{out}}(0); \qquad F(0) = \pi I^{\mathrm{out}}(0); \qquad K(0) = \frac{1}{6}I^{\mathrm{out}}(0), \\
\rightarrow &J(0) = \frac{F(0)}{2\pi} \equiv \frac{F_0}{2\pi}; \qquad (8.17) \\
&K(0) = \frac{F(0)}{6\pi} \equiv \frac{F_0}{6\pi}
\end{aligned}$$

(in radiative equilibrium), which gives us the constant of integration in eqtn. (8.13):

$$K(\tau) = \frac{F_0}{4\pi}\tau + \frac{F_0}{6\pi}, = \frac{J(\tau)}{3} \quad \text{(from eqtn. 8.16)} \qquad (8.18)$$

whence

$$J(\tau) = \frac{3F_0}{4\pi}\left(\tau + \frac{2}{3}\right), = S(\tau) \quad \text{(from eqtn. 8.11} \Leftarrow \text{8.5).} \qquad (8.19)$$

This gives us the required source function for a grey atmosphere (eqtn. 8.11 ⇐ 8.5), under the condition of radiative equilibrium (eqtn. 8.12 ⇐ 8.8) in the Eddington (two-stream) approximation (eqtns. 8.14–8.16). Note the important result that, in this case, $S(\tau)$ is a linear function of optical depth (i.e., $S = a + b\tau$) – note also that this is not true in general!

A more rigorous treatment of the grey problem, relaxing the Eddington two-stream approximation in favour of a proper angle integration (Chandrasekhar 1957), gives only a slightly different result:

$$J(\tau) = \frac{3F_0}{4\pi}\left[\tau + q(\tau)\right] \tag{8.20}$$

where $q(\tau)$, called the Hopf function, is a slowly varying function which ranges from 0.577 at $\tau = 0$ to 0.710 as $\tau \to \infty$ (compare with the factor $2/3 \simeq 0.67$ appearing in the two-stream solution).

8.3.2 Temperature structure

We have seen that, in radiative equilibrium,

$$\int_0^\infty k_\nu J_\nu \, \mathrm{d}\nu = \int_0^\infty k_\nu S_\nu \, \mathrm{d}\nu, \tag{8.5}$$

so for grey opacities and LTE ($S_\nu = B_\nu$) we have

$$\begin{aligned}\int_0^\infty J_\nu \, \mathrm{d}\nu &= \int_0^\infty B_\nu \, \mathrm{d}\nu, \\ &= \sigma T^4/\pi.\end{aligned}$$

That is, the frequency-integrated mean intensity at some optical depth τ is

$$J(\tau) = \frac{\sigma}{\pi} T^4(\tau)$$

Comparing this to eqtn. (8.19), and using $F_0 = \sigma T_{\rm eff}^4$, gives us the temperature profile of the grey atmosphere:

$$T^4(\tau) = \frac{3}{4}\left(\tau + \frac{2}{3}\right) T_{\rm eff}^4. \tag{8.21}$$

[Relaxing the two-stream again gives only a slightly different result:

$$T(\tau) = \left\{\frac{3}{4}\left[\tau + q(\tau)\right]\right\}^{1/4} T_{\rm eff} \tag{8.22}$$

where $q(\tau)$ is the Hopf function.]

Note that

- larger τ corresponds to larger T – temperature increases with increasing depth in the atmosphere (although $T_{\rm eff}$ is constant, T is not!);

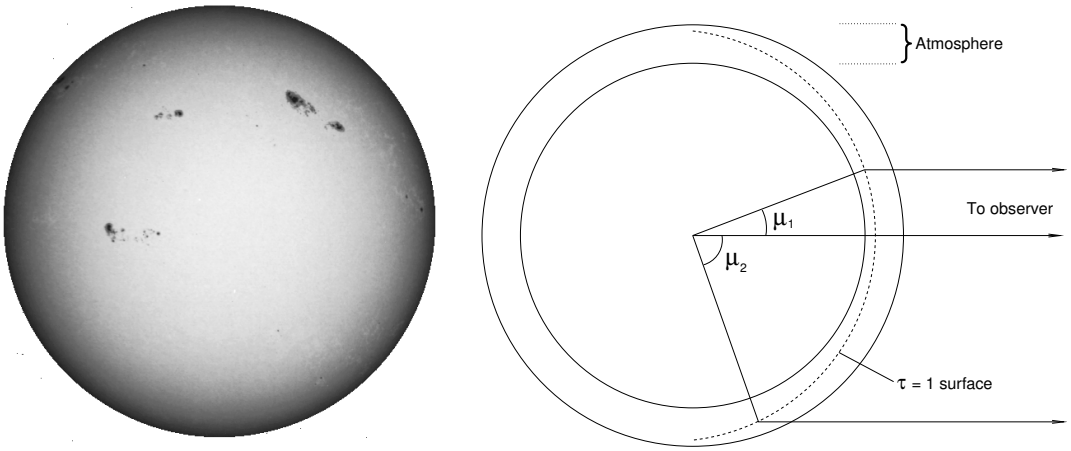


Figure 8.1: Left, white-light image of the Sun, showing limb darkening; right, limb-darkening geometry. The "$\tau = 1$ surface" illustrates the locus of constant *line of sight* optical depth; the corresponding *radial* optical depth is smaller at the limb than at the centre of the disk, so we see deeper into the atmosphere at the centre.
[The angles are labelled with 'μ' ($= \cos\theta$), so *numerically* $\mu_1 > \mu_2$.]

- $T(\tau) = T_{\rm eff}$ at $\tau = 2/3$;[1]
- $T(\tau)$ scales linearly with the effective temperature $T_{\rm eff}$ (for the grey atmosphere).

It's important to recognize that a change in temperature with depth isn't inconsistent with radiative equilibrium; although the mean *intensity* of radiation, J, increases with depth (reflected in the increasing temperature), the radiative *flux*, F (the nett outwards-directed component of the intensity), is constant. Physically, the 'throttling' of the outward flow of radiation depends on the opacity – which is why we use optical depth (rather than physical depth) as a measure of radial location in the atmosphere.

8.4 Limb darkening and the Eddington–Barbier Relation

If we observe a spatially resolved star, like the Sun (Fig. 8.1), then at every point on the disk we see down to a *line-of-sight* optical depth of $\tau \simeq 1$. However, this corresponds to different radial physical depths – we see more deeply into the atmosphere at the centre of the disk than we do at the edge. As demonstrated in §8.3.2, deeper means hotter (hence higher surface brightness), so the limb (edge) of the star appears darker then the centre – an effect known as limb darkening.

We know that in the Eddington approximation, the source function is a linear function of optical depth for the grey atmosphere (eqtn. 8.19). We can make a modest generalization of this specific result by

[1] Stars with significantly extended atmospheres may have ill-defined linear radii. Based on the grey-atmosphere result, 'the' (wavelength-dependent!) radius – and hence $T_{\rm eff}$ – is conventionally defined at $\tau = 2/3$.

taking the first two terms in a Taylor expansion of S_ν around $\tau_\nu = 0$, thereby also approximating the *frequency-dependent* source function as a linear function of optical depth:

$$S_\nu(\tau_\nu) = a_\nu + b_\nu \tau_\nu, \tag{8.23}$$

where τ_ν is, as usual, the optical depth measured *radially* inwards. The specific intensity at the surface[2] (radial optical depth $\tau_\nu = 0$) is

$$I_\nu^{\rm out}(\mu, 0) = \int_0^\infty S_\nu \exp\left\{\frac{-\tau_\nu}{\mu}\right\} \frac{{\rm d}\tau_\nu}{\mu} \tag{7.21}$$

or, setting $\tau_\nu/\mu = x$,

$$\begin{aligned} I_\nu(\mu, 0) &= \int_0^\infty S_\nu \exp\{-x\}\ {\rm d}x, \\ &= a_\nu \int_0^\infty \exp\{-x\}\ {\rm d}x + b_\nu \mu \int_0^\infty x \exp\{-x\}\ {\rm d}x \end{aligned}$$

(using eqtn. 8.23). The integrals are both of the standard form

$$\int_0^\infty x^n \exp\{-x\}\ {\rm d}x,$$

(with $n = 0, 1$) which evaluates to n! Thus, recalling that $0! = 1$,

$$I_\nu(\mu, 0) = a_\nu + b_\nu \mu. \tag{8.24}$$

Equation (8.24) shows that the intensity varies from $I_\nu \simeq a_\nu + b_\nu$ at the centre of the stellar disk to $\sim a_\nu$ at the limb;[3] that is, the limb is darkened relative to the centre (for +ve b_ν). Limb darkening always occurs when the source function decreases outward with decreasing optical depth τ_ν; the consequence is that if we look toward the limb of a star we see systematically higher (and cooler) layers, which are less and less bright.

This is schematically illustrated in Fig. 8.1. The penetration depth of our line of sight to unit effective *line of sight* optical depth is smaller at larger θ. When we look toward $\theta = \pi/2$ we see layers closer to the top of the atmosphere, where the *radial* optical depth τ_ν is smaller; but here S_ν is smaller, giving rise to limb darkening.

Comparing eqtn. (8.24) with (8.23), we see that the emergent ($\tau = 0$) intensity at some angle μ is

$$I_\nu(\mu, 0) \simeq S_\nu(\tau_\nu = \mu);$$

[2] This is the intensity emerging at $\tau = 0$, but of course this emergent radiation originates over a range of optical depths.

[3] In practical applications, eqtn. (8.24) is often cast in the form $I(\mu)/I(1) = 1 - u(1 - \mu)$, where the linear limb-darkening coefficient is $u = b/(a+b)$. Limb darkening should not be confused with Lambert's cosine law, which is quite separate (applying to reflecting, not emitting, bodies).

that is, the emergent intensity corresponds to the source function at $\tau = 1$ *along the line of sight*. This results is known as the Eddington–Barbier relation.[4] For a resolved stellar disk, this means that mapping $I_\nu(0)$ at different angles across the disk allows the source function S_ν (and hence the temperature stratification) to be determined observationally as a function of optical depth, at least over the range $\tau = 0{:}1$.

[4]Some authors distinguish between the 'Eddington–Barbier relation', which is an exact result for linear source functions, and the 'Eddington–Barbier approximation' for more general cases.

Section 9

Constructing non-grey model atmospheres

In this section we'll consider how to construct more-realistic models, discarding the (rather poor) grey approximation. Recall that by 'constructing a model atmosphere' we mean computing a self-consistent run of physical quantities as a function of depth in the atmosphere. Those 'physical quantities' describe the state of the matter (e.g., temperature, pressure) *and* the radiation, normally under some simplifying assumptions; and 'depth' is, in practice, the Rosseland mean *optical* depth, τ.[1] The "simplifying assumptions" may include hydrostatic equilibrium, time independence, plane-parallel geometry, and/or LTE; although state-of-the-art computer codes may relax any or all of those approximations, we'll retain them for the time being.

9.1 Construction of a LTE model atmosphere: basic equations and approximate (iterative) method of solution

9.1.1 Key equations

We begin by reviewing six basic equations which, given our assumptions, are sufficient, in principle, to compute a model atmosphere:

1. the definition of the frequency-dependent optical depth, relating it to physical depth:

$$\mathrm{d}\tau_\nu = -k_\nu \,\mathrm{d}s = -\kappa_\nu \rho \,\mathrm{d}s \tag{4.6}$$

2. hydrostatic equilibrium (an assumption in this development):

$$\frac{\mathrm{d}P}{\mathrm{d}\tau_\nu} = \frac{g}{\kappa_\nu} \tag{2.3}$$

[1] Or the optical depth at some reference frequency of choice; the requirement is only for some unique correspondence between linear depth and 'the' optical depth.

3. the equation of state; e.g.,

$$P = nkT = \frac{\rho kT}{\mu m_H} \tag{2.4}$$

4. the formal solution for the radiation field (i.e., the integral solution of the radiative transfer problem expressed in terms of the mean intensity J_ν); for plane-parallel geometry, in shorthand form,

$$J_\nu(\tau_\nu) = \frac{1}{2}\int_0^\infty S_\nu(t_\nu)\, E_1|\tau_\nu - t|\, \mathrm{d}t_\nu, \qquad \equiv \Lambda[B_\nu] \text{ in LTE} \tag{7.23}$$

5. radiative equilibrium (implying flux conservation)

$$\int_0^\infty k_\nu J_\nu \,\mathrm{d}\nu = \int_0^\infty k_\nu S_\nu \,\mathrm{d}\nu \tag{8.5}$$

6. the radiative flux per unit area, F_0, expressed in terms of the effective temperature, $T_{\rm eff}$:

$$F_0 = \sigma T_{\rm eff}^4 \,. \tag{8.2}$$

Broadly speaking, the first three equations (let's call them 'Set 1') describe quantities related to the matter (P, ρ, τ_ν), while the last three equations (let's call them 'Set 2') give the properties related to the radiation field (T, J_ν, F).

These two sets of equations are not independent: they are strongly coupled to each other, due to the strong coupling between matter and radiation. However, there is no direct simultaneous solution of the whole system of six equations. For this reason, it's necessary to seek a solution through an iterative scheme.

9.1.2 Solution philosophy

In order to solve the complete system of six equations we need to specify three parameters (not six, because the equations are coupled, not independent). The parameters which normally chosen to characterize a plane-paralle model atmosphere are the effective temperature ($T_{\rm eff}$), the surface gravity[2] (g), and the chemical composition.[3] The first two are related to the fundamental stellar parameters mass, radius, and luminosity, through

$$T_{\rm eff}^4 = \frac{L}{\sigma 4\pi R^2} \qquad g = \frac{GM}{R^2}.$$

[2] In practice, 'the' gravity is customarily expressed as $\log(g)$, the base-10 logarithm of the surface gravity, in cgs units [i.e., cm s^{-2}].

[3] Not really a single free parameter, of course, though it's often reasonable to start by assuming solar abundances, perhaps with a global scaling to the metallicity.

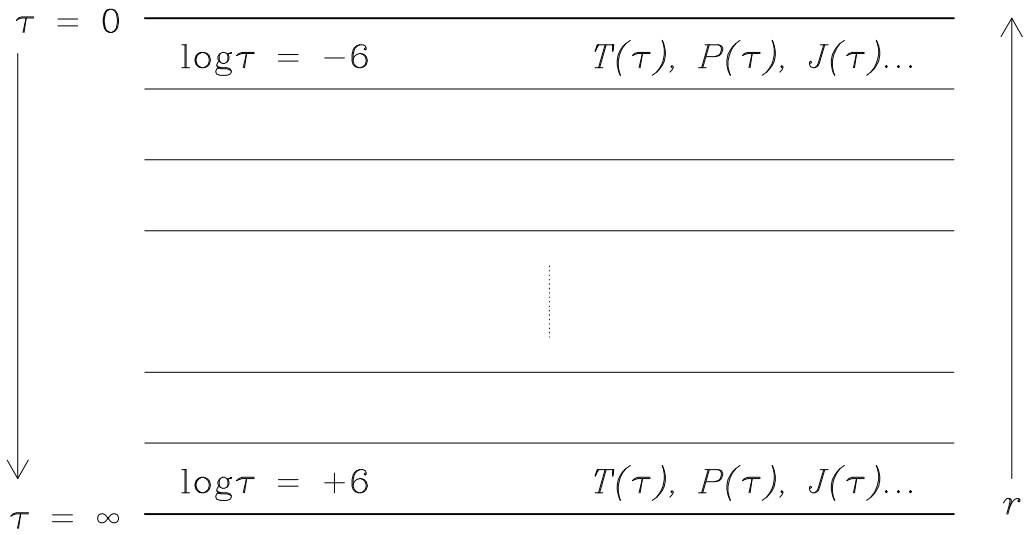


Figure 9.1: Model atmospheres are necessarily evaluated on a fixed grid of points distributed vertically.

For most stars (certainly for main-sequence stars), the atmosphere is thin compared to the stellar radius, and has negligible mass compared to the total mass. Consequently, it's normally safe to assume that the gravity is constant through the atmosphere.

We need to solve the equations for every depth point in the atmosphere. Because the various integrations (for example) need to be conducted numerically, we necessarily represent the atmosphere by a finite number of depth points in order to do this. For a plane-parallel, hydrostatic atmosphere model, typically $\sim 10^2$ depth points would be chosen, at more or less uniform steps in $\log(\tau)$, over a range from very optically thin to very optically thick (say, $\log(\tau) \simeq -6{:}{+6}$; Fig. 9.1).

Conceptually, we then proceed as follows:

Step 0:

Specify $T_{\rm eff}$, $\log_{10}(g)$, and abundances that define the model.

Step 1:

Make a first guess at the temperature profile, $T(\tau)$, and the opacities (e.g., from a grey-atmosphere model; or from a previously calculated model with similar parameters).

Step 2:

(a) Using the temperature profile, integrate the equation of hydrostatic equilibrium, to get pressure (hence density) as a function of depth.

[Some small pressure can be assigned to the outermost layer, and the equation of hydrostatic equilibrium then integrated inwards.]

(b) Knowing $T(\tau)$, $P(\tau)$ (hence $\rho(\tau)$), update the opacities (from lookup tables; or by using Saha-Boltzmann, or statistical-equilibrium calculations, to get level populations).

(c) conduct a formal solution of the radiation field (at each depth in the atmosphere), eqtn. (7.23).

Step 3:

The initial guess for temperature given in step 1 is (very!) unlikely to be correct; hence the computed radiation field will surely not meet the condition of flux conservation (or equivalently, given our set of assumptions, radiative equilibrium).

We therefore need to compute some correction to the radiation field (i.e., for the temperature profile). We apply this correction to update our temperature profile, and go back to Step 2. We continue in this manner until we arrive at a self-consistent solution, with a radiation field that satisfies radiative equilibrium throughout the atmosphere.

Several such iterative temperature-correction schemes exist; we'll review two.

9.2 Λ iteration

Named for the formal solution for the mean intensity, $J_\nu(\tau) = \Lambda_\tau[S_\nu(t)]$, this is the simplest temperature-correction scheme. Although it is useful for illustrating the principles of such schemes, we should note from the outset that it is very inefficient in practice (and consequently not used in real-world codes).

Recall that the condition of radiative equilibrium led us to

$$\int_0^\infty k_\nu J_\nu \, \mathrm{d}\nu = \int_0^\infty k_\nu S_\nu \, \mathrm{d}\nu. \tag{8.5}$$

This is true at every depth in the atmosphere; i.e.,

$$\int_0^\infty k_\nu J_\nu(\tau) \, \mathrm{d}\nu = \int_0^\infty k_\nu S_\nu(\tau) \, \mathrm{d}\nu. \tag{9.1}$$

for any τ. This raises two points requiring clarification:

1. Remember, eqtns. (8.5/9.1) do *not* imply that $J_\nu = S_\nu$ – we need to evaluate $J_\nu(\tau)$ by integrating $S_\nu(t)$ over a range of directions and optical depths (eqtn. 7.23).

2. The source function (and the mean intensity) are functions of temperature, which is a function of physical depth in the atmosphere; e.g., in LTE $S_\nu(\tau) = B_\nu(T(\tau))$. In the context of a temperature-correction scheme, we therefore require that the *optical* depth identified in eqtn. (9.1) correspond to a particular *physical* depth.[4] Conceptually, this is readily accomplished by thinking of τ as, say, the Rosseland mean optical depth (although other options are available; e.g., choosing some specific reference frequency at which to evaluate τ_ν).

For some initial guess at the temperature at some optical depth, $T_1(\tau)$, radiative equilibrium will not be satisfied, because the temperature used to evaluate the source function, S_ν, will not be not correct. To improve that guess, we can update the temperature employed on the right-hand side of eqtn. (9.1) by using a first-order Taylor expansion;[5] that is, we set

$$\begin{aligned}\int_0^\infty k_\nu S_\nu[T_2(\tau)]\,\mathrm{d}\nu &= \int_0^\infty k_\nu S_\nu[T_1(\tau)]\,\mathrm{d}\nu \;+\; \frac{\partial}{\partial T}\int_0^\infty k_\nu S_\nu[T_1(\tau)]\,\mathrm{d}\nu \times [T_2(\tau) - T_1(\tau)] \quad (+\text{ higher-order terms}) \\ &\equiv \int_0^\infty k_\nu S_\nu[T_1(\tau)]\,\mathrm{d}\nu \;+\; \int_0^\infty k_\nu \frac{\partial S_\nu[T_1(\tau)]}{\partial T}\,\mathrm{d}\nu \times \delta T(\tau)\end{aligned}$$

(where $T_2(\tau)$ is our update from $T_1(\tau)$ at depth τ; and we have assumed that k_ν can be adequately evaluated at the *current* temperature estimate, T_1). We can now write eqtn. (9.1) as

$$\int_0^\infty k_\nu J_\nu(\tau)\,\mathrm{d}\nu = \int_0^\infty k_\nu S_\nu[T_1(\tau)]d\nu \;+\; \delta T(\tau)\int_0^\infty k_\nu \frac{\partial S_\nu[T_1(\tau)]}{\partial T}d\nu$$

from which $\delta T(\tau)$, the required correction to the current guess at the temperature at depth τ, is

$$\delta T(\tau) = \frac{\int_0^\infty k_\nu \left[J_\nu(\tau) - S_\nu[T_1(\tau)]\right] d\nu}{\int_0^\infty k_\nu \frac{\partial S_\nu}{\partial T} d\nu} \tag{9.2}$$

That is, in the framework of the iterative process outlined in Section 9.1.2 we can correct/update the initial temperature at (every) depth τ by setting $T_2 = T_1 + \delta T$; go back to Step 2; and repeat the process, until convergence is reached. In LTE, $S_\nu = B_\nu$, and the source function and its derivative with respect to T can be evaluated analytically; otherwise, this would have to be done numerically.

Because $J_\nu = \Lambda[S_\nu]$, this iterative process is conventionally referred to as as 'Λ iteration'. It is conceptually simple, but has significant disadvantages:

1. The method gives suitable corrections at the stellar surface, but in the deeper layers the photon mean free path is short, and both J_ν and S_ν tend to B_ν. Consequently, the numerator in eqtn. (9.2) vanishes at all frequencies, and the correction δT tends to zero as the optical depth tends to infinity. As a result, convergence at depth is indefinitely slow.

[4] Not true in general: eqtn. (4.6), $\mathrm{d}\tau_\nu = -k_\nu\,\mathrm{d}s$, tells us that a fixed *optical* depth generally occurs at different *physical* depths at different frequencies, because of the frequency dependence of the opacity.

[5] Recall $f(x)$ about a: $f(x) \simeq f(a) + f'(a)(x-a)\ldots$

2. A major weakness of Λ iteration is that it is a local method; that is, it takes no account of the effect that the correction $\delta T(\tau)$ at depth τ has on the mean intensity $J_\nu(\tau')$ in some adjacent layer at depth τ'. A better method would be sensitive to the *global* status of the atmosphere; that is, we would prefer to use a correction in an integral form.

3. Furthermore, because radiative equilibrium is enforced *locally* (rather than globally), the method can lead to $\mathrm{d}F/\mathrm{d}\tau = 0$, as required, but not at the desired value for the flux F_0. More generally, the method can stabilize at an incorrect solution – a very bad numerical property.

As a result of these limitations, Λ iteration is not used as a practical solution method – although often wheeled out as a simple demonstration of how to solve an atmospheric structure in principle, it is also a good example of how *not* to do it in practice. However, towards the end of the 20th century, it was realised that its convergence properties could be dramatically improved through the introduction of modified, 'accelerated' or 'approximate', lambda operators. These so-called ALI methods are now a mainstay of modern stellar-atmosphere computer programs.

Nevertheless, other techniques remain in use, and are (arguably) easier to demonstrate, including (e.g.) Avrett-Krook and Unsöld–Lucy iteration; we'll review the latter as a practical alternative to simple Λ iteration.

9.3 Unsöld–Lucy iteration

The Unsöld–Lucy method incorporates constraints on both the absolute value of the flux and its (lack of) depth dependence. With minor modifications, this method is embodied in the 'state of the art' modelling code `PHOENIX`.

Although we want to relax the grey-atmosphere approximation, it's still convenient to avoid the full frequency dependence of opacities by defining several frequency-integrated, flux-weighted forms (where all quantities with subscripts are to be understood to be functions of [optical] depth in the atmosphere). This is tolerable because at this stage we're not really concerned with the frequency-dependent spectrum, but just the overall radiative energy transport. We define the following:

$$\text{Planck mean opacity:} \qquad k_\mathrm{P} = \frac{\int_0^\infty k_\nu^\mathrm{a} B_\nu \,\mathrm{d}\nu}{\int_0^\infty B_\nu \,\mathrm{d}\nu}, \qquad \equiv \frac{\int_0^\infty k_\nu^\mathrm{a} B_\nu \,\mathrm{d}\nu}{B};$$

$$\text{Eddington-flux mean opacity:} \qquad k_\mathrm{H} = \frac{\int_0^\infty k_\nu H_\nu \,\mathrm{d}\nu}{\int_0^\infty H_\nu \,\mathrm{d}\nu}, \qquad \equiv \frac{\int_0^\infty k_\nu H_\nu \,\mathrm{d}\nu}{H};$$

$$\text{intensity mean opacity:} \qquad k_\mathrm{J} = \frac{\int_0^\infty k_\nu^\mathrm{a} J_\nu \,\mathrm{d}\nu}{\int_0^\infty J_\nu \,\mathrm{d}\nu}, \qquad \equiv \frac{\int_0^\infty k_\nu^\mathrm{a} J_\nu \,\mathrm{d}\nu}{J},$$

with corresponding (frequency-independent) optical-depth increments $\mathrm{d}\tau_\mathrm{P}$, $\mathrm{d}\tau_\mathrm{H}$, and $\mathrm{d}\tau_\mathrm{J}$ ($= -k_\mathrm{P}\,\mathrm{d}r$, $-k_\mathrm{H}\,\mathrm{d}r$, and $-k_\mathrm{J}\,\mathrm{d}r$), where $k_\nu = k^\mathrm{a} + k_\nu^\mathrm{s}$ is the sum of 'true' and scattering opacities (cf. eqtn. 7.11).

9.3.1 Zeroth moment

Starting from the zeroth moment of the transfer equation (integrating the transfer equation over solid angle, as in our discussion of Milne's first equation; Section 8.2.1) we previously obtained the flux derivative in the form

$$\frac{\mathrm{d}H_\nu(\tau_\nu)}{\mathrm{d}r} = k_\nu S_\nu(\tau_\nu) - k_\nu J_\nu(\tau_\nu), \tag{8.4}$$

(recalling that the Eddington flux is $H_\nu = F_\nu/4\pi$). Using eqtn. (7.11) for S_ν and integrating over frequency, then at some (frequency-independent) depth τ_P in the atmosphere

$$\int_0^\infty \frac{\mathrm{d}H_\nu(\tau_\mathrm{P})}{\mathrm{d}r}\,\mathrm{d}\nu = \int_0^\infty k_\nu^\mathrm{a} B_\nu\left(T(\tau_\mathrm{P})\right)\,\mathrm{d}\nu - \int_0^\infty k_\nu^\mathrm{a} J_\nu(\tau_\mathrm{P})\,\mathrm{d}\nu.$$

Using the intensity- and Planck-mean volume opacities on the right-hand side this becomes

$$\frac{\mathrm{d}H(\tau_\mathrm{P})}{\mathrm{d}r} = k_\mathrm{P} B\left(T(\tau_\mathrm{P})\right) - k_\mathrm{J} J(\tau_\mathrm{P}).$$

Dividing both sides by k_P and rearranging we obtain

$$B\left(T(\tau_\mathrm{P})\right) = \frac{k_\mathrm{J}}{k_\mathrm{P}} J(\tau_\mathrm{P}) - \frac{\mathrm{d}H(\tau_\mathrm{P})}{\mathrm{d}\tau_\mathrm{P}}. \tag{9.3}$$

In principle, eqtn. (9.3) allows us to compute the frequency-integrated Planck source function at depth τ_P; or, essentially equivalently, the temperature structure temperature $T(\tau_\mathrm{P})$. It has the desireable property that all the terms are frequency-averaged or frequency-integrated (so we don't have to explicitly evaluate them frequency by frequency).

However, to use eqtn. (9.3) in practice, knowing that we want a radiative-equilibrium temperature structure in which the radiative flux H is constant with depth,[6] we see that we require a useful expression for $J(\tau_\mathrm{P})$. To achieve that, we look to the first moment.

9.3.2 First moment

From the first moment of the transfer equation we saw that

$$\frac{\mathrm{d}K_\nu(\tau_\nu)}{\mathrm{d}\tau_\nu} = \frac{F_\nu(\tau_\nu)}{4\pi} \equiv H_\nu(\tau_\nu); \tag{8.7}$$

but in the Eddington (two-stream) approximation $K_\nu = J_\nu/3$ (eqtn. 8.16), and so

$$\frac{\mathrm{d}J_\nu(\tau_\nu)}{\mathrm{d}\tau_\nu} = 3H_\nu(\tau_\nu).$$

[6] Hence the $\mathrm{d}H(\tau_\mathrm{P})/\mathrm{d}\tau_\mathrm{P}$ term goes to zero when we achieve a solution with a correct, radiative-equilibrium, temperature structure. However, we can't ignore the term at this stage, because it will be non-zero for any incorrect 'first guess' temperature structure.

Using $\mathrm{d}\tau_\nu = -k_\nu \mathrm{d}r$, and integrating over frequency, then at some Planck-mean optical depth τ_P

$$\int_0^\infty \frac{\mathrm{d}J_\nu(\tau_\mathrm{P})}{\mathrm{d}r}\,\mathrm{d}\nu = -3\int_0^\infty k_\nu H_\nu(\tau_\mathrm{P})\,\mathrm{d}\nu,$$

or, using the flux-mean opacity on the right-hand side

$$\frac{\mathrm{d}J(\tau_\mathrm{P})}{\mathrm{d}r} = -3k_\mathrm{H}H(\tau_\mathrm{P}). \tag{9.4}$$

Dividing both sides of eqtn. (9.4) by $-k_\mathrm{P}$ and integrating over optical depth leads to

$$J(\tau_\mathrm{P}) = \int_0^{\tau_\mathrm{P}} 3\frac{k_\mathrm{H}}{k_\mathrm{P}}H(t)\,\mathrm{d}t + J(0)$$

where $J(0)$ is a constant of integration corresponding to the mean intensity at the surface. Recalling that

$$J(0) = F(0)/2\pi, = 2H(0) \tag{8.17}$$

in the Eddington two-stream approximation, we obtain

$$J(\tau_\mathrm{P}) = \int_0^{\tau_\mathrm{P}} 3\frac{k_\mathrm{H}}{k_\mathrm{P}}H(t)\,\mathrm{d}t + 2H(0). \tag{9.5}$$

This is our required expression for $J(\tau_\mathrm{P})$, the mean intensity at some depth τ_P in the atmosphere.

9.3.3 The correction

Combining eqtns. (9.3) and (9.5) gives, in effect, the temperature as a function of Planck mean optical depth, in terms of the Eddington flux – which, for given T_eff, we know.

$$B\left(T(\tau_\mathrm{P})\right) = \frac{k_\mathrm{J}}{k_\mathrm{P}}\left[\int_0^{\tau_\mathrm{P}} 3\frac{k_\mathrm{H}}{k_\mathrm{P}}H(t)\,\mathrm{d}t + 2H(0)\right] - \frac{\mathrm{d}H(\tau_\mathrm{P})}{\mathrm{d}\tau_\mathrm{P}}, \qquad = \sigma T^4(\tau_\mathrm{P})/\pi. \tag{9.6}$$

As in our outline of Λ iteration, some initial trial solution for the temperature structure $T_1(\tau_\mathrm{P})$ will, in general, predict an initial set of fluxes $H_1(\tau_\mathrm{P})$ that vary with depth; while in radiative equilibrium the correct solution should give constant flux H for all τ_P.

We therefore need to evaluate the correction required to the temperature – or, equivalently, the correction to frequency-integrated Planck function, $B\left(T(\tau_\mathrm{P})\right)$ in eqtn. (9.6), which we can translate directly into a temperature correction. We write the correction as

$$\Delta B\left(T(\tau_\mathrm{P})\right) = B\left(T_2(\tau_\mathrm{P})\right) - B\left(T_1(\tau_\mathrm{P})\right)$$

where our first-guess solution is $B\left(T_1(\tau_\mathrm{P})\right)$ and the updated estimate, after adding this correction, is $B\left(T_2(\tau_\mathrm{P})\right)$.

Writing eqtn. (9.6) for $B(T_1)$ and for $B(T_2)$ and subtracting gives

$$\Delta B\left(T(\tau_\mathrm{P})\right) = \frac{k_\mathrm{J}}{k_\mathrm{P}}\left[3\int_0^{\tau_\mathrm{P}} \frac{k_\mathrm{H}}{k_\mathrm{P}}\Delta H(t)\,\mathrm{d}t + 2\Delta H(0)\right] - \frac{\mathrm{d}(\Delta H(\tau_\mathrm{P}))}{\mathrm{d}\tau_\mathrm{P}}$$

where the 'ΔH' terms are corrections $H_2 - H_1$, and we have assumed that the ratios k_J/k_P, k_H/k_P are the *same* for the new estimate as for the first one.[7]

We know what the target value[8] is for the Eddington flux – it's just $\sigma T_{\rm eff}^4/(4\pi)$, which is therefore always our guess at H_2. Similarly, the target gradient, $\mathrm{d}(H_2(\tau_P))/\mathrm{d}\tau_P$ is zero, so

$$\frac{\mathrm{d}(\Delta H(\tau_P))}{\mathrm{d}\tau_P} = \left(\frac{\mathrm{d}(H_2(\tau_P))}{\mathrm{d}\tau_P}\right) - \frac{\mathrm{d}(H_1(\tau_P))}{\mathrm{d}\tau_P}$$
$$= B\left(T_1(\tau_P)\right) - \frac{k_J}{k_P} J_1(\tau_P). \tag{9.3}$$

Our required temperature correction is therefore

$$\Delta B\left(T(\tau_P)\right) = \frac{k_J}{k_P}\left[3\int_0^{\tau_P} \frac{k_H}{k_P}\Delta H(t)\,\mathrm{d}t + 2\Delta H(0)\right] - B\left(T_1(\tau_P)\right) + \frac{k_J}{k_P} J_1(\tau_P) \tag{9.7}$$

Eqtn. (9.7) allows us to compute the desired correction to the Planck function (or temperature structure), which can then be applied iteratively as follows:

0. Obtain a first estimate of the temperature structure $T(\tau)$ (e.g., from a grey-atmosphere solution);

1. From $T(\tau)$, compute $B_\nu\left(T(\tau)\right)$, and thence k_P (ab initio, or, more realistically, from a precomputed set of opacities $k_\nu(\rho, T)$).

2. With the source function in hand, compute the mean intensities $J(\tau_P)$ from the formal solution (eqtn. 7.23; cf. item 4 in §9.1.1, above), and thence k_J.

3. Similarly, $H_1(\tau_P)$ (hence k_H) and $H_1(0)$ can be computed (eqtns. 7.24, 7.25). We know the true frequency-integrated flux – it is $H = F/4\pi = \sigma T_{\rm eff}^4/4\pi$ (at all depths, in radiative equilibrium) – so we can write the required corrections to the current estimates, $H_1(\tau_P)$, as

   $$\Delta H(\tau_P) = H - H_1(\tau_P).$$

   at each depth τ_P (including $\tau_P = 0$).

4. Substituting into eqtn. (9.7) gives depth-dependent corrections to $B\left(T_1(\tau_P)\right)$, hence updated values, $B\left(T_2(\tau_P)\right)$ – which translates into an updated temperature structure, $T_2(\tau_P)$. Return to step 1 until convergence is achieved.

[7] The opacities are updated *after* we have a new estimate of temperature structure, so we're always using values from the previous iteration. Nevertheless, we can safely assume that this will provide a pretty good estimate of the opacity *ratios,* which we can expect to be less sensitive to temperature than are the separate values.

[8] The iterative process won't converge on it immediately, because we don't have self-consistent opacities, so it remains a 'target' through the process.

Points to note:

- At depth, $J \rightarrow B$ (and $k_{\rm J} \rightarrow k_{\rm P}$). Hence the $\mathrm{d}(\Delta H)/\mathrm{d}\tau_{\rm P}$ term, eqtn. (9.3), tends to zero – just as for Λ iteration.
- However, the first [bracketed] term in eqtn. (9.7) gives rapid convergence at large optical depth, because the target H is known exactly.

Because eqtn. (9.7) uses an exact calculation of the the flux error to determine the correction, convergence is rapid (although it isn't achieved in a single iteration because the calculation is exact for only approximate estimates of physical parameters). This also ensures that the solution converges to the correction solution (i.e., correct F_0).

Section 10

Line formation

The strength and shape of absorption lines in stellar spectra vary according to the chemical abundances and physical conditions in the line-forming regions. In order to study this problem, we modify the transfer equation to account separately for line and continuum opacities; if there is strong continuum opacity, then the line must be formed high in the atmosphere, but if the continuum opacity is weaker, the line can be formed deeper.

We therefore introduce

- $k_\nu^{\mathrm{L}}, k_\nu^{\mathrm{C}}$: the line and continuum absorption coefficients
- $j_\nu^{\mathrm{L}}, j_\nu^{\mathrm{C}}$: the line and continuum emission coefficients

The total optical depth at frequency ν is

$$\mathrm{d}\tau_\nu = (k_\nu^{\mathrm{L}} + k_\nu^{\mathrm{C}})\,\mathrm{d}s$$

and the separate (but cospatial) source functions for the line and continuum are

$$S_\nu^{\mathrm{L}} = \frac{j_\nu^{\mathrm{L}}}{k_\nu^{\mathrm{L}}} \qquad \text{and} \qquad S_\nu^{\mathrm{C}} = \frac{j_\nu^{\mathrm{C}}}{k_\nu^{\mathrm{C}}}.$$

We can therefore write the total source function as

$$\begin{aligned} S_\nu &= \frac{j_\nu^{\mathrm{L}} + j_\nu^{\mathrm{C}}}{k_\nu^{\mathrm{L}} + k_\nu^{\mathrm{C}}} \\ &= \frac{(k_\nu^{\mathrm{L}}/k_\nu^{\mathrm{C}})S_\nu^{\mathrm{L}} + S_\nu^{\mathrm{C}}}{1 + k_\nu^{\mathrm{L}}/k_\nu^{\mathrm{C}}} \\ &= \frac{S_\nu^{\mathrm{L}} + (k_\nu^{\mathrm{C}}/k_\nu^{\mathrm{L}})S_\nu^{\mathrm{C}}}{1 + k_\nu^{\mathrm{C}}/k_\nu^{\mathrm{L}}}. \end{aligned}$$

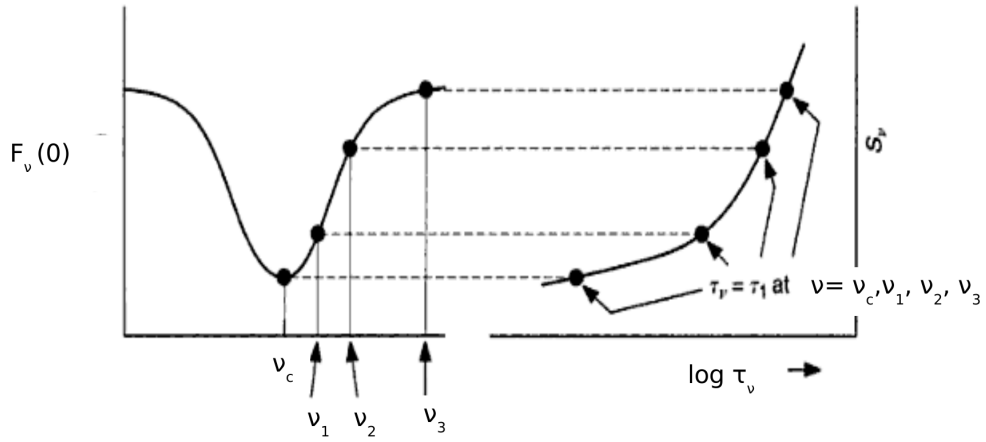


Figure 10.1: The mapping between the source function and the line profile; the decrease of the source function outward through the photosphere produces absorption lines.

(It's not simply the sum of the source functions, because, e.g., the line emissivity 'sees' continuum opacity).

There is is a close connection between the source function and the observed (i.e., $\tau_\nu = 0$) flux in the line, $F_\nu(0)$, as we've already seen for the continuum in our discussion of the grey atmosphere. For illustrative purposes, we can adopt a simple linear source function, inspired by the result obtained in the Eddington approximation for the grey atmosphere, but allowing for a frequency dependence:

$$S_\nu(\tau_\nu) \simeq aF_\nu(0)\,(\tau_\nu + b)\,. \qquad \text{(cf. 8.19)}$$

Evidently, the source function S_ν equals the surface flux $F_\nu(0)$ for some particular value of τ_ν – say, $\tau_\nu^{\rm S}$. For our illustrative linear source function,

$$\frac{S_\nu}{F_\nu(0)} = 1 = a(\tau_\nu^{\rm S} + b), \text{ so } \tau_\nu^{\rm S} = (1 - ab)/a.$$

To get some idea of roughly where this occurs we can take numerical values from the grey-atmosphere solution ($a = 3/4\pi$, $b = 2/3$; eqtn. 8.19) to find $\tau_\nu^{\rm S} \simeq 3.5$.

The line absorption coefficient $k_\nu^{\rm L}$ necessarily varies across the line, being largest at the line centre.[1] Therefore, the condition $\tau_\nu = \tau_\nu^{\rm S}$ is met nearer the surface of the photosphere closer to the line centre; that is, the emergent flux in the line originates higher in the atmosphere.[2] Hence, if the source function decreases outwards (as is usually the case), an absorption line is formed (see Fig.10.1).

[1]The line absorption profile is usually a Voigt function, being the convolution of a Gaussian (arising from thermal broadening) and a Lorentzian (pressure broadening).

[2]In reality, of course, the emergent flux at any frequency in the line profile –or continuum – comes from a *range* of depths. Furthermore the observed line profile may be modified by a variety of macroscopic mechanisms (i.e., process which operate on linear scales greater than the photon mean free path), such as rotation. The mapping between S_ν and $F_\nu(0)$ holds only for the 'intrinsic' line profile, in the absence of these mechanisms.

10.1 Source function for the line

The transition that is responsible for a line may undergo the *radiative* processes of absorption (photoexcitation from the lower state ℓ to the upper state u), spontaneous emission, and stimulated emission.

The emissivity in the line is

$$j_\nu^{\rm L} = n_u A_{ul}\, \psi(\nu)\, h\nu.$$

where n_u is the number density of potential emitters, $A_{u,\ell}$ is the Einstein coefficient for spontaneous emission,[3] $\psi(\nu)$ is some function (whose area is normalized to unity) describing the frequency dependence of the spontaneous emission, and $h\nu$ is just the energy of an emitted photon.

If the frequency dependence of the absorption and stimulated emission is $\phi(\nu)$ (again normalized to unit area), then the corresponding line absorption coefficient is

$$k_\nu^{\rm L} = n_\ell\, B_{\ell u}\, \phi(\nu)\, h\nu - n_u\, B_{u\ell}\, \phi(\nu)\, h\nu$$

where n_ℓ is the number density of absorbers in the lower energy state, and $B_{\ell u}$, $B_{u\ell}$ are the Einstein coefficients appropriate to photoexcitation and stimulated emission.

Dividing the two expression gives the generalized line source function,

$$S_\nu^{\rm L} = \frac{j_\nu^{\rm L}}{k_\nu^{\rm L}} = \frac{n_u\, A_{u\ell}\, \psi(\nu)}{n_\ell\, B_{\ell u}\, \phi(\nu) - n_u\, B_{u\ell}\, \phi(\nu)}.$$

The Einstein coefficients are related through

$$\begin{aligned} B_{u\ell}\, g_u &= B_{\ell u}\, g_\ell \\ A_{u\ell} &= \frac{2h\nu^3}{c^2} B_{u\ell} \end{aligned}$$

(where g_u and g_ℓ are the statistical weights of levels u and ℓ), allowing us to re-write the generalized source function for the line as

$$S_\nu^{\rm L} = \frac{2h\nu^3}{c^2}\left[\frac{n_\ell}{n_u}\frac{g_u}{g_\ell} - 1\right]^{-1} \frac{\psi(\nu)}{\phi(\nu)}. \tag{10.1}$$

[3] $A_{u,\ell}$ describes the probability of spontaneous decay per unit time. The probability that an atom emits a photon in the time interval $\mathrm{d}t$, frequency interval $\mathrm{d}\nu$, and solid angle $\mathrm{d}\omega$ is $A_{u,\ell}\, \mathrm{d}t\, \mathrm{d}\omega\, \mathrm{d}\nu\, \psi(\nu)$ and if the initial number density of particles in the excited state is n_u, then $\mathrm{d}n_u/\mathrm{d}t = -A_{u,\ell} n_u(t)$

The concealed complexity in this superficially straightforward formalism is in the evaluation of the population ratios n_ℓ/n_u. In statistical equilibrium the number density of ions with electrons in some level j is

$$\frac{\mathrm{d}n_j}{\mathrm{d}t} = \sum_{i\neq j}\left(n_i r_{ij} - n_j r_{ji}\right) = 0$$

where r_{ij} is the rate at which electrons move from state i to state j as a consequence of all processes (radiative and collisional),

$$r_{ij} = 4\pi A_{ij} + 4\pi B_{ij}\int_0^\infty J_\nu \phi(\nu)\,\mathrm{d}\nu + Cij.$$

That is, the level populations n depend on the mean intensity, J_ν, and hence on the source function (eqtn. 7.21). . .but the source function depends on the level populations (eqtn. 10.1).

Only if collisional process dominate is this interdependence broken; then Boltzmann equilibrium applies, and

$$\frac{n_u}{n_\ell} = \frac{g_u}{g_\ell}\exp\left\{\frac{-h\nu}{kT}\right\}.$$

Furthermore, in equilibrium, on average, then frequency by frequency there must be an absorbed photon for every emitted photon, and $\phi(\nu) = \psi(\nu)$. In this case, eqtn. (10.1) becomes

$$\begin{aligned} S_\nu &= \frac{2h\nu^3}{c^2}\left\{\exp\left(\frac{h\nu}{kT}\right) - 1\right\}^{-1} \\ &\equiv B_\nu(T) \end{aligned}$$

The line is then formed in LTE and we recover the result that the line source function is the Planck function. LTE is a crude but useful approximation for line-formation computations and has been used extensively in modelling of spectral lines (though this is becoming less commonplace with increasing computing power).

10.2 Computation of a weak-line profile

Recalling that $S_\nu(\tau_\nu) = F_\nu(0)$ for some value $\tau_\nu = \tau_\nu^{\mathrm{S}}$ ($\simeq 3.5$ according to our earlier back-of-the-envelope estimate), we can write the normalized line profile directly in terms of the source function:

$$\frac{F_\nu}{F_{\mathrm{c}}} \simeq \frac{S_\nu(\tau_\nu^{\mathrm{T}} = \tau_\nu^{\mathrm{S}})}{S_\nu(\tau_\nu^{\mathrm{C}} = \tau_\nu^{\mathrm{S}})} \tag{10.2}$$

where F_ν, F_c are the observed and continuum fluxes;[4] $S_\nu(\tau_\nu = \tau_\nu^S)$ is the source function where the optical depth τ_ν has the numerical value τ_ν^S; and the total (combined line+continuum) source function at this optical depth is

$$\begin{aligned} S_\nu(\tau_\nu^T = \tau_\nu^S) &\equiv S_\nu(\tau_\nu^L + \tau_\nu^C = \tau_\nu^S), \\ &= S_\nu(\tau_\nu^C = \tau_\nu^S - \tau_\nu^L) \end{aligned}$$

(by simple algebra). We can develop the right-hand side as a Taylor expansion around the point $\tau_\nu^C = \tau_\nu^S$:

$$S_\nu(\tau_\nu^T = \tau_\nu^S) \approx S_\nu(\tau_\nu^C = \tau_\nu^S) + \left.\frac{dS_\nu}{d\tau_\nu^C}\right|_{\tau_\nu^S} \times \left(-\tau_\nu^L\right) + \ldots$$

Eqtn. (10.2) then becomes

$$\frac{F_\nu}{F_c} \simeq \frac{S_\nu(\tau_\nu^C = \tau_\nu^S)}{S_\nu(\tau_\nu^C = \tau_\nu^S)} - \frac{\tau_\nu^L}{S_\nu(\tau_\nu^C = \tau_\nu^S)} \times \left.\frac{dS_\nu}{d\tau_\nu^C}\right|_{\tau_\nu^S} + \ldots$$

In the weak-line approximation $\tau_\nu^L \ll \tau_\nu^C$; so, taking just these first two terms,

$$\begin{aligned} \frac{F_\nu}{F_c} &\simeq 1 - \frac{\tau_\nu^L}{S_\nu(\tau_\nu^C = \tau_\nu^S)} \times \left.\frac{dS_\nu}{d\tau_\nu^C}\right|_{\tau_\nu^S} \\ &= 1 - \tau_\nu^L \left.\frac{d\ln S_\nu}{d\tau_\nu^C}\right|_{\tau_\nu^S} \end{aligned} \tag{10.3}$$

If the ratio of line to continuum opacities, k_ν^L/k_ν^C, doesn't vary strongly with depth then

$$\frac{\tau_\nu^L}{\tau_\nu^C} \simeq \frac{k_\nu^L}{k_\nu^C}$$

(for any value of τ). Furthermore, $\tau_\nu^C \approx \tau_\nu^S$ in the weak-line limit. Thus

$$\tau_\nu^L \simeq \tau_\nu^C \frac{k_\nu^L}{k_\nu^C}, \simeq \tau_\nu^S \frac{k_\nu^L}{k_\nu^C}.$$

With this approximation, eqtn. (10.3) becomes

$$\begin{aligned} \frac{F_\nu}{F_c} &\approx 1 - \tau_\nu^S \left.\frac{d\ln S_\nu}{d\tau_\nu^C}\right|_{\tau_\nu^S} \frac{k_\nu^L}{k_\nu^C} \\ &= 1 - \text{constant} \times \frac{k_\nu^L}{k_\nu^C}\,; \quad \text{i.e.,} \end{aligned} \tag{10.4}$$

$$1 - \frac{F_\nu}{F_c} \propto \frac{k_\nu^L}{k_\nu^C}$$

[4] 'Continuum flux' is to be understood as the flux that *would* be observed if no line was present.

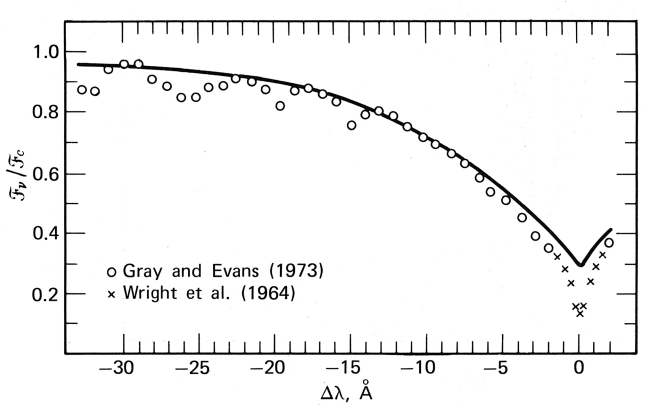


Figure 10.2: Fig. 13.15 of Gray's book. The observed H_γ, profile is shown by the circles. (Several weak lines are also present, including at −33, −26, −20, and −15Å.) The calculated LTE profile is shown by the line. The theoretical profile agrees adequately in the wings; in the core, however, the observed line is significantly deeper than the calculated one.

Equation (10.4) represents the approximate line profile of a 'weak' line ($\tau_\nu^{\rm L} \ll \tau_\nu^{\rm C}$). It shows that:

- The shape of the line profile for a weak line is proportional to $k_\nu^{\rm L}$. In fact, it mimics directly the frequency dependence of $k_\nu^{\rm L}$ if the continuum opacity $k_\nu^{\rm C}$ varies little across the line profile (as is usually the case).

- More generally, the strength of the line profile is proportional to $k_\nu^{\rm L}/k_\nu^{\rm C}$, which means it can be increased not only by increasing $k_\nu^{\rm L}$, but also by decreasing $k_\nu^{\rm C}$ at the line frequency.

  [Intuitively, the absorption-line strength must depend on both the line opacity and the amount of material doing the absorbing. If the continuum opacity is high, we can see only a relatively short distance into the atmosphere, so the amount of material capable of line absorption is relatively small, which is why the *line* strength depends on the *continuum* opacity.]

- Furthermore, in LTE a spectral line can never go to zero intensity. Recall, the observed flux (including at the line centre) is equal to the source function at some finite optical depth ($\sim 3^1/_2$ for our rough estimate). Because $S_\nu(\tau) = B_\nu(T(\tau))$ in LTE, and $T(\tau)$ is always greater than zero (e.g., $\gtrsim 0.8T_{\rm eff}$ in the grey-atmosphere approximation; eqtn. 8.21), it follows that the line-centre flux is also always greater than zero.

In general (for both strong and weak lines), the line profile is strongly dependent on T, P, and on the chemical composition of the material (gas) in the line-emitting region, as these quantities enter the evaluation of $k_\nu^{\rm L}$. Conversely, photospheric absorption lines provide potentially powerful diagnostics of these atmospheric parameters.

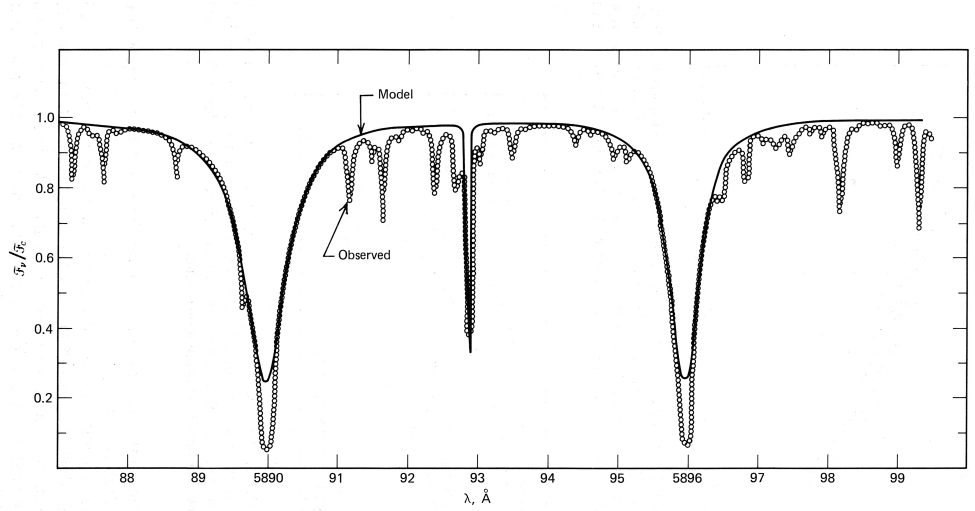


Figure 10.3: Fig. 13.16 of Gray's book. Observations of the solar sodium-D lines are compared with an LTE model. Serious discrepancies are seen in the core. (Almost all the weaker lines arise from telluric water vapour.)

10.3 Breakdown of LTE

Sometimes the LTE approximation can be acceptably good, sometimes really poor. We can anticipate circumstances under which we can be confident that it won't be satisfactory (e.g., when excitation and ionization are certainly not determined collisionally – low-density nebulae, for example) but the only way to tell for sure if it's adequate is to compute the problem in both LTE and non-LTE (or sometimes, and less reliably, to compare the expectations of a model with observations).

LTE is most likely to be an acceptable approximation when the rate of collisionally-induced transitions greatly exceeds radiatively-induced transitions, or when the loss of radiation is small with respect to the radiation field in the volume (i.e. if κ is large).

We can therefore expect LTE to be a poor approximation

- In the outer photospheric layers, where radiation is free to escape into space.
- In the cores of strong lines, formed in the outer layers of the photospheres
- In O stars, where radiative rates exceed collisional rates by a large factors throughout the atmosphere.
- Similarly, in supergiants where the atmospheric density is very low (and hence collisional rates low).

Fig 10.2 shows observations of H-alpha in θ Leo, compared to a theoretical profile computed in LTE. The figure shows that the fit to the wings (formed deep in the atmosphere) is good, but the observed line core is deeper than the theoretical one. This is a common failure of LTE computations and in fact, for this particular line, back in 1970 and 1972 Auer and Mihalas have found that a better agreement is reached by using non-LTE computations.

Also, in Vega and Sirius a better agreement with non-LTE computations is found in the analysis of the Paschen continuum, Balmer jump and Balmer lines (exclusive to the core). A similar situation occurs with other lines. For instance (Fig.10.3), the sodium D-lines in the Sun are well fitted by LTE models in the wings, but not in the cores.

The reality can be even more complex: non-LTE calculations generally show deeper cores than LTE ones, but still sometimes they do not fit. The cause can be multiple:

- part of the discrepancy may be due to non thermal motions. Solar-D lines observations show their profile varies with time and with position on the solar disk. For other stars, the spatial information is lost and we need to deal with average properties.
- Also, saturated lines are stronger than predicted $\Rightarrow$ this may be due to turbulence effects.
- Many stars show spectral lines broadened by large amounts. Ex lines of δCas are broadened by $\sim$ 130 km/s, nearly two order of magnitudes more than what is expected from thermal broadening $\Rightarrow$ rotational Doppler broadening, effect from macroturbulence.

10.4 P-Cygni profiles

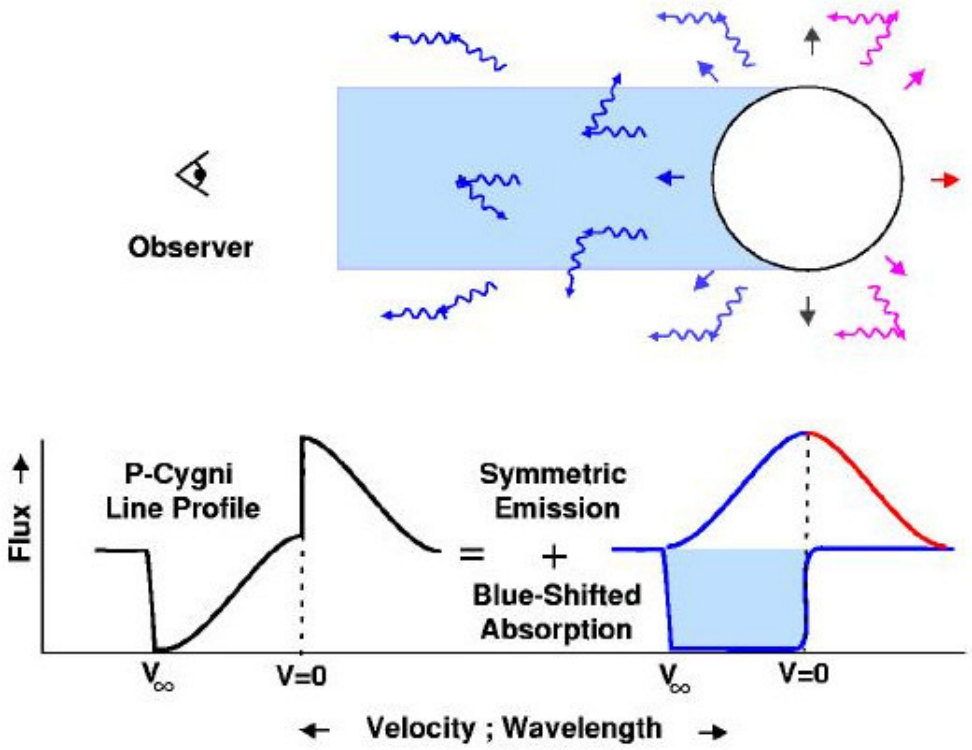


Figure 10.4: Formation of P-Cygni profiles.

P-Cygni profiles consist of a blue-shifted absorption component and a red-shifted emission components; they are characteristic of mass loss. The simplest mechanism for line formation (and the one which applies to ultraviolet P-Cygni profiles in most OB stars) is resonance-line scattering.

Ions in the (relatively low-density) outflow are normally in the ground state – i.e., every electron is in the lowest available energy level. This is, of course, a strongly non-LTE situation, which arises because particle collisions are generally of almost no importance in established ionization and excitation equilibria – photoionization, photoexcitation, and radiative recombination dominate. We also have to abandon hydrostatic equilibrium!

A resonance line is formed by the transition of an outer electron from the ground state to the first excited energy level, or vice versa. Resonance-line scattering occurs when a photon of the correct energy photo-excites an electron from the ground state, followed very rapidly by spontaneous decay (typically on timescales of order 10^{-6} s or less). The photon must have the correct wavelength in the rest frame of the scatterer; because the scatterer is moving with respect to the star (and the observer), this 'correct wavelength' will be Doppler shifted from the laboratory wavelength.

Fig. 10.4 illustrates the actual line formation process. There is a radial outflow of matter which

accelerates fron essentially zero velocity at the photosphere to a maximum, or terminal, velocity, v_∞ (perhaps ~2000 km s^{-1}), at large distance from the star.

Photons that are originally emitted in the direction of the observer may interact with the column of gas seen in projection against the photosphere, and be scattered out of the line of sight, producing, in effect, an absorption line. This gas is flowing away towards the observer, producing blue-shifted absorption over the range of velocities that occur in the wind, from 0 to v_∞. (Note that the line can – and often does – go to zero intensity; LTE does not apply here!)

However, photons are not destroyed in the scattering process, and so for every photon scattered out of the line of sight in the column of gas seen projected against the photosphere, there must be a photon scattered into the line of sight elsewhere in the wind.[5] The range of velocities in this 'halo' around the star is from $+v_\infty$ approaching the observer to $-v_\infty$ receding, with the extra emission is Doppler-shifted over this same range. The combination of absorption from zero velocity to $+v_\infty$ and emission from $+v_\infty$ to $-v_\infty$ produces the P-Cygni profile.

(The astute student will notice that the nett effect from the observer's perspective is that some photons have been shifted from the blue part of the profile to red part of the profile – thereby losing energy. It is this 'lost' energy that goes into accelerating the outflow as it moves away from the star; some of the outward momentum of the photons is transferred to the scattering particles as they interact.)

[5] Photons are not strictly conserved, because some are scattered back into the photosphere.

PHAS0036 – Part II, Stellar Structure & Evolution

Section 11

Introduction to Stellar Structure and Evolution [review of PHAS0018 material]

11.1 Motivation

In this part of the course, we consider aspects of the internal operation of (principally) single stars: their structure and evolution. Our overarching aim in this is to interpret observations such as the Hertzsprung–Russell diagrams shown in Fig. 11.1

For the present purposes, we use a working definition of a star as an isolated body that is bound by self-gravity, and which radiates energy supplied by an internal source. Self-gravity ensures that the star is approximately spherical (rotation introduces centrifugal forces which, for sufficiently fast rotation, may introduce distortions); the internal source of energy is nuclear fusion for most of the stellar lifetime (although for, e.g., white dwarfs, stored thermal energy is responsible for the observed luminosity).

The essence of stellar structure is the competition between the force of gravity, which always wants to make a star collapse, and the outward force of pressure. For almost the entire lifetime of a star, these forces are in balance; the star is in (or very close to) hydrostatic equilibrium, but as internal energy is released, the internal composition, and hence structure, must evolve. Thus ‘stellar structure’ and ‘evolution’ are intimately linked.

11.2 Stellar-structure equations

For reference, we remind ourselves of the basic equations of stellar structure:

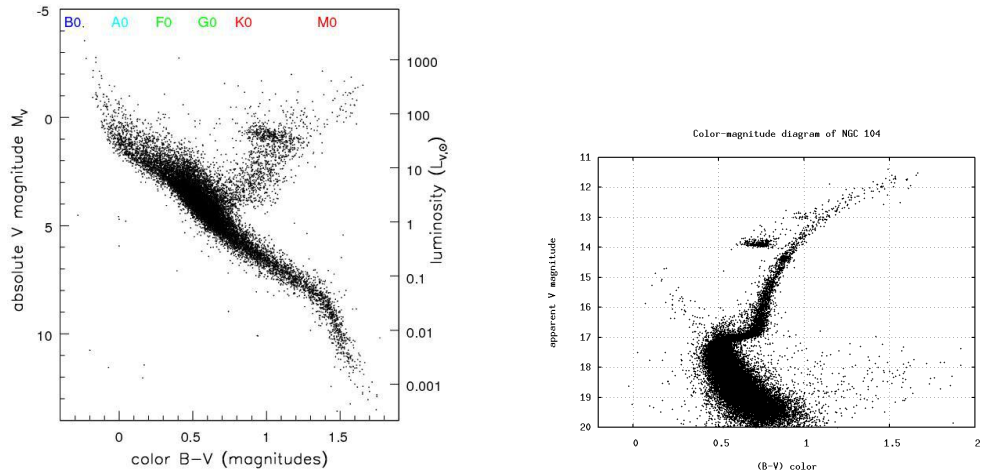


Figure 11.1: Hertzsprung–Russell (colour–magnitude) diagrams. Left, Hipparcos volume-limited sample (stars of different ages); right, HST observations of the globular cluster 47 Tuc (coeval sample).

11.2.1 Hydrostatic (pressure) equilibrium

Using standard terminology elaborated at the end of this section, hydrostatic equilibrium (HSE) is described by

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} = -\rho(r)\, g(r) \tag{2.1}$$

or

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} + \rho(r)\, g(r) = 0$$

The principal sources of pressure throughout a 'normal' (non-degenerate) star are gas pressure, and radiation pressure.[1] We will take the corresponding equations of state to be, in general,

$$\begin{aligned} P_\mathrm{G} &= nkT; \\ &= (\rho kT)/(\mu m(\mathrm{H})) \qquad (2.4) \\ P_\mathrm{R} &= \frac{1}{3}aT^4 \qquad (3.21) \end{aligned}$$

for number density n at temperature T, density ρ; μ is the mean molecular weight, and m(H) the hydrogen mass; a is the radiation constant, $a = 4\sigma/c$; with σ the Stefan-Boltzmann constant, and k Boltzmann's constant.

[1] Electron degeneracy pressure is important in white dwarfs, and neutron degeneracy pressure in neutron stars.

11.2.2 Mass Continuity

For spherical symmetry, the equation of mass continuity for static configurations is

$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \rho(r) \tag{2.5}$$

[In a spherically symmetric flow, such as may apply in a stellar wind with mass-loss rate $\dot{M}$, we instead have

$$\dot{M} = 4\pi r^2 \rho(r) v(r)$$

where $\rho(r), v(r)$ are the density and (radial) flow velocity at radius r.]

11.2.3 Energy continuity

$$\frac{\mathrm{d}L}{\mathrm{d}r} = 4\pi r^2 \rho(r) \epsilon(r) \tag{2.7}$$

where

- r is radial distance measured from the centre of the star
- $P(r)$ is the total pressure at radius r
- $\rho(r)$ is the density at radius r
- $g(r)$ is the gravitational acceleration at radius r
- $m(r)$ is the mass contained with radius r
- $L(r)$ is the total energy transported through a spherical surface at radius r
- $\epsilon(r)$ is the energy generation rate per unit mass at radius r

The stellar radius is R, the stellar mass is $M \equiv m(R)$, and the emergent luminosity $L \equiv L(R)$ (dominated by radiation at the stellar surface).

11.3 Virial theorem

In its most general form, the virial theorem expresses the relationship between the total kinetic energy of a system and its total potential energy. For our purposes, it relates the gravitational potential energy Ω and thermal energy U in a 'virialized system' (such as a star):

$$2U + \Omega = 0 \tag{11.1}$$

or $U = -\Omega/2$; thus the total energy is

$$E = U + \Omega = \Omega/2,$$

(which must be negative for the system to be bound).

11.4 Timescales

Several timescales are relevant to the study of stellar structure and evolution:

- The free-fall, or dynamical timescale, on which departures from hydrostatic equilibrium are restored in stellar systems,

$$\begin{aligned} t_{\rm dyn} &\simeq \sqrt{\frac{r^3}{Gm(r)}} \\ &\simeq 1/\sqrt{(G\rho)} \qquad\qquad \sim 2000 \text{ s for the Sun} \end{aligned} \tag{11.2}$$

- The Kelvin–Helmholtz timescale, the time taken for the gravitational potential energy to be radiated away:

$$t_{\rm KH} = \frac{|\Omega|}{L}$$

where

$$\begin{aligned} \Omega &= \int_0^R -G\frac{16\pi^2}{3} r^4 \overline{\rho}^2(r)\,{\rm d}r \\ &\simeq -\frac{16}{15}\pi^2 G\overline{\rho}^2 R^5 \\ &= -\frac{9}{15}\frac{GM^2}{R} \end{aligned}$$

(assuming $\overline{\rho}(r) = \overline{\rho}(R)$). The Kelvin–Helmholtz timescale for the Sun ($\overline{\rho} = 1.4 \times 10^3$ kg m^{-3}, $\Omega = 2.2 \times 10^{41}$ J) is $t_{\rm KH} \simeq 10^7$ yr.

- The thermal timescale is properly defined as

$$t_{\rm th} = \frac{U(\odot)}{\rm L_\odot}, \tag{11.3}$$

which (from the virial theorem) is $\sim {}^1\!/_2 t_{\rm KH}$. In practice, the factor 2 difference is of little importance for these order-of-magnitude timescales, and it is commonplace to take

$$t_{\rm th} \simeq t_{\rm KH}$$

- The nuclear timescale is a measure of how long it takes the reservoir of nuclear energy to be released,

$$t_{\rm N} = \frac{f_{\rm N} M c^2}{L} \tag{11.4}$$

where $f_{\rm N}$ is just the fraction of the rest mass available to the relevant nuclear process. In the case of hydrogen burning the fractional 'mass defect' is 0.007, so we might expect

$$t_{\rm N} = \frac{0.007 {\rm M}_\odot c^2}{\rm L_\odot} (\simeq 10^{11} \text{ yr for the Sun}).$$

However, in practice, only the core of the Sun – about ~10% of its mass – takes part in hydrogen burning, so $f_{\rm N} \simeq 10^{-3}$, and the nuclear timescale for hydrogen burning is $\sim 10^{10}$ yr for the Sun. Other evolutionary stages have their respective (shorter) timescales.

Overall, $t_{\rm dyn} < t_{\rm KH} \simeq t_{\rm th} < t_{\rm N}$. We will see that each of these timescales is appropriate, in turn, in the evolution of stars.

11.5 Pressure and temperature in the cores of stars

We can use the foregoing results immediately in order to obtain some very crude indicators of conditions in the cores of stars.

11.5.1 Central pressure (1)

The equation of hydrostatic equilibrium gives a rough order-of-magnitude estimate of pressures in stellar interiors; by letting $\mathrm{d}r = R_*$ (for an *approximate* solution!) then

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} \tag{2.1}$$

becomes

$$\frac{P_{\rm C} - P_{\rm S}}{R_*} \simeq \frac{GM_*\overline{\rho}}{R_*^2}$$

where P_C, P_S are the central and surface pressures; and since $P_C \gg P_S$,

$$P_C \simeq \frac{GM_*\overline{\rho}}{R_*} \tag{11.5}$$

$$= \frac{3}{4\pi}\frac{GM_*^2}{R_*^4}$$

$$\simeq 3 \times 10^{14} \text{ Pa } (=\text{N m}^{-2}) \tag{11.6}$$

(or about 10^{10} atmospheres) for the Sun. Because of our very crude approximation in integrating eqtn. (2.1) we expect this to be an underestimate (more nearly an average than a central value), and reference to more-detailed models[2] shows that it indeed falls short by about two orders of magnitude. Nevertheless, this simple calculation does serve to demonstrate high core values, along with a $\sim M^2/R^4$ dependence of P_C.

11.5.2 Central pressure (2)

Another estimate can be obtained by dividing the equation of hydrostatic equilibrium (2.1) by the equation of mass continuity (2.5):

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r}\bigg/\frac{\mathrm{d}m}{\mathrm{d}r} \equiv \frac{\mathrm{d}P(r)}{\mathrm{d}m} = \frac{-Gm(r)}{4\pi r^4}$$

Integrating over the entire star,

$$-\int_0^{M_*}\frac{\mathrm{d}P(r)}{\mathrm{d}m}\,\mathrm{d}m = P_C - P_S = \int_0^{M_*}\frac{Gm(r)}{4\pi r^4}\,\mathrm{d}m.$$

Evidently, because $R_* \geq r$, it must be the case that

$$\int_0^{M_*}\frac{Gm(r)}{4\pi r^4}\,\mathrm{d}m \geq \underbrace{\int_0^{M_*}\frac{Gm(r)}{4\pi R_*^4}\,\mathrm{d}m}_{=\frac{GM_*^2}{8\pi R_*^4};} \tag{11.7}$$

that is,

$$P_C > \frac{GM_*^2}{8\pi R_*^4} \quad (+P_S, \text{ but } P_S \simeq 0)$$

$$> 4.5 \times 10^{13} \text{ Pa } (=\text{N m}^{-2})$$

for the Sun. This is a weaker estimate than, but is consistent with, eqtn. (11.5), and shows the same overall scaling of $P_C \propto M_*^2/R_*^4$.

[2] A more accurate result will be obtained later on from polytropic models; eqtn. (14.13)

11.5.3 Central temperature

For a perfect gas,

$$P = nkT = \frac{\rho kT}{\mu m(\mathrm{H})} \tag{2.8}$$

but

$$P_\mathrm{C} \simeq \frac{GM_*\overline{\rho}}{R_*} \tag{11.5}$$

so

$$\begin{aligned} T_\mathrm{C} &\simeq \frac{\mu m(\mathrm{H})}{k}\frac{GM_*}{R_*}\frac{\overline{\rho}}{\rho_\mathrm{C}} \\ &\simeq 2.3\times10^7\,\mu\frac{\overline{\rho}}{\rho_\mathrm{C}}\ \mathrm{K} \end{aligned} \tag{11.8}$$

for the Sun – which is suprisingly close to the results of detailed calculations ($\sim 1.5{\times}10^7$ K). Note that at these temperatures the gas is fully ionized, and the perfect gas equation is an excellent approximation.

11.5.4 Mean temperature

We can use the Virial Theorem to obtain a limit on the mean temperature of a star. We have

$$\begin{aligned} U &= \int_V \frac{3}{2}kT(r)\,n(r)\,\mathrm{d}V \\ &= \int_V \frac{3}{2}kT(r)\,\frac{\rho(r)}{\mu m(\mathrm{H})}\,\mathrm{d}V \\ &= \int_0^{M_*} \frac{3}{2}kT(r)\frac{\rho(r)}{\mu m(\mathrm{H})}\frac{\mathrm{d}m}{\rho(r)} \end{aligned} \tag{11.9}$$

and

$$\begin{aligned} -\Omega &= \int_0^{M_*}\frac{Gm(r)}{r}\,\mathrm{d}m \\ &> \int_0^{M_*}\frac{Gm(r)}{R_*}\,\mathrm{d}m \\ &> \frac{GM_*^2}{2R_*} \end{aligned} \tag{11.10}$$

From the Virial Theorem, $2U = -\Omega$ (eqtn. 11.1), so

$$\frac{3k}{\mu m(\mathrm{H})}\int_0^{M_*} T\,\mathrm{d}m > \frac{GM_*^2}{2R_*}$$

The integral represents the sum of the temperatures of the infinitesimal mass elements contributing to the integral; the mass-weighted average temperature is

$$\begin{aligned}\overline{T} &= \frac{\int_0^{M_*} T \, \mathrm{d}m}{\int_0^{M_*} \mathrm{d}m} \\ &= \frac{\int_0^{M_*} T \, \mathrm{d}m}{M_*} \\ &> \frac{GM_*}{2R_*} \frac{\mu m(\mathrm{H})}{3k}\end{aligned}$$

For the Sun, this evaluates to $\overline{T}(\odot) > 2.3 \times 10^6$ K (using $\mu = 0.61$), i.e., $kT \simeq 200$ eV – comfortably in excess of the ionization potentials of hydrogen and helium (and enough to substantially ionize the most abundant metals), justifying the assumption of complete ionization in evaluating μ.

11.5.5 Solar values

Many very detailed models of the Sun's structure have been constructed, some in the contexts of helioseismology or of studying the solar-neutrino problem; for reference, we give the results one of these detailed models, namely Bahcall's standard model `bp2004stdmodel.dat`. This has

$$\begin{aligned}T_\mathrm{c} &= 1.570 \times 10^7 \text{ K} \\ \rho_\mathrm{c} &= 1.531 \times 10^5 \text{ kg m}^{-3} \\ P_\mathrm{c} &= 2.351 \times 10^{16} \text{ N m}^{-2}\end{aligned}$$

with 50% (95%) of the solar luminosity generated in the inner 0.1 (0.2) $\mathrm{R}_\odot$.

Section 12

Energy transport – I. Radiation

Stars are self-evidently hotter inside than outside (so that, e.g., nuclear fusion may occur), so there must be an energy flow. We are familiar with three basic mechanisms of energy transport:
– radiation
– convection
– conduction

In the context of stellar astrophysics, conduction is important only under the degenerate conditions found in white dwarfs and neutron stars (since gases in general are poor conductors). For 'normal' stars, therefore, the key processes transporting energy are radiation and convection.

Radiative transport: Energy is transported by photons. In stellar interiors the opacities are high, and the mean free path correspondingly low – about 1 mm in the case of the Sun (Box 12.3). In this sense, the radiation doesn't flow outwards, but rather diffuses outwards.

Convective transport: If the radiation is unable to escape a layer at a rate that matches the energy input, then 'something's got to give'. What gives is the static nature of the layer: convection is initiated and starts to transport energy. This suggests that hydrostatic equilibrium breaks down; but the dynamical timescale is short compared to the flow timescale, so in practice HSE continues to be an excellent approximation.

The nett energy flux is, under most circumstances, simply the sum of radiative and convective terms,

$$L(r) = L_{\rm rad}(r) + L_{\rm cnv}(r)$$

In order to determine under what circumstances convection is important, we first evaluate how much energy can be transported by radiation alone.

12.1 The equation of radiative transfer in stellar interiors

In optically thick environments – in particular, stellar interiors – radiation is often the most important transport mechanism, but, to repeat, for large opacities the radiant energy doesn't flow directly outwards; instead, it diffuses slowly outwards.

To express this transport quantitatively, the same general principles may be applied as led to the equation of radiative transfer in plane-parallel stellar atmospheres,

$$\mu \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = I_\nu - S_\nu. \tag{7.14}$$

[In the interior the photon mean free path is (very!) short compared to the radius, so 'plane parallel' is fine.]

We recall that, in general, the intensity, I_ν, at some position is direction-dependent; i.e., is $I_\nu(\theta, \phi)$ (although the explicit angular dependence is generally dropped for economy of nomenclature); the same is true in principle of the source function, although in practice any such dependence is negligible in stellar interiors.[1]

We proceed exactly as in our discussion of the Milne equations for radiative equilibrium (§8.2.1). That is, we multiply eqtn. (7.14) by $\mu \equiv \cos\theta$ and integrate over solid angle, using $\mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi = -\mathrm{d}\mu\,\mathrm{d}\phi$; then

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int_0^{2\pi} \int_{-1}^{+1} \mu^2 I_\nu(\mu, \phi)\,\mathrm{d}\mu\,\mathrm{d}\phi = \int_0^{2\pi} \int_{-1}^{+1} \mu I_\nu(\mu, \phi)\,\mathrm{d}\mu\,\mathrm{d}\phi - \int_0^{2\pi} \int_{-1}^{+1} \mu S_\nu(\mu, \phi)\,\mathrm{d}\mu\,\mathrm{d}\phi;$$

The radiation field in the interior is axially symmetric, (i.e., no azimuthal dependence), so $\int_0^{2\pi} \mathrm{d}\phi = 2\pi$ on both sides, and cancels, whence

$$\frac{\mathrm{d}}{\mathrm{d}\tau_\nu} \int_{-1}^{+1} \mu^2 I_\nu(\mu)\,\mathrm{d}\mu = \int_{-1}^{+1} \mu I_\nu(\mu)\,\mathrm{d}\mu - \int_{-1}^{+1} \mu S_\nu(\mu)\,\mathrm{d}\mu, \tag{8.6}$$

which we've already seen (in §8.2.1).

Working through the terms in this equation from right to left, we can safely assume that the source function is isotropic (since we expect the emissivity and opacity to be locally isotropic in the interior), so the 'S_ν' term comes out of the integral, which then evalautes to zero.

The next (i.e., first) term on the right-hand side should be familiar; recall that

$$\frac{1}{2} \int_{-1}^{+1} \mu I_\nu(\mu)\,\mathrm{d}\mu = H_\nu = \frac{F_\nu}{4\pi} \tag{3.11}$$

[1] The radiation quantities in eqtn. (7.14) are, of course, implicitly also functions of depth (τ or r).

where H_ν is the Eddington flux, or first-order moment of the radiation field. This is progress: the flux is the rate of energy transport, which is what we're seeking to evaluate.

Finally, the left-hand term in eqtn. (8.6) contains the second-order moment of the radiation field, or 'K integral'

$$\frac{1}{2}\int_{-1}^{+1}\mu^2 I_\nu(\mu)\,\mathrm{d}\mu \equiv K_\nu. \tag{3.12}$$

Using eqtns. (3.11) and (3.12) for the first two terms in eqtn. (8.6), we recover a previous result,

$$\frac{\mathrm{d}K_\nu}{\mathrm{d}\tau_\nu} = \frac{F_\nu}{4\pi}. \tag{8.7}$$

So far, our analysis has closely followed our discussion for stellar atmospheres. However, in interiors the radiation field is locally isotropic to a very good approximation (see Box 12.2), so in this instance – *unlike the stellar-atmosphere case* – we can take I_ν out of the integral in eqtn. (3.12), whence

$$\begin{aligned} K_\nu &= \frac{1}{2} I_\nu \left.\frac{\mu^3}{3}\right|_{-1}^{+1} \\ &= \frac{1}{3} I_\nu \qquad\qquad \left[\equiv \frac{1}{3} J_\nu \text{ for isotropy}\right], \end{aligned} \tag{3.13}$$

so, from eqtn. (8.7),

$$\frac{1}{3}\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = \frac{F_\nu}{4\pi}. \tag{12.1}$$

Because the photon mean free paths are very short, conditions in the interior are very close to local thermodynamic equilibrium (LTE). We therefore set $I_\nu = B_\nu(T)$, the Planck function; and by definition, $\mathrm{d}\tau_\nu = -k_\nu\,\mathrm{d}r$ (where the minus arises because the optical depth is measured inwards, and decreases with increasing r). Making these substitutions into eqtn. (12.1); introducing a term $\mathrm{d}T/\mathrm{d}T$; and integrating over frequency,

$$\int_0^\infty F_\nu\,\mathrm{d}\nu = -\frac{4\pi}{3}\int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\frac{\mathrm{d}T}{\mathrm{d}r}\mathrm{d}\nu \tag{12.2}$$

To simplify this further, we recall that the Rosseland mean opacity, $\bar{k}_\mathrm{R}$ ($= \bar{\kappa}_\mathrm{R}\rho$),[2,3] is defined by

$$\frac{1}{\bar{k}_\mathrm{R}}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu = \int_0^\infty \frac{1}{k_\nu}\frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu, \tag{5.1}$$

which can be evaluated separately, as a function of temperature and density.

[2]Remember, opacity may be expressed in several ways, most commonly as 'per unit mass' or 'per unit volume'. We use k to denote opacity per unit volume, and κ where reference is made to opacity per unit mass; thus $k = \kappa\rho$.

[3]The Rosseland mean opacity represents the harmonic mean of k_ν, weighted by $\mathrm{d}B_\nu(T)/\mathrm{d}T$. This weighting factor is small for very low and very high frequencies, and peaks at $\nu_\mathrm{p} = 4kT/h$.

Furthermore, given that

$$\int_0^\infty \pi B_\nu \, \mathrm{d}\nu = \sigma T^4$$

we also have

$$\begin{aligned}\int_0^\infty \frac{\mathrm{d}B_\nu(T)}{\mathrm{d}T}\mathrm{d}\nu &= \frac{\mathrm{d}}{\mathrm{d}T}\int_0^\infty B_\nu(T)\mathrm{d}\nu \\ &= \frac{4\sigma T^3}{\pi}\end{aligned}$$

(at given T) so that eqtn. (12.2) can be written as

$$\int_0^\infty F_\nu \, \mathrm{d}\nu = -\frac{4\pi}{3}\frac{1}{\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r}\frac{4\sigma T^3}{\pi}; \tag{12.3}$$

that is, the total (frequency-integrated) radiant energy flux is

$$F = -\frac{4}{3}\frac{acT^3}{\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r} \tag{12.4}$$

where a is the radiation constant, $4\sigma/c$ (and the minus sign simply means that the energy flows in the opposite direction to the temperature gradient).

Equation (12.4) is our adopted form for the radiative flux, or transport of energy by radiation. It's essential property is that $F \propto \mathrm{d}T/\mathrm{d}r$ (with no direct dependence on the rate of energy generation!); that is, radiative diffusion is completely analogous to Fourier's law of thermal conduction. It may be applied in environments where the photon mean free path is short compared to the scales over which physical parameters (notably temperature) change; it therefore becomes inappropriate as the stellar surface is approached, where a more detailed approach to radiative transfer is required.

Box 12.1. The radiative energy density is $U = aT^4$ (eqtn. 3.16), so that $\mathrm{d}U/\mathrm{d}T = 4aT^3$, and we can express eqtn. (12.4) as

$$\begin{aligned}F &= \int_0^\infty F_\nu \, \mathrm{d}\nu \\ &= -\frac{c}{3\bar{k}_\mathrm{R}}\frac{\mathrm{d}T}{\mathrm{d}r}\frac{\mathrm{d}U}{\mathrm{d}T} \\ &= -\frac{c}{3\bar{k}_\mathrm{R}}\frac{\mathrm{d}U}{\mathrm{d}r}\end{aligned}$$

This 'diffusion approximation' shows explicitly how the radiative flux relates to the energy-density gradient; the constant of proportionality, $c/3\bar{k}_\mathrm{R}$, is called the diffusion coefficient. The larger the opacity, the less the flux of radiative energy, as one might intuitively expect.

12.2 The radiative temperature gradient

The stellar luminosity at some radius r is given by

$$L(r) = 4\pi r^2 \int_0^\infty F_\nu \, \mathrm{d}\nu$$

so, from eqtn. (12.4)

$$L(r) = -\frac{16\pi}{3} \frac{r^2}{\overline{k}_\mathrm{R}} \frac{\mathrm{d}T}{\mathrm{d}r} acT^3, \tag{12.5}$$

For given L, we can simply rearrange eqtn. (12.5) to express the temperature gradient where energy transport is radiative:

$$\frac{\mathrm{d}T}{\mathrm{d}r} = -\frac{3}{16\pi} \frac{\overline{k}_\mathrm{R}}{r^2} \frac{L(r)}{acT^3} = -\frac{3}{16\pi} \frac{\overline{\kappa}_\mathrm{R}\rho(r)}{r^2} \frac{L(r)}{acT^3}. \tag{12.6}$$

We'll need this in a slightly different form later on, so combining this result with hydrostatic equilibrium,

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2}, \tag{2.1}$$

we obtain

$$\frac{\mathrm{d}T}{\mathrm{d}P} = \frac{3\overline{\kappa}_\mathrm{R} L(r)}{16\pi \, acT^3 \, Gm(r)} \tag{12.7}$$

or equivalently, since $\frac{\mathrm{d}\ln x}{\mathrm{d}x} = \frac{1}{x}$,

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} = \frac{3\overline{\kappa}_\mathrm{R} L(r) P}{16\pi \, acT^4 \, Gm(r)} \tag{12.8}$$

(a form that we'll use in Section 13.2).

12.3 Von Zeipel's law (not in lectures)

From eqtn. (12.4),

$$F \propto \frac{T^3}{\overline{\kappa}_\mathrm{R}\rho} \frac{\mathrm{d}T}{\mathrm{d}r} \tag{12.9}$$

$$\propto \frac{T^3}{\overline{\kappa}_\mathrm{R}\rho} \frac{\mathrm{d}T}{\mathrm{d}\psi} \frac{\mathrm{d}\psi}{\mathrm{d}r} \tag{12.10}$$

where ψ is the gravitational potential (and hence $\mathrm{d}\psi/\mathrm{d}r$ is the local gravity,[4] g). In hydrostatic equilibrium (see eqtn. 2.1)

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho(r)g(r) \propto \rho\frac{\mathrm{d}\psi}{\mathrm{d}r} \tag{2.1}$$

so that the pressure P is a function of the potential ψ – and hence the density must also be a function of ψ.[5] For an equation of state of the general form

$$T = T(P, \rho) \tag{12.11}$$

we therefore see that T must also be a function of ψ,

$$T = T(\psi). \tag{12.12}$$

The coefficient of $\mathrm{d}\psi/\mathrm{d}r$ in eqtn. (12.10) is therefore a function of ψ alone, whence

$$F \propto \frac{\mathrm{d}\psi}{\mathrm{d}r} \propto g \tag{12.13}$$

or, equivalently,

$$T_{\mathrm{eff}} \propto g^{0.25} \tag{12.14}$$

which is known as von Zeipel's law. Although it relies on the assumption of radiative energy transport by diffusion, which breaks down in a stellar atmosphere, the atmosphere is usually very thin compared to the radiative envelope, so even the *surface* flux can be expected to obey eqtn. (12.14) for stars in hydrostatic equilibrium and for which energy transport through the outer envelope is radiative.

Von Zeipel's law is of particular interest for close binary stars and rapidly rotating single stars. In either case, the local gravity, and hence the local temperature, can vary over the stellar surface (which is conventionally defined by a constant potential). Although increasing gravity results in increasing flux, the practical effects have come to be known as gravity *darkening*, because rapid rotation, or a close companion star, both serve to reduce a star's local gravity (and hence reduce the temperature locally).

It's of interest that von Zeipel also demonstrated that a rotating star *cannot* be simultaneously in strict hydrostatic and radiative equilibrium, undermining the basis of his 'law'. In practice, as shown by Eddington and by Sweet, rotation induces circulation currents in the stellar interior; however, these currents are sufficiently slow as to not lead to significant departures from hydrostatic equilibrium (the circulation timescales are long compared to the dynamical timescales discussed in Section 11.4), and gravity darkening is observed to occur in practice.

[4] In circumstances where von Zeipel's law is important, gravity is, in general, not a central force, so we should actually set $g = \nabla\psi$; but the central-force approximation is adequate for our purposes (and the correct general result is obtained).

[5] Since ρ is a scalar, the gradients of P and ψ are everywhere parallel.

12.4 Mass–Luminosity Relationship

We can now put together our basic stellar-structure relationships to demonstrate a scaling between stellar mass and luminosity. From hydrostatic equilibrium,

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} \quad \rightarrow \quad P \propto \frac{M}{R}\rho \tag{11.5}$$

but our (gas) equation of state is $P = (\rho kT)/(\mu m(\mathrm{H}))$, so

$$T \propto \frac{\mu M}{R}$$

(and $\mathrm{d}T/\mathrm{d}R \propto R^{-2}, \propto T/R$). For stars in which the dominant energy transport is radiative, we have

$$L(r) \propto \frac{r^2}{\bar{k}_\mathrm{R}} \frac{\mathrm{d}T}{\mathrm{d}r} T^3 \qquad \propto \frac{r^2}{\bar{\kappa}_\mathrm{R}\rho(r)} \frac{\mathrm{d}T}{\mathrm{d}r} T^3 \tag{12.5}$$

so at the surface ($r = R$)

$$L \propto \frac{RT^4}{\bar{\kappa}_\mathrm{R}\rho}.$$

From mass continuity (or just by inspection) $\rho \propto M/R^3$, giving

$$\begin{aligned} L &\propto \frac{R^4 T^4}{\bar{\kappa}_\mathrm{R} M} \\ &\propto \frac{R^4}{\bar{\kappa}_\mathrm{R} M} \left(\frac{\mu M}{R}\right)^4 ; \end{aligned}$$

i.e.,

$$L \propto \frac{\mu^4}{\bar{\kappa}_\mathrm{R}} M^3.$$

This simple dimensional analysis (which makes no assumptions about energy sources) yields a dependency which is in quite good agreement with observations; for solar-type main-sequence stars, the empirical mass–luminosity relationship is $L \propto M^{3.5}$. Small discrepancies between the simple analysis and nature should come as no surprise, as our two underpinning assumptions – radiative energy transport and negligible radiation-pressure support –are never fully realised in real stars.

Box 12.2. It may not be immediately obvious that the intensity of the radiation field in stellar interiors is, essentially, isotropic; after all, outside the energy-generating core, the full stellar luminosity is transmitted across any spherical surface of radius r. However, if this flux is small compared to the local mean intensity, then isotropy of the latter is justified.

The total radiative flux passing through a sphere of radius r (outside the energy-generating core) must equal the total radiative flux at R (the surface); that is,

$$\begin{aligned} F(r) &= \frac{4\pi R^2 \sigma T_{\rm eff}^4}{4\pi r^2} \\ &= \sigma T_{\rm eff}^4 \frac{R^2}{r^2} \end{aligned}$$

while in the interior the mean intensity is

$$J_\nu(r) \simeq B_\nu(T(r)) = \sigma T^4(r).$$

The ratio of flux to mean intensity is thus

$$\frac{F(r)}{J(r)} = \left(\frac{T_{\rm eff}}{T(r)}\right)^4 \left(\frac{R}{r}\right)^2 .$$

The temperature rises rapidly below the surface of a star, so this ratio is always small; for example, in the Sun, $T(r) \simeq 0.5$ MK at $r = 0.9\mathrm{R}_\odot$, whence $F/J \simeq 10^{-8}$ – that is, the radiation field is isotropic to around 1 part in a hundred million.

Equivalently, the average temperature gradient from the centre of the Sun (for example) to the surface is

$$\frac{\Delta T}{\Delta r} = \frac{T_{\rm c} - T_{\rm eff}}{R_\odot} \simeq 10^{-2}\ \mathrm{K\ m^{-1}} \tag{12.15}$$

The photon mean free path is $\ell = 1/\kappa \simeq 1$ mm (§7.2; numerical value from detailed models), so the temperature change over this distance is of order 10^{-5} K. The radiant energy density is $U = aT^4$, so the relative anisotropy $\Delta U/U = 4\Delta T/T \simeq 10^{-11}$ at 10^6 K.

Nonetheless, although the anisotropy is very small, the nett radiation flux is large – in fact, equal to the stellar luminosity.

Box 12.3: Diffusion timescale for radiative transport

Deep inside stars the radiation field is very close to black body. For a black-body distribution the average photon energy is

$$\overline{E} = U/n \simeq 4 \times 10^{-23} T \qquad [\text{J photon}^{-1}]. \tag{3.18}$$

The core temperature of the Sun is $T_c \simeq 1^1/_2 \times 10^7$ K (cf. eqtn. 11.8), whence $\overline{E} = 3.5$ keV – i.e., photon energies are in the X-ray regime.

Light escaping the surface of the Sun ($T_{\rm eff} \simeq 5770$K) has a mean photon energy $\sim 3 \times 10^3$ smaller, in the optical.

The source of this degradation in the mean energy is the coupling between radiation and matter. Photons obviously don't flow directly out from the core, but rather they diffuse through the star, travelling a distance of order the local mean free path, ℓ, before being absorbed and re-emitted in some other direction (a 'random walk').[a] The mean free path depends on the opacity of the gas:

$$\ell = 1/k_\nu = 1/(\kappa_\nu \rho) \tag{7.5}$$

where k_ν is the volume opacity (dimensions of area per unit volume, or SI units of m^{-1}) and κ_ν is the mass opacity (area per unit mass, so units of $\mathrm{m}^2\ \mathrm{kg}^{-1}$).

After $n_{\rm sc}$ scatterings the *radial* distance travelled is, on average, $\sqrt{n_{\rm sc}}\ell$ (it's a statistical, random-walk process). Thus to travel a distance $\mathrm{R}_\odot$ we require

$$n_{\rm sc} = \left(\frac{\mathrm{R}_\odot}{\ell}\right)^2. \tag{12.16}$$

Solar-structure models give an average mean free path of $\ell \simeq 1$ mm (incidentally, justifying the LTE approximation in stellar interiors); with $\mathrm{R}_\odot \simeq 7 \times 10^8$ m we have

$$n_{\rm sc} \simeq 5 \times 10^{23}.$$

The total distance travelled by a (fictitious) photon travelling from the centre to the surface is $n_{\rm sc} \times \ell \simeq 5 \times 10^{20}$ m ($\sim 10^{12} \mathrm{R}_\odot$!), and the time to diffuse to the surface is $(n_{\rm sc} \times \ell)/c \simeq 5 \times 10^4$ yr.

[More-detailed calculations[b] give 17×10^4 yr; why? Simply because the volume opacity in the inner regions of the Sun is greater than the average value, so the mean free path there is shorter. Because of the 'square root' nature of the diffusion, a region with twice the volume opacity takes four times longer to pass through, while a region with half the volume opacity takes only 1.414 times shorter; so any non-uniformity in the volume opacity inevitably leads to a longer total diffusion time than the simple estimate based on a constant mean free path.]

[a] Of course, no single photon really travels from the core to the surface, and the mean photon energy is steadily degraded moving outwards; but for heuristic purposes we can at least consider a photon emitted at the surface to be the $n^{\rm th}$-generation descendant of a photon generated in the core.

[b] Mitalas & Sills, ApJ, 401, 759, (1992). Even these 'more-detailed' calculations are quite crude estimates, ignoring any frequency dependence.

Section 13

Energy transport – II. Convection

For convection to occur, then clearly there must be some temperature gradient (in the case of stars, a radial temperature gradient). We have seen that, where energy transport is radiative, the temperature gradient is given by

$$\frac{\mathrm{d}T}{\mathrm{d}r} = -\frac{3}{16\pi}\frac{\bar{k}_\mathrm{R}}{r^2}\frac{L(r)}{acT^3} \tag{12.6}$$

that is, high opacity leads to large temperature gradients (as we might expect intuitively; the opacity blocks the flow of radiant energy from hotter to cooler regions, so there's a strongh temperature gradient across the opaque region). If the energy flux isn't contained by the temperature gradient, we have to invoke another mechanism for energy transport: convection (since conduction is negligible in ordinary stars.) Under what circumstances will this arise? Karl Schwarzschild[1] (1906) developed a now-standard criterion for determining when convection will occur.

13.1 The Schwarzschild criterion

To follow Schwarzschild's reasoning, we suppose that we start with a stellar envelope in radiative equilibrium and that, through some minor perturbation, a cell (or bubble, or blob) of gas is displaced upwards within the star. Our basic physical assumptions are that the cell rises slowly enough that it remains in pressure equilibrium (essentially, that it moves subsonically, so that hydrostatic equilibrium is maintained), but fast enough that it doesn't exchange energy with its surroundings (i.e., that the cell behaves adiabatically).

As the cell rises into a lower-pressure regime, it will expand in order to remain in pressure equilibrium with its surroundings; this takes place on a dynamical timescale that is set by the speed of sound and the

[1] Perhaps better known for finding the first exact solutions to the field equations of Einstein's general theory of relativity, leading to the 'Schwarzschild radius' for the event horizon of nonrotating black holes.

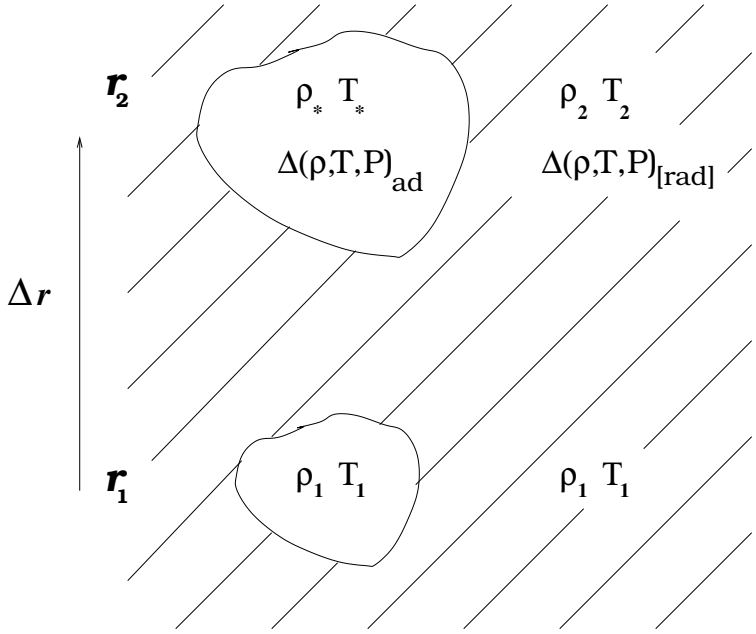


Figure 13.1: A (potentially) convective 'blob' in a stellar envelope, rising from radius r_1 to r_2, illustrating the principles discussed in Section 13.1.

linear scale of the perturbation. However, the adiabatic cell will not necessarily remain in thermal equilibrium with its surroundings; that is, the pressure in the cell will match that of the surrounding gas, but, individually, the density and temperature need not. If the gas in the cell is less dense than its surroundings, then simple buoyancy comes into play; the cell will continue to rise, and convective motion occurs.[2] If the cell is more dense than its surroundings, it will sink back down, and the system is stable against convection.

The essence of the Schwarzschild criterion is, therefore, that instability occurs (rising cell less dense than surroundings) if the adiabatic density decrease is greater than in the ambient medium:

$$|\Delta\rho|_{\rm ad} > |\Delta\rho|_{\rm amb}$$

(where the 'ad', 'amb' subscripts indicate adiabatic and ambient conditions, and we take absolute values just so as not to have to worry about the meaning of inequalities of negative quantities should they arise). In the limit of small changes, the density change can be written as

$$\Delta\rho = \left(\frac{{\rm d}\rho}{{\rm d}r}\right)\Delta r$$

and so we can express this statement of instability as

$$\left|\frac{{\rm d}\rho}{{\rm d}r}\right|_{\rm ad} > \left|\frac{{\rm d}\rho}{{\rm d}r}\right|_{\rm amb} \qquad (13.1)$$

[2] Another way of looking at this is that the entropy (per unit mass) of the blob is conserved, so the star is unstable if the ambient entropy per unit mass decreases outwards in the star.

(since Δr is the same for the cell and the ambient gas; Fig. 13.1).

Although eqtn. (13.1) captures the essential physics behind the Schwarzschild criterion, it more usual to work in terms of temperature gradients (rather than density gradients). To this end, we use our assumption of hydrostatic (pressure) equilibrium; the pressures inside the cell and in the ambient gas are equal at both r_1 and r_2, so the changes in pressure must also be equal:

$$\Delta P_{\rm ad} = \Delta P_{\rm amb}.$$

However, $P \propto \rho T$ (equation of state, eqtn. 2.4), so any increase in density must be matched by a decrease in temperature, whence the temperature-gradient equivalent to eqtn. (13.1) (i.e., the condition for convective instability) is

$$\left|\frac{{\rm d}T}{{\rm d}r}\right|_{\rm ad} < \left|\frac{{\rm d}T}{{\rm d}r}\right|_{\rm amb}. \tag{13.2}$$

Finally, for practical applications it is convenient to express the temperature gradient in terms of pressure change (rather than radius change)/ Again invoking the equation of hydrostatic equlibrium

$$\frac{{\rm d}P(r)}{{\rm d}r} = -\rho(r)\, g(r) \tag{2.1}$$

and the (gas-pressure) equation of state,

$$P = (\rho k T)/(\mu m({\rm H})) \tag{2.4}$$

we can write

$$\begin{aligned}
\left|\frac{{\rm d}T}{{\rm d}r}\right| &\equiv \left|\frac{{\rm d}T}{{\rm d}P}\frac{{\rm d}P}{{\rm d}r}\right| \\
&= \left|\frac{{\rm d}T}{{\rm d}P}\, g\rho\right| && \text{(HSE)} \\
&= \left|\frac{{\rm d}T}{{\rm d}P}\right|\, g\frac{\mu m({\rm H})}{kT}P && \text{(EOS)} \\
&= \left|\frac{{\rm d}\ln T}{{\rm d}\ln P}\right|\, g\frac{\mu m({\rm H})}{k}.
\end{aligned} \tag{13.3}$$

Substituting this into eqtn. (13.2) we obtain the condition for instability in yet another (and widely used) form,

$$\left|\frac{{\rm d}\ln T}{{\rm d}\ln P}\right|_{\rm ad} < \left|\frac{{\rm d}\ln T}{{\rm d}\ln P}\right|_{\rm amb},$$

frequently written using the more compact notation

$$\nabla_{\rm ad} < \nabla \tag{13.4}$$

(where ∇ is understood to mean $\nabla_{\rm amb}$). This is the form in which the Schwarzschild criterion for convective instability is often expressed. It tells us that if a *superadiabatic* temperature gradient exists (i.e., if the actual temperature gradient exceeds the adiabatic value, which is given by eqtn. 13.5, below), then convection will occur.

13.2 When do superadiabatic temperature gradients actually occur?

Since large – potentially superadiabatic – temperature gradients will arise if the opacity is high (eqtn. 12.6), one circumstance under which convection may occur is when the opacity is too high for radiative transport to be efficient. To demonstrate this analytically we appeal to thermodynamic arguments.

Under the adiabatic conditions appropriate to our rising cell,

$$PV^{\gamma} = \text{constant}$$

where $\gamma = C_P/C_V$, the ratio of specific heats at constant pressure and constant volume. Thus, for a gas cell of constant mass (for which $V \propto \rho^{-1}$),

$$\begin{aligned} P &\propto \rho^{\gamma}; && \text{but also} \\ P &\propto \rho T, && \text{whence} \\ P^{\gamma-1} &\propto T^{\gamma} \end{aligned} \tag{2.4}$$

or

$$\left|\frac{\mathrm{d}(\ln T)}{\mathrm{d}(\ln P)}\right|_{\mathrm{ad}} = \frac{\gamma - 1}{\gamma} \tag{13.5}$$

(which is one way of expressing the adiabatic temperature gradient). The Schwarzschild criterion for convective instability,

$$\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| > \left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right|_{\mathrm{ad}} \qquad [\equiv\ \nabla > \nabla_{\mathrm{ad}}]\,, \tag{13.4}$$

can therefore be written as

$$\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| > \frac{\gamma - 1}{\gamma} \qquad \left[\equiv\ \nabla > \frac{\gamma - 1}{\gamma}\right]. \tag{13.6}$$

For example, convective energy transport must occur if the temperature gradient for radiative transport is superadiabatic; i.e., if

$$\frac{3\bar{\kappa}_{\mathrm{R}} L(r) P}{16\pi a c T^4 G m(r)} > \frac{\gamma - 1}{\gamma}. \tag{13.7}$$

(using eqtn. 12.8)

13.2.1 Physical conditions associated with convection

Equations(13.7) identifies several ways in which convection may, in principle, be induced – that is, to achieve a superadiabatic temperature gradient.

(i) If the luminosity is large, then the radiative flux may not be able to transport all the energy. An example is the cores of massive stars, where the radiation flux per unit area, $L(r)/4\pi r^2$, can be very large, driving convection (in spite of quite small, electron-scattering, opacities).

(ii) Where the opacity $\bar{\kappa}_{\rm R}$ is too great to allow the radiation to flow at an equilibrium rate (e.g., protostars).

(iii) In the envelopes of cool stars, where the adiabatic exponent γ can approach unity, and hence $(\gamma - 1)/\gamma$ can become very small (the γ effect; cf. eqtn. 13.7).

For a monatomic ideal gas (representative of stellar interiors),[3] $\gamma = 5/3$ and so $(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = 0.4$ (eqtn. 13.5); but under changing conditions of ionization this exponent changes. For a simple pure-hydrogen composition it can be shown that

$$\left|\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}\right| = \frac{2 + X(1-X)\left((5/2) + E_1/(kT)\right)}{5 + X(1-X)\left((5/2) + E_1/(kT)\right)^2}$$

where

$$X = \frac{n_{\rm e}}{n_{\rm P} + n({\rm H}^0)}$$

is the degree of ionization and E_1 is the ionization potential. For $X = 0$ or 1, this recovers $(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = 0.4$; but in regions of *partial* ionization lower values apply, with a minimum at $X = 0.5$ [$(\mathrm{d}\ln T/\mathrm{d}\ln P)_{\rm ad} = (\gamma - 1)/\gamma = 0.07$] which occurs (e.g.) near the base of the solar photosphere.

Additionally, rearranging and differentiating the equation of state $P = nkT = (\rho kT)/(\mu m({\rm H}))$ gives

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} = 1 + \frac{\mathrm{d}\ln \mu}{\mathrm{d}\ln P} - \frac{\mathrm{d}\ln \rho}{\mathrm{d}\ln P}, \tag{13.8}$$

which demonstrates that large gradients in μ (i.e., in composition and/or degree of ionization) can also influence whether or not convection occurs (separately to the effect of composition on the opacities). Hence

(iv) Convection is facilitated by a strong μ gradient (eqtn. 13.8); this can come about when there is a change in ionization (so the number of particles per unit mass changes) or composition (such as may occur at the boundary between the core and envelope).

The switch from radiative core/convective envelope to convective core/radiative envelope occurs on the main sequence at masses only very slightly more than the Sun's. This is related to the core energy-generation mechanism, as the principal hydrogen-burning process switches from proton-proton chains (which generate energy at a rate that can be transported radiatively) to CNO processing.

[3] For stars supported by radiation pressure, the EOS obeys a 'gas law' with $\gamma = 4/3$, and $(\mathrm{d}\ln T/\mathrm{d}\ln P) = 0.25$

13.3 Convective energy transport: mixing-length 'theory'

So far, we have only tested whether or not convection is likely to occur; we have not addressed how energy is actually transported by this mechanism – i.e., we don't yet know what the convective energy flux actually is. Unfortunately, convection is a complex, hydrodynamic process. Although much progress is being made in numerical modelling of convection over short timescales, at present it's not feasible to model convection in detail in stellar-evolution codes as a matter of routine, because of the vast disparities between convective (~dynamical) and evolutionary (~nuclear) timescales. Instead, we appeal to simple parameterizations of convection, of which mixing-length 'theory' is the most venerable, and still the most widely applied.

We suppose that
(i) the envelope becomes convectively unstable at some radius r_0, and that a convective cell then rises, hydrostatically and adiabatically, through some characteristic distance ℓ – the 'mixing length';
(ii) the excess thermal energy of the cell is then released into the ambient medium; and
(iii) the cooled cell sinks back down (or, from another perspective, is displaced downwards by the next rising cell).
Because we are moving energy from deeper to higher regions, the temperature gradient is shallower for the cell than the pure radiative case (the upper layers are hotter than they would otherwise be).

From hydrostatic equilibrium (eqtn. 2.1) and the perfect gas equation (eqtn. 2.4) we have

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -gP\frac{\mu m(\mathrm{H})}{kT}, \qquad \text{or}$$
$$\frac{\mathrm{d}P}{P} = -g\frac{\mu m(\mathrm{H})}{kT}\mathrm{d}r, \equiv -\frac{\mathrm{d}r}{H}. \tag{13.9}$$

The solution of eqtn. (13.9) is

$$P = P_0 \exp(-r/H)$$

where H, the pressure scale height, is the vertical distance over which the pressure changes by a factor e. The mixing length, ℓ, is conveniently expressed in terms of this scale height; typically, we expect $\ell \simeq H$, but, since the detailed physics is not well captured, a scaling factor (or fudge factor!) is usually introduced, whereby

$$\ell = \alpha H,$$

with $\alpha \sim 0.5$–1.5.

For simplicity (and given the weakness of other assumptions), we suppose that ℓ is the same for all cells, and that the velocity of all cells is also the same.

For a cell moving with velocity v the flux of energy across unit area is given by the mass flux times the heat energy per unit mass:

$$F_{\mathrm{conv}} = \rho v \times C_{\mathrm{P}}\Delta T \tag{13.10}$$

where C_P is the specific heat at constant pressure and ΔT is the temperature difference between cell and surroundings. To progress we need an estimate of the velocity v, which we obtain by considering the buoyancy force,

$$f_\mathrm{b} = -g\Delta\rho. \tag{13.11}$$

Here $\Delta\rho$ is the density difference between the cell and ambient gas which we can determine from the equation of state ($P = (\rho kT)/(\mu m(\mathrm{H}))$, eqtn. 2.4),

$$\frac{\Delta P}{P} = \frac{\Delta\rho}{\rho} + \frac{\Delta T}{T} - \frac{\Delta\mu}{\mu}.$$

In hydrostatic (pressure) equilibrium $\Delta P = 0$, whence

$$\frac{\Delta\rho}{\rho} = \frac{\Delta\mu}{\mu} - \frac{\Delta T}{T}$$

or

$$\begin{aligned}\Delta\rho &= -\rho\frac{\Delta T}{T}\left(1 - \frac{\Delta\mu}{\mu}\frac{T}{\Delta T}\right),\\ &\rightarrow -\rho\frac{\Delta T}{T}\left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right)\end{aligned}$$

so the buoyancy force, eqtn. (13.11), is

$$f_\mathrm{b} = g\rho\frac{\Delta T}{T}\left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right)$$

but force equals mass (per unit volume) times acceleration,

$$= \rho\frac{\mathrm{d}v}{\mathrm{d}t}$$

so

$$\frac{\mathrm{d}v}{\mathrm{d}t} = g\frac{\Delta T}{T}\left(1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T}\right) \tag{13.12}$$

where the excess temperature of the cell compared to the ambient gas can be written

$$\Delta T = \left\{\left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_\mathrm{amb} - \left|\frac{\mathrm{d}T}{\mathrm{d}r}\right|_\mathrm{ad}\right\} \times \Delta r. \tag{13.13}$$

For constant acceleration over distance ℓ, starting from rest, the final velocity is given by Torricelli's equation,

$$v = \sqrt{\left\{2\frac{\mathrm{d}v}{\mathrm{d}t}\ell\right\}}$$

so, substituting eqtn. (13.13) for ΔT in eqtn. (13.12) (setting $\Delta r = \ell$, and neglecting a factor $\sqrt{2}$ by way of a rough correction to go from final velocity to mean velocity), the required velocity is

$$v = \left\{ \frac{g}{T} \left| 1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T} \right| \right\}^{1/2} \left\{ \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{amb}} - \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{ad}} \right\}^{1/2} \times \ell.$$

We can now rewrite eqtn. (13.10) as

$$F_{\mathrm{conv}} = \rho C_{\mathrm{P}} \left\{ \frac{g}{T} \left| 1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T} \right| \right\}^{1/2} \left\{ \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{amb}} - \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right|_{\mathrm{ad}} \right\}^{3/2} \times \ell^2.$$

Rearranging the equation of state, eqtn. (2.4),

$$\begin{aligned} \left| \frac{\mathrm{d}T}{\mathrm{d}r} \right| &= \frac{g\mu m(\mathrm{H})}{k} \left| \frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} \right| \\ &= \frac{T}{H} \left| \frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} \right|, \qquad = \frac{T}{H} |\nabla| \end{aligned}$$

and so

$$F_{\mathrm{conv}} = \rho C_{\mathrm{P}} \alpha^2 T \left\{ gH \left| 1 - \frac{\mathrm{d}\ln\mu}{\mathrm{d}\ln T} \right| \right\}^{1/2} \{ |\nabla|_{\mathrm{amb}} - |\nabla|_{\mathrm{ad}} \}^{3/2}$$

which is our final formulation. Although it looks rather unwieldy, this form is useful in that everything on the right-hand side is 'known' at each step in an iterative solution of a stellar-structure model.

In calculating actual temperature structures in stellar envelopes, we require the total energy flux to obey

$$F = F_{\mathrm{rad}} + F_{\mathrm{conv}} = \sigma T_{\mathrm{eff}}^4$$

(outside the energy-generating core). In practice, in numerical modelling, the initial temperature structure is calculated on the basis of radiative transfer only ($F_{\mathrm{rad}} = \sigma T_{\mathrm{eff}}^4$), then a correction $\Delta T(r)$ computed iteratively, for given α, if the Schwarzschild criterion indicates convective transport occurs.

13.4 Closing remarks

XXXVelocities, timescales, mixing length

Mixing-length theory was developed by Ludwig Prandtl (1875–1953), a pioneer in fluid dynamics, and was introduced to stellar astrophysics in the 1950s by Erika Böhm-Vitense. There are several formulations that differ in technical details, but they all share the same essential concepts, and the same fundamental strengths and weaknesses:

The basic weakness is that the theory is purely local and stationary. The α parameter is normally adjusted to produce models that are in agreement with observations, and in general only a global average

can be established (any dependencies on mass, depth, and composition are indeterminate, and the empirical adjustment will tend to compensate for any other shortcomings in, or failings of, the models).

In reality, convection is turbulent, and is associated with a spectrum of cell sizes. This is a complex hydrodynamic problem, so the basic strength of mixing-length theory is that it is straightforward to implement computationally (and that it appears to do a reasonable job in practice).

Section 14

Analytical stellar-structure models: polytropes and the Lane–Emden equation

14.1 Introduction

We've already assembled a set of equations that embody the basic principles governing stellar structure; these are

$$\frac{\mathrm{d}m(r)}{\mathrm{d}r} = 4\pi r^2 \rho(r) \qquad \text{Mass continuity;} \qquad (2.5)$$

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = \frac{-Gm(r)\rho(r)}{r^2} \qquad \text{Hydrostatic equilibrium;} \qquad (2.1)$$

$$\frac{\mathrm{d}L(r)}{\mathrm{d}r} = 4\pi r^2 \rho(r)\varepsilon(r) \qquad \text{Energy continuity.} \qquad (2.7)$$

Our aim is to use these to investigate (or to predict) the properties of real stars. To do this in detail we also need descriptions of the quantities P, ε, and $\bar{k}_\mathrm{R}$ (pressure, energy-generation rate, and Rosseland mean opacity), which enter these equations, directly or indirectly. These are each functions (often complex functions) of density, temperature, and composition, and in modern work are computed explicitly, separately from the stellar-structure problem itself. However, progress in analytical models can be made simply by adopting an appropriate equation of state – discussed in the next section.

14.2 Polytropes

Historically, the idea of polytropes arose through modelling of fully convective gases, in which the gas is completely turned over – hence polytropic ('poly', many, much; tropos, to turn).

In the context of the Schwarzschild criterion (§13.1), convection may be assumed to occur rapidly enough that a cell of gas doesn't exchange heat with its surroundings, i.e., is adiabatic, so that PV^γ is constant. We can generalize this adiabatic equation of state to a polytropic equation of state, such that the pressure is assumed to be proportional to density to some power (equivalent to PV^γ = constant, since $V \propto 1/\rho$ for fixed mass):

$$P = K\rho^\gamma \equiv K\rho^{(n+1)/n}, \tag{14.1}$$

where K is the polytropic constant (of proportionality),
$\gamma(\equiv (n+1)/n)$ is the polytropic exponent, and
$n(\equiv 1/(\gamma - 1)$ is the polytropic index (not to be confused with number-density n).

[In the case of adiabatic processes, γ is given by the ratio of specific heats, $C_\mathrm{P}/C_\mathrm{V}$; but in general, γ need not have this direct physical interpretation. Indeed, eqtn. (14.1) need not necessarily even be taken to be an equation of state (the Lane–Emden equation has applicability outside stellar structures) – it simply expresses an assumption regarding the relationship between pressure and density.]

A sphere of gas satisfying eqtn. (14.1) is said to be a 'polytrope of index n'. The importance of polytropes is that they allow simple solutions to the equations of stellar structure; the Swiss astronomer Jacob Robert Emden[1] first developed the astrophysical applications,[2] although the essential physics had been described by the American astronomer Jonathan Lane in 1870. Superficially, the price is paid for the simplifications offered by polytropic models is the apparent decoupling of pressure from temperature in eqtn. (14.1); however, this turns out to be less restrictive than one might initially suppose (as we shall see in Section 14.4). Of course, pressure is not actually independent of temperature (a perfect-gas equation of state can hold in parallel with the polytropic EOS); rather, a relationship between density and temperature in a form that allows the latter to be subsumed is implied.

14.3 The Lane–Emden Equation

We can now proceed with using results we've assembled so far to investigate stellar structure, using polytropic models. The Lane–Emden equation represents a solution of the most basic of the equations of stellar structure for such models, expressed in a dimensionless form.

We start with the equation of hydrostatic equilibrium,

$$\frac{\mathrm{d}P(r)}{\mathrm{d}r} = -\frac{Gm(r)\rho(r)}{r^2}; \tag{2.1}$$

[1] Emden was Karl Schwarzschild's brother-in-law, and Martin Schwarzschild's uncle.

[2] Emden 1907: *Gaskugeln: Anwendungen der mechanischen Wärmetheorie auf kosmologische und meteorologische Probleme* ('Gas balls: Applications of the mechanical heat theory to cosmological and meteorological problems').

rearranging and differentiating gives

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r}\right) = -G\frac{\mathrm{d}m(r)}{\mathrm{d}r}$$

whence, from mass continuity ($\mathrm{d}m/\mathrm{d}r = 4\pi r^2\rho(r)$; eqtn. 2.5)

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{\rho}\frac{\mathrm{d}P}{\mathrm{d}r}\right) = -4\pi G\rho, \tag{14.2}$$

which already embodies the critical aspects of the Lane–Emden equation.[3]

Eqtn. (14.2) describes, in principle, how P, ρ vary with radius r in a star, but leaves us with one equation for these two unknowns. We seek to eliminate one or other of these; therefore, having dealt with 'mechanical' issues, we now need to turn to the thermal structure. It might appear that we can't solve hydrostatic equilibrium without knowing something about the pressure, i.e., the temperature, which in turn suggests needing to know about energy generation processes, opacities, and other complexities. However, we can find interesting results by adopting a polytropic relationship between pressure and density,

$$P = K\rho^{\gamma} = K\rho^{(n+1)/n} \tag{14.1}$$

(thereby sidestepping the temperature dependence of pressure). Introducing eqtn. (14.1) into (14.2) gives

$$\frac{K}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{r^2}{\rho}\frac{\mathrm{d}\rho^{\gamma}}{\mathrm{d}r}\right) = -4\pi G\rho. \tag{14.3}$$

We now have one equation with just one unknown; given some boundary conditions, we can therefore now solve for $\rho(r)$. If we wanted to do so now, we'd just use some sort of numerical integration. Before the era of electronic computers, however, this approach was discouragingly laborious, so the pioneer investigators of stellar structures sought to reduce eqtn. (14.3) to a non-dimensional form that could be solved once and for all (for a given polytropic index).

We proceed in this direction by defining dimensionless scaled variables ξ and θ in place of the physical variables r and ρ:

$$\xi = r/\alpha, \tag{14.4}$$

$$\theta^n(\xi) = \rho(r)/\rho_{\mathrm{c}}, \tag{14.5}$$

[3] It's also, essentially, Poisson's equation in spherical co-ordinates; from hydrostatic equilibrium,

$$\frac{1}{\rho}\frac{\mathrm{d}P(r)}{\mathrm{d}r} = -g(r), = \frac{\mathrm{d}\psi}{\mathrm{d}r}$$

whence $\nabla^2\psi = 4\pi G\rho$.

where α is some constant having units of length, ρ_c is the central density, and our task now is to determine θ as a function of ξ (equivalent to finding ρ as a function of r). Substituting our scaled variables, eqtn. (14.3) becomes

$$\frac{K}{(\alpha\xi)^2}\frac{\mathrm{d}}{\mathrm{d}(\alpha\xi)}\left\{\frac{(\alpha\xi)^2}{\rho_c\theta^n}\frac{\mathrm{d}\left(\{\rho_c\theta^n\}^{(n+1)/n}\right)}{\mathrm{d}(\alpha\xi)}\right\} = -4\pi G\rho_c\theta^n$$

which looks moderately intimidating; but since

$$\frac{\mathrm{d}}{\mathrm{d}\xi}(\theta^{n+1}) = (n+1)\theta^n\frac{\mathrm{d}\theta}{\mathrm{d}\xi}$$

it reduces to

$$\frac{(n+1)K\rho_c^{(1-n)/n}}{4\pi G\alpha^2}\frac{1}{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right) = -\theta^n.$$

This is still 'symbol soup' to some extent, but by setting the constant α (which is freely selectable, provided its dimensionality – length – is preserved) to be

$$\alpha \equiv \left\{\frac{(n+1)K\rho_c^{(1-n)/n}}{4\pi G}\right\}^{1/2}, \tag{14.6}$$

we obtain a compact second-order differential equation relating (scaled) radius to (scaled) density:

$$\frac{1}{\xi^2}\frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right) = -\theta^n \quad \left[= -\frac{\rho}{\rho_c}\right] \tag{14.7}$$

which is the standard form of the Lane–Emden equation.[4]

14.3.1 Solutions

The Lane–Emden equation is, in essence, a dimensionless form of eqtn. (14.3), and its solutions give us density as a function of radius, for appropriate boundary conditions. Because it is a second-order differential equation, we need two boundary conditions in order to solve it.

A first is obtained trivially from the definition of θ; at the core,

$$\theta(\xi = 0)\left[= \frac{\rho}{\rho_c}(r = 0)\right] = 1.$$

Furthermore, the density gradient at the core must be zero (if we go 'inwards' the density increases, then after we pass through the centre it decreases); $\mathrm{d}\rho/\mathrm{d}r = 0$ as $r \to 0$, so

$$\left.\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right|_{\xi=0} = 0$$

(where the suffix indicates that the derivative is evaluated at $\xi = 0$).

[4] If desired, we can expand this to

$$\frac{1}{\xi^2}\left(2\xi\frac{\mathrm{d}\theta}{\mathrm{d}\xi} + \xi^2\frac{\mathrm{d}^2\theta}{\mathrm{d}\xi^2}\right) + \theta^n = \frac{\mathrm{d}^2\theta}{\mathrm{d}\xi^2} + \frac{2}{\xi}\frac{\mathrm{d}\theta}{\mathrm{d}\xi} + \theta^n = 0.$$

With these boundary conditions, analytical solutions describing the entire run of θ with ξ are possible for polytropic indexes $n = 0$, 1 and 5 (i.e., $\gamma = \infty, 2$, and 1.2); these are, respectively,[5]

$$\begin{aligned} \theta(\xi) &= 1 - \xi^2/6 & \left(\xi_1 = \sqrt{6}\right. \\ &= \sin\xi/\xi & \left. = \pi \right. \\ &= \left(1 + \xi^2/3\right)^{-1/2} & \left. = \infty \right) \end{aligned}$$

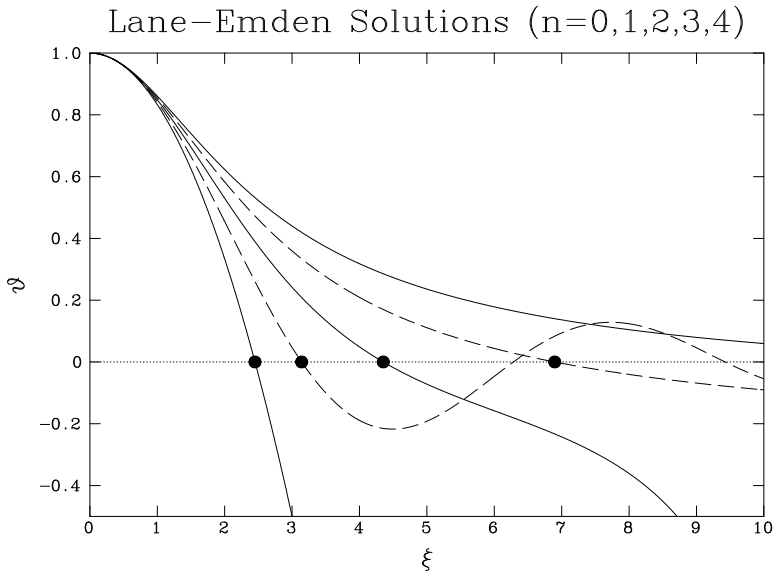
where ξ_1 is the first root of θ (i.e., the smallest positive value of ξ for which $\theta = 0$) – that is, the rescaled radius of the star, R/α.

Solutions of the Lane–Emden equation for other values of n can't be expressed as analytical functions of ξ in this way, but are straightforward to compute by simple numerical integration. These solutions are normally given just in terms of surface ('ξ_1') values of ξ and $\mathrm{d}\theta/\mathrm{d}\xi$ (or $-\xi^2\,\mathrm{d}\theta/\mathrm{d}\xi$), as in the following table (which I calculated using a very simple numerical integration):

Solutions of the Lane–Emden Equation

n	ξ_1	$-\xi_1^2 \left.\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right\vert_{\xi_1}$
0.0	2.4495	4.8990
0.5	2.7527	3.7887
1.0	3.1416	3.1416
1.5	3.6538	2.7141
2.0	4.3529	2.4111
2.5	5.3553	2.1872
3.0	6.8968	2.0182
3.5	9.5358	1.8906
4.0	14.972	1.7972
4.5	31.836	1.7378
5.0	∞	1.7321

[5] Details of the solutions are available at, e.g., `http://mathworld.wolfram.com/Lane-EmdenDifferentialEquation.html`



Solutions of the Lane–Emden equation for several values of the polytropic index n (increasing left to right). Black dots show the first root of θ (ξ_1, corresponding to the stellar radius) for each value of the polytropic index n.

A polytrope with index $n = 0$ has a uniform density (recall $\rho/\rho_c = \theta^n$), while a polytrope with index $n = 5$ has an infinite radius. In general, the larger the polytropic index, the more centrally condensed the density distribution; and only polytropes with $n \leq 5$ are bound systems (Section 14.5). A polytrope with index $n = \infty$ has $P(= K\rho^{(1+1/n)}) = K\rho$ and is a so-called 'isothermal sphere', a self-gravitating, isothermal body – used to analyse collisionless systems of stars (in particular, globular clusters).

A thermodynamic aside

Any thermodynamic process that obeys $P = K\rho^\gamma$ is called a polytropic process. By inspection,
$\gamma = 0$ corresponds to an isobaric (constant-pressure) process, and
$\gamma = \infty$ is an isochoric (constant-volume) process.
If the process involves a constant ratio of energy transfer as heat to energy transfer as work ($\delta Q/\delta W$ = constant), then the perfect-gas equation of state applies ($P_G = nkT$; eqtn. 2.4), in which case
$\gamma = 1$ is an isothermic process, and
$\gamma = C_P/C_V$ is an adiabatic process.

A perfect-gas EOS can be applied in polytropic models of stars. Then

$$\frac{\rho}{\rho_\mathrm{c}} = \left(\frac{P}{P_\mathrm{c}}\right)^{1/\gamma} = \left(\frac{\rho T}{\rho_\mathrm{c} T_\mathrm{c}}\right)^{1/\gamma} \text{; i.e.,}$$

$$\left(\frac{\rho}{\rho_\mathrm{c}}\right)^{(\gamma-1)/\gamma} = \left(\frac{T}{T_\mathrm{c}}\right)^{1/\gamma} = \theta^{1/\gamma}, \text{ or}$$

$$\frac{T}{T_\mathrm{c}} = \theta \tag{14.8}$$

Thus θ can be seen as reflecting density *or* temperature (or pressure) as a function of radius.

14.3.2 Relating scaled variables to physical stellar parameters

The solution of Lane–Emden equation depends only on the polytropic index, n, but is expressed in terms of scaled parameters. If we want to obtain astrophysically more interesting solutions (in terms of actual stellar masses, radii, etc.) then we need two physical parameters to transform from the scaled ones. We might choose actual numerical values for K and ρ_c, for example; we can then evaluate:

- The stellar radius,

$$R \equiv \alpha \xi_1 \tag{14.4}$$

$$= \left\{\frac{K(n+1)}{4\pi G}\right\}^{1/2} \rho_\mathrm{c}^{(1-n)/(2n)} \xi_1 \tag{14.9}$$

  (from eqtn. 14.6).

- The stellar mass,

$$M = \int_0^R 4\pi r^2 \rho \, \mathrm{d}r$$

  but $r = \alpha\xi$ (eqtn. 14.4) and $\rho = \rho_\mathrm{c}\theta^n$ (eqtn. 14.5), so

$$M = 4\pi\alpha^3 \rho_\mathrm{c} \int_0^{\xi_1} \xi^2 \theta^n \, \mathrm{d}\xi;$$

  then using the Lane–Emden equation, (14.7), for θ^n,

$$M = -4\pi\alpha^3 \rho_\mathrm{c} \int_0^{\xi_1} \frac{\mathrm{d}}{\mathrm{d}\xi}\left(\xi^2 \frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right) \mathrm{d}\xi$$

$$= 4\pi \left\{\frac{K(n+1)}{4\pi G}\right\}^{3/2} \rho_\mathrm{c}^{(3-n)/(2n)} \left\{-\xi_1^2 \left.\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right|_{\xi_1}\right\} \tag{14.10}$$

- The mean stellar density,

$$\begin{aligned}\overline{\rho} &= \frac{3M}{4\pi R^3} \\ &= \rho_c \frac{3}{\xi_1^3}\left\{-\xi_1^2 \left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\},\end{aligned} \tag{14.11}$$

  (where we use our previous results, eqtns. 14.10 and 14.9 for M and R).

- The central pressure can be expressed trivially as

$$P_c = K\rho_c^{(n+1)/n}, \tag{14.1}$$

Should we want these results expressed in terms of stellar mass and radius, we need to find K, ρ_c in terms of these parameters. We first rearrange eqtn. (14.11) to obtain

$$\rho_c = \frac{M}{4\pi R^3}\xi_1 \left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^{-1},$$

or, numerically,

$$\rho_c = 4.70 \times 10^2\, \xi_1 \left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^{-1} \left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-3} \text{kg m}^{-3}$$

Eliminating ρ_c between eqtns (14.9) and (14.10) we obtain K in terms of M and R:

$$\begin{aligned}K &= \frac{G}{n+1}(4\pi)^{1/n}\,\xi_1^{-(n+1)/n}\left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^{(1-n)/n} M^{(n-1)/n}R^{(3-n)/n}, \\ &= \frac{G}{n+1}\left[\frac{4\pi}{\xi_1^{(n+1)}}\left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^{(1-n)}\right]^{1/n} M^{(n-1)/n}R^{(3-n)/n}.\end{aligned} \tag{14.12}$$

Then the central pressure, in terms of mass and radius, is

$$\begin{aligned}P_c &= \left[4\pi(n+1)\left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^2\right]^{-1}\frac{GM^2}{R^4} \\ &= 8.96 \times 10^{13}\left[(n+1)\left\{-\left.\frac{d\theta}{d\xi}\right|_{\xi_1}\right\}^2\right]^{-1}\left(\frac{M}{M_\odot}\right)^2\left(\frac{R}{R_\odot}\right)^{-4} \text{Pa}\end{aligned} \tag{14.13}$$

(matching the functional form $P_c \propto M^2/R^4$ found previously from simpler considerations; eqtn. 11.6).

The core temperature can be estimated from the central pressure and density, *if* we know the equation of state. In general, it isn't safe to assume that the dominant pressure is gas pressure; radiation pressure

also needs to be considered, and we'll address this in Section 14.4.2. However, if (but only if) gas pressure dominates, then we can just use the perfect-gas equation of state,

$$\begin{aligned} T_\mathrm{C} &\simeq \frac{\mu m(\mathrm{H})}{k}\frac{P_\mathrm{c}}{\rho_\mathrm{c}} \\ &= \frac{\mu m(\mathrm{H})}{(n+1)k}\left\{-\xi\left.\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right|_{\xi_1}\right\}^{-1}\frac{GM}{R} \\ &= \frac{2.29\times10^7\mu}{(n+1)}\left\{-\xi\left.\frac{\mathrm{d}\theta}{\mathrm{d}\xi}\right|_{\xi_1}\right\}^{-1}\left(\frac{M}{\mathrm{M}_\odot}\right)\left(\frac{R}{\mathrm{R}_\odot}\right)^{-1}\ \mathrm{K} \end{aligned}$$

14.4 Astrophysical Solutions

We've done a lot of algebra; what about the physical interpretation of all this? Recall that the Lane–Emden equation is the solution of the equations of hydrostatic equilibrium and mass continuity, for a polytropic equation of state, expressed in dimensionless form. We might therefore reasonably ask what, if any, types of star can be reasonably modelled as polytropes.

14.4.1 Stars with fully convective energy transport

The first law of thermodynamics states that the change in internal energy of a system, $\mathrm{d}U$, is given by the heat added to the system, $\mathrm{d}Q$, less the work done by the system, $\mathrm{d}W$:

$$\mathrm{d}U = \mathrm{d}Q - \mathrm{d}W.$$

For fully convective stars, all the convective cells are supposed to be adiabatic, so $\mathrm{d}Q \equiv 0$; and for a quasistatic process $\mathrm{d}W = P\mathrm{d}V$, whence

$$\mathrm{d}U = -P\mathrm{d}V.$$

For an ideal gas[6]

$$\begin{aligned} P &= nkT & &= \frac{N}{V}kT & &\text{(where } N = nV\text{) and so} \\ U &= \frac{3}{2}NkT & &\to \frac{3}{2}VP \end{aligned}$$

[6] A gas is close to ideal if it is fully ionized, or entirely neutral. If the gas is partially ionized, then some energy may go into ionization/dissociation, and $U \neq {}^3\!/_2 NkT$.

so that

$$\begin{aligned}
\mathrm{d}U &= -P\mathrm{d}V; \\
\mathrm{d}\left(\frac{3}{2}PV\right) &= -P\mathrm{d}V; \\
\frac{3}{2}(P\mathrm{d}V + V\mathrm{d}P) &= -P\mathrm{d}V; \\
\frac{5}{2}P\mathrm{d}V &= -\frac{3}{2}V\mathrm{d}P; \\
\frac{\mathrm{d}P}{P} &= -\frac{5}{3}\frac{\mathrm{d}V}{V}; \\
P &\propto V^{-5/3}, \qquad \text{but } V \propto \rho^{-1} \text{ so} \\
P &\propto \rho^{5/3}. \qquad (14.14)
\end{aligned}$$

That is, fully convective stars are approximately polytropic, with $n\,(= 1/(\gamma - 1)) = 3/2$.

or...

In a fully convective star, $P \simeq P_{\rm G}$ (i.e., radiation pressure is unimportant); then, from $P = (\rho kT)/\mu m(\mathrm{H})$, we have $T \propto (P/\rho)$.

If we have an adiabatic temperature gradient

$$\left(\frac{\partial T}{\partial P}\right) \equiv \nabla_{\rm ad} \rightarrow P \propto T^{1/\nabla_{\rm ad}}$$

so

$$\begin{aligned}
P^{1-(1/\nabla_{\rm ad})} &\propto \rho^{-1/\nabla_{\rm ad}}, \text{ or} \\
P &= K\rho^{1/(1-\nabla_{\rm ad})}
\end{aligned}$$

(and for a perfect gas, $P \propto \rho^{5/3}$).

14.4.2 The 'Eddington Standard Model'

The total pressure P is the sum of gas pressure $P_{\rm G}$ and radiation pressure $P_{\rm R}$,[7] where

$$P_{\rm G} = \frac{\rho}{\mu m(\mathrm{H})}kT, \qquad \text{and} \qquad (2.4)$$

$$P_{\rm R} = \frac{1}{3}aT^4. \qquad (3.21)$$

We define

$$P_{\rm G} \equiv \eta P \qquad \text{(or, equivalently,} \qquad P_{\rm R} \equiv (1-\eta)P); \qquad (14.15)$$

[7] Plus degeneracy pressure, which is negligible for normal stars.

we can therefore write

$$\begin{aligned} P &= \frac{P_{\rm G}}{\eta} \\ &= \frac{\rho k T}{\eta \mu m({\rm H})}. \end{aligned}$$

Rearranging for T, substituting the result into eqtn. (3.21), and inserting the result into eqtn. (14.15) gives

$$P^4 \left(\frac{\eta \mu m({\rm H})}{\rho k} \right)^4 = \frac{3(1-\eta)}{a} P$$

i.e.,

$$P = \left\{ \frac{k}{\mu m({\rm H})} \right\}^{4/3} \left\{ \frac{3(1-\eta)}{a\eta^4} \right\}^{1/3} \rho^{4/3}; \tag{14.16}$$

thus if η and μ are constant throughout a star then

$$P \equiv K \rho^{4/3} \tag{14.17}$$

and the star is a polytrope with polytropic index $n = 3$. The behaviour of $n = 3$ polytropes was investigated by Eddington, and is embodied in the 'Eddington Standard Model'.

The assumption of constant μ through the star is likely to be tolerable, at least for main-sequence stars, but is constant η plausible? Since $P_{\rm R} = \frac{1}{3} a T^4$ (eqtn. 3.21), we have

$$\frac{{\rm d}T}{{\rm d}P_{\rm R}} = \frac{3}{4aT^3};$$

but we know that radiative energy transport is described by

$$\frac{{\rm d}T}{{\rm d}P} \equiv \frac{{\rm d}T}{{\rm d}P_{\rm R}} \frac{{\rm d}P_{\rm R}}{{\rm d}P} = -\frac{3 \bar{\kappa}_{\rm R} L(r)}{16\pi\, acT^3\, Gm(r)}, \tag{12.7}$$

so

$$\frac{{\rm d}P_{\rm R}}{{\rm d}P} = \frac{\bar{\kappa}_{\rm R} L(r)}{4\pi c G m(r)}.$$

Integrating gives[8]

$$\frac{P_{\rm R}}{P} \simeq \frac{\bar{\kappa}_{\rm R} L}{4\pi c G M}$$

[8]The astute student may note the connection with the Eddington limit to luminosity: $L_{\rm Edd} = (4\pi c G M)/\kappa$, eqtn. (3.22)

assuming negligible surface pressure and that $\overline{\kappa}_{\rm R}$ is constant with radius r (not unreasonable for hot stars). That is, for stars where energy transport is radiative, there is a constant ratio of gas pressure to radiation pressure (i.e., constant η), and so the Eddington model is indeed applicable.

[Kramers opacity law has $\overline{\kappa}_{\rm R} \propto T^{-3.5}$ while the proton–proton chain leads to $L/M \propto T^{+3.5}$; thus $\overline{\kappa}_{\rm R} L/M \simeq$ constant for a range of stars of different masses.]

Some results for the Eddington model ($n = 3$, constant η) are set in Problem Sheet 4. We note one additional result, namely that from eqtns. (14.16) and (14.17) we have

$$K \propto \left\{ \frac{(1-\eta)}{\eta^4} \right\}^{1/3}$$

[where $P_{\rm R} = (1-\eta)P$], while from eqtn. (14.10), with $n = 3$

$$K \propto M^{2/3};$$

that is,

$$\left\{ \frac{(1-\eta)}{\eta^4} \right\} \propto M^2.$$

This shows us that as M increases, η decreases (the left-hand side is a monotonically decreasing function of η) – that is, with increasing mass, radiation pressure becomes increasingly important. While stars like the Sun are largely supported by gas pressure, the most massive stars are almost entirely supported by radiation.

14.4.3 Degenerate stars as polytropes

White dwarfs are supported by electron degeneracy pressure (or 'Fermi pressure'), and neutron stars by neutron-degeneracy pressure, so we anticipate an equation of state that is independent of temperature. The uncertainty principle tells us that, in general,

$$\Delta p \Delta x \geq \hbar/2$$

where p is the momentum and x is some spatial co-ordinate. Heuristically, using electron degeneracy as an example, the electron momentum must be of the order of, or at least scale with, the uncertainty, i.e.,

$$p_{\rm e} \propto \hbar/\Delta x.$$

If the electron density is $n_{\rm e}$ then the average spacing between electrons is of order

$$\Delta x \propto n_{\rm e}^{-1/3}, \qquad \text{so}$$
$$p_{\rm e} \propto \hbar n_{\rm e}^{1/3}$$

Non-relativistic case. The electron momentum is $p_e = m_e v$; i.e.,

$$v \propto n_e^{1/3}.$$

The (electron-degeneracy) pressure – momentum per unit area per unit time – is

$$P_e \sim p \cdot n_e \cdot v; \qquad \text{that is,}$$
$$P_e \propto n_e^{5/3}, \quad \propto \rho^{5/3},$$

a polytrope of index $n = 3/2$. Detailed calculations give the constant of proportionality.

$$P_{e,n} = \frac{h^2}{20 m_e m(\mathrm{H}) \mu_e} \left(\frac{3}{\pi m(\mathrm{H}) \mu_e} \right)^{2/3} \rho^{5/3}$$
$$\equiv K_1 \rho^{5/3} \tag{14.18}$$

(where μ_e is the mean number of nucleons per electron).

Relativistic case. In the relativistic limit $v \simeq c$, so

$$P_e \sim n_e \cdot c \cdot p; \qquad \text{i.e.,}$$
$$P_e \propto \rho^{4/3},$$

or in detail

$$P_{e,r} = \frac{hc}{8 m(\mathrm{H}) \mu_e} \left(\frac{3}{\pi m(\mathrm{H}) \mu_e} \right)^{1/3} \rho^{4/3},$$
$$\equiv K_2 \rho^{4/3} \tag{14.19}$$

so that relativistic degenerate stars are also polytropic, though with a different polytropic index ($n = 3$).

XXX Compare Maxwell-Boltzmann distribution for momentum (velocity) in a non-degenerate gas with the distribution for degenerate gas (w/Fermi momentum).

14.4.4 Mass–radius relation: application to white dwarfs

In § 14.3.2, the central density, ρ_c, was eliminated between equations (14.9) and (14.10) to give

$$M^{(n-1)/n} R^{(3-n)/n} = K \left/ \left\{ \left[\frac{4\pi}{\xi_1^{(n+1)}} \left(-\left. \frac{\mathrm{d}\theta}{\mathrm{d}\xi} \right|_{\xi_1} \right)^{(1-n)} \right]^{1/n} \right\} \right. \tag{14.12}$$

The right-hand side has a constant numerical value that depends only on the polytropic index. We see that polytropes therefore follow a mass–radius relationship,

$$M \propto R^{(n-3)/(n-1)}.$$

We know that non-relativistic white dwarfs are polytropes with $P \propto \rho^{5/3}$ (i.e., $n = 3/2$; Section 14.4.3), whence $M \propto R^{-3}$. Using eqtn. (14.18) in eqtn. (14.12), and inserting numerical values, we obtain a quantitative mass-radius relation for non-relativistic white dwarfs,

$$R = \frac{K_1}{0.4242 G M^{1/3}}, \text{ or}$$
$$\frac{R}{R_\odot} \simeq 10^{-2} \frac{M}{M_\odot}^{-1/3},$$

showing that when a star like our Sun becomes a white dwarf it will have about the same radius as that of the Earth (~6000 km), and hence a density of $\sim 10^9$ kg m^{-3}.

With increasing mass, the radius decreases, and the density rises, eventually entering the relativistic regime. (As the particles get squeezed into smaller and smaller volumes, the uncertainty principle implies that the velocities are larger and larger.) Then

$$P = P_{\rm e,r} = K_2 \rho^{4/3},$$

where

$$K_2 = \frac{hc}{8 m(\mathrm{H}) \mu_\mathrm{e}} \left(\frac{3}{\pi m(\mathrm{H}) \mu_\mathrm{e}} \right)^{1/3} \rho^{4/3} \tag{14.19}$$

Using eqtn. (14.19) in eqtn. (14.12), and inserting numerical values,

$$\begin{aligned} M &= \left(\frac{K_2}{0.3639 G} \right)^{3/2} \\ &= 1.142 \times 10^{31} \mu_\mathrm{e}^{-2} \text{ kg}, \\ &= 5.74 \mu_\mathrm{e}^{-2} \mathrm{M}_\odot \end{aligned}$$

This is the 'Chandrasekhar mass', the maximum mass that can be supported by relativistic degeneracy pressure (recall, we obtained eqtn. 14.19 by setting $v \simeq c$, so this is a strong upper limit); for a helium white dwarf the number of nucleons per electron is $\mu_\mathrm{e} \simeq 2$, and so $M_{\rm Ch} \simeq 1.44 \mathrm{M}_\odot$. A star more massive than the Chandrasekhar limit that has no other significant sources of pressure will continue to collapse, to a neutron star (supported by neutron degeneracy pressure) or to a black hole.[9]

14.4.5 An alternative look at the mass-radius relation: white dwarfs [not for lectures]

The kinetic energy of a particle of mass m, velocity v, is

$$E_\mathrm{p} = \frac{1}{2} m v^2;$$

[9] The equation of state for nuclear matter is not well established, so that the upper limit for a neutron-star mass – the 'Tolman–Oppenheimer–Volkoff limit' – is less certain than is the Chandrasekhar limit. The most accurate determinations of neutron-star masses (from binary-star systems) are all remarkably close to $1.4 \mathrm{M}_\odot$, but with a range of ~ 1–$2 \mathrm{M}_\odot$.

and the uncertainty principle tells us that

$$\Delta x \Delta p \geq \hbar/2$$

where p is the momentum and x is some spatial co-ordinate such that

$$\Delta x^3 = V/N$$

for a star of volume V containing N particles. That is, the uncertainty principle shows us that

$$p^2 \propto \hbar^2 N^{2/3}/R^2$$

whence the total kinetic energy of the entire star ($\propto N$) is

$$E_* \propto (\hbar^2 N^{5/3})/(mR^2)$$

The star is in equilibrium when the kinetic energy equals the gravitational energy; that is, when

$$(\hbar^2 N^{5/3})/(mR^2) \propto M^{5/3}/R^2 \simeq GM^2/R,$$

or

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

as before.

As the mass increases the radius gets smaller, so we must eventually enter the relativistic regime (Δx decreases so Δp increases). The relativistic form for the energy in the limit $v \to c$ is

$$E_\mathrm{p} = pc;$$

then repeating the arguments from above leads to

$$M^{2/3} \approx \hbar c m_\mathrm{p}^{-4/3}/G \simeq 1.4\mathrm{M}_\odot.$$

That is, the maximum mass of a star that can be supported by electron degeneracy pressure depends only on physical constants, and not on any other property of the star.

14.5 Binding energy [not for lectures]

It's of interest also to investigate the total energy of a polytrope. We start by considering the gravitional potential energy; this is given in the usual way by

$$\begin{aligned} \Omega &= -\int_\mathrm{c}^\mathrm{s} \frac{Gm}{r}\,\mathrm{d}m \\ &\equiv -\int_\mathrm{c}^\mathrm{s} \frac{G}{2r}\,\mathrm{d}(m^2) \end{aligned} \tag{14.20}$$

(where the c, s limits mean 'centre' and 'surface'). Integrating by parts,

$$\Omega = -\frac{GM^2}{2R} - \int_\mathrm{c}^\mathrm{s} \frac{Gm^2}{2r^2}\,\mathrm{d}r,$$

but from hydrostatic equlibrium, $\mathrm{d}r = -(r^2\mathrm{d}P)/(Gm\rho)$, so this can be written as

$$\Omega = -\frac{GM^2}{2R} + \int_\mathrm{c}^\mathrm{s} \frac{\cancel{Gm}m}{2\cancel{r^2}} \frac{\cancel{r^2}}{\cancel{Gm}} \frac{\mathrm{d}P}{\rho}. \tag{14.21}$$

We now use the polytropic equation of state,

$$P = K\rho^\gamma = K\rho^{(n+1)/n}, \tag{14.1}$$

to obtain

$$\frac{\mathrm{d}P}{\mathrm{d}\rho} = K\frac{n+1}{n}\rho^{1/n};$$

we also use the trick of writing eqtn. (14.1) as

$$\frac{P}{\rho} = K\rho^{1/n}$$

to write

$$\begin{aligned}
\frac{\mathrm{d}\,(P/\rho)}{\mathrm{d}\rho} &= \frac{K}{n}\rho^{(1-n)/n} \\
\frac{\mathrm{d}\,(P/\rho)}{\cancel{\mathrm{d}\rho}} &= \frac{\mathrm{d}P}{\rho\cancel{\mathrm{d}\rho}}\frac{1}{n+1}; \qquad \text{i.e.,} \\
\frac{\mathrm{d}P}{\rho} &= (n+1)\,\mathrm{d}\,(P/\rho)
\end{aligned}$$

Substituting back into eqtn. (14.21) gives us

$$\Omega = -\frac{GM^2}{2R} + \frac{n+1}{2}\int_\mathrm{c}^\mathrm{s} m\,\mathrm{d}\,(P/\rho);$$

integrating by parts for a second time,

$$= -\frac{GM^2}{2R} + \left[\frac{n+1}{2}m\frac{P}{\rho}\right]_\mathrm{c}^\mathrm{s} - \frac{n+1}{2}\int_\mathrm{c}^\mathrm{s} \frac{P}{\rho}\,\mathrm{d}m.$$

The term in [brackets] vanishes ($m = 0$ at the centre, $P/\rho = 0$ at the surface), so, using mass continuity (eqtn. 2.5),

$$\begin{aligned}
\Omega &= -\frac{GM^2}{2R} - \frac{n+1}{2}\int_\mathrm{c}^\mathrm{s} \frac{P}{\rho}4\pi r^2\rho\,\mathrm{d}r \\
&= -\frac{GM^2}{2R} - \frac{n+1}{2}\int_\mathrm{c}^\mathrm{s} P\frac{4\pi}{3}\,\mathrm{d}r^3.
\end{aligned}$$

Integrating by parts yet again,

$$\Omega = -\frac{GM^2}{2R} - \left[\frac{n+1}{2}\frac{4\pi}{3}Pr^3\right]_\mathrm{c}^\mathrm{s} + \frac{n+1}{6}\int_\mathrm{c}^\mathrm{s} 4\pi r^3\,\mathrm{d}P.$$

The second term again vanishes ($r_\mathrm{c} = 0$, $P_\mathrm{s} = 0$); then, using the equation of hydrostatic equilibrium once more,

$$\Omega = -\frac{GM^2}{2R} - \frac{n+1}{6}\int_\mathrm{c}^\mathrm{s} 4\pi r^3\frac{Gm\rho}{r^2}\,\mathrm{d}r,$$

and mass continuity,

$$\Omega = -\frac{GM^2}{2R} - \frac{n+1}{6}\int_c^s \frac{Gm}{r}\,\mathrm{d}m$$

– but the integral is the definition of Ω (eqtn. 14.20) that we started with! That is,

$$\Omega = -\frac{GM^2}{2R} - \frac{n+1}{6}\Omega, \qquad \text{or}$$
$$\Omega = \left(\frac{3}{n-5}\right)\frac{GM^2}{R}.$$

We can then appeal to the virial theorem to obtain the total energy,

$$E = \Omega + U = \frac{1}{2}\Omega = \left(\frac{3}{n-5}\right)\frac{GM^2}{2R}.$$

We see that for $n > 5$ the energy is $E > 0$, meaning the system is unbound; only polytropes with $n < 5$ are gravitationally bound, and hence potentially of interest as stellar models.

14.5.1 Summary of ~polytropic stars

To summarize, the following systems are reasonably approximated as polytropes:
– convective stars

$$P \propto \rho^{5/3} \qquad n = 1.5. \tag{14.14}$$

– stars with constant η (~radiative stars)

$$P \propto \rho^{4/3} \qquad n = 3; \tag{14.17}$$

– non-relativistic white dwarfs,

$$P_{\mathrm{e,n}} \propto \rho^{5/3} \qquad n = 1.5; \tag{14.18}$$

– relativistic white dwarfs,

$$P_{\mathrm{e,r}} \propto \rho^{4/3} \qquad n = 3; \tag{14.19}$$

Section 15

Pre-main-sequence evolution

THIS SECTION UNDER DEVELOPMENT

15.1 Introduction: review

We now turn to evolution of stars. We'll divide this into three broad sections, namely
– Pre-main-sequence evolution;
– Main-sequence evolution; and
– Post-main-sequence evolution.

We'll adopt a definition of 'main-sequence evolution' as being phases when core hydrogen burning is powering the stars. In this section, we'll consider the stages of development prior to this.

The general principles of star formation (and associated planetary-system formation) can be traced back as far as the 18th-century ideas of Emanuel Swedenberg, Immanual Kant, and Pierre-Simon Laplace. Their 'nebular hypothesis' already contained the key concepts that underpin current star-formation theories: a cloud of gas may collapse under gravity, forming a spinning disk (through conservation of angular momentum) with high central densities that eventually lead to the emergence of a star.

We can directly observe objects representing the different stages of the star-formation process. On scales of parsecs and up, we see the giant molecular clouds which may be subject to gravitational collapse. This gives rise to high-density 'molecular cores' (on sub-parsec scales). As a zeroth-order approximation, we can view these gravitationally collapsing gas clouds as being in more or less free-fall collapse, radiating away much of their initial potential energy on a dynamical timescale (so that collapse is approximately isothermal).

15.2 Initial cloud collapse: Jeans mass (should be PHAS0018 review)

Observationally, we know that stars form from interstellar matter, typically in groups (clusters and associations). The essential physical process is evidently one of gravitational collapse, which occurs on a dynamical timescale. However, because the densities are low, even in the 'dense' clouds associated with star formation ($n \sim 10^4$ cm^{-3}, $\sim 10^{-17}$kg m^{-3}), even this timescale is quite long ($\sim 10^6$ yr).

We'll consider collapse from a simple, idealised perspective. The internal (thermal) energy of our initial system is

$$U = \frac{3}{2}NkT$$

where N is the total number of particles, nV. The gravitational potential energy is

$$\Omega = -f\frac{GM^2}{R}$$

where f is a factor, of order unity, that depends on the mass distribution; for uniform density, $f = 3/5$ (box 15.1; more centrally condensed systems yield larger values of f). So, the Virial theorem ($2U = -\Omega$) for an initially uniform gas cloud becomes

$$3NkT = \frac{3}{5}\frac{GM^2}{R}$$

If the equality actually holds (i.e., if the 'virial ratio', $|U/\Omega|$ is 0.5), then the system is in virial equilibrium (and nothing much happens). If the left-hand side is larger, thermal energy exceeds the gravitational potential energy, and the cloud disperses. If the right-hand side is greater, gravity wins, and the cloud collapses under gravity. Thus our condition for collapse is therefore

$$3NkT < \frac{3}{5}\frac{GM^2}{R}. \tag{15.1}$$

For mean density $\overline{\rho}$ and cloud mass M, the corresponding characteristic length scale is

$$R = \left(\frac{3M}{4\pi\overline{\rho}}\right)^{1/3}$$

(smaller clouds collapse) and, substituting $N = M/(\mu m(\mathrm{H}))$, eqtn. (15.1) can be written as

$$M > \left(\frac{5kT}{G\mu m(\mathrm{H})}\right)^{3/2}\left(\frac{3}{4\pi\overline{\rho}}\right)^{1/2}, \equiv M_\mathrm{J}. \tag{15.2}$$

This is the constraint on cloud mass required for collapse. This sort of argument was first put forward by Sir James Jeans, and the critical mass $M_\mathrm{J}$is therefore known as the Jeans mass, M_J. We see that $M_\mathrm{J} \propto T^{3/2}\rho^{-1/2}$; hence the easiest clouds to collapse are cold, dense ones (typically, dense molecular clouds).

[Jeans himself derived this limiting mass by supposing that a cloud collapses if the sound-crossing time is greater than the free-fall time. Jeans' original argument was actually flawed, but his general results still provides a useful rule of thumb indicating whether or not a given system is liable to collapse.]

Sizes, timescales, luminosities

We can also arrange eqtn. (15.1) to solve for the corresponding radius,

$$R_{\rm J} = \left(\frac{15}{4\pi} \frac{kT}{G\mu m({\rm H})\overline{\rho}} \right)^{1/2}$$

where $R_{\rm J}$ is called the Jeans length – the radius of a uniform-density cloud that contains the Jeans mass. At the specified density and temperature, larger (i.e., more massive) clouds will collapse.

From the Virial theorem, half the change in gravitational potential energy is radiated away. A cloud in free-fall is not virialized, but if completely optically thin will radiate away *all* the released gravitational potential energy. In any event, we can get an order-of-magnitude handle on the luminosity of the collapsing cloud as

$$\frac{1}{2}\frac{{\rm d}\Omega}{{\rm d}t}, = \frac{1}{2}\frac{{\rm d}\Omega}{{\rm d}R}\frac{{\rm d}R}{{\rm d}t}, \simeq \frac{GM_{\rm J}^2}{R_{\rm J}^2}\frac{{\rm d}R}{{\rm d}t}$$

(from $\Omega \simeq GM^2/R$). To get a crude estimate of ${\rm d}R/{\rm d}t$, we just divide the Jeans length by the free-fall time (which is $(G\rho)^{-1/2}$; eqtn. 11.2):

$$\frac{{\rm d}R}{{\rm d}t} \simeq R_{\rm J}\sqrt{G\rho},$$

whence the luminosity is, roughly,

$$L \simeq \frac{GM_{\rm J}^2}{R_{\rm J}}(G\rho)^{1/2}$$

Using these results, we find that a Jeans-mass molecular cloud with $n \simeq 10^4$ m^{-3}, $T \simeq 20$ K, $\mu \simeq 2$ has
$M_{\rm J} \simeq 20$ M$_\odot$,
$R_{\rm J} \simeq 0.2$ pc ($\sim 10^7$ R$_\odot$),
$\tau_{\rm ff} \simeq 10^6$ yr (independent of temperature), and
$L \simeq 10^{-3}$ L$_\odot$ (independent of density).
Such objects are cold and faint, and may be seen as dark clouds ('Bok globules', named for Bart Bok, a Dutch-American astronomer who conducted early investigations in this field).

15.3 Protostars: initial phases

15.3.1 Implications for star formation: fragmentation

An astrophysically interesting implication of our Jeans-mass analysis follows from considering changing circumstances during collapse. We've just seen that

$$M_{\rm J} \propto T^{3/2}\rho^{-1/2}.$$

As the cloud collapses, then ρ must increase (same mass in smaller volume); and we'd expect the temperature to increase (partial conversion of gravitational potential energy to thermal energy).

The Jeans mass will therefore change during collapse. If it were to *decrease,* then any substructure within the region with mass in excess of the new M_J will start its own distinct collapse within the larger region – that is, the cloud will fragment whenever $T^{3/2}\rho^{-1/2}$ decreases.[1]

During the initial phases of collapse, conditions will not be too far from isothermal – rising temperatures promotes molecular hydrogen into excited rotational levels. Subsequent de-excitation results in the emission of photons, mainly at long wavelengths – IR and mm – where cloud is transparent. The escaping photons carry away energy, cooling the cloud. Since $T \sim$ constant, the increase in density results in fragmentation.

Once the (core of) the cloud becomes optically thick, conditions switch from roughly isothermal to roughly adiabatic. The increasing gas pressure – together with turbulence – stops collapse, and fragmentation.

The assumption of adiabaticity, together with the perfect-gas and polytropic equations, lead to $T \propto \rho^{(\gamma-1)}$. Fragmentation requires decreasing $T^{3/2}\rho^{-1/2}$, corresponding to an adiabatic index $\gamma < 4/3$. If the collapsing gas cloud can radiate away enough thermal energy to remain approximately isothermal, $\gamma = 1$, and we expect fragmentation to occur; at later phases, the increasing density will result in greater optical depths, and γ will increase.

Fragmentation is (one reason) why giant molecular clouds don't collapse to single supermassive stars; but the physics poses some problems, too. Naïve considerations based on the foregoing arguments suggest that fragmentation should set an upper limit on stellar masses of order 10 $M_\odot$, but we know that in reality stars form to much higher masses ($\sim 10^2$ $M_\odot$). The origin of high-mass stars is therefore a major focus of research into star formation.

What actually causes a molecular cloud to collapse? Observationally, star formation is observed to be triggered by density waves in spiral galaxies, or sequentially by (presumably) supernova blast waves.

And do be aware that collapse is by no means isotropic; even if we start with uniform density, the inevitable consequence of collapse is that the central regions quickly become denser.

Conservation of angular momentum means that collapse will be favoured along the 'rotation' axis (magnetic fields may also have an influence). Hence the fragment flattens as it collapses, giving rise to a disk-like configuration with a well-defined, dense central condensation: our protostar.

The accepted picture is that the disk may accrete further material from the surrounding molecular cloud, and feeds material onto the protostar.

[1] Hoyle 1953 ApJ 118 513. This is a considerable simplification of the relevant physics; a more complete analysis would include the effects of turbulence and magnetic fields.

Start with prestellar core, ~ 0.1 pc, $n \sim 10^5$ cm^{-3}, $T \sim 10$K. Not always easy to distinguish true prestellar cores from transient density enhancements. 'Larson's laws' (1969–1981):

1. Velocity dispersion (line width) is proportional to cloud size: $\sigma \propto r^{0.4}$. This is the signature expected of *turbulence* (and is similar to the Kolmogoroff Law for turbulence of small eddies super-imposed on larger ones).
2. Velocity dispersion is proportional to cloud mass: $\sigma \propto M^{0.2}$, and
3. Cloud size is inversely proportional to density: $n \propto r^{-1.1}$.

At densities of $n \sim 10^{12}$cm^{-3} ($\rho \sim 10^{-13}$g cm^{-3}) the core becomes optically thick to its own sub-mm emission, and switches from ~isothermal to ~adiabatic behaviour. The Jeans mass at this density defines a minimum mass, called the opacity limit for fragmentation (Low and Lynden-Bell 1976; Rees 1976); it corresponds to a mass of ~ 0.01M$_\odot$. Collapse now proceeds on a Kelvin-Helmholtz timescale (fragmentation stops as the stars starts to heat).

Protostellar objects classified according to SEDs (related to long-wavelength excess over ~b-b radiation, resulting from reprocessing). Relative lifetimes from relative frequencies.

Class 0: central object still has much of its final mass to accrete; surrounding envelope still substantial. $T \lesssim 70$K, cold bb flux distribution peaking at $\sim 100\mu$m, $\tau \lesssim 10^5$ yr.

Class I: central object nearly complete (more mass in protostar than surrounding accretion zone), envelope settling onto disk; large IR excess. $T \lesssim 650$K, peaking at ~10s of μm, $\tau \sim$ few10^5 yr.

Class II: $T \lesssim 2800$K, emission peaks in $\sim \mu$m band, with a large IR excess (from disk). Envelope fully is more or less settled onto the disk. At ~2000K, moilecular hydorgen dissociates. Evolution becomes almost isothermal again, as most of the released potential energy goes into dissociating H_2 (rather than heating the gas). As a result of this 'second collapse' phase, the protostar collapses rapidly, to $\sim$ stellar densities. Now categorised as a 'pre-main-sequence' (PMS) object. The result is a classical T-Tauri star (CTT; or their more massive counterparts, Herbig Ae/Be stars), $\tau \sim$ few10^6 yr. A large fraction of CTTs show (protoplanetary) accretion discs.

Class III: modest IR excess, weak-lined T-Tauri (WLTT) stars, , $\tau \sim$ few10^7 yr.

Returning to the simple picture, toward the end of this contraction phase, central temperatures reach $\sim 10^4$ K, and hydrogen begins to ionize. This is a phase change, and during phase changes, temperature remains constant. As a result, further contraction of the cloud rapidly increases its density and pressure, and the cloud goes into a rapidly accelerating free fall which reduces its size by a factor of ten or more, hence decreasing its volume by a thousand or more times, and increasing its density by a thousand or more times, on a timescale of decades. Within those few years, the clouds changes from a large, cool

cloud of gas, radiating only infrared radiation, to a much smaller, denser, hotter cloud of ionized gas, radiating much larger amounts of infrared and visible light – a protostar.

The Virial Theorem tells us how a virialized system responds to changing conditions, namely

$$\Delta U = -\frac{\Delta\Omega}{2}.$$

Thus as a gas cloud contracts it gets hotter (U becomes more positive as Ω becomes more negative; that is, as the system becomes more tightly bound). Only half the change in Ω has been accounted for; the remaining energy is 'lost' – in the form of radiation. This occurs on a Kelvin–Helmholtz timescale.

As the temperature rises, first molecules (principally H_2) dissociate, then ionization of hydrogen and helium occurs. Eventually, hydrostatic equilibrium is established as a result of rising pressure, and the collapsing gaseous condensation has become a protostar.

We can therefore roughly estimate the properties of the protostar by supposing that all the available gravitational potential energy initially released in collapse (from infinity to some protostellar radius $R_{\rm ps}$) is used in dissociation and ionization; that is, that

$$\frac{GM^2}{R_{\rm ps}} \simeq \frac{M}{m({\rm H})}\left(\frac{X}{2}\cdot\chi({\rm H}_2) + X\cdot\chi({\rm H}) + \frac{Y}{4}\cdot\chi({\rm He})\right) \tag{15.3}$$

(neglecting some factors, such as μ, of order unity), where
$X, Y(\simeq 1-X)$ are the abundances by mass of hydrogen and helium,
$\chi({\rm H}_2)$ is the dissociation energy of molecular hydrogen (4.5 eV), and
$\chi({\rm H}), \chi({\rm He})$ are the ionization potentials of hydrogen and helium (13.6 eV, and 24.6+54.4 eV), giving

$$\begin{aligned}\frac{R_{\rm ps}}{{\rm R}_\odot} &\simeq \frac{40}{(1-0.2X)}\frac{M}{{\rm M}_\odot} \\ &\sim 50 \text{ for a solar-mass star.}\end{aligned} \tag{15.4}$$

In reality, additional energy is lost through other mechanisms (radiation, outflows etc.), and this 'back of the envelope' calculation significantly overestimates $R_{\rm ps}$. Nevertheless, it provides us with a simple, if crude, estimate of the maximum radius for a protostar as it begins its evolution. We can also estimate the average internal temperature, from the Virial theorem,

$$\begin{aligned}\overline{T} &\simeq \frac{\mu m({\rm H})}{3k}\frac{GM}{R_{\rm ps}} \\ &\sim 10^5 \text{ K.}\end{aligned}$$

Even though the core is hotter than this mean value, temperatures are not high enough, at this stage, to ignite hydrogen fusion.

The opacity is high at this stage (largely due to H^-), as is the luminosity. As a consequence, the system is effectively fully convective, and can therefore be well approximated by a polytrope with $n = 1.5$ (Section 14.4.1).

15.4 Protostellar evolution: Hayashi tracks

The relevant behaviour of convective protostars in this phase was investigated by the Japanese astronomer Chushiro Hayashi. We investigate this behaviour through a polytropic model.

15.4.1 Interior properties

For a polytrope of index n, obeying the perfect gas equation,

$$P = K\rho^{(n+1)/n}, \tag{14.1}$$

$$= \frac{\rho}{\mu m(\mathrm{H})} kT \tag{2.4}$$

if gas pressure dominates (a safe assumption). Eliminating ρ between eqtns. (14.1) and (2.4) we obtain

$$P = K^{-n}\left(\frac{k}{\mu m(\mathrm{H})}\right)^{(1+n)} T^{(1+n)};$$

that is,

$$P = C_1 T^{(1+n)} \tag{15.5}$$

where

$$C_1 = K^{-n}\left(\frac{k}{\mu m(\mathrm{H})}\right)^{(1+n)}$$

is a constant for a given model.

From our mass-radius relation for polytropes,

$$R^{(3-n)} M^{(n-1)} \propto K^n, \tag{14.12}$$

we have, for $n = 1.5$,

$$K \propto M^{1/3} R$$

whence

$$\begin{aligned} C_1 &\propto \left(M^{1/3} R\right)^{-n}, \\ &\propto M^{-1/2} R^{-3/2} \end{aligned}$$

for $n = 1.5$. So, finally, from eqtn. (15.5),

$$P = C_2 M^{-1/2} R^{-3/2} T^{5/2} \tag{15.6}$$

for $n = 1.5$, where C_2 is a constant (for given mean molecular weight).

15.4.2 Boundary condition

To solve for C_2, the constant of proportionality in eqtn. (15.6), we consider an outer boundary condition – the photosphere. The photospheric radius is defined by optical depth $\tau = 2/3$, measured radially inwards (§8.3.2); that is, since

$$\tau(R) \equiv \frac{2}{3} = \int_R^\infty \kappa(r)\,\rho(r)\,\mathrm{d}r,$$

or, assuming constant opacity in the atmosphere,

$$\frac{2}{3} = \kappa \int_R^\infty \rho(r)\,\mathrm{d}r.$$

From hydrostatic equilibrium,

$$\begin{aligned} P(R) &= \int_R^\infty g(r)\,\rho(r)\,\mathrm{d}r \\ &\simeq \frac{GM}{R^2}\int_R^\infty \rho(r)\,\mathrm{d}r \\ &= \frac{GM}{R^2}\frac{2}{3\kappa}. \end{aligned}$$

We've assumed constant κ in the atmosphere of a given protostar, but the opacity will vary with time for any one object, and between different objects, according to their individual photospheric pressures and temperatures. We adopt the usual power-law dependence of opacity,

$$\kappa = \kappa_0 P^p T^q, \tag{5.5}$$

so

$$P(R) = \frac{GM}{R^2}\frac{2}{3\kappa_0 P^p T^q}.$$

This has P on both sides; rearranging,

$$P(R) = \left(\frac{GM}{R^2}\frac{2}{3\kappa_0}T_{\rm eff}^{-q}\right)^{1/(1+p)} \tag{15.7}$$

(where $T \equiv T_{\rm eff}$ at the photosphere).

15.4.3 Solution

At the photosphere, both eqtn. (15.6) and (15.7) are true; i.e.,

$$\left(\frac{GM}{R^2}\frac{2}{3\kappa_0}T_{\rm eff}^{-q}\right)^{1/(1+p)} = C_2 M^{-1/2} R^{-3/2} T_{\rm eff}^{5/2}.$$

So, for any given mass, there is a single-valued relationship between R and $T_{\rm eff}$; but $L = 4\pi R^2 \sigma T_{\rm eff}^4$, so this is equivalent to a single-valued relationship between $T_{\rm eff}$ and L – that is, a track in the HR diagram. Such tracks are called Hayashi tracks. We can see that the physical quantities M, $T_{\rm eff}$, and R are all represented to some power, so on taking logs it must be that

$$\ln T_{\rm eff} = A \ln L + B \ln M + \text{constant} \tag{15.8}$$

(where use has been made of $L \propto R^2 T_{\rm eff}^4$); some lengthy but straightforward algebra shows that

$$A = \frac{0.75p - 0.25}{5.5p + q + 1.5}, \qquad B = \frac{0.5p + 1.5}{5.5p + q + 1.5}.$$

Opacity calculations indicate $p \simeq 1$, $q \simeq 3$ whence $A \simeq 0.05$, $B \simeq 0.2$. From eqtn. (15.8),

$$\frac{\partial \ln L}{\partial \ln T_{\rm eff}} = 1/A;$$

since A is small, the track must be very steep (i.e., nearly constant temperature) in the HR diagram.[2]

We also see that

$$B = \frac{\partial \ln T_{\rm eff}}{\partial \ln M}$$

is positive, so that the tracks move slightly to the left (i.e., higher temperature) with increasing mass; but the dependence is weak, so all fully-convective stars lie on almost the same 'Hayashi track' (or Hayashi line). For a given mass and chemical composition, no fully convective star can lie to the right of the Hayashi track (because convection is the most efficient available means of energy transport). The region to the right of the Hayashi track is the Hayashi 'forbidden zone' ($T_{\rm eff} \lesssim 4$ kK).

15.5 Protostellar evolution: Henyey tracks

As the contracting protostar descends the Hayashi track the internal temperature continues to rise until ionization is complete, and the opacity drops in the core. Other than for the lowest-mass stars ($M \lesssim 0.5 \mathrm{M}_\odot$), which remain fully convective right onto the zero-age main sequence (ZAMS), this fall in opacity allows energy to be transported radiatively in the interior. The star then moves away from the Hayashi track, to higher $T_{\rm eff}$, following 'Henyey' track.

Consider the equation of radiative energy transport

$$L(r) \propto \frac{r^2}{\bar{\kappa}_{\rm R}\rho} \frac{{\rm d}T}{{\rm d}r} T^3; \tag{12.5}$$

[2] More-detailed calculations give small negative values for A; that is the temperature increases (slightly) with decreasing L.

reviewing terms on the right-hand side, we have

$$\rho \propto \frac{M}{R^3},$$
$$\frac{\mathrm{d}T}{\mathrm{d}r} \simeq \frac{T_\mathrm{c}}{R},$$

and we adopt

$$\bar{\kappa}_\mathrm{R} \propto \rho T^{-3.5}$$

(Kramers' opacity). Finally, the core temperature scales as

$$T_\mathrm{c} \propto \frac{M}{R}. \tag{11.8}$$

Inserting these results into eqtn. (12.5) yields

$$L \propto M^{5.5} R^{-0.5}$$

and, since $L \propto R^2 T_\mathrm{eff}^4$,

$$T_\mathrm{eff} \propto R^{-5/8} M^{11/8}. \tag{15.9}$$

That is, for given mass M, the luminosity and temperature increase as the star shrinks, with $L \propto R^{-0.5}$, $T_\mathrm{eff} \propto R^{-5/8}$. Although the power-law dependences of L and T_eff on R are numerically similar (whence $L \propto T_\mathrm{eff}^{4/5}$), in a typical HR diagram the temperature axis ranges over about one order of magnitude, while the luminosity axis may vary over five or more orders of magnitude; the result is that Henyey tracks appear as more or less horizontal features (Fig. 15.1).

15.6 Protostar to star

Until fusion ignites, the relevant timescale is the Kelvin–Helmholtz timescale (since the star is radiating gravitational potential energy). This timescale is short, and the contracting protostar is still shrouded in the dusty molecular cloud from which it formed. The increasing core temperature eventually results in ignition of hydrogen burning; the protostar is now a star, on the zero-age main sequence (ZAMS).

Of course, the true circumstances are more complex in detail than the simple Hayashi picture; protostars show circumstellar accretion disks, and outflows such as jets and stellar winds. Magnetic fields also play a role. Material falling onto the protostar generates accretion luminosity, potentially through shocks. This extra energy loss results in protostellar radii smaller than simple estimates (eqtn. 15.4). Furthermore, the most massive stars may ignite hydrogen burning at the core while still accreting at the surface.

Stars more massive than $\sim 5\mathrm{M}_\odot$ become stable against convection very quickly; most of their pre-main-sequence evolution is on the Henyey branch, while stars with $M \lesssim 0.5\mathrm{M}_\odot$ never become stable against convection, and evolve vertically onto the MS as fully convective stars.

comes a larger fraction of the total pressure and electron scattering replaces free-free and bound-free absorption as the main contributor to interior opacity; the adiabatic temperature gradient in the convective core and the radiative gradient outside of this core approach one another more closely so that the changes in structure accompanying variations in the core mass become less radical.

As in the case of $M = 1.5\ M_\odot$, the point at which luminosity attains its final minimum, before increasing during the main hydrogen-burning stages, may be defined as the

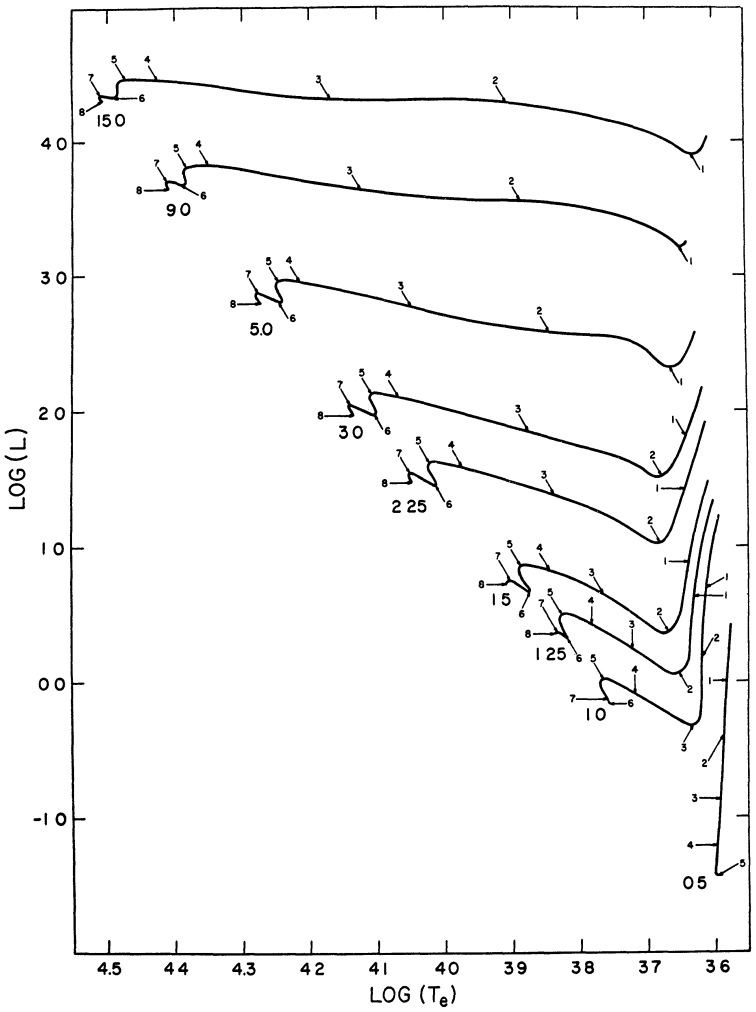


FIG. 17.—Paths in the Hertzsprung-Russell diagram for models of mass ($M/M_\odot$) = 0.5, 1.0, 1.25, 1.5, 2.25, 3 0, 5.0, 9.0, and 15.0. Units of luminosity and surface temperature are the same as those in Fig. 1

Figure 15.1: The approach to the main sequence for (proto)stars of different masses. From Iben, ApJ, 141, 993 (1965).

Box 15.1: Gravitational potential energy of a uniform-density sphere

Consider a spherical shell of thickness dr at radius r within a sphere of density ρ, radius R. The shell mass is

$$\mathrm{d}m = 4\pi r^2 \rho \, \mathrm{d}r$$

and the mass within the shell is

$$M(r) = \frac{4}{3}\pi r^3 \rho$$

so the gravitational potential energy of the shell is

$$\begin{aligned} \mathrm{d}\Omega(r) &= -G\frac{M(r)\,\mathrm{d}m}{r} \\ &= -\frac{16}{3}G\pi^2\rho^2 r^4 \mathrm{d}r. \end{aligned}$$

The total PE is therefore

$$\begin{aligned} \Omega &= \int_0^R \mathrm{d}\Omega(r) \\ &= -\frac{16}{3}G\pi^2\rho^2 \int_0^R r^4 \mathrm{d}r \\ &= -\frac{16}{15}G\pi^2\rho^2 R^5 \end{aligned}$$

but $M = (4/3)\pi R^3 \rho$, so

$$\Omega = -\frac{3}{5}\frac{GM^2}{R}$$

Section 16

The main sequence: homologous models

We know that stars on the main sequence share many basic characteristics, as a consequence of their common energy source. We also know that many properties vary along the main sequence, i.e., vary with mass. Both aspects can be accommodated by supposing that stars along the main sequence are essentially just scaled version of each other. This will give us insight into (e.g.) main-sequence mass–luminosity relationships.

16.1 Transformed structure equations

The basic equations of stellar structure are normally cast in such a way as to describe the run of physical properties with radius; but mass is the more fundamental physical property of a star (the radius of a solar-mass star will change by orders of magnitude over its lifetime, while its mass remains more or less constant), so for practical purposes it is customary to reformulate these structure equations in terms of mass as the independent variable. We start by simply inverting eqtn. (2.5), the equation of mass continuity, giving

$$\frac{\mathrm{d}r}{\mathrm{d}m(r)} = \frac{1}{4\pi r^2 \rho(r)}; \tag{16.1}$$

and the remaining basic structure equations are just multiplied by eqtn. (16.1), giving

$$\frac{\mathrm{d}P(r)}{\mathrm{d}m(r)} = \frac{-Gm(r)}{4\pi r^4} \qquad \text{(hydrostatic equilibrium; eqtn. 2.2)} \tag{16.2}$$

$$\frac{\mathrm{d}L(r)}{\mathrm{d}m(r)} = \varepsilon(r), \quad \simeq \varepsilon_0 \, \rho(r) \, T^{\beta}(r) \qquad \text{(energy continuity; eqtn. 2.7)} \tag{16.3}$$

$$\frac{\mathrm{d}T(r)}{\mathrm{d}m(r)} = -\frac{3\overline{\kappa}_{\mathrm{R}} L(r)}{64\pi^2 r^4 acT^3(r)} \qquad \text{(radiative energy transport; eqtn. 12.6)} \tag{16.4}$$

(where all the radial dependences have been shown explicitly; we could, equivalently, use the dependence on mass $m(r)$).

16.2 Homologous models

Homologous stellar models are defined such that their properties scale in the same way with fractional mass $\mathtt{x} \equiv m(r)/M$. That is, for some property X (which might be temperature, or density, etc.), a plot of $X(\mathtt{x})$ vs. $\mathtt{x}$ is the same for all homologous models. Such models may reasonably be used to investigate zero-age main-sequence stars which have uniform chemical composition and which are in thermal, hydrostatic, and radiative equilibrium.

Without any loss of generality, we can write the actual variables r, ρ, etc., as the product of some function of relative mass, and some reference value of the variable; i.e.,

$$r(\mathtt{x}) = f_r(\mathtt{x})\, R_\# \qquad \rho(\mathtt{x}) = f_\rho(\mathtt{x})\, \rho_\#$$
$$P(\mathtt{x}) = f_P(\mathtt{x})\, P_\# \qquad T(\mathtt{x}) = f_T(\mathtt{x})\, T_\#$$
$$L(\mathtt{x}) = f_L(\mathtt{x})\, L_\#$$

The 'reference values' will have the appropriate dimensions; e.g., $R_\#$ must have units of length. So, for example, $R_\#$ *might* be the stellar radius – but not necessarily. It *could* be that, say, $R_\# = 2R_*$ (in which case $f_r(1)$ would have the value 0.5, to ensure that $r(1) = R_*$, as required). Both the functions f and the reference values $X_\#$ are, as yet, unspecified in detail; our only supposition is that the functions are the same from star to star (while the reference values are different).

Now consider hydrostatic equilibrium in the mass-dependent form expressed in the preceding subsection:

$$\frac{\mathrm{d}P(r)}{\mathrm{d}m(r)} = \frac{-Gm(r)}{4\pi r^4}. \tag{16.2}$$

Since $\mathtt{x} \equiv m(r)/M$, we have $\mathrm{d}m = M\mathrm{d}\mathtt{x}$; and also $\mathrm{d}P(\mathtt{x}) = P_\#\, \mathrm{d}f_P(\mathtt{x})$ (from the definition of $P(\mathtt{x})$ just given). Hence

$$\begin{aligned}\frac{\mathrm{d}P(r)}{\mathrm{d}m(r)} \rightarrow \frac{P_\#}{M}\frac{\mathrm{d}f_P(\mathtt{x})}{\mathrm{d}\mathtt{x}}, &= \frac{-Gm(r)}{4\pi r^4},\\ &= \frac{-GM\mathtt{x}}{4\pi\, f_r^4(\mathtt{x})\, R_\#^4},\\ &= \frac{GM}{R_\#^4} \times \frac{-\mathtt{x}}{4\pi\, f_r^4(\mathtt{x})}.\end{aligned}$$

Here the terms are collected such that first group of quantities has dimensions while the second does not. It must therefore be the case that

$$\frac{\mathrm{d}f_P}{\mathrm{d}\mathtt{x}} = \frac{-\mathtt{x}}{4\pi f_r^4}, \qquad P_\# = \frac{GM^2}{R_\#^4}. \tag{16.5}$$

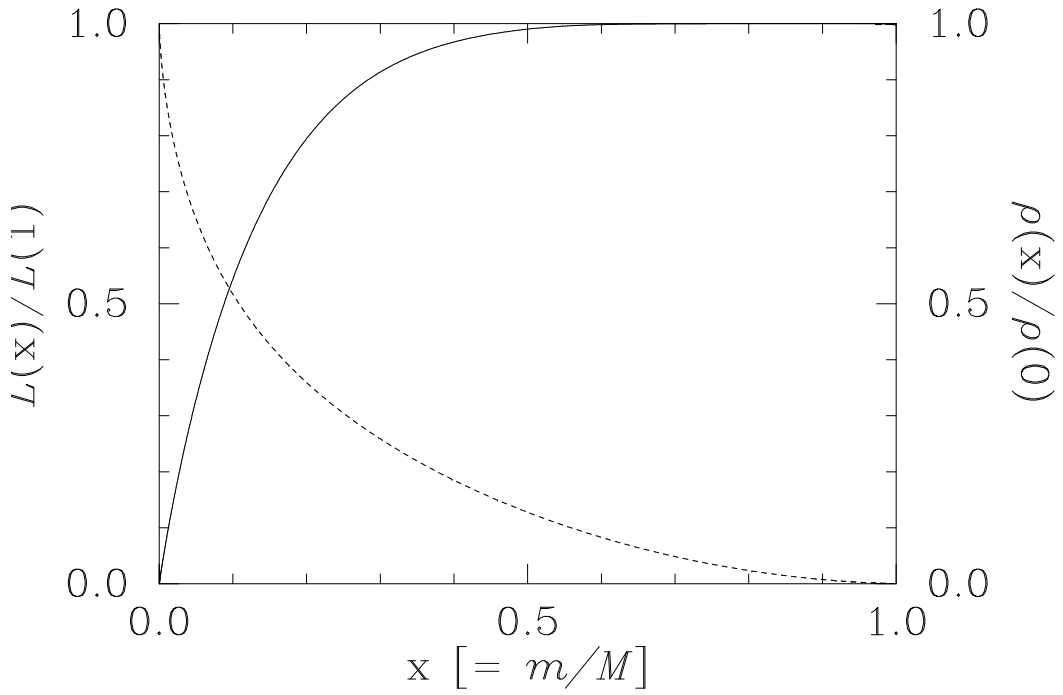


Figure 16.1: The variation of normalized luminosity (solid line) and normalised density (dashed line) as a function of normalized mass. The curves shown are actually from a specific, detailed model of the Sun; the principle of homology is, essentially, that these curves apply equally well to any other (main-sequence) star, regardless of its mass.

Repeating this reasoning for the other structure equations gives

$$\frac{\mathrm{d}f_r}{\mathrm{d}x} = \frac{1}{4\pi f_r^2 f_\rho}, \qquad \rho_\# = \frac{M}{R_\#^3}, \tag{16.6}$$

$$\frac{\mathrm{d}f_L}{\mathrm{d}x} = f_\rho^\alpha \, f_T^\beta \qquad L_\# = \varepsilon_0 \rho_\# \, T_\#^\beta \, M \tag{16.7}$$

$$\frac{\mathrm{d}f_T}{\mathrm{d}x} = \frac{3 f_L}{64\pi^2 \, f_r^4 f_T^3}, \qquad L_\# = \frac{ac}{\bar{\kappa}_\mathrm{R}} \frac{(T_\# R_\#)^4}{M}, \tag{16.8}$$

$$f_\rho = \frac{f_P}{f_T} \qquad T_\# = \frac{\mu m(\mathrm{H})}{k} \frac{P_\#}{\rho_\#} \tag{16.9}$$

where eqtns. (16.5)–(16.8) follow from eqtns. (16.2), (16.1), (16.3), and (16.4), respectively; eqtn. (16.9) follows from the perfect gas equation, $P_\mathrm{G} = (\rho k T)/(\mu m(\mathrm{H}))$ (eqtn. 2.4). The left-hand set are differential equations in the five functions f (which we will explore no further), and the right-hand set gives us five equations for five unknowns, for given mass (and 'known' β, ε_0, μ, $\bar{\kappa}_\mathrm{R}$). These equations can therefore be solved simultaneously to show how things like luminosity and radius vary with mass – for homologous stars in which the energy transport is radiative (eqtns. 16.4/16.8) and gas pressure dominates (eqtns. 2.4/16.9).

16.3 Results

16.3.1 Mass–luminosity relationship

Inserting eqtns. (16.5) and (16.6) into (16.9) gives us

$$T_\# = \frac{\mu m(\mathrm{H})}{k}\frac{GM}{R_\#}; \tag{16.10}$$

and using that in eqtn. (16.8) yields

$$L_\# = ac\left(\frac{Gm(\mathrm{H})}{k}\right)^4 \frac{\mu^4}{\bar{\kappa}_\mathrm{R}} M^3. \tag{16.11}$$

Recalling that $L(\mathtt{x}) = f_L(\mathtt{x})L_\#$, what this means is that the flux through some shell containing a fraction $\mathtt{x}$ of the total mass varies as (total) mass cubed; in particular, the surface luminosity is

$$L_* \propto \frac{\mu^4}{\bar{\kappa}_\mathrm{R}} M^3$$

where we can adopt a power-law form for the opacity, $\bar{\kappa}_\mathrm{R} \simeq \kappa_0 \rho^p T^q$ (eqtn. 5.5). For stars more massive than $\sim 2\mathrm{M}_\odot$, electron scattering dominates the opacity, and $p = q = 0$, whence

$$L \propto M^3$$

for given μ.[1]

For lower-mass stars Kramers' law ($p = 1$, $q = -3.5$) is a better approximation. Substituting $\rho \propto M/R^3$ and $T \propto \mu M/R$ (so $\bar{\kappa}_\mathrm{R} \propto \mu^{-3.5} M^{-2.5} R^{+0.5}$), and anticipating that $R \propto M^{0.4}$ for low-mass stars (§16.3.2; so $\bar{\kappa}_\mathrm{R} \propto \mu^{-3.5} M^{-2.3}$), we find

$$L \propto \mu^{7.5} M^5$$

These turn out to be really pretty good approximations to the (ZA)MS; only for $M \lesssim 0.7\mathrm{M}_\odot$ do the predictions start to break down, as the mass of the convective envelope becomes significant.

16.3.2 Mass–radius relationship

Eqtns. (16.7) and (16.11) give us

$$\begin{aligned}
\rho_\# &= ac\left(\frac{Gm(\mathrm{H})}{k}\right)^4 \frac{\mu^4}{\bar{\kappa}_\mathrm{R}} M^3 \frac{1}{\varepsilon_0 T_\#^\beta M} \\
&= \frac{ac}{\varepsilon_0}\left(\frac{Gm(\mathrm{H})}{k}\right)^4 \frac{\mu^4}{\bar{\kappa}_\mathrm{R}} \frac{M^2}{T_\#^\beta}, \quad = \frac{M}{R_\#^3} \quad \text{(from eqtn. 16.6), so} \\
M &\propto \frac{\bar{\kappa}_\mathrm{R}}{\mu^4}\frac{T_\#^\beta}{R_\#^3}.
\end{aligned}$$

[1] As stars burn hydrogen to helium, the global average value of μ increases, leading to an increase in L.

Substituting our previous result for $T_\#$ (eqtn. 16.10),

$$M \propto \frac{\overline{\kappa}_\mathrm{R}}{\mu^4}\frac{(\mu M)^\beta}{R_\#^{(\beta+3)}}, \quad \text{whence}$$
$$R_\# \propto \overline{\kappa}_\mathrm{R}^{1/(\beta+3)}\mu^{(\beta-4)/(\beta+3)}M^{(\beta-1)/(\beta+3)}$$

Low-mass, solar-type, stars burn hydrogen through the proton-proton chain, for which $\varepsilon \propto \rho T^4$ (i.e., $\beta \simeq 4$), whence

$$R \propto M^{0.43};$$

for stars more massive than $\sim 1.5\mathrm{M}_\odot$, the CNO cycle operates, with $\beta \simeq 18$ (depending somewhat on the core temperature), giving

$$R \propto M^{0.8}$$

16.3.3 Luminosity–temperature relationship

Neglecting $\mu, \overline{\kappa}_\mathrm{R}$ terms, then from §16.3.2

$$R \propto M^{(\beta-1)/(\beta+3)}$$

but also

$$L \propto R^2 T_\mathrm{eff}^4$$

whence

$$L \propto M^{2(\beta-1)/(\beta+3)}T_\mathrm{eff}^4.$$

We also have

$$L \propto M^3$$

(§16.11, again neglecting μ and $\overline{\kappa}_\mathrm{R}$ terms fort simplicity); eliminating M we obatin

$$T_\mathrm{eff}^4 \propto L^{1-\frac{2(\beta-1)}{3(\beta+3)}}.$$

The outcome is that we can predict the slopes of the main sequence in the H–R diagram; for $\beta = 4, 18$ we obtain

$$\log L = 5.6 \log T_\mathrm{eff} + C_4,$$
$$= 8.7 \log T_\mathrm{eff} + C_{16}$$

As before, these relations break down for low-mass stars, where the radiative-transport assumption breaks down; but otherwise, these are quite reasonable representation of reality.

Section 17

Stellar evolution

17.1 Mass limits for stars

We showed in Section 14.4 that the central conditions of a star (density, pressure, temperature) are readily estimated for a polytropic model, for a given polytropic index n. These allow us to estimate the minimum mass for which core hydrogen burning can take place ($T_c \simeq 10^6$ K); this turns out to be $\sim 0.1 M_\odot$.

There's also an upper limit to the possible mass of a star. Classically, this is determined by the Eddington limit. The inward force of gravity is proportional to mass, while the outward radiation-pressure force scales with luminosity, which increase with mass to the power ~3; thus there must be a point at which radiation overcomes gravity, and the star cannot be stable. This limit is generally thought to be at $\sim 150 M_\odot$. Only a few dozen stars are known that are more massive than $100 M_\odot$, although the most massive star has a reported mass of $\sim 200 M_\odot$ (with a ZAMS mass of perhaps $300 M_\odot$). Possibly this mass estimate is in error; but also, there are ways round the Eddington limit (e.g., inhomogeneous atmospheres that allow radiation to 'leak out' through lower-density routes).

17.2 Evolution on the main sequence

Within this allowed range of $\sim 10^{-1}$–10^2 $M_\odot$, a star, by our definition, spends most of its lifetime on the main sequence, burning hydrogen to helium in the core. However, even on the main sequence there are different evolutionary behaviours, depending on mass. In this context, it's convenient to divide the MS into the 'upper' main sequence ($M \gtrsim 2 M_\odot$, $T_{\rm eff} \gtrsim 10^4$ K) and 'lower' main sequence:

> Lower main sequence. Spectral types ~F and later. Energy generation is through the proton-proton chain, with radiative cores and convective envelopes (though stars less massive

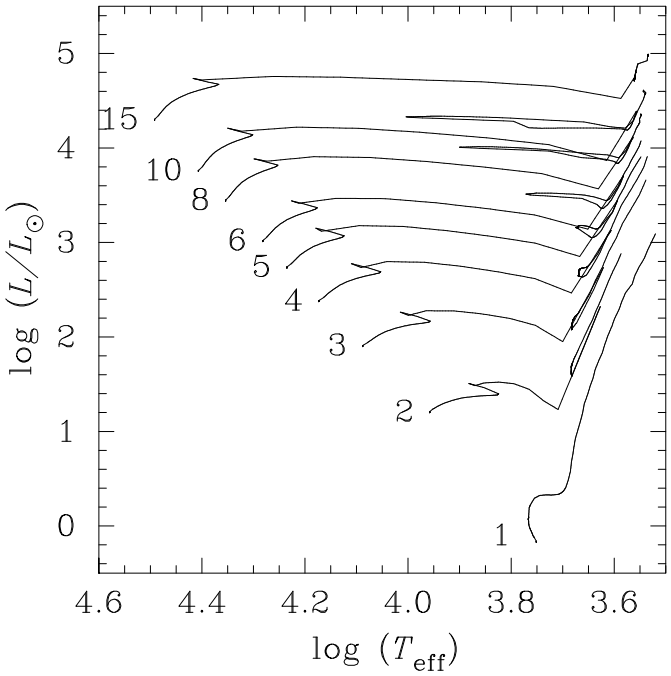


Figure 17.1: Evolutionary tracks for stars of different initial masses, starting at the ZAMS (based on calculations by the Geneva group). The tracks of intermediate-mass stars, with initial masses of ~2–10$M_\odot$, all show broadly similar characteristics.

than ~0.25$M_\odot$are fully convective). The extent of the envelope varies with mass; at 1$M_\odot$ the envelope accounts for ~3% by mass (~30% by radius), increasing to ~40% at 0.5$M_\odot$.

Upper main sequence. Energy generation is through the CNO cycle; the high energy fluxes result in convective cores, but lower envelope opacities result in radiative envelopes. The fractional core mass decreases with total mass, because of the strong dependence of energy generation rate on temperature (i.e., mass); ~0.17M at 3$M_\odot$, rising to ~0.38M at 15$M_\odot$.

[Not for lectures:]

Homology implies for two stars 1, 2

$$\frac{r_1(\mathbf{x})}{R_1} = \frac{r_2(\mathbf{x})}{R_2}, \tag{17.1}$$

$$\frac{m_1(\mathbf{x})}{M_1} = \frac{m_2(\mathbf{x})}{M_2} \tag{17.2}$$

whence we can obtain the density scaling from

$$\frac{\mathrm{d}m_2}{\mathrm{d}r_2} = \frac{\mathrm{d}m_1}{\mathrm{d}r_2}\frac{M_2}{M_1} \qquad \text{(from eqtn. 17.2)}$$

$$= \frac{\mathrm{d}m_1}{\mathrm{d}r_1}\frac{R_1}{R_2}\frac{M_2}{M_1} \qquad \text{(from eqtn. 17.1).} \tag{17.3}$$

We also have

$$\frac{\mathrm{d}m_2}{\mathrm{d}r_2} = 4\pi r_2^2 \rho_2 \qquad \text{(from mass continuity, eqtn. 2.5)}$$

$$= 4\pi r_1^2 \left(\frac{R_2}{R_1}\right)^2 \rho_2 \qquad \text{(from eqtn. 17.1),} \qquad (17.4)$$

so, from eqtns 17.3 and 17.4,

$$\frac{\mathrm{d}m_1}{\mathrm{d}r_1} = 4\pi r_1^2 \left(\frac{M_1}{M_2}\right)\left(\frac{R_2}{R_1}\right)^3 \rho_2,$$

$$= 4\pi r_1^2 \rho_1$$

or

$$\frac{\rho_2}{\rho_1} = \rho_1 \left(\frac{M_2}{M_1}\right)\left(\frac{R_1}{R_2}\right)^3 \qquad (17.5)$$

(where 'ρ_1' = $\rho_1(\mathbf{x})$', etc.). Mass continuity and hydrostatic equilibrium give

$$\frac{\mathrm{d}P}{\mathrm{d}m} = \frac{-Gm}{4\pi r^4} \qquad (16.2)$$

$$(17.6)$$

so

$$\frac{\mathrm{d}P_1}{\mathrm{d}m_1} \equiv \frac{\mathrm{d}P_2}{\mathrm{d}m_2}\frac{\mathrm{d}m_2}{\mathrm{d}m_1}\frac{\mathrm{d}P_1}{\mathrm{d}P_2}$$

$$= \frac{-Gm_2}{4\pi r_2^4}\left(\frac{M_2}{M_1}\right)\frac{\mathrm{d}P_1}{\mathrm{d}P_2}$$

$$= \frac{-Gm_1}{4\pi r_1^4}; \qquad \text{that is,}$$

$$\frac{\mathrm{d}P_2}{\mathrm{d}P_1} = \left(\frac{M_2}{M_1}\right)\left(\frac{m_2}{m_1}\right)\left(\frac{r_1}{r_2}\right)^4$$

$$= \left(\frac{M_2}{M_1}\right)^2 \left(\frac{R_1}{R_2}\right)^4.$$

Integrating (and taking the surface pressure to be zero),

$$\frac{P_2}{P_1} = \left(\frac{M_2}{M_1}\right)^2 \left(\frac{R_1}{R_2}\right)^4. \qquad (17.7)$$

Finally, for an equation of state

$$P_1 = (\rho_1 k T_1)/(\mu_1 m(\mathrm{H})) \qquad (2.4)$$

we find, from eqtns. (17.5) and (17.7),

$$\frac{T_2}{T_1} = \frac{\mu_2}{\mu_1}\frac{M_2}{M_1}\frac{R_1}{R_2} \qquad (17.8)$$

(provided that μ_1/μ_2 is independent of radius).

Evolution on the main sequence is driven by the slow (nuclear-timescale) conversion of hydrogen to helium in the core, with a consequent increase in mean molecular weight. To maintain the pressure (i.e., to maintain hydrostatic equilibrium) the core contracts, causing the core temperature to rise (eqtn. 17.8). Because of the temperature dependence of energy generation, this results in an increase in luminosity, and the star moves upwards in the HRD. (Now half-way through its main-sequence lifetime, the Sun is currently about 30% more luminous than it was on the zero-age main sequence.) The shrinking core is also accompanied by expansion of the envelope, and a drop in $T_{\rm eff}$.

Precise details of main-sequence evolution depend on the effectiveness of mixing in the core (corresponding to our division between 'low-mass' and 'high-mass' stars). This can be seen in the different evolutionary tracks shown in Fig. 17.1; for stars on the lower main sequence ($M \lesssim 2{\rm M}_\odot$), there is little or no mixing in the radiative core, and the mean molecular weight in the centre builds up relatively quickly. Energy production becomes concentrated in a 'thick shell' around the centre. Consequently, there is a gradual transition to subsequent evolutionary phases where fusion occurs in a shell around a core consisting of helium (and heavier elements).

More massive stars have fully mixed (convective) cores, so exhaust hydrogen at roughly the same time throughout the core. In the final stages of core hydrogen fusion the entire star contracts to try to maintain energy generation, increasing $T_{\rm eff}$ and producing a short-lived 'blue hook' in the HRD. Again, as core hydrogen is exhausted, the star establishes a hydrogen-burning shell around the core.

The main-sequence lifetime is

$$\tau \propto M/L,$$

or $\tau \propto M^{(1-x_5)}$ for our homologous models (i.e., $\tau \propto M^{-2}$ for upper main sequence stars); hence more massive stars have shorter lifetimes than less mass ones. The main-sequence lifetime of the Sun is $\sim 10^{10}$ yr, while that of a massive star is much less ($\sim 10^7$ yr for $\sim 20{\rm M}_\odot$).

The evolution of the most massive (O- and B-type) stars are also affected by mass loss through stellar winds. A $25{\rm M}_\odot$ star with a main-sequence lifetime of 6×10^6 yr losing mass at a rate of $\sim 10^{-6}$ ${\rm M}_\odot$ yr^{-1} will lose a quarter of its ZAMS mass in this way. The paradoxical consequence of losing mass is that the star's main-sequence lifetime is extended.

17.3 Mass dependence of subsequent evolution

The discussion of subsequent evolution can be structured in several ways – for example, according to the minimum mass required for a star to form its first degenerate core. We can therefore elaborate our simple 'high-mass/loss-mass' division to discuss this subsequent evolution.

First degenerate cores as a function of mass

Mass range ($M_\odot$)	Category	First degenerate core
≤2	Low	He
2–8	Intermediate	C/O
8–11	High	O/Ne/Mg
≥11	High	–

- Low-mass stars develop a degenerate helium core while ascending the red-giant branch (discussed below; Section 17.4). Core contraction is slow, leading to a slow ascent of the RGB. Core helium ignition takes place explosively in a thermonuclear runway; in this 'helium flash' enormous energy ($\sim 10^{11} L_\odot$!) is generated for a few seconds, lifting the degeneracy; the core expands and helium burning becomes stable.

  Helium burns to a C/O core, which never gets hot enough to ignite. The star ends its life as a C/O white dwarf.

- Intermediate-mass stars are discussed in detail below. They have core temperatures high enough to ignite helium without going through a degenerate phase. The star again ends as a white dwarf of degenerate C/O.

- High-mass stars are hot enough to ignite core carbon burning before developing O/Ne/Mg degenerate cores. For initial masses in excess of $\sim 11 M_\odot$ subsequent stages of nuclear fusion can occur, all the way to Fe; these stars end their lives as core-collapse supernovae.

17.4 Evolution off the main sequence: intermediate-mass stars

As an example of post-main-sequence evolution, we'll discuss in detail intermediate-mass stars, which illustrate a number of features of interest (specifically, a $4 M_\odot$ star; Fig. 17.2).

17.4.1 Red-giant branch: shell hydrogen burning

When core hydrogen is exhausted (the 'terminal age main sequence', or TAMS; point `B` in Fig. 17.2) the star has, essentially, an inert helium core surrounded by a hydrogen-burning shell. The core mass increases (gaining mass from the shell), and contracts under gravity. As the core contracts, the response of the envelope is to expand: the 'mirror principle' (or 'shell-burning law').

All numerical stellar-evolution models predict this transition, and yet we lack a simple, didactic physical explanation:

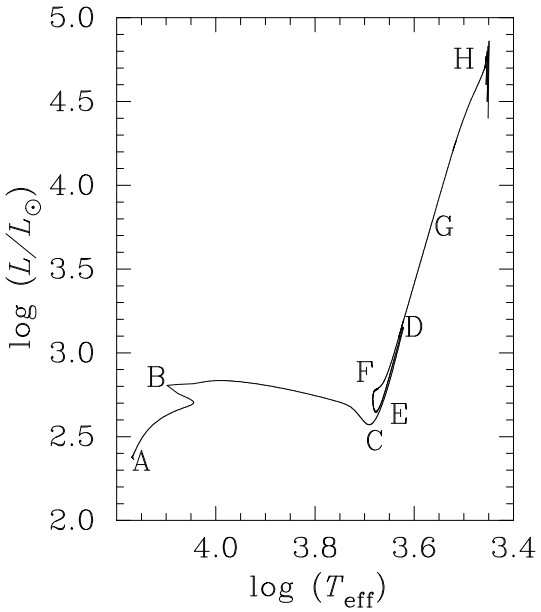


Figure 17.2: Evolutionary track for a 4-$M_\odot$ star (calculated using `EZ-Web`, schematically extended to include the thermal-pulsing AGB at `H`).

"Why do some stars evolve into red giants though some do not? This is a classic question that we consider to have been answered only unsatisfactorily." – *D. Sugimoto & M. Fujimoto (ApJ, 538, 837, 2000)*

Evolutionary timescales for a 4-$M_\odot$ star

(based on Geneva models)

Phase	Fig. 17.2	t (yr)	Δt	Energy source
Main sequence	A–hook	1.62+8	–	Core H burning
	hook–B	1.65+8	5.30+6	
Hertzsprung gap	B–C	1.66+8	0.97+6	Shell H burning
RGB	C–D	1.66+8	0.66+6	
CHeB	D–E	1.74+8	8.11+6	Core He + shell H
CHeB	E–F	1.92+8	1.77+7	
AGB	F–G	1.94+8	1.78+6	Shell He + pulsing

Mirror principle

Whenever a star has an active shell-burning source, the shell acts as a 'mirror' between the core and envelope; core contraction leads to envelope expansion, and vice versa.

To maintain thermal equilibrium, the burning shell must remain at approximately constant temperature due to the thermostatic action of nuclear burning. Contraction of

the burning shell would entail heating, so the burning shell must also remain at roughly constant radius. As the core contracts, the shell density, and therefore pressure, must decrease; hence the pressure of the overlying envelope must decrease – so the layers above the shell must expand.

We simplify the stellar structure into an inner core and an outer envelope, with masses and radii M_c, M_e and $R_c, R_e (= R_*)$ respectively. At the end of core hydrogen burning, we suppose that core contraction happens quickly – faster than the Kelvin–Helmholtz timescale, so that the virial theorem holds; and thermal and gravitational potential energy are conserved to a satisfactory degree of approximation. We formalize this supposition by writing

$$\Omega + 2U = \text{constant} \qquad \text{(Virial theorem)}$$

$$\Omega + U = \text{constant} \qquad \text{(Energy conservation)}$$

These two equalities can only hold simultaneously if both Ω and U are individually constant, summed over the whole star. In particular, the total gravitational potential energy is constant

Stars are centrally condensed, so we make the approximation that $M_c \gg M_e$; then, adding the core and envelope,

$$|\Omega| \simeq \frac{GM_c^2}{R_c} + \frac{GM_cM_e}{R_*}$$

We are interested in evolution, so we take the derivative with respect to time,

$$\frac{\mathrm{d}|\Omega|}{\mathrm{d}t} = 0 = -\frac{GM_c^2}{R_c^2}\frac{\mathrm{d}R_c}{\mathrm{d}t} - \frac{GM_cM_e}{R_*^2}\frac{\mathrm{d}R_*}{\mathrm{d}t}$$

i.e.,

$$\frac{\mathrm{d}R_*}{\mathrm{d}R_c} = -\frac{M_c}{M_e}\left(\frac{R_*}{R_c}\right)^2.$$

The negative sign demonstrates that as the core contracts, the envelope must expand – a good rule of thumb throughout stellar evolution, and, in particular, what happens at the end of the main sequence for solar-type stars.

As the radius increases, the effective temperature falls, and the star moves rightwards across the HRD. The temperature and density gradients between core and envelope are initially shallow, and the shell is quite extensive (in mass); this phase is referred to as ‘thick-shell burning’. As the core contracts and the envelope expands, these gradients increase, and the shell occupies less mass. During this ‘thin-shell burning’, a significant part of the energy goes into expanding the envelope, leading to a drop in luminosity.

The increasing opacity that accompanies cooling temperatures favours convection, and the star approaches the Hayashi track (Fig. 17.2, point C). As the core continues to shrink, and the envelope expands in response, further decreases in T_{eff} are not possible (the star can’t transport energy efficiently enough), so the luminosity increases and the star ascends the ‘red-giant branch’ (RGB).

The transition from main sequence to red giant is rapid, so few stars are observed in this region of the HRD (points B–C) – the so-called 'Hertzsprung gap'. The expansion C–D occurs on a thermal timescale, so the hydrogen shell-burning phase is short-lived for intermediate-mass stars (it is much longer for lower-mass stars).

First dredge-up

As an intermediate-mass star ascends the RGB, it develops an extensive convective envelope which reaches down to the hydrogen-burning shell (echoing the fully convective phase of the Hayashi track), bring CNO-processed material to the surface - the first dredge-up. Because the C–N cycle reaches equilibrium before the O–N cycle, C–N processed material is exposed at the surface; surface N is typically enriched by a factor ∼2, C is depleted by ∼30% (and O is unchanged). The observation of this CN-processed material is important evidence that it is the CNO cycle (not proton-proton burning) that occurs in the H-burning shell.

Stellar winds

Red giants are observed to lose mass in the form of slow winds ($v \simeq 5$–30 km s^{-1}, $\dot{M} \sim 10^{-8}$ M$_\odot$ yr^{-1}). Depending on luminosity, the star can lose several tenths of a solar mass through this wind.

The end of the RGB phase: core helium ignition.

The helium 'ashes' of shell hydrogen burning increase the mass of the core, which contracts under increasing pressure. According to the virial theorem, half the energy released by the gravitational contraction of the core is radiated away, and half goes into heating the core. Eventually the core temperature is high enough to ignite core helium burning in the triple-α process ($T \sim 10^8$ K; point D).

For stars in the mass range[1] $M \sim 0.8$–2.3M$_\odot$ the inert helium core is degenerate. Once it reaches a mass of ∼0.45M$_\odot$, the temperature is high enough for helium ignition in a 'helium flash'. This is a thermonuclear runaway (because the degenerate equation of state is independent of temperature, the temperature – hence energy generation rate – initially rises without any increase in pressure), raising the core luminosity to as much as 10^{10}–10^{11}L$_\odot$, but only briefly (a few minutes), before the degeneracy is lifted. The helium flash doesn't have a disruptive effect on the star, and isn't directly observable; the energy initially goes into lifting the degeneracy and expanding the core (and some is lost through neutrino emission). Remaining energy is absorbed by the envelope, and is released slowly (on a thermal timescale).

[1] In lower-mass stars, the pressures are not high enough to produce degeneracy; for higher-mass stars, the gas pressure remains high enough to prevent degeneracy.

Because the helium flash takes place at approximately constant *core* mass, the maximum luminosity at the stellar surface generated by this process (which exceeds that of other RGB phases) is reasonably insensitive to the initial mass or chemical composition of the star. Consequently, the brightest bolometric magnitude reached by the star as a result of the helium flash is nearly constant. The brightness of the 'tip of the red-giant branch' (TRGB) in the HR diagram of a sufficiently large sample of stars can therefore be used as a 'standard(ish) candle', and hence a useful distance indicator.

17.4.2 Horizontal branch: core helium (+shell hydrogen) burning

The star now has two energy sources: the helium-burning core, and the hydrogen-burning shell.

The core expands as a result of the energy input (heating) and the density in the adjacent hydrogen shell drops. The shell contributes most (∼70%) of the energy, so the drop in density leads to a fall in total energy generation (luminosity), while the shrinkage of the envelope (mirror law) leads to to a rise in effective temperature – the star moves down and to the left in the HRD (to point `E`).

For our $4\mathrm{M}_\odot$ example, homology indicates $L_* \propto M^3\mu^4/\bar{\kappa}_\mathrm{R}$ (§16.3.1); that is, a core-helium-burning star is expected to be $\gtrsim$50× more luminous than a core-hydrogen-burning star of the same mass ($\Delta M_\mathrm{Bol} \gtrsim 4^\mathrm{m}$), just because of the difference in mean molecular weight. Of course, the internal structures of a main-sequence star and the evolved counterpart are quite different, so they are not really at all homologous; but nonetheless, the analysis gives a qualitatively, and even semi-quantitatively, correct answer.

The star is now in a relatively stable stage of core-helium (+shell-hydrogen) burning (`E`–`F`). In the HRDs of globular clusters (coeval low-metallicity systems), these stars form a 'horizon branch', as a result of bluewards loops at constant luminosity (not obvious for the 4-$\mathrm{M}_\odot$ track shown in Fig. 17.2, but clear for the 6–10-$\mathrm{M}_\odot$ tracks in Fig. 17.1); this relatively long-lived phase (∼10% of the main-sequence lifetime) is therefore called the horizontal branch.[2]

The 'Red Clump'

The stars at the red end of the horizontal branch (typically of higher metallicity than globular-cluster stars) form a 'red clump' (RC). This is a visually striking feature of the colour—magnitude diagrams of intermediate-age star clusters and nearby galaxies, and is defined by a group of core-helium-burning stars which all have similar absolute magnitudes, making the RC another potential standard candle.

The sizable convective envelopes result from either a moderately high metallicity or, more generally, a significant buffer of mass (a few 0.1 M) above the H-burning shell. Low-mass stars without such a

[2]The horizontal branch of the globular cluster 47 Tuc (Fig. 11.1) is evident at $V \simeq 13.8$, $(B-V) \simeq 0.7$.

buffer, or with extremely low metallicities, will burn He at higher effective temperatures, defining the horizontal branch. Red-clump stars are associated with younger, more metal-rich populations than those associated with the Horizontal Branch, although the RC is sometimes referred to as the red extremity of the HB.

Metallicity is the key factor explaining that difference (low metallicity stars have lower opacities and are hotter), but other factors (composition, age, rotation, envelope mass) appear to affect the morphology of the HB.

17.4.3 Early asymptotic giant branch (E-AGB): shell helium burning [G]

Eventually core helium is exhausted (point F), leaving an inert core composed principally of carbon and oxygen. Helium continues to burn outside the core, in a thick shell. Core contraction is accompanied by expansion of the envelope, and the drop in density and temperature quenches the shell hydrogen burning; the sole energy source is now the helium-burning shell. Cooling of the envelope is accompanied by an increase in opacity, and the envelope becomes strongly convective; the star moves back onto the Hayashi track, and ascends the asymptotic giant branch (AGB; G). For intermediate-mass stars, the contracting CO core becomes degenerate (Table H). The AGB phase lasts about 10% of the horizontal-branch lifetime.

Paczynski relation.

The star's luminosity at this stage is largely determined not by the total mass, but by the mass of the degenerate core (although the latter is a function of the former, of course). We know that there exists a mass–radius relationship for degenerate bodies (Section 14.4.4); thus the pressure, and hence the energy-generation rate, in a shell outside the core is principally dependent on the core mass. Paczynski (1970) quantified this idea; a modern version (from Langer) is

$$L_* \simeq 5.9 \times 10^4 \, (M_\mathrm{c}/\mathrm{M}_\odot - 0.52) \, \mathrm{L}_\odot$$

(where M_c is the core mass).

Stars with higher initial masses will have higher-mass, smaller cores, and hence higher AGB luminosities. Since, ultimately, the luminosity of the helium-burning shell drives the envelope expansion, stars with higher core masses also evolve to larger radii ($\sim 300 \mathrm{R}_\odot$ for the model shown in Fig. 17.2).

Second dredge-up.

The convective envelope reaches down to the now-dormant hydrogen-burning shell, cycling the CNO-processed material to the surface. In this second dredge-up, surface abundances of He and N are

enhanced, and C and O depleted. (The second dredge-up only occurs for stars with initial masses $\gtrsim$3–4$M_\odot$; at lower masses, the convective zone doesn't extend down to the hydrogen discontinuity.)

Stellar winds.

The stellar-wind mass-loss rate increases dramatically during the AGB phase, reaching $\sim 10^{-8}$–$\sim 10^{-6}$ $M_\odot$ yr^{-1}; this wind is believed to be powered by radiation pressure on dust forming in the cool outer regions of the atmosphere.

17.4.4 Thermal pulsing AGB (TP-AGB); mostly shell H burning [H]

As the helium-burning shell uses up the available helium fuel, its mass decreases; the overlying hydrogen layers contract, heating a shell of hydrogen to the point where it can ignite. Two shell sources now exist, but because their rates of energy generation are so different, the situation is unstable, leading to a phase of thermal pulsing. During most of this phase, the energy comes from hydrogen shell burning, with relatively brief helium shell 'flashes'.

Schematically, the cycle proceeds as follows:

1. We start with a degenerate CO core and a helium-burning shell.
2. As the star ascends the AGB, the 'ash' from the helium-burning shell adds mass to the CO core, driving the luminosity up (Paczynski relation). The helium-shell mass is reduced (by He burning, and the continuing second dredge-up); when it reaches ~0.02$M_\odot$, the temperature at the H/He interface is sufficient to ignite shell hydrogen burning.

   The structure of the star is now:

   a degenerate CO core;

   a helium-burning shell;

   a hydrogen-burning shell; and

   an outer (hydrogen-rich) envelope.
3. Burning in the helium shell is unstable because of the thinness of the shell, coupled with the strong temperature dependence of the 3α reaction. Essentially, the energy liberated raises the pressure, initiating expansion; once the expansion has progressed far enough, cooling sets in and the helium burning stops (Iben & Renzini 1983).
4. The hydrogen shell continues to burn, increasing the mass of the underlying helium layer between the CO core and the H-burning shell. H-shell burning is providing practically all the luminosity; this phase lasts $\sim 10^3$–10^5 yr, depending on core mass (smaller cores lead to longer cycles).

5. The helium layer between the core and hydrogen-burning shell is mildly degenerate. As mass is added from the hydrogen-burning shell, the pressure and temperature rise in this zone.

6. Eventually, helium is reignited explosively in a 'helium flash' when the temperature reaches $\sim 10^8$ K. The fusion raises the temperature, which increases the reaction rate, which raises the temperature... The helium-burning layer doesn't expand at first because of the degeneracy (pressure support isn't thermal), so there is no regulation of the rate of fusion (hence a 'flash'), but then the temperature reaches the point where the degeneracy is fully lifted.

7. Considerable energy is generated in the flash ($\sim 10^8 L_\odot$ for ~a year), but much of it goes into expanding the helium zone. The helium-burning region stabilises and stable burning ensues for $\sim 10^2$ yr. As a result of expansion and cooling, shell hydrogen burning is extinguished, which restores the conditions at the end of the E-AGB phase (step 1); and the cycle repeats.

The thermal pulse cycle can repeat many times, but is barely noticeable at the surface of the star. The helium shell is dormant for more than 99% of the cycle; hydrogen shell burning is the main luminosity source averaged over the time.

Third dredge-up.

The large energy flux above the helium shell at step 7 above demands convective energy transport. If this merges with the convective outer envelope, the products of shell helium burning (notably carbon) can be exposed at the surface, in a third dredge-up.[3] This can lead to the formation of 'carbon stars' (surface C/O> 1; cp. solar, C/O$\simeq$ 0.4).

However, for sufficiently massive stars ($M \gtrsim 4$–$5 M_\odot$), the base of the convective envelope can reach $T \gtrsim 5 \times 10^7$ K, hot enough to convert dredged-up carbon to nitrogen through the CN cycle ('hot bottom burning').

Furthermore, ^{14}N is a product of CNO burning in the hydrogen shell. During a thermal pulse, this can be burnt to neon through

$$^{14}\mathrm{N}(\alpha,\gamma)^{18}\mathrm{F}(\beta^+,\nu)^{18}\mathrm{O}(\alpha,\gamma)^{22}\mathrm{Ne}$$

and then, for stars with cores masses $\gtrsim 1 M_\odot$(i.e., high enough temperatures),

$$^{22}\mathrm{Ne}(\alpha,\mathrm{n})^{25}\mathrm{Mg},$$

generating a flux of neutrons at every thermal pulse. This provides 'fuel' for the *s* process; and the repeated application of neutron fluxes generates heavier and heavier elements. The third dredge-up can

[3]The dredge-up of helium-burning products is always called a third dredge-up, even if a second dredge-up did not occur; and multiple third dredge-ups can occur in a star.

(and does) reveal such elements, including, famously, ^{99}Tc, a radioactive isotope[4] with a half-life of only $\sim 2 \times 10^5$ yr.

Mass loss.

TP-AGB stars show dusty, massive, slow stellar winds, with $\dot{M}$ reaching up to $\sim 10^{-4}$ $M_\odot$ yr^{-1}. Since this can continue for $\sim 10^5$ yr, the star can lose a substantial fraction of its original mass through such winds.

The dust composition depends on the C/O ratio (i.e., on the nature of the third dredge-up). If C/O<1 (by number), then all the carbon is locked into the stable CO molecule, and the remaining oxygen forms silicate dust; if C/O>1, then all the oxygen is locked into CO, and the dust is carbon-rich (containing, e.g., SiC and C_nH_n molecules). The mechanism responsible for the substantial mass loss is not well understood; radiation pressure must play a role, but pulsational instabilities may be involved (many AGB stars are long-period variables).

17.4.5 Post-AGB evolution

The mass of the convective envelope decreases steadily through the AGB phase, through shell burning (removes mass from the bottom of the envelope) and stellar winds (removes mass from the top of the envelope). When the mass of the hydrogen envelope becomes very small ($\sim 10^{-3}$–$10^{-2}M_\odot$), convection can no longer be sustained and the envelope contracts into radiative equilibrium.

Hydrogen shell burning is still taking place, so the luminosity is unchanged (following the Paczynski relation); the star therefore leaves the AGB, moving leftwards in the HRD. As it gets hotter it ionizes the circumstellar envelope generated principally through AGB mass loss, and develops a fast, radiation-driven wind ($\dot{M} \simeq 10^{-7}$ $M_\odot$ yr^{-1}, $v \simeq 2000$ km s^{-1}). These two effects produce the diversity of planetary nebulae observed.

When the envelope mass falls to $\sim 10^{-5}M_\odot$, the H-burning shell dies out, and the remnant core becomes a cooling white dwarf at $\sim$3–10$\times 10^4$ K, slowly radiating away its thermal energy over $\sim$a Hubble time.[5]

17.4.6 Summary

To review the key evolutionary stages of an intermediate-mass star:

Main-sequence stars: powered by core hydrogen burning; longest evolutionary phase.

[4] Technetium was discovered in red-giant spectra by Paul Merrill in 1952.

[5] In some cases a late thermal pulse can bring the star back as a 'born-again' AGB star.

RGB stars: powered by shell hydrogen burning. First dredge-up.

Horizontal-branch stars: powered by shell hydrogen burning and core helium burning. Second-longest evolutionary phase (~10% of MS lifetime).

Early-AGB stars: powered by shell helium burning. Second dredge-up (of CNO processed material from dormant H-burning shell).

TP-AGB stars: thermally pulsing. Hydrogen shell burning is the main energy source, with repeating brief helium shell flashes (thermal pulses). Third dredge-up (of He-burning products and *s*-process elements).

Planetary nebula; white dwarf.

Appendix A

SI units

A.1 Base units

Astronomers have been rather bad at using the SI system; for understandable reasons, they prefer their own 'custom' units ($M_\odot$, pc, etc.), and often adopt cgs units (centimetre, gram, second) in preference to the mks units (metre, kilogram, second) that are adopted in the SI system. Nevertheless, students, in particular, really should strive to use SI, which (among other advantages) greatly simplifies treatments of electricity.[1]

The seven irreducible base units of the SI system [and their cgs counterparts, where different] are:

Quantity	SI Unit		cgs equivalent	
length	metre	m	[centimetre	cm = 10^{-2} m]
mass	kilogram	kg	[gram	g = 10^{-3} kg]
time	second	s		
electric current	ampere	A	[Biot	bi = 10^{-1} A]
amount of substance	mole	mol		
luminous intensity	candela	cd		
thermodynamic temp.	kelvin	K	[degree Celsius	°C = K − 273.15]

[1] In SI, electric current is defined in terms of the directly measurable magnetic force it exerts, and charge is then defined as current multiplied with time.
In cgs 'electrostatic units', the unit of charge (or statcoulomb), is defined by the quantity of charge which gives a force constant of 1 in Coulomb's law. That is, for two point charges, each with charge 1 statcoulomb, separated by 1 centimetre, the electrostatic force between them is one dyne. This also has the effect of making electric charge dimensionless (and not requiring a fundamental unit).

A.2 Derived units

'Derived quantities' can be defined in terms of the seven base quantities. There are 20 derived quantities which are not dimensionless and which, for convenience, have named units; these are tabulated overleaf.

The units of angle and solid angle are, formally, simply the number 1 (being ratios of dimensionally identical quantities). Nonetheless, these two further derived quantities have named units, as the lack of units could easily be confusing. They are:

- radian (rad): the unit of angle is the angle subtended at the centre of a circle by an arc of the circumference equal in length to the radius of the circle (so there are 2π radians in a circle).

- steradian (sr): the unit of solid angle is the solid angle subtended at the centre of a sphere of radius r by a portion of the surface of the sphere having an area r^2 (so there are 4π steradians on a sphere).

Many other derived quantities in more or less common use don't have special names for their units; some are given in the tables which follow (in a few cases, these quantities do have named units in the cgs system). A number of other convenient units are not directly derived from the SI base units, but can nonetheless be expressed in terms of those units, and are recognized by the guardians of the SI system. Important examples for astrophysics include:

- The minute (m=60 s), hour (h = 3600 s), and day (d = 86 400 s).
  (The year is not an admitted unit, as it varies in length; for rough calculations it may be adequate to assume 1 yr $\simeq$ 365.25 d.)

- the degree ($^\circ = {}^{2\pi}/_{360}$ rad), arcminute ($' = {}^{2\pi}/_{21\,600}$ rad), and arcsecond ($'' = {}^{2\pi}/_{1.296\times10^6}$ rad)

- the atomic mass unit (amu = $1.66053886 \times 10^{-27}$ kg)

- the electron volt (eV = $1.60217646 \times 10^{-19}$ J)

- the ångström (Å = 10^{-10} m = 0.1 nm)

- the astronomical unit (au = 1.49598×10^{11} m) and
  the parsec (pc = $3.08568025 \times 10^{16}$ m).

Named Derived SI units

Quantity	Unit			Base units	Equivalent cgs unit		
angular measure	radian	rad	1				
solid angle	steradian	sr	1				
frequency	hertz	Hz	s^{-1}	s^{-1}			
force, weight	newton	N	m kg s^{-2}	m kg s^{-2}	dyne	dyn	10^{-5} N
pressure, stress	pascal	Pa	N m^{-2}	m^{-1} kg s^{-2}			
energy, work, heat	joule	J	N m	m^2 kg s^{-2}	erg	erg	10^{-7} J
power, radiant flux	watt	W	J s^{-1}	m^2 kg s^{-3}			
electric charge or electric flux	coulomb	C	A s	A s			
voltage, electrical potential difference, electromotive force	volt	V	W A^{-1} = J C^{-1}	m^2 kg s^{-3} A^{-1}			
electric capacitance	farad	F	C V^{-1}	m^{-2} kg^{-1} s^4 A^2			
electric resistance, impedance, reactance	ohm	Ω	V A^{-1}	m^2 kg s^{-3} A^{-2}			
electrical conductance	siemens	S	$Ω^{-1}$	m^{-2} kg^{-1} s^3 A^2			
magnetic flux	weber	Wb	J A^{-1}	m^2 kg s^{-2} A^{-1}	maxwell	Mx	10^{-8} Wb
magnetic field	tesla	T	V s m^{-2} = Wb m^{-2} = N A^{-1} m^{-1}	kg s^{-2} A^{-1}	gauss	G	10^{-4} T
inductance	henry	H	V s A^{-1} = Wb A^{-1}	m^2 kg s^{-2} A^{-2}			
Celsius temperature	Celsius	°C	t(°C) = t(K) − 273.15	K			
luminous flux	lumen	lm	cd sr	cd			
illuminance	lux	lx	lm m^{-2}	m^{-2} cd	phot	ph	10^4 lx
radioactivity (decays per unit time)	becquerel	Bq	Hz	s^{-1}			
absorbed dose (of ionizing radiation)	gray	Gy	J kg^{-1}	m^2 s^{-2}			
equivalent dose (of ionizing radiation)	sievert	Sv	J kg^{-1}	m^2 s^{-2}			
catalytic activity	katal	kat	mol s^{-1}	mol s^{-1}			

Unnamed Derived SI units		
Quantity	Units	Dimensions
area	m^2	m^2
volume	m^3	m^3
speed, velocity	$m\ s^{-1}$	$m\ s^{-1}$
acceleration	$m\ s^{-2}$	$m\ s^{-2}$
jerk	$m\ s^{-3}$	$m\ s^{-3}$
angular velocity	$rad\ s^{-1}$	s^{-1}
momentum, impulse	N s	$kg\ m\ s^{-1}$
angular momentum	N m s	$kg\ m^2\ s^{-1}$
torque, moment of force	N m	$kg\ m^2\ s^{-2}$
wavenumber	m^{-1}	m^{-1}
mass density	$kg\ m^{-3}$	$kg\ m^{-3}$
heat capacity, entropy	$J\ K^{-1}$	$kg\ m^2\ s^{-2}\ K^{-1}$
specific heat capacity, specific entropy	$J\ K^{-1}\ kg^{-1}$	$m^2\ s^{-2}K^{-1}$
specific energy	$J\ kg^{-1}$	$m^2\ s^{-2}$
energy density	$J\ m^{-3}$	$kg\ m^{-1}\ s^{-2}$
surface tension	$N\ m^{-1} = J\ m^{-2}$	$kg\ s^{-2}$
heat flux density, irradiance	$W\ m^{-2}$	$kg\ s^{-3}$
thermal conductivity	$W\ m^{-1}\ K^{-1}$	$kg\ m\ s^{-3}\ K^{-1}$
diffusion coefficient	$m^2\ s^{-1}$	$m^2\ s^{-1}$
dynamic viscosity[1]	$Pa\ s = N\ s\ m^{-2}$	$kg\ m^{-1}\ s^{-1}$
kinematic viscosity[2]	$m^2\ s^{-1}$	$m^2\ s^{-1}$
electric charge density	$C\ m^{-3}$	$m^{-3}\ A\ s$
electric current density	$A\ m^{-2}$	$A\ m^{-2}$
conductivity	$S\ m^{-1}$	$kg^{-1}\ m^{-3}\ s^3\ A^2$
permittivity	$F\ m^{-1}$	$kg^{-1}\ m^{-3}\ s^4\ A^2$
permeability	$H\ m^{-1}$	$kg\ m\ s^{-2}\ A^{-2}$
electric field strength	$V\ m^{-1}$	$kg\ m\ s^{-3}\ A^{-1}$
magnetic field strength[3]	$A\ m^{-1}$	$A\ m^{-1}$
luminance[4]	$cd\ m^{-2}$	$cd\ m^{-2}$
cgs named units:		
[1]poise	$P = 0.1\ Pa\ s$	
[2]stokes	$St = 10^{-4}\ m^2\ s^{-1}$	
[3]oersted	$Oe = \frac{1000}{4\pi}\ A\ m^{-1}$	
[4]stilb	$sb = 10^4\ cd\ m^{-2}$	

A.3 Prefixes

The SI system also specifies that names of multiples and submultiples of units are formed by means of the following prefixes:

Multiplying Factor	Prefix	Symbol	Multiplying Factor	Prefix	Symbol
10^{24}	yotta	Y	10^{-1}	deci	d
10^{21}	zetta	Z	10^{-2}	centi	c
10^{18}	exa	E	10^{-3}	milli	m
10^{15}	peta	P	10^{-6}	micro	μ
10^{12}	tera	T	10^{-9}	nano	n
10^{9}	giga	G	10^{-12}	pico	p
10^{6}	mega	M	10^{-15}	femto	f
10^{3}	kilo	k	10^{-18}	atto	a
10^{2}	hecto	h	10^{-21}	zepto	z
10^{1}	deca	da	10^{-24}	yocto	y

Multiple prefixes may not be used, even for the kilogram (unique among SI base units in having one of these prefixes as part of its name), for which the prefix names are used with the unit name 'gram', and the prefix symbols are used with the unit symbol 'g'; e.g, 10^{-6} kg = 1 mg (not 1 μkg).

With this exception, any SI prefix may be used with any SI unit (whether base or derived, including the degree Celsius and its symbol °C). Note that use of 'micron' for the μm persists very widely (almost universally!) in astrophysics, although the approved SI name is the micrometre (not to be confused with a micrometer, which is a measuring instrument; one good reason to use English spelling in preference to 'American English'…).

According to SI rules, these prefixes strictly represent powers of 10, and should not be used to represent the powers of 2 commonly found in computing applications. Thus one kilobyte (1 kbyte) is 1000 bytes – and not 2^{10} bytes = 1024 bytes. In an attempt to resolve this ambiguity, prefixes for binary multiples have been recommended by the International Electrotechnical Commission for use in information technology (though they're achieving acceptance only very slowly):

Factor	Name	Symbol	Origin		
2^{10}	kibi	Ki	'kilobinary',	$(2^{10})^1$	kilo, $(10^3)^1$
2^{20}	mebi	Mi	'megabinary',	$(2^{10})^2$	mega, $(10^3)^2$
2^{30}	gibi	Gi	'gigabinary',	$(2^{10})^3$	giga, $(10^3)^3$
2^{40}	tebi	Ti	'terabinary',	$(2^{10})^4$	tera, $(10^3)^4$
2^{50}	pebi	Pi	'petabinary',	$(2^{10})^5$	peta, $(10^3)^5$
2^{60}	exbi	Ei	'exabinary',	$(2^{10})^6$	exa, $(10^3)^6$

A.4 Writing conventions

For those really interested in the details, here are some of the more important elements of recommended writing style for SI units:

- Symbols are written in upright, roman type ('m' for metres, 'l' for litres).
- Units are written without a capital (other than where the rules of punctuation require it), as are their corresponding symbols, except for symbols derived from the name of a person; thus "the symbol for the coulomb is 'C' ". However, some American-speaking countries use 'L' for 'litre' (to avoid potential confusion with numeric '1').
- Names of units take plurals according to the usual rules of grammar; e.g., 20 kilograms, 40 henries. 'Hertz', 'lux', and 'siemens' have the same form in the singular and the plural. Symbols of units are not pluralised ('20 kg', not '20 kgs'), thereby avoiding potential confusion with the second ('s').
- A space should separates a number and its unit ('20 kg', not '20kg'). Exceptions are the symbols for degrees, arcminutes, and arcseconds (°, ′, ″), which should be contiguous with the number (e.g., 20° 15′).
- Symbols do not have an appended full stop (other than where the rules of punctuation require it; specifically, at the end of a sentence).
- Commas should not be used to break up long runs of digits, though spaces may be used (3.141 592 654, not 3.141,592,654).

A.5 SI vs. cgs: some notes on units of electromagnetism

Thanks to the French physicist Charles-Augustin de Coulomb, we know that the electrostatic force between two charged objects is proportional to the product of the charges, $q_1 q_2$, and inversely proportional to the square of the distance d between them; i.e.,

$$F = k_E \frac{q_1 q_2}{d^2} \tag{A.1}$$

where k_E is some constant of proportionality, the dimensions and numerical value of which depends on how we define 'charge' (and 'distance').

A wire carrying a current is a simple example of a moving charge that generates a magnetic field. For two long, parallel wires separated by distance d, the force per unit length is proportional to the product of the currents they carry, $I_1 I_2$, and inversely proportional to the perpendicular separation of the wires:

$$\frac{F}{L} = k_M \frac{I_1 I_2}{d} \tag{A.2}$$

where k_M is again a constant of proportionality with a value that depends on what we mean by 'current'.

Of course, these are just two different sides of the same coin: electromagnetism. A current is just a flow of charge, and EM theory tells us that

$$\frac{k_E}{k_M} = \frac{c^2}{2}$$

so if we know (or define) one of k_E and k_M, then we know the other.

The critical point is that the cgs and SI systems make different choices for the constants.

A.5.1 The cgs system.

In the cgs system, Coulomb's law, eqtn. (A.1), is just

$$F = \frac{q_1 q_2}{d^2};$$

that is k_E is implicitly defined as unity (and is dimensionless); thus

$$\frac{F}{L} = \frac{2}{c^2}\frac{I_1 I_2}{d}.$$

The (derived) unit of charge is the electrostatic unit (esu), defined as

$$\text{esu} \equiv \sqrt{\text{dyne} \times \text{cm}^2} = \text{g}^{1/2}\text{cm}^{3/2}\text{s}^{-1}$$

A.5.2 The SI system.

The SI system, on the other hand, is constructed from eqtn. (A.2), which lends itself to use of the ampere, where one ampere is defined to be the constant current that will produce an attractive force of 2×10^{-7} newtons per metre of length between two straight, parallel conductors of infinite length and negligible circular cross section placed one metre apart in a vacuum.

Then

$$k_M = \frac{\mu_0}{2\pi}$$

where

$$\mu_0 \equiv 4\pi \times 10^{-7}\text{N A}^{-2}$$

is the permeability of free space, requiring a definition as a consequence of choosing a base unit for current. We can now write Coulomb's law in the SI system, since

$$k_E = \frac{c^2}{2}k_M = \frac{\mu_0 c^2}{4\pi} = 8.99 \times 10^9 \text{N m}^2\ (\text{A s})^{-2};$$

that is, charge, in the SI system, has units of ampere seconds (i.e., coulombs).

We also define the permittivity of free space,

$$\epsilon_0 \equiv \frac{1}{\mu_0 c^2}.$$

Thus in SI units we write Coulomb's Law as

$$F = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{d^2}.$$

A.6 May 2019

Originally the SI system was based largely on arbitrary references – for example, the metre was originally defined (in 1793) as one ten-millionth of the distance from the equator to the North Pole, but was redefined in terms of a prototype metre bar. Similarly, until very recently 'the' kilogram was, essentially, defined by a cylinder of platinum-iridium alloy kept in a laboratory in France.

One problem (of several) with this reference system is that it's not constant. By definition, the mass of the reference kilogram was always one kilogram – even if a lump gets knocked out of it. And the bar that defined the length of the metre was changed in 1889 (with a concomitant change in the unit of length).

This in turn meant that any measurement of a physical constant, like the speed of light in vaccuum, was liable to yield a value that may be subject to change, if our reference for length, or speed, changed.

In recognition of this, there has been a slow migration from the philosophical viewpoint of measuring physical constants (which we believe are effectively *truly constant*) in terms of arbitrarily defined units to instead defining the units in terms of physical constants. We can recognise an illustration of this trend in the 1983 redefinition of the metre in terms of the distance travelled by light in vacuum in 1/299792458 of a second.

But what's "a second"? We're familiar with it as 1/3600th part of an hour, which is 1/24th of a day; but the length of the day is not constant. . .. So in the metric system, the second was originally defined as 1 part in 86400 of the 'mean solar day'; but then we have to define a mean solar day. The second is now defined in terms of a basic physical process: it is defined as "the duration of 9,192,631,770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the caesium-133 atom."

A third SI base unit, the candela, is similarly defined in terms of basic physical constants: it is the luminous intensity, in a given direction, of a source that emits monochromatic radiation of frequency

540×10^{12} Hz[2] with a radiant intensity in that direction of 1/683 watt per steradian.

May 2019 marks the completion of this migration, with a fundamental revision to the construction of the SI system – moving away completely from material artefacts (such as the prototype kilogram in Paris), and instead defining all the base units in terms of a set of constants of nature.

Rather than defining the Plank constant in terms of the kilogram, the kilogram will be defined in terms of the Planck constant. Of course, to do this we need to adopt some specific value for the Planck constant (expressed in SI units). But we can do this, just as the metre is defined in terms of the speed of light; we just need to agree on a numerical value. And by choosing a value that agrees with the best measurements from the 'old philosophy', we ensure continuity in our metrology.

Similarly, the ampere will no longer be used to measure the charge on the electron; rather, the charge on the electron will define the ampere.

And the Kelvin will be defined by Boltzmann's constant (and not by the triple point of water).

The final fundamental unit of the SI base system is Avagadro's number (the number of elementary particles in one mole of substance[3]).

These changes are formalised in the Draft Resolution to the 26th meeting of the General Conference on Weights and Measures, reproduced overleaf.

[This is functionally a diplomatic, not scientific, meeting, with binding powers. "Delegates at the meeting of the CGPM represent the Member States of the BIPM and the Associates of the CGPM, and each delegate must therefore be designated by their State as being entitled to represent it and to act on its behalf. This credential must be issued by and bear the signature of an authority with power to bind the State, such as the Head of State, the Minister of Foreign Affairs (or the head of the competent diplomatic mission) or the Minister responsible for the questions discussed by the Conference."]

[2]This is the value quoted on the web site of the Bureau International des Poids et Mesures, `https://www.bipm.org/en/publications/si-brochure/candela.html`. I do not know why it isn't written in the conventional style, as 5.40×10^{14} Hz.

[3]The mass of one mole of a substance is equal to that substance's molecular weight. For example, the mean molecular weight of water is 18.015 atomic mass units (amu), so one mole of water weight 18.015 grams.

Draft Resolution A

On the revision of the International System of Units (SI)

The General Conference on Weights and Measures (CGPM), at its 26th meeting,

considering

- the essential requirement for an International System of Units (SI) that is uniform and accessible world-wide for international trade, high-technology manufacturing, human health and safety, protection of the environment, global climate studies and the basic science that underpins all these,
- that the SI units must be stable in the long term, internally self-consistent and practically realizable being based on the present theoretical description of nature at the highest level,
- that a revision of the SI to meet these requirements was proposed in Resolution 1 adopted unanimously by the CGPM at its 24th meeting (2011) that laid out in detail a new way of defining the SI based on a set of seven defining constants, drawn from the fundamental constants of physics and other constants of nature, from which the definitions of the seven base units are deduced,
- that the conditions set by the CGPM at its 24th meeting (2011), confirmed at its 25th meeting (2014), before such a revised SI could be adopted have now been met,

decides that, effective from 20 May 2019, the International System of Units, the SI, is the system of units in which:

- the unperturbed ground state hyperfine transition frequency of the caesium 133 atom $\Delta\nu_{\mathrm{Cs}}$ is 9 192 631 770 Hz,
- the speed of light in vacuum c is 299 792 458 m/s,
- the Planck constant h is $6.626\,070\,15 \times 10^{-34}$ J s,
- the elementary charge e is $1.602\,176\,634 \times 10^{-19}$ C,
- the Boltzmann constant k is $1.380\,649 \times 10^{-23}$ J/K,
- the Avogadro constant N_{A} is $6.022\,140\,76 \times 10^{23}$ mol^{-1},
- the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz, K_{cd}, is 683 lm/W,

where the hertz, joule, coulomb, lumen, and watt, with unit symbols Hz, J, C, lm, and W, respectively, are related to the units second, metre, kilogram, ampere, kelvin, mole, and candela, with unit symbols s, m, kg, A, K, mol, and cd, respectively, according to Hz = s^{-1}, J = m^2 kg s^{-2}, C = A s, lm = cd m^2 m^{-2} = cd sr, and W = m^2 kg s^{-3}.

notes the consequences as set out in Resolution 1 adopted by the CGPM at its 24th meeting (2011) in respect of the base units of the SI and confirms these in the following Appendices to this Resolution, which have the same force as the Resolution itself,

invites the International Committee for Weights and Measures (CIPM) to produce a new edition of its Brochure entitled "*The International System of Units*" in which a full description of the revised SI will be given.

Appendix 1. Abrogation of former definitions of the base units

It follows from the new definition of the SI described above that, effective from 20 May 2019:

- the definition of the second in force since 1967/68 (13th meeting of the CGPM, Resolution 1) is abrogated,
- the definition of the metre in force since 1983 (17th meeting of the CGPM, Resolution 1) is abrogated,
- the definition of the kilogram in force since 1889 (1st meeting of the CGPM, 1889, 3rd meeting of the CGPM, 1901) based upon the mass of the international prototype of the kilogram is abrogated,
- the definition of the ampere in force since 1948 (9th meeting of the CGPM) based upon the definition proposed by the CIPM (1946, Resolution 2) is abrogated,
- the definition of the kelvin in force since 1967/68 (13th meeting of the CGPM, Resolution 4) is abrogated,
- the definition of the mole in force since 1971 (14th meeting of the CGPM, Resolution 3) is abrogated,
- the definition of the candela in force since 1979 (16th meeting of the CGPM, Resolution 3) is abrogated,
- the decision to adopt the conventional values of the Josephson constant K_{J-90} and of the von Klitzing constant R_{K-90} taken by the CIPM (1988, Recommendations 1 and 2) at the request of the CGPM (18th meeting of the CGPM, 1987, Resolution 6) for the establishment of representations of the volt and the ohm using the Josephson and quantum Hall effects, respectively, is abrogated.

Appendix 2. Status of constants previously used in the former definitions

It follows from the new definition of the SI described above, and from the recommended values of the 2017 special adjustment of the Committee on Data for Science and Technology (CODATA) on which the values of the defining constants are based, that effective from 20 May 2019:

- the mass of the international prototype of the kilogram $m(K)$ is equal to 1 kg within a relative standard uncertainty equal to that of the recommended value of h at the time this Resolution was adopted, namely 1.0×10^{-8} and that in the future its value will be determined experimentally,
- the vacuum magnetic permeability μ_0 is equal to $4\pi \times 10^{-7}$ H m^{-1} within a relative standard uncertainty equal to that of the recommended value of the fine-structure constant α at the time this Resolution was adopted, namely 2.3×10^{-10} and that in the future its value will be determined experimentally,

– the thermodynamic temperature of the triple point of water T_{TPW} is equal to 273.16 K within a relative standard uncertainty closely equal to that of the recommended value of k at the time this Resolution was adopted, namely 3.7×10^{-7}, and that in the future its value will be determined experimentally,

– the molar mass of carbon 12, $M(^{12}C)$, is equal to 0.012 kg mol^{-1} within a relative standard uncertainty equal to that of the recommended value of $N_A h$ at the time this Resolution was adopted, namely 4.5×10^{-10}, and that in the future its value will be determined experimentally.

Appendix 3. The base units of the SI

Starting from the new definition of the SI described above in terms of fixed numerical values of the defining constants, definitions of each of the seven base units are deduced by taking, as appropriate, one or more of these defining constants to give the following set of definitions, effective from 20 May 2019:

– The second, symbol s, is the SI unit of time. It is defined by taking the fixed numerical value of the caesium frequency $\Delta\nu_{Cs}$, the unperturbed ground-state hyperfine transition frequency of the caesium 133 atom, to be 9 192 631 770 when expressed in the unit Hz, which is equal to s^{-1}.

– The metre, symbol m, is the SI unit of length. It is defined by taking the fixed numerical value of the speed of light in vacuum c to be 299 792 458 when expressed in the unit m/s, where the second is defined in terms of $\Delta\nu_{Cs}$.

– The kilogram, symbol kg, is the SI unit of mass. It is defined by taking the fixed numerical value of the Planck constant h to be $6.626\,070\,15 \times 10^{-34}$ when expressed in the unit J s, which is equal to kg m^2 s^{-1}, where the metre and the second are defined in terms of c and $\Delta\nu_{Cs}$.

– The ampere, symbol A, is the SI unit of electric current. It is defined by taking the fixed numerical value of the elementary charge e to be $1.602\,176\,634 \times 10^{-19}$ when expressed in the unit C, which is equal to A s, where the second is defined in terms of $\Delta\nu_{Cs}$.

– The kelvin, symbol K, is the SI unit of thermodynamic temperature. It is defined by taking the fixed numerical value of the Boltzmann constant k to be $1.380\,649 \times 10^{-23}$ when expressed in the unit J K^{-1}, which is equal to kg m^2 s^{-2} K^{-1}, where the kilogram, metre and second are defined in terms of h, c and $\Delta\nu_{Cs}$.

– The mole, symbol mol, is the SI unit of amount of substance. One mole contains exactly $6.022\,140\,76 \times 10^{23}$ elementary entities. This number is the fixed numerical value of the Avogadro constant, N_A, when expressed in the unit mol^{-1} and is called the Avogadro number.

  The amount of substance, symbol n, of a system is a measure of the number of specified elementary entities. An elementary entity may be an atom, a molecule, an ion, an electron, any other particle or specified group of particles.

– The candela, symbol cd, is the SI unit of luminous intensity in a given direction. It is defined by taking the fixed numerical value of the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz, K_{cd}, to be 683 when expressed in the unit lm W^{-1}, which is equal to cd sr W^{-1}, or cd sr kg^{-1} m^{-2} s^3, where the kilogram, metre and second are defined in terms of h, c and $\Delta\nu_{Cs}$.

Appendix B

Constants

B.1 Some physical constants

Speed of light*	c	2.99792458×10^{8} m s^{-1}
Universal gravitational constant	G	$6.67384\,[80] \times 10^{-11}$ m^3 kg^{-1} s^{-2} (= N m^2 kg^{-2})
Planck constant*	h	$6.62607015 \times 10^{-34}$ m^2 kg s^{-1} (=J s)
Boltzmann constant*	k	1.380649×10^{-23} m^2 kg s^{-2} K^{-1} (=J K^{-1})
Stefan-Boltzmann constant	σ	$5.670373\,[21] \times 10^{-8}$ W m^{-2} K^{-4}
Radiation constant	$a = 4\sigma/c$	$7.565731\,[28] \times 10^{-16}$ J m^{-3} K^{-4}]
Unified atomic mass unit	amu, u	$1.660538921\,[73] \times 10^{-27}$ kg
Hydrogen mass	m(H)	1.00794 amu
Proton mass	$m_{\rm P}$	$1.672621777\,[74] \times 10^{-27}$ kg
		= 1.007276466812 [90] u
Electron mass	$m_{\rm e}$	$9.10938291\,[40] \times 10^{-31}$ kg
Electron charge	e	$1.602176565\,[35] \times 10^{-19}$ C
	$\frac{\pi e^2}{m_{\rm e} c}$	2.654×10^{-6} m^2 s^{-1}

*Exact value (defining the SI system)

The main source of these values is the National Institute of Standards and Technology (NIST) listing of the 2010 recommendations of International Council for Science Committee on Data for Science and Technology (CODATA); bracketed values are uncertainties in the least significant digits. The hydrogen mass allows for a 0.015% fractional abundance of deuterium. XXXNeeds May 2019 revision!

B.2 Some astronomical constants

Astronomical unit	AU	$1.49597870800 \times 10^{11}$ m (exact)
Parsec	pc	$3.08567758 \times 10^{16}$ m ($10 \times 60 \times 60 \times 360/(2\pi)$ AU)

The AU is defined as a specific value by the International Astronomical union. The 'official' (IAU recommended) symbol for the astronomical unit is au, but AU is more widely used

B.2.1 Solar parameters

No solar parameter is a 'constant' in principle, because of variability on evolutionary, or shorter, timescales, but pragmatism dictates adoption of specific values. At its 2015 General Assembly, the International Astronomical Union adopted a set of 'nominal conversion constants', including reference values for solar parameters

The solar constant is the (very slightly variable) energy flux from the Sun measured at a distance of 1 AU; numerically,

$$\text{solar constant, } C_\odot = 1360.8\,[5]\ \text{J m}^{-2}\ \text{s}^{-1}.$$

according to Kopp & Lean (2011; Geophys. Res. Letters, 38, L01706); earlier determinations clustered around ~ 1368 J m^{-2} s^{-1}. By reference to $L = 4\pi d_\odot^2 C_\odot$ (where $d_\odot \simeq 1$ au), the 'nominal solar luminosity' is then *defined* to be

$$\mathcal{L}_\odot^{\rm N} = 3.828 \times 10^{26}\ \text{W}$$

The nominal solar radius is

$$\mathcal{R}_\odot^{\rm N} = 6.957 \times 10^{8}\ \text{m};$$

(following Haberreiter et al. 2008, ApJ, 675, L53). Since $L = 4\pi R^2 \sigma T_{\rm eff}^4$, we find the (nominal) effective temperature of the Sun to be

$$\mathcal{T}_{\rm eff\odot}^{\rm N} = 5772\ \text{K}$$

The solar mass follows from equating centrifugal and gravitational accelerations of the Earth in orbit,

$$\frac{\text{M}_\oplus v_\oplus^2}{d_\odot} = \frac{G\text{M}_\odot\text{M}_\oplus}{d_\odot^2}.$$

Because G is one of the least precisely determined constants, the uncertainty on $G\text{M}_\odot$ is around five orders of magnitude less than the uncertainty on $\text{M}_\odot$. The adopted nominal solar mass is therefore expressed as as

$$(\mathcal{GM})_\odot^{\rm N} = 1.327\,124\,4 \times 10^{20}\ \text{m}^3\ \text{s}^{-2}$$

such that an object's mass can be expressed in nominal solar masses, $\mathcal{M}_\odot$, by taking the ratio $(GM)_{\rm object}/(\mathcal{GM})_\odot^{\rm N}$.

Where a specific numerical value is required, a reasonable choice is

$$\mathrm{M}_\odot = 1.98855\,[24] \times 10^{30}\ \mathrm{kg}.$$

The mean solar density is then

$$\begin{aligned}\overline{\rho}_\odot &= \frac{\mathrm{M}_\odot}{{}^4\!/_3 \pi \mathrm{R}_\odot^3} \\ &= 1.4 \times 10^3\ \mathrm{kg\ m^{-3}}\end{aligned}$$

(i.e., about the same as liquid water on the surface of the Earth). Finally, the mean number density is

$$\overline{n}_\odot = \frac{\overline{\rho}}{\mu m(\mathrm{H})} \simeq 1.4 \times 10^{30}\ \mathrm{m^{-3}}$$

(using a mean molecular weight of $\mu \simeq 0.61$).

Appendix C

Black-body radiation

In astrophysics, a radiation field can often be usefully approximated by that of a 'black body', for which the intensity is given by the Planck function:

$$I_\nu = B_\nu(T) = \frac{2h\nu^3}{c^2}\left\{\exp\left(\frac{h\nu}{kT}\right) - 1\right\}^{-1} \quad [\mathrm{J\ m^{-2}\ s^{-1}\ Hz^{-1}\ sr^{-1}}]; \text{ or} \tag{C.1}$$

$$I_\lambda = B_\lambda(T) = \frac{2hc^2}{\lambda^5}\left\{\exp\left(\frac{hc}{\lambda kT}\right) - 1\right\}^{-1} \quad [\mathrm{J\ m^{-2}\ s^{-1}\ m^{-1}\ sr^{-1}}] \tag{C.2}$$

(where $B_\nu\,\mathrm{d}\nu = B_\lambda\,\mathrm{d}\lambda$).

C.1 Derivation

The Planck function can be obtained from consideration of Bose-Einstein statistics (photons are bosons!), by counting the degenerate states for a given energy. The energy density per unit frequency (i.e., the specific energy density) is $4\pi B_\nu/c$ (§3.6), which we can express as

$$\frac{4\pi}{c}B_\nu = D_\nu\, f(\nu)\, E \tag{C.3}$$

where D_ν is the number of degenerate states per unit volume for a given energy E, and $f(\nu)$ is the probability of occupying that energy.

For a finite volume V characterized by length L, the energy states that a boson (photon) can occupy are quantized. From Bose-Einstein statistics, and using $E = h\nu$, we have

$$f(\nu) = \left[\exp\left\{\frac{h\nu}{kT}\right\} - 1\right]^{-1}.$$

To count the number of degenerate states that all have the same energy for some differential frequency interval $\mathrm{d}\nu$, we can calculate the volume these states occupy in phase space. Since $E = h\nu$, the surface

area of a shell of constant energy is $4\pi\nu^2$; and since each quantized state occupies a volume $(c/L)^3$ in phase space, the total number of states is given by:

$$D_\nu V \,\mathrm{d}\nu = 2\frac{4\pi\nu^2 \,\mathrm{d}\nu}{(c/L)^3}$$

where $V = L^3$ and the factor 2 arises because photons can have two different polarizations. Substituting these expressions back into eqtn. C.3 gives

$$B_\nu = \frac{2h\nu^3}{c^2}\left[\exp\left\{\frac{h\nu}{kT}\right\} - 1\right]^{-1}.$$

C.2 Flux

We have seen that

$$F_\nu = 2\pi \int_{-1}^{+1} I_\nu(\mu)\mu \,\mathrm{d}\mu. \tag{3.8}$$

If we have a surface radiating like a black body then $I_\nu = B_\nu(T)$, and there is no μ dependence, other than that the energy is emitted over the limits $0 \le \mu \le 1$; thus the physical flux for a black-body radiator is given by

$$\begin{aligned} F_\nu = F_\nu^+ \quad &= 2\pi \int_0^{+1} B_\nu(T)\mu \,\mathrm{d}\mu \quad = B_\nu \left.\frac{2\pi\mu^2}{2}\right|_0^{+1} \\ &= \pi B_\nu. \end{aligned} \tag{C.4}$$

(cp. eqtn. 3.10: $F_\nu = \pi I_\nu$)

The total radiant energy flux is obtained by integrating eqtn. (C.4) over frequency,

$$\begin{aligned} \int_0^\infty F_\nu \,\mathrm{d}\nu &= \int_0^\infty \pi B_\nu \,\mathrm{d}\nu \\ &= \int_0^\infty \frac{2\pi h\nu^3}{c^2}\left\{\exp\left(\frac{h\nu}{kT}\right) - 1\right\}^{-1} \mathrm{d}\nu. \end{aligned} \tag{C.5}$$

We can solve this by setting $x = (h\nu)/(kT)$ (whence $\mathrm{d}\nu = [{}^{kT}/_{h}]\ \mathrm{d}x$), so

$$\int_0^\infty F_\nu \,\mathrm{d}\nu = \left(\frac{kT}{h}\right)^4 \frac{2\pi h}{c^2} \int_0^\infty \frac{x^3}{\exp(x) - 1}\,\mathrm{d}x$$

The integral on the right-hand side is now a standard form, which has the solution $\pi^4/15$; hence

$$\int_0^\infty F_\nu \,\mathrm{d}\nu = \left(\frac{k\pi}{h}\right)^4 \frac{2\pi h}{15c^2} T^4 \tag{C.6}$$

$$\equiv \sigma T^4 \tag{C.7}$$

where σ is the Stefan-Boltzmann constant,

$$\sigma = \frac{2\pi^5 k^4}{15h^3c^2} = 5.67 \times 10^{-5} \quad [\mathrm{J\ m^{-2}\ K^{-4}\ s^{-1}}].$$

C.3 Approximate forms

There are two important approximations to the Planck function which follow directly from eqtn. C.1:

$$B_\nu(T) \simeq \frac{2h\nu^3}{c^2}\left\{\exp\left(\frac{h\nu}{kT}\right)\right\}^{-1} \qquad \text{for } \frac{h\nu}{kT} \gg 1 \tag{C.8}$$

(Wien approximation), and

$$B_\nu(T) \simeq \frac{2\nu^2 kT}{c^2} \qquad \text{for } \frac{h\nu}{kT} \ll 1 \tag{C.9}$$

(Rayleigh-Jeans approximation; $\exp(h\nu/kT) \simeq 1 + h\nu/kT$).

The corresponding wavelength-dependent versions are, respectively,

$$B_\lambda(T) \simeq \frac{2hc^2}{\lambda^5}\left\{\exp\left(\frac{hc}{\lambda kT}\right)\right\}^{-1},$$
$$B_\lambda(T) \simeq \frac{2ckT}{\lambda^4}.$$

The Wien approximation to the Planck function is very good at wavelengths shortwards of and up to the peak of the flux distribution; but one generally needs to go something like $\sim 10\times$ the peak wavelength before the long-wavelength Rayleigh-Jeans approximation is satisfactory.

C.4 Wien's Law

Wien's displacement law (not to be confused with the Wien approximation!) relates the black-body temperature to the wavelength of peak emission. To find the peak, we differentiate eqtn. (C.2) with respect to wavelength, and set to zero:

$$\frac{\partial B}{\partial \lambda} = 8hc\left(\frac{hc}{\lambda^7 kT}\frac{\exp\{hc/\lambda kT\}}{(\exp\{hc/\lambda kT\}-1)^2} - \frac{1}{\lambda^6}\frac{5}{\exp\{hc/\lambda kT\}-1}\right) = 0$$

whence

$$\frac{hc}{\lambda_{\max} kT}\left(1 - \exp\{-hc/\lambda_{\max} kT\}\right)^{-1} - 5 = 0$$

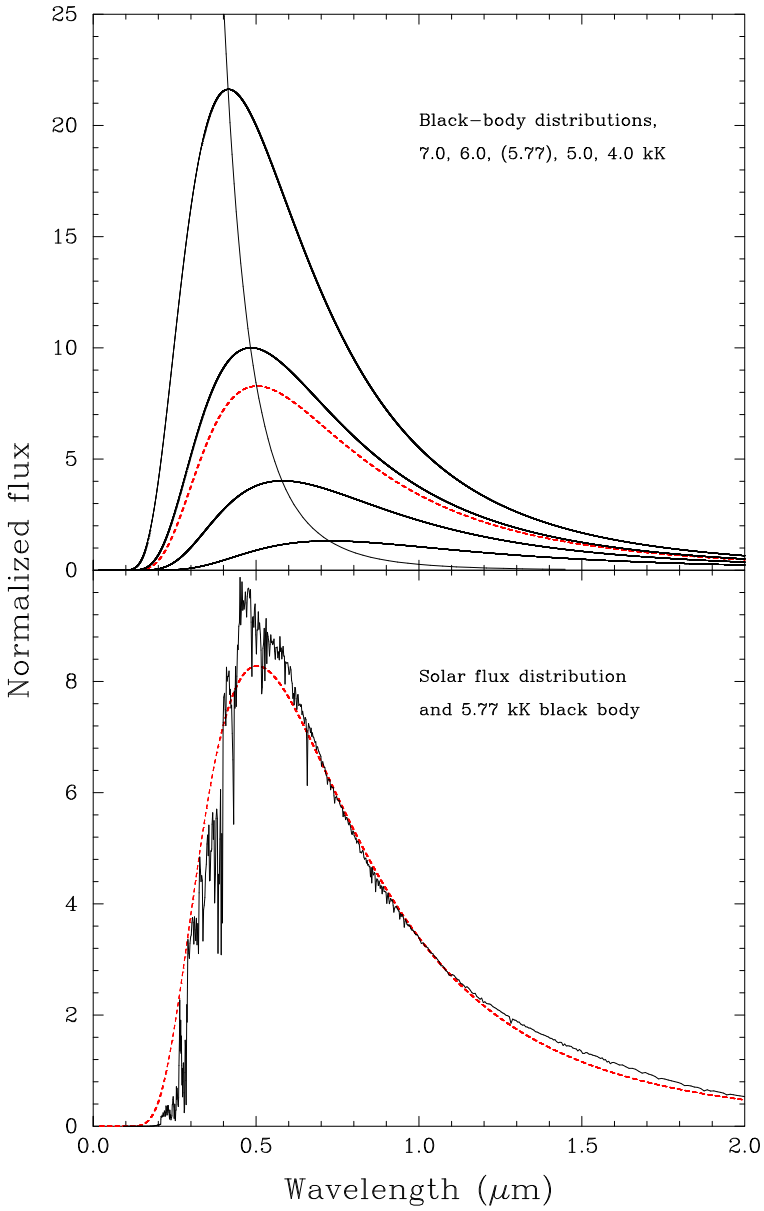


Figure C.1: Upper panel: Flux distributions for black bodies at several different temperatures. A hotter black body radiates more energy at all wavelengths than a cooler one, but the increase is greater at shorter wavelengths. The peak of the black-body distribution migrates blueward with increasing temperature, in accordance with Wien's law (also plotted).

Lower panel: Flux distribution for the Sun (actually, a Kurucz solar model) compared with a black-body distribution at the same temperature. The black body is a reasonable, though far from perfect, match to the model, the main differences arising because of line blocking in the sun at short wavelengths. This energy must come out elsewhere, and appears as an excess over the black body at long wavelengths.

(Flux units are 10^7 J m^{-2} s^{-1} μm^{-1}.)

An analytical solution of this equation can be obtained in terms of the Lambert W function; we merely quote the result,

$$\frac{\lambda_{\max}}{\mu\mathrm{m}} = \frac{2898}{T/\mathrm{K}}$$

We therefore expect the Sun's output to peak around 500 nm (for $T_{\mathrm{eff}} = 5772$ K) – near where the human eye has peak sensitivity, for obvious evolutionary reasons.

Appendix D

LTE: Saha equation

Fairly obviously, in Local Thermodynamic Equilibrium (LTE) it is assumed that all thermodynamic properties in a small volume have the thermodynamic equilibrium values at the local values of temperature and pressure.

Specifically, this applies to quantities such as the occupation numbers of atoms, the opacity, emissivity, etc. The LTE assumption is equivalent to stating that

1. the electron and ion velocity distributions are Maxwellian,

$$\frac{\mathrm{d}n(v)}{\mathrm{d}v} = n\left(\frac{m}{2\pi kT_\mathrm{k}}\right)^{3/2} \exp\left\{\frac{-mv^2}{2kT_\mathrm{k}}\right\}$$

   for number density n of particles of mass m at kinetic temperature T_k;

2. the photon source function is given by the Planck function at the local temperature (i.e., $S_\nu = B_\nu,$ and $j_\nu = k_\nu B_\nu$).

3. the excitation equilibrium is given by the Boltzmann equation

$$\frac{n_j}{n_i} = \frac{g_j}{g_i} \exp\left\{\frac{-(E_j - E_i)}{kT}\right\} \tag{D.1}$$

4. the ionization equilibrium is given by the Saha equation

$$\frac{n_\mathrm{e} n_{2,1}}{n_{1,i}} = \frac{2g_{2,1}}{g_{1,i}} \exp\left\{\frac{-\chi_{1,i}}{kT}\right\} \frac{(2\pi m_\mathrm{e} kT)^{3/2}}{h^3} \tag{D.2}$$

   where $1, i, 2, 1$ denote levels $i, 1$ in ionization stages 1, 2.

One might augment this list with the perfect-gas equation of state, $P = nkT$, but since this applies under many circumstances where LTE doesn't hold, it's not usually mentioned in this context.

If a process is purely collisional, conditions are, naturally, determined on a purely local basis locally, and LTE applies. We have already encountered one such situation where LTE is a good approximation: free-free emission results from a purely collisional process, justifying our adoption of $S_\nu = B_\nu$ (Section **??**).

If radiation plays a role, then provided the photon and particle mean free paths are short compared to the length scales over which conditions change significantly (i.e., if the opacity is high), then we can again expect LTE to be a reasonable assumption; this is a good approximation in stellar interiors.

In stellar atmospheres the LTE approximation may be a poor one, as photon mean free paths are typically larger than those of particles. Thus one region can be affected by the radiation field in another part of the atmosphere (e.g., a deeper, hotter region). As a rule of thumb, therefore, LTE is a poor approximation if the radiation field is important in establishing the ionization and excitation equilibria (as in hot stars, for example). It's more likely to be acceptable when particle densities are high and the radiation field is relatively weak; for stars, this means higher gravities (i.e., main-sequence stars rather than supergiants) and cooler effective temperatures. When LTE breaks down, we have a 'non-LTE' (nLTE) situation, and level populations must be calculated assuming statistical equilibrium (section J.0.2).

D.1 The Saha Equation

The Boltzmann Equation gives the relative populations of two bound levels i and j, in some initial (or 'parent') ionization stage '1':

$$\frac{n_{1,j}}{n_{1,i}} = \frac{g_{1,j}}{g_{1,i}} \exp\left\{\frac{-(E_{1,j} - E_{1,i})}{kT}\right\} \tag{D.1}$$

where $E_{1,i}$ & $E_{1,j}$ are the level energies (measured from the ground state, $E_{1,1} = 0$), and $g_{1,i}$ & $g_{1,j}$ are their statistical weights ($2J + 1$, where J is the total angular-momentum quantum number).

To generalize the Boltzmann eqtn. to deal with collisional ionization to the next higher (or 'daughter') ionization stage '2', we identify the upper level j with a continuum state; $n_{1,j}$, the number of parent ions in excitation state j, then equates with $n_{2,1}(v)$, the number of ionized atoms where the detached electron has velocity v.

(Note that ionization stages '1' and '2' always represent any two consecutive stages – for example, H^0 and H^+, or C^{2+} and C^{3+}.)

The total statistical weight of the ionized system is given by the combined statistical weights[1] of the newly created ion and the electron, i.e., $g_2 g_e(v)$; while the relevant energy is the sum of the ionization energy and the kinetic energy of the free electron. Thus we have

$$\frac{n_{2,1}(v)}{n_{1,i}} = \frac{g_2 g_e(v)}{g_{1,i}} \exp\left\{\frac{-(\chi_{1,i} + \frac{1}{2}m_e v^2)}{kT}\right\} \tag{D.3}$$

where $\chi_{1,i} = E_\infty - E_{1,i}$ is the ionization potential for level i in the parent species.

An aside: The statistical weight of a free electron. The statistical weight of a free electron is just the probability of finding it in a specific cell of 'phase space'. Since the state of a free particle is specified by three spatial coördinates x, y, z and three momentum coördinates $p(x)$, $p(y)$, $p(z)$, the number of quantum states (for which the statistical weights are each 1) in an element of phase space,

$$\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \cdot \mathrm{d}p(x)\,\mathrm{d}p(y)\,\mathrm{d}p(z) = \mathrm{d}N$$

is given by

$$g_e(v) = \frac{2\,\mathrm{d}N}{h^3} = \frac{2}{h^3}\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \cdot \mathrm{d}p(x)\,\mathrm{d}p(y)\,\mathrm{d}p(z)$$

where h is Planck's constant and the factor 2 arises because the electron has two possible spin states. The statistical weight per unit volume is thus

$$\frac{2}{h^3}\mathrm{d}p(x)\,\mathrm{d}p(y)\,\mathrm{d}p(z)$$

for a single electron. However, there may be other free electrons, from other ions, which occupy some of the available states in the element of phase space $\mathrm{d}N$. If the number density of electrons is n_e then the effective volume available to a collisionally ejected electron is reduced by a factor $1/n_\mathrm{e}$. Thus the statistical weight available to a single free electron is

$$g_e(v) = \frac{2\,\mathrm{d}p(x)\,\mathrm{d}p(y)\,\mathrm{d}p(z)}{n_\mathrm{e}h^3}.$$

Furthermore, if the velocity field is isotropic, the 'momentum volume' can be replaced simply by its counterpart in spherical coördinates,

$$\mathrm{d}p(x)\,\mathrm{d}p(y)\,\mathrm{d}p(z) = 4\pi p^2\,\mathrm{d}p$$

Using these results we can write eqtn. (D.3) as

$$\frac{n_\mathrm{e} n_{2,1}(v)}{n_{1,i}} = \frac{2g_i}{h^3 g_{1,i}} \exp\left\{\frac{-\left(\chi_{1,i} + \frac{1}{2}m_\mathrm{e}v^2\right)}{kT}\right\} 4\pi p^2\,\mathrm{d}p$$

but the momentum $p = m_\mathrm{e}v$; i.e., $\frac{1}{2}m_\mathrm{e}v^2 = p^2/(2m_\mathrm{e})$, whence

$$\frac{n_\mathrm{e} n_{2,1}(v)}{n_{1,i}} = \frac{2g_2}{h^3 g_{1,i}} \exp\left\{\frac{-\chi_{1,i}}{kT}\right\} \exp\left\{\frac{-p^2}{2m_\mathrm{e}kT}\right\} 4\pi p^2\,\mathrm{d}p.$$

Since we're interested in the ionization balance (not the velocity distribution of the ionized electrons), we integrate over velocity to obtain the total number of daughter ions:

$$\frac{n_\mathrm{e} n_{2,1}}{n_{1,i}} = \frac{2g_2}{h^3 g_{1,i}} \exp\left\{\frac{-\chi_{1,i}}{kT}\right\} 4\pi \int_0^\infty p^2 \exp\left\{\frac{-p^2}{2m_\mathrm{e}kT}\right\} \mathrm{d}p.$$

We can then use result of a standard integral,

$$\int_0^\infty x^2 \exp\left(-a^2x^2\right) \mathrm{d}x = \sqrt{\pi}/\left(4a^3\right)$$

to obtain

[1] The statistical weight is a form of probability, and the probability of 'A and B', P(A+B), is the product P(A)P(B).

$$\frac{n_e n_{2,1}}{n_{1,i}} = \frac{2g_2}{g_{1,i}} \exp\left\{\frac{-\chi_{1,i}}{kT}\right\} \frac{(2\pi m_e kT)^{3/2}}{h^3} \tag{D.2}$$

This is one common form of the Saha Equation (often expressed in terms of the ground state of the parent ion, $n_{1,1}$).

D.2 Partition functions

The version of the Saha equation given in eqtn. (D.2) relates populations in single states of excitation for each ion. Generally, we are more interested in the ratios of number densities of different ions summed over all states of excitation – i.e., the overall ionization balance. We determine this by defining the partition function as

$$U = \sum_n g_n \exp\left(-E_n/kT\right)$$

(an easily evaluated function of T), whence

$$\frac{n_e n_2}{n_1} = \frac{2U_2}{U_1} \frac{(2\pi m_e kT)^{3/2}}{h^3} \exp\left\{\frac{-\chi_1}{kT}\right\} \tag{D.4}$$

where we use χ_1, the ground-state ionization potential of the parent atom, as it is to this that the partition function is referred (i.e., $E_{1,1} \equiv 0$).

Since the electron pressure is $P_e = n_e kT$ we can also express the Saha equation in the form

$$\frac{n_2}{n_1} = \frac{2U_2}{U_1} \frac{(2\pi m_e)^{3/2}}{h^3} \frac{(kT)^{5/2}}{P_e} \exp\left\{\frac{-\chi_1}{kT}\right\} \tag{D.5}$$

D.2.1 An illustration: hydrogen

The Balmer lines of hydrogen, widely observed as absorption lines in stellar spectra, arise through photoexcitation from the $n = 2$ level of neutral hydrogen. To populate the $n = 2$ level, we might suppose that we need temperatures such that $kT \simeq E_{1,2} = 10.2$eV; i.e., $T \simeq 10^5$K. However, the Hα line strength peaks in A0 stars, which are much cooler than this ($T \sim 10^4$K). Why? Because we need to consider ionization as well as excitation. We therefore need to combine the Saha and Boltzmann equations to obtain the density of atoms in a given state of excitation, for a given state of ionization.

We express the Boltzmann equation, eqtn. (D.1), in terms of the partition function U:

$$\frac{n_{1,2}}{n_1} = \frac{g_{1,2}}{U_1} \exp\left(\frac{-E_{1,2}}{kT}\right)$$

where n_1 is the number density of H^0 atoms in all excitation states and $E_{1,2}$ is the excitation energy of the $n = 2$ level (10.2 eV); that is,

$$n_{1,2} = \frac{g_{1,2}}{U_1} \exp\left(\frac{-E_{1,2}}{kT}\right) n_1.$$

However, the total number of hydrogen nuclei is $n(\mathrm{H}) = n_1 + n_2 = n_1(1 + {}^{n_2}/_{n_1})$; that is, $n_1 = n(\mathrm{H})(1 + {}^{n_2}/_{n_1})^{-1}$. Using this, and ${}^{n_2}/_{n_1}$ from eqtn. (D.4), we find

$$\begin{aligned} n_{1,2} &= \frac{g_{1,2}}{U_1} \exp\left(\frac{-E_{1,2}}{kT}\right)\left(1 + \left[\frac{2U_2}{n_1 U_1}\frac{(2\pi m_\mathrm{e} kT)^{3/2}}{h^3} \exp\left\{\frac{-\chi_1}{kT}\right\}\right]\right)^{-1} n(\mathrm{H}) \\ &= \frac{g_{1,2}}{U_1} \exp\left(\frac{-E_{1,2}}{kT}\right)\left(1 + \left[\frac{2U_2}{n_1 U_1}\frac{(2\pi m_\mathrm{e})^{3/2}}{h^3}\frac{(kT)^{5/2}}{P_\mathrm{e}} \exp\left\{\frac{-\chi_1}{kT}\right\}\right]\right)^{-1} n(\mathrm{H}) \end{aligned}$$

We can now see why the Balmer lines peak around 10^4K: while higher temperatures give larger populations $n_{1,2}/n(\mathrm{H}^0)$, they give smaller populations $n(\mathrm{H}^0)/n(\mathrm{H})$. The overall result is that $n_{1,2}/n(\mathrm{H})$ peaks around 10kK.

The Saha equation also gives an explanation of why supergiant stars are cooler than main-sequence stars of the same spectral type. Spectral types are defined by ratios of lines strengths; e.g., O-star subtypes are defined by the ratio (He II λ4542)/(He I λ4471), which in turn traces the ratio $\mathrm{He}^+/\mathrm{He}^0$. Of course, higher temperatures increase the latter ratio. However, a supergiant star has a lower surface gravity (and atmospheric pressure) than a main-sequence star. From eqtn. (D.5) we see that a lower pressure at the same temperature gives rise to a larger ratio n_2/n_1 (essentially, the rate of recombinations is lower), so for two stars of the same temperature, the supergiant has an earlier spectral type (or, equivalently, at the same spectral type the supergiant is cooler).

Appendix E

Magnitude systems

The brightness of astronomical sources is almost universally measured in 'magnitudes'. Practising astronomers are so familiar with this system that they often overlook its idiosyncracies: it has no units, runs the 'wrong way' (brighter sources have smaller magnitudes), and is logarithmic in nature. These oddities arise for historical reasons, dating back to when Hipparchus and Ptolemy first quantified stellar brightnesses. They assigned the brightest stars to 'the first magnitude' (the handful of very brightest stars are assigned negative magnitudes today), and sequentially fainter stars larger magnitudes; the faintest stars visible to the naked eye under good conditions are around magnitude 6.

The magnitude scale in modern use dates back to the 1850s, when N.R. Pogson proposed that a magnitude *difference* of 5 magnitudes should correspond to a brightness (flux) *ratio* of exactly 100, whence

$$m_\lambda = -2.5 \log F_\lambda + C_\lambda. \tag{E.1}$$

The normalizing constant C_λ has to be determined observationally, by measuring the flux for stars of known magnitude. This is a challenging task, not only because of the difficulty in measuring the fluxes, but also because 'the' magnitude of a star is also an observed quantity; it is not only subject to to measurement errors, but is also fundamentally arbitrary. For example, Pogson initially calibrated his system by choosing to define Polaris as being exactly magnitude 2.0; we now know that Polaris is slightly variable in brightness, and modern systems typically derive ultimately from a standard whereby Vega was assigned magnitude 0.0.[1]

Variable; colour dependence.

Different observers have different colour sensitivities; and even a single observer has different colour sensitivities as a function of time and brightness (the Purkinje effect).

[1] Vega is too bright to be directly observed with modern instruments on large telescopes, and subsequent refinements include development of network of standards. Within that framework, 'the' magnitude of every star becomes an observational quantity; Vega has $V \simeq 0.03$.

Magnitudes also expressed in frequency (AB magnitudes; Oke & Gunn 1983)

Appendix F

Initial Mass Function

The Initial Mass Function (IMF) describes the distribution of stellar masses at birth. Observationally, this has to be derived indirectly; typically, the present-day luminosity function (PDLF; the number of stars as a function of luminosity) is converted to a present-day mass function (PDMF), from which the IMF is inferred. (Of course, the PDMF differs from the IMF, if only because high-mass stars 'burn out' quickly and disappear from the PDMF while low-mass stars live on.) The first step is subject to numerous difficulties in assembling suitable datasets and in understand biasses and selection effects; while subsequent stages in the analysis require evolutionary models that take into account binary evolution and a host of other factors.

F.1 Formulating the IMF

We define the initial mass spectrum (or mass number function), $\Phi(M)$, such that $\Phi(M)\,\mathrm{d}M$ is the probability of creating a star in the mass range M_1 to M_2; that is, the number of stars born with masses between M and $M + \mathrm{d}M$ is

$$N(M_1, M_2) = \int_{M_1}^{M_2} \Phi(M)\,\mathrm{d}M$$

(modulo some overall scaling factor). Differentiating emphasizes that this mass spectrum is just the number of stars per unit mass interval:

$$\frac{\mathrm{d}N}{\mathrm{d}M} = \Phi(M).$$

As well of the *number* of stars between some mass limits, it's often convenient to think in terms of $\mathfrak{M}$, the total stellar *mass* within those limits. We get this simply by multiplying the number of stars by the

mass per star:

$$\mathfrak{M}(M_1, M_2) = \int_{M_1}^{M_2} M\Phi(M)\,\mathrm{d}M \equiv \int_{M_1}^{M_2} \xi(M)\,\mathrm{d}M$$

where the function $\xi(M)$ is commonly cited as 'the' Initial Mass Function. Although $\xi(M)$ represents the initial mass distribution, it can readily be related back to the number of stars:

$$\xi(M) \;=\; M\Phi(M) \;=\; M\,\frac{\mathrm{d}N}{\mathrm{d}M} \;=\; M\frac{\mathrm{d}(\ln M)}{\mathrm{d}M}\,\frac{\mathrm{d}N}{\mathrm{d}(\ln M)} \;=\; \frac{\mathrm{d}N}{\mathrm{d}(\ln M)}$$

Warning! The symbol ξ is almost always employed in discussions of the mass function, but may be used to refer the mass number function, here designated as Φ.

F.2 Representation

The seminal study of the IMF was published by Edwin Salpeter,[1] who found that, empirically, the IMF can be reasonably well approximated by a power law,

$$\xi(M) \simeq \xi_0 M^{-\Gamma}$$

(where the constant ξ_0 simply reflects the total number – or mass – of stars under consideration). With this formulation, the mass spectrum is just

$$\begin{aligned}\Phi(M) &= \frac{\xi(M)}{M} = \xi_0 M^{-\Gamma-1} \\ &\equiv \Phi_0 M^{-\alpha}.\end{aligned}$$

The 'Salpeter IMF', originally established over the mass range $0.4 \lesssim M/\mathrm{M}_\odot \lesssim 10$, has $\Gamma = 1.35$ ($\alpha = 2.35$). Subsequent studies suggest that little revision this power-law index is required, and also indicate that it is a reasonable approximation to higher masses than Salpeter was originally able to study. (For lower-mass stars and brown dwarfs the slope appears to flatten, to $\Gamma \simeq 0.3$.) Moreover, the IMF appears remarkably invariant in a wide range of environments (e.g., high or low density, or high or low metallicity); an extensive review[2] concluded that "there is no clear evidence that the IMF varies strongly and systematically as a function of initial conditions", although others may dispute that conclusion.

The Salpeter IMF is therefore a satisfactory approximation for many purposes, and certainly for order-of-magnitude estimates. For more detailed work, a 'broken power law' may be used; a formulation by Kroupa is widely employed.

[1] 1955ApJ...121..161S

[2] Bastian, Covey & Meyer; 2010ARA&A..48..339B.

F.3 Some implications

The number of stars between any two mass limits is

$$\begin{aligned} N(M_1, M_2) &= \int_{M_1}^{M_2} \Phi_0 M^{-\alpha} \, \mathrm{d}M, \\ &= \frac{\xi_0}{1.35} \left(M_1^{-1.35} - M_2^{-1.35} \right) \end{aligned}$$

for a Salpeter IMF; and the integrated mass within these limits is

$$\begin{aligned} \mathfrak{M}(M_1, M_2) &= \int_{M_1}^{M_2} \xi_0 M^{-\Gamma} \, \mathrm{d}M \\ &= \frac{\xi_0}{0.35} \left(M_1^{-0.35} - M_2^{-0.35} \right); \end{aligned}$$

the negative exponent shows that the total mass is dominated by low-mass stars (for any $\Gamma > 1$).

For a mass–luminosity relationship $L \propto M^\beta$, the integrated luminosity over the same mass range is

$$\begin{aligned} \mathfrak{L}(M_1, M_2) &\propto \int_{M_1}^{M_2} M^\beta M^{-\Gamma} \, \mathrm{d}M, \\ &\propto M_2^{2.15} - M_1^{2.15} \end{aligned}$$

for a Salpeter IMF and $\beta \simeq +3.5$ (§12.4); that is, in contrast to the mass, the luminosity is dominated by high-mass stars (for any $\beta > \Gamma$).

PHAS 0018/2112 PART V: STARS – II

Appendix G

Nuclear reactions in stars

G.1 Introduction

In general, nuclear processes in stars involve fission of a nucleus, or (more usually in 'normal' evolutionary phases), the fusion of two nuclei. Through all these processes, key physical quantities are conserved:

- the baryon number (the number of protons, neutrons, and their antiparticles);
- the lepton number (electrons, neutrinos, related light particles, and their antiparticles);
- charge; and
- total mass–energy.

Consider two types of nuclei, A and B, number densities $n(A)$, $n(B)$. The rate at which a particular (nuclear) reaction occurs between particles moving with relative velocity v is

$$r(v) = n(A)\, n(B)\, v\, \sigma(v) \tag{G.1}$$

(per unit volume per unit time) where $\sigma(v)$ is the cross-section for the reaction. Of course, we need to integrate over velocity to get the total reaction rate:

$$\begin{aligned} r &= n(A)\, n(B) \int v\, \sigma(v)\, f(v)\, \mathrm{d}v \\ &\equiv n(A)\, n(B)\, \langle \sigma(v)\, v \rangle \qquad [\ \mathrm{m}^{-3}\ \mathrm{s}^{-1}] \end{aligned} \tag{G.2}$$

where $f(v)$ is the (Maxwellian) velocity distribution, and the angle brackets denote a weighted average (i.e., the integral in the first part of eqtn. G.2).

Since the reaction destroys A (and B), we have

$$\frac{\partial n(A)}{\partial t} = -n(A)\, n(B)\ \langle \sigma(v)\, v \rangle\,; \tag{G.3}$$

and the number density of species A falls with time as

$$n(A,t) = n_0(A)\ \exp\{-n(B)\ \langle \sigma(v)\, v \rangle\, t\} \tag{G.4}$$

which defines a characteristic (e-folding) timescale

$$\tau = \frac{1}{n(B)\ \langle \sigma(v)\, v \rangle}. \tag{G.5}$$

Finally, the total energy generated through this reaction, per unit mass per unit time, is

$$\begin{aligned} \varepsilon &= \frac{Q\, r}{\rho} \\ &= \frac{n(A)\, n(B)}{\rho} Q\, \langle \sigma(v)\, v \rangle \qquad [\mathrm{J\ kg^{-1}\ s^{-1}}] \end{aligned} \tag{G.6}$$

where Q is the energy produced per reaction and ρ is the mass density.

G.2 Tunnelling

Charged nuclei experience Coulomb repulsion at intermediate separations, and nuclear attraction at small separations. In stellar cores the high temperatures give rise to high velocities, and increased probability of overcoming the Coulomb barrier. For nuclear charge Z (the atomic number), the energy needed to overcome the Coulomb barrier is

$$E_\mathrm{C} \simeq \frac{Z_1 Z_2 e^2}{r_0} \tag{G.7}$$

$$(\simeq 2 \times 10^{-13}\ \mathrm{J}, \qquad \simeq 1\ \mathrm{MeV, for}\ Z_1 = Z_2 = 1) \tag{G.8}$$

where $r_0 \simeq 10^{-15}$m is the radius at which nuclear attraction overcomes Coulomb repulsion for proton pairs.

In the solar core, $T_\mathrm{c} \sim 1.5 \times 10^7$K; that is, $E(= {}^3\!/_2 kT) \simeq$ keV, or $\sim 10^{-3} E_\mathrm{C}$. This energy is only sufficient to bring protons to within $\sim 10^3 r_0$ of each other; this is much too small to be effective, so reactions only occur through a process of "quantum tunneling" (barrier penetration). In this temperature regime the rate of nuclear energy generation is well approximated by a power-law dependence on temperature,

$$\varepsilon \simeq \varepsilon_0 \rho T^\alpha \tag{G.9}$$

where $\alpha \simeq 4.5$ for proton-proton reactions in the Sun [Section G.4; $\varepsilon_0 \propto n^2(\mathrm{H})$], and $\alpha \simeq 18$ for CN processing [Section G.5; $\varepsilon_0 \propto n(\mathrm{H}) n(\mathrm{C,N})$].

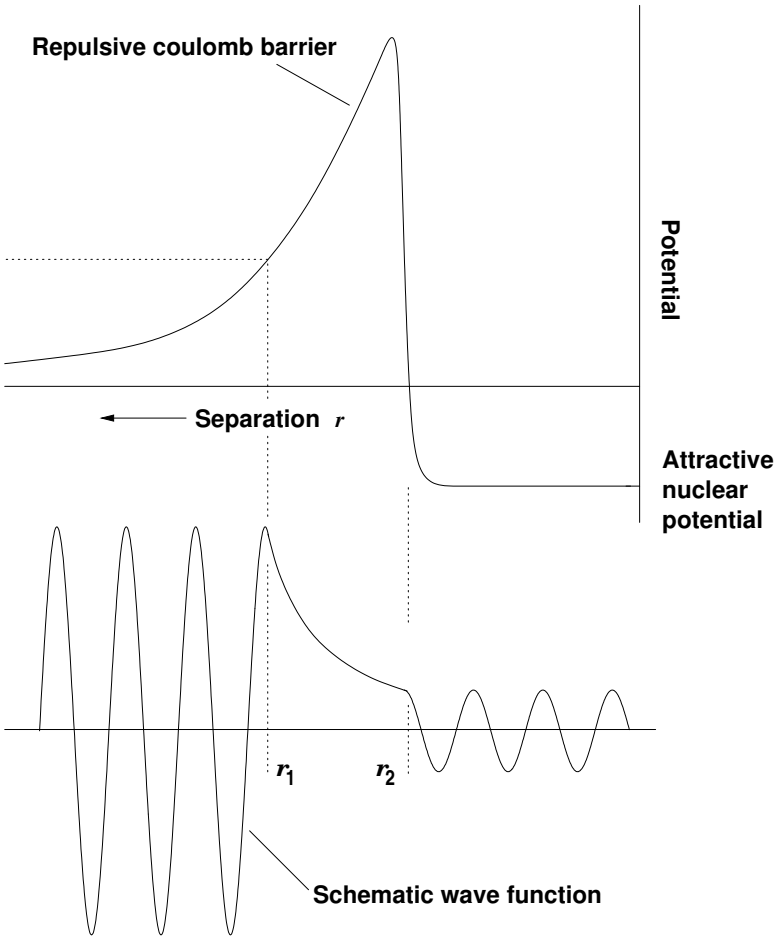


Figure G.1: Upper section: a schematic plot of the potential between two charged nucleons as a function of separation. At 'large' separations ($\gtrsim 10^{-15}$ m), the repulsive Coulomb force is given by eqtn. (G.8); classically, particles cannot come closer than the point r_1 at which the relative kinetic energy corresponds to the repulsive potential. Quantum-mechanical tunneling allows the nucleons to approach closer, to separation r_2, at which point the strong nuclear force dominates.
The lower panel expresses this tunnelling schematically. The (square of the) amplitude of the wave function is a measure of the probability of a particle being in a particular location; the amplitude of the wave function decreases exponentially between r_1 and r_2, but does not fall to zero. (See Aside 7.1 for further details.)

[Note that eqtn. (G.9) characterizes the rate of energy generation per unit *mass* (or, if you like, per nucleon). Although density appears here as a simple linear multiplier, reference to eqtn. G.6 reminds us that, like nearly all 'collisional' processes, the energy generation rate per unit *volume* – or the probability of a given nucleus undergoing fusion – depends on density *squared*.]

Aside 7.1: The Gamow Peak

As illustrated in Fig. G.1, 'tunnelling' can occur to allow fusion to occur at particle energies which classical mechanics would indicate to be too low to overcome the Coulomb barrier. For higher temperatures (and larger kinetic energies), particles will come closer together (r_1 approaches r_2), the decay of the wave function is reduced, and so the amplitude of the wave function in the region $r < r_2$ becomes larger – that is, the tunnelling probability increases as the kinetic energy of the incoming nucleus

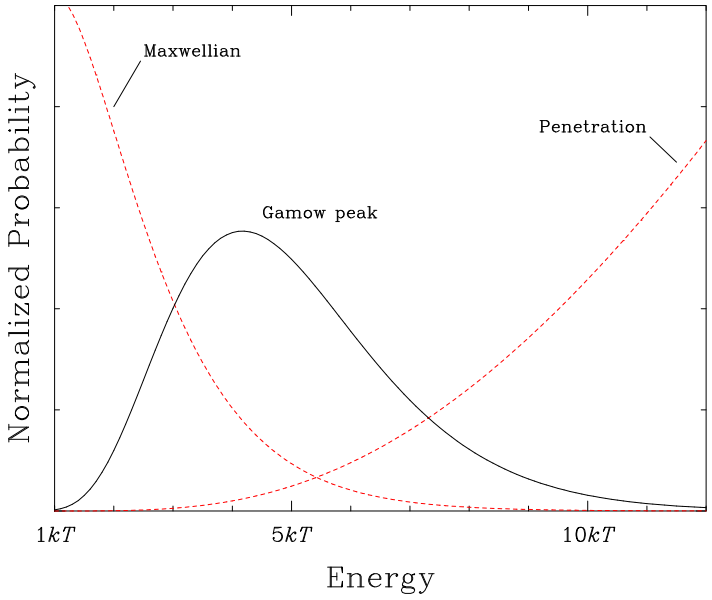


Figure G.2: The main energy-dependent factors determining two-body reaction rates are the numbers of reagents as a function of energy (the Maxwellian velocity distribution) and the tunnelling probability of penetration. The product of these two terms gives the energy distribution of particles participating in fusion. These factors are illustrated here, on arbitrary vertical scales, for the fusion of two protons in the solar core (Gamow energy $E_G = 290kT$ for $T = 2\times10^7$ K; $E_0 = 4.2kT$, $1/e$ width $\Delta = 4.8kT$). See Aside 7.1.

increases.

Obtaining the probability of barrier penetration, p_p, for given energy, is a standard problem in wave mechanics. We simply quote the result that the probability of penetration varies exponentially with the ratio of kinetic energy to barrier size,

$$p_p \propto \exp\left\{-\left(\frac{E_G}{E}\right)^{1/2}\right\} \tag{A-7.1}$$

with the 'Gamow energy' E_G (unnamed and written as b^2 in some sources) given by

$$E_G = 2m_R c^2 (\pi\alpha Z_1 Z_2)^2 \quad (= 493 \text{ keV for proton-proton fusion}), \tag{A-7.2}$$

where α is the fine structure constant,

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \simeq \frac{1}{137}. \tag{A-7.3}$$

and m_R is the 'reduced mass',

$$m_R = \frac{m_1 m_2}{m_1 + m_2}$$

for particles of mass m_1, m_2 ($\simeq A_1 m(\mathrm{H})$, $A_2 m(\mathrm{H})$) of charge Z_1, Z_2. (Using the reduced mass means that velocities and kinetic energies are measured with reference to the centre of mass of the particles involved.)

The fusion cross-section $\sigma(v)$ (eqtn G.1) is evidently dependent on this penetration probability. We also expect it to depend on the effective size, or 'target area', of the particles; this geometrical factor is proportional to $\pi\lambda^2$, where λ is the de Broglie wavelength, $\lambda^2 \propto 1/E$. The intrinsic properties of the nuclei must also be involved; these will be constant, or slowly varying functions of energy, in most circumstances (although resonances may occur). We therefore write the total reaction cross-section in the form

$$\sigma(E) = \frac{S(E)}{E} \exp\left\{-\left(\frac{E_G}{E}\right)^{1/2}\right\} \tag{A-7.4}$$

where $S(E)$ encapsulates the nuclear physics of the actual fusion process.

At any given temperature, the number of particles in a Maxwellian velocity distribution falls off exponentially with increasing energy (eqtn. **??**); that is, the probability of encountering a particle with energy E at kinetic temperature T is

$$f(E)\,\mathrm{d}E = \frac{2}{\sqrt{\pi}}\frac{E}{kT}\exp\left\{-\frac{E}{kT}\right\}\frac{\mathrm{d}E}{(kTE)^{1/2}} \tag{A-7.5}$$

These two competing factors – the increasing probability of penetration with increasing energy (eqtn. A-7.1) and the decreasing number of particles with increasing energy (eqtn. A-7.5) – mean that there is a limited range of energies at which most reactions occur. This is illustrated in Fig. G.2; the product of the two exponential terms leads to the 'Gamow peak', where the probability of fusion occuring is at a maximum.[1]

To explore this in greater detail, we write the reaction rate per particle pair, eqtn. G.2, as

$$\langle\sigma(v)\,v\rangle = \int_0^\infty \sigma(E) v f(E)\mathrm{d}E$$

where $\sigma(E)$, v are particle cross-sections and velocities at energy E; from eqtns. (A-7.4) and (A-7.5), and using $E = \frac{1}{2}m_\mathrm{R}v^2$,

$$\langle\sigma(v)\,v\rangle = \int_0^\infty \frac{S(E)}{E}\exp\left\{-\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right\}\sqrt{\frac{2E}{m_\mathrm{R}}}\frac{2}{\sqrt{\pi}}\frac{E}{kT}\exp\left\{-\frac{E}{kT}\right\}\frac{\mathrm{d}E}{(kTE)^{1/2}} \tag{A-7.6}$$

$$= \left(\frac{8}{\pi m_\mathrm{R}}\right)^{1/2}\frac{1}{(kT)^{3/2}}\int_0^\infty S(E)\exp\left\{-\frac{E}{kT}-\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right\}\mathrm{d}E \tag{A-7.7}$$

at some fixed temperature T. Eqtn. (A-7.7) is the integral over the Gamow peak; the larger the area, the greater the reaction rate.

The Gamow peak is appropriately named in that it is indeed quite strongly peaked; it is therefore a reasonable approximation to take the $S(E)$ term as locally constant. In that case, the integrand peaks at energy E_0, when

$$\frac{\mathrm{d}}{\mathrm{d}E}\left\{\frac{E}{kT}+\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right\} = \frac{1}{kT}-\frac{1}{2}\left(\frac{E_\mathrm{G}}{E_0^3}\right)^{1/2} = 0;$$

i.e.,

$$E_0 = \left(\frac{kT\sqrt{E_\mathrm{G}}}{2}\right)^{2/3}. \tag{A-7.8}$$

$$= \left[\sqrt{2}(\pi\alpha kc)^2 m_\mathrm{R}(Z_1Z_2T)^2\right]^{1/3}$$

E_0, the location of the Gamow peak, is the most effective energy for thermonuclear reactions; it greatly exceeds kT, the typical thermal energy, but falls well below the Gamow energy of the Coulomb barrier.

There is no simple analytical solution for the width of the peak, but one common (and reasonable) approach is to approximate the exponential term in the integral (eqtn. A-7.7) with a gaussian centred on E_0. Conventionally, in this context 'the' width is not characterized by the gaussian 'σ' parameter, but rather by Δ, the full width at $1/e$ of the peak value (so $\Delta \equiv 2\sqrt{2}\sigma$); thus we need to solve for

$$\exp\left\{-\frac{E}{kT}-\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right\} \simeq C\exp\left\{-\left(\frac{E-E_0}{\Delta/2}\right)^2\right\}. \tag{A-7.9}$$

Requiring the two sides to be equal at $E = E_0$ we immediately find

$$C = \exp\left\{-\frac{E_0}{kT}-\left(\frac{E_\mathrm{G}}{E_0}\right)\right\},$$

$$= \exp\left\{-\frac{3E_0}{kT}\right\} \qquad \text{(from eqtn. A-7.8)}$$

[1] Clearly, the area under the Gamow peak determines the total reaction rate.

while requiring the curvatures (second derivatives) on either side of eqtn. A-7.9 to be equal gives, after some algebra,

$$\Delta = \sqrt{\frac{16}{3} E_0 kT}.$$

The total reaction rate depends on the integrated area under the Gamow peak; again using a gaussian approximation to the peak, and constant $S(E)$ across the peak, then from eqtn. (A-7.7), we have

$$\langle \sigma(v)\, v \rangle = \left(\frac{8}{\pi m_R}\right)^{1/2} \frac{S(E_0)}{(kT)^{3/2}} \exp\left\{-\frac{3E_0}{kT}\right\} \int_0^\infty \exp\left\{-\left(\frac{E - E_0}{\Delta/2}\right)^2\right\} \mathrm{d}E, \qquad \text{(A-7.10)}$$

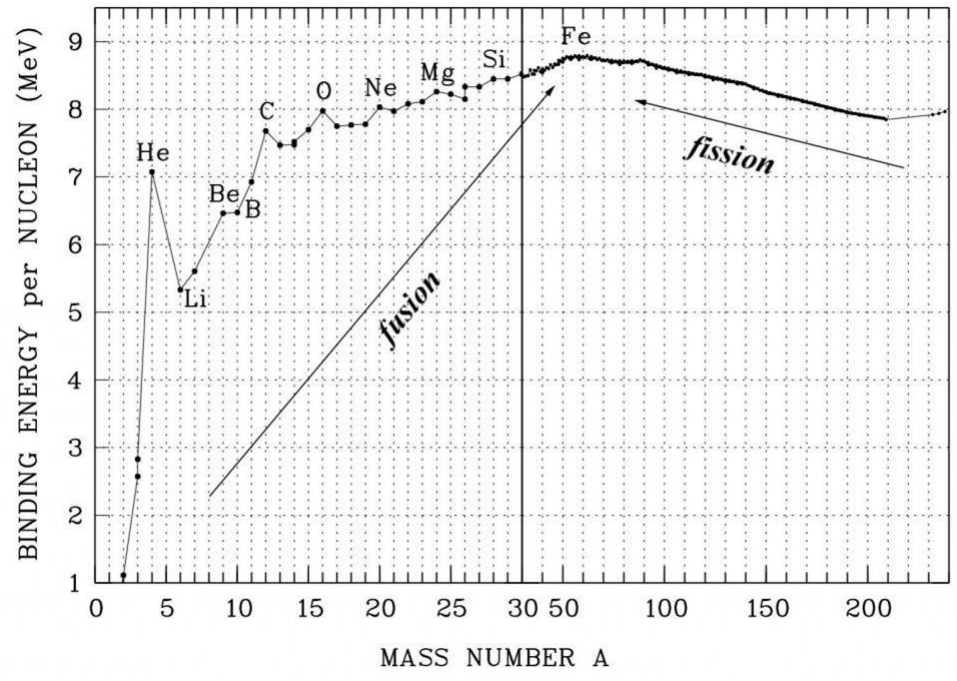
$$\simeq \left(\frac{8}{\pi m_R}\right)^{1/2} \frac{S(E_0)}{(kT)^{3/2}} \exp\left\{-\frac{3E_0}{kT}\right\} \frac{\Delta\sqrt{\pi}}{2}$$

(where, in order to perform the integration analytically, the limits have been extended from $0/+\infty$ to $-\infty/+\infty$; the error thus introduced is negligible provided that $E_0 > \Delta/2$). Bowers & Deeming give a mathematical development from this point which leads to a demonstration that $\varepsilon \simeq \varepsilon_0 \rho T^\alpha$ (eqtn. G.9).

Furthermore, substituting eqtn. (A-7.8) into eqtn. (A-7.7) we obtain

$$\langle \sigma\, v \rangle \propto \exp[-(E_G/kT)^{1/3}].$$

G.3 The mass defect and nuclear binding energy



The mass of any nucleus is less than the sum of the separate masses of its protons and neutrons. The *binding energy* of a particular isotope is the energy corresponding to the 'missing' mass (or *mass defect*), and is the energy produced in forming that isotope from its raw ingredients; equivalently, it is

the amount of energy needed to break it up into protons and neutrons.[2] The binding energy peaks in the iron group, with ^{62}Ni the most tightly-bound nucleus, followed by ^{58}Fe and ^{56}Fe;[3] this is the basic reason why iron and nickel are very common metals in planetary cores, since they are produced as end products in supernovae.

For atomic masses $A \gtrsim 60$, energy release is through fission (generally involving much less energy).

For a nucleus with Z protons, $N(= A - Z)$ neutrons, and mass $m(Z, N)$ the binding energy is therefore

$$Q(Z, N) = \left[Z m_{\rm p} + N m_{\rm n} \; - \; m(Z, N) \right] \, c^2 \tag{G.10}$$

(where $m_{\rm p}$, $m_{\rm n}$ are the proton, neutron masses), and the binding energy per baryon is

$$Q(Z, N)/(Z + N).$$

Converting 'MeV per baryon' to 'J kg^{-1}', we find that burning protons into helium yields

$$\text{H} \rightarrow \text{He:} \qquad 6.3 \times 10^{14} \text{ J kg}^{-1}$$

but

$$\text{H} \rightarrow \text{Fe:} \qquad 7.6 \times 10^{14} \text{ J kg}^{-1};$$

that is, burning H to He alone releases 83% of the total nuclear energy available per nucleon.

Physical processes

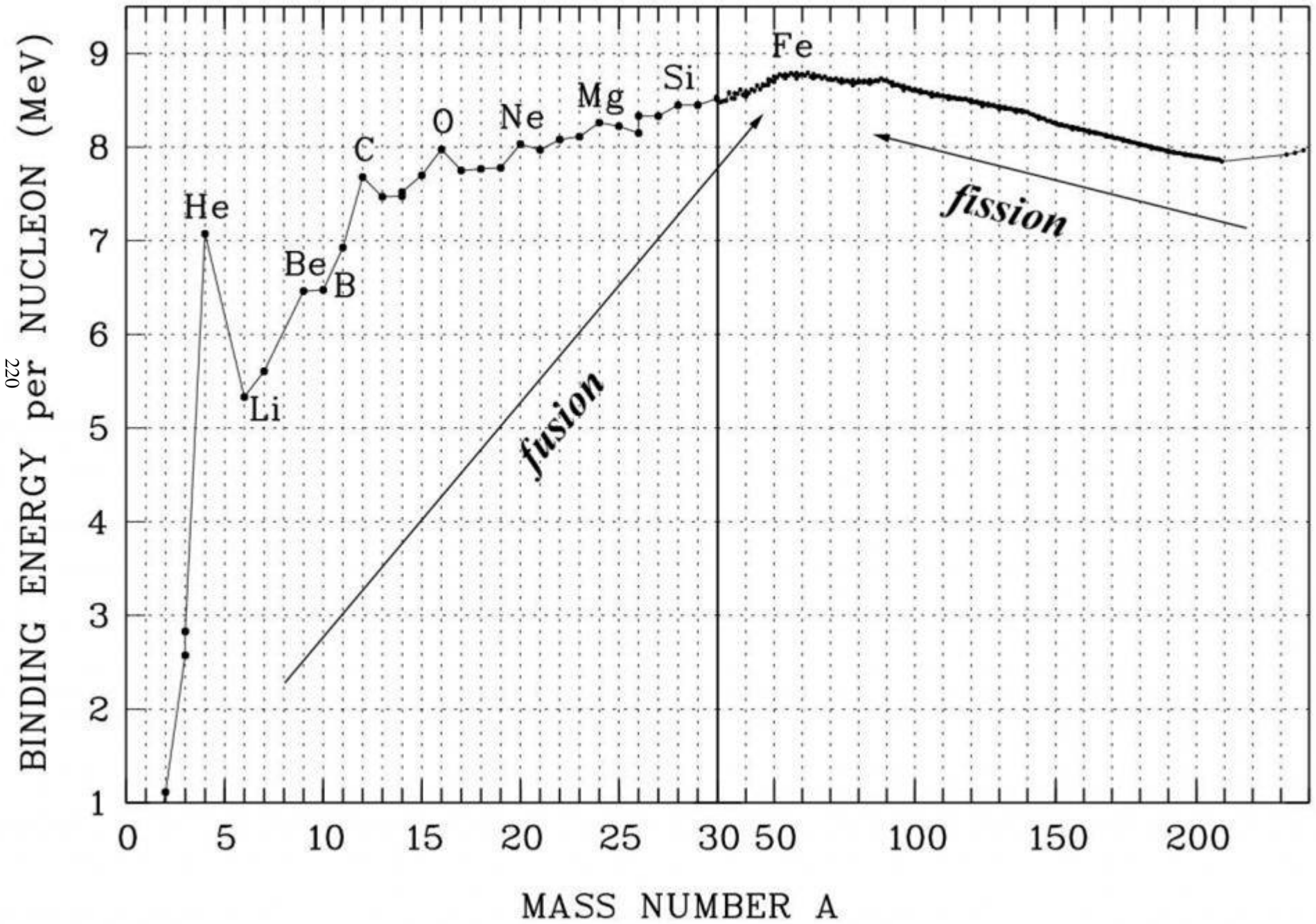
To do –

Nuclear models (liquid-drop, shell)

Line of stability (neutron, proton drip lines)

[2] The binding energy explains why the masses of the proton and neutron are both larger than the 'atomic mass unit', or amu; the amu is defined to be $^1/_{12}$ the mass of ^{12}C, but each nucleon in that isotope has given up almost 1% of its mass in binding energy.

[3] Many sources cite ^{56}Fe as the most tightly bound nucleus; see M.P. Fewell, Am.J.Phys., 63, 653, 1995 for a discussion which lays the blame for this misconception squarely at the door of astrophysicists!


BINDING ENERGY per NUCLEON (MeV)
MASS NUMBER A
1
2
3
4
5
6
7
8
9
0
5
10
15
20
25
30
50
100
150
200
He
Li
Be
B
C
O
Ne
Mg
Si
Fe
fusion
fusion
fission

G.4 Hydrogen burning – I: the proton–proton (PP) chain

G.4.1 PP–I

Step	Process			Energy Release		Solar Timescale
(1)	$p + p$	$\rightarrow$	$^2D + e^+ + \nu_e$	1.44 MeV	†	7.9×10^9 yr
(2)	$^2D + p$	$\rightarrow$	^3He	5.49 MeV		1.4s
				6.92 MeV	×2	
(3a)	$^3He + ^3He$	$\rightarrow$	$^4He + p + p$	12.86 MeV		2.4×10^5 yr
				26.72 MeV		

†Includes 1.02 MeV from $e^+ + e^- \rightarrow 2\gamma$

Reaction (1) is very slow because it involves the *weak interaction,*[4] which is required to operate during the short period when protons are close together.

Reactions (2) and (3a) involve the *strong interaction* and in consequence are much faster.

[Note that reaction (3a) is preferred to

$$^3He + p \quad \rightarrow \quad ^4He + e^+ + \nu_e,$$

even though protons vastly outnumber ^{3}He particles, because this again involves the weak interaction (the ν_e is the giveaway).]

Reaction (1) occurs twice for each ^{4}He production, each time generating an electron neutrino with energy 0.26 MeV. These leave the Sun without further interaction, so the energy available for heating is 26.2 MeV ($26.72 - 2 \times 0.26$ MeV).

G.4.2 PP–II, PP–III

There are two principal secondary channels in the proton-proton chain, each catalysed by a pre-existing α particle (^{4}He nucleus):

PP–II (follows steps 1 & 2, which yield 6.92 MeV):

[4] i.e., involves β decay; in this case β^+ decay, $p^+ \rightarrow n^0 + e^+ + \nu_e$ (cp. β^- decay, $n^0 \rightarrow p^+ + e^- + \overline{\nu}_e$). To go on to form a deuterium, this decay must occur within $\sim 10^{-15}$ m of another proton; that is, step (1) is itself actually a two-step process.

Step	Process			Energy Release		Solar Timescale
(1)	$p + p$	$\rightarrow$	$^2D + e^+ + \nu_e$	1.44 MeV	†	7.9×10^9 yr
(2)	$^2D + p$	$\rightarrow$	^3He	5.49 MeV		1.4s
(3b)	$^3He + ^4He$	$\rightarrow$	^7Be	1.59 MeV		9.2×10^5 yr
(4b)	$^7Be + e^-$	$\rightarrow$	$^7Li + \nu_e$	0.86 MeV		0.39 yr
(5b)	$^7Li + p$	$\rightarrow$	$^4He + ^4He$	17.35 MeV		570s
				26.72 MeV		

†Includes 1.02 MeV from $e^+ + e^- \rightarrow 2\gamma$

In this case, neutrino losses average 0.80 MeV.

PP–III (follows steps 1, 2, and 3b):

Step	Process			Energy Release		Solar Timescale
(1)	$p + p$	$\rightarrow$	$^2D + e^+ + \nu_e$	1.44 MeV	†	7.9×10^9 yr
(2)	$^2D + p$	$\rightarrow$	^3He	5.49 MeV		1.4s
(3*b*)	$^3He + ^4He$	$\rightarrow$	^7Be	1.59 MeV		9.2×10^5 yr
(4*c*)	$^7Be + p$	$\rightarrow$	^8B	0.14 MeV		66 yr
(5*c*)	^8B	$\rightarrow$	$^8Be^* + e^+ + \nu_e$	16.04 MeV	†	1 s
(6*c*)	$^8Be^*$	$\rightarrow$	$^4He + ^4He$	3.30 MeV		10^{-16} s
				26.72 MeV		

†Includes 1.02 MeV from $e^+ + e^- \rightarrow 2\gamma$

Neutrino losses here are 7.2 MeV on average, predominantly through step (5*c*).[5]

In the Sun, ~91% of reactions go through (3*a*); ~9% end at (5*b*); and ~0.1% end at (6*c*).

G.5 Hydrogen burning – II: the CNO cycle

Because the first reaction in the PP chain is so slow (7.9×10^9 yr), under certain circumstances it is possible for reactions involving (much less abundant) heavier nuclei, acting as catalysts, to proceed faster than PP. The larger charges (and masses) of these heavier particles imply that higher temperatures

[5]It is the high-energy neutrinos from this reaction that were famously search for by experimentalist Raymond Davis and his partner theoretician John Bahcall; the failure to detect them in th expected numbers became known as the 'Solar Neutrino Problem'. The 'problem' is now resolved through better understanding of neutrino physics – the electron neutrinos (the only type of neutrino detectable in tyhe 1960s, '70s, and '80s) 'oscillate' to other neutrino flavours.

are required. Of these processes, the CNO, or CNO-I, cycle[6] is the most important:

Step	Process			Energy Release		Solar Timescale
(1)	$^{12}_{6}\mathrm{C}\ \boxed{+\mathrm{p}}$	$\rightarrow$	$^{13}_{7}\mathrm{N}$	1.94 MeV		1.3×10^7 yr
(2)	$^{13}_{7}\mathrm{N}$	$\rightarrow$	$^{13}_{6}\mathrm{C} + \mathrm{e}^+ + \nu_\mathrm{e}$	2.22 MeV	†	7 m
(3)	$^{13}_{6}\mathrm{C}\ \boxed{+\mathrm{p}}$	$\rightarrow$	$^{14}_{7}\mathrm{N}$	7.55 MeV		2.7×10^6 yr
(4)	$^{14}_{7}\mathrm{N}\ \boxed{+\mathrm{p}}$	$\rightarrow$	$^{15}_{8}\mathrm{O}$	7.29 MeV		3.2×10^8 yr
(5)	$^{15}_{8}\mathrm{O}$	$\rightarrow$	$^{15}_{7}\mathrm{N} + \mathrm{e}^+ + \nu_\mathrm{e}$	2.76 MeV	†	82 s
(6*a*)	$^{15}_{7}\mathrm{N}\ \boxed{+\mathrm{p}}$	$\rightarrow$	$^{12}_{6}\mathrm{C}\ \boxed{+^4\mathrm{He}}$	4.96 MeV		1.1×10^5 yr
				26.72 MeV		

†Includes 1.02 MeV from $\mathrm{e}^+ + \mathrm{e}^- \rightarrow 2\gamma$

As in PP, we have created one ^{4}He from four protons, with release of some 26.7 MeV in the process; the neutrinos carry off 1.71 MeV for every α particle created, so 25 MeV is available to heat the gas. Although steps (2) and (5) both involve the weak interaction, they proceed faster than reaction (1) of the PP chain, since the nucleons involved are already bound to each other (which allows more time for the weak interaction to occur).

The cycle starts *and finishes* with ^{12}C, which acts as a catalyst.[7] However, during CNO cycling, the overall abundances nonetheless change – why is this?

Step (4), ^{14}N + p, is more than 10× slower than the next-slowest reaction (step (1), ^{12}C + p). It therefore acts as a 'bottleneck', with a build-up of ^{14}N at the expense of ^{12}C until the reaction rates[8] of steps (1) and (4) are equal (these depending on the number densities of reagents; eqtn. (G.2)). The equilibrium condition that reaction rates are equal determines the abundances, which can be compared to 'solar' abundances:

	CN cycle	Solar	
$n(^{12}\mathrm{C})/n(^{13}\mathrm{C})$	4	89	
$n(^{14}\mathrm{N})/n(^{15}\mathrm{N})$	2800	250	[^{15}N reduced by step (6a)]
$n(^{14}\mathrm{N} + ^{15}\mathrm{N})/n(^{12}\mathrm{C} + ^{13}\mathrm{C})$	21	0.3	[^{14}N increased by step (3)]

at $T \sim 1.3 \times 10^7$K (the solar-core temperature; the timescale required to establish equilibrium is set by

[6]Sometimes called the 'carbon cycle', although this risks confusion with cycling of carbon between the Earth's atmosphere, biosphere, hydrosphere, which also goes by that name. The CNO-I and CNO-II cycles together constitute the 'CNO bi-cycle' Where do the CNO nuclei come from? The answer is that they were created in previous generations of stars, in processes shortly to be described.

[7]Note that given ordering is arbitrary – the cycle can be considered as beginning at any point [e.g., starting at step (4), ending at (3)].

[8]Recall that reaction rates depend on both timescales and reagent abundances – cf. eqtnG.1

the slowest reaction, and so is $\sim 10^8$ yr at this temperature). These anomalous abundance patterns are a clear signature of CN processing if the products are brought to the stellar surface.

We can similarly evaluate equilibrium abundances for PP processing; for $T \simeq 1.3 \times 10^7$ K,

$$n(^2\mathrm{D})/n(^1\mathrm{H}) = 3 \times 10^{-17}$$
$$n(^3\mathrm{He})/n(^1\mathrm{H}) = 10^{-4}$$
$$(= 10^{-2} \text{ at } 8 \times 10^6 \mathrm{K})$$

G.5.1 CNO-II

There are a number of subsidiary reactions to the main CNO cycle, particularly involving oxygen. The CNO-II bi-cycle accounts for about 1 in 2500 ^{4}He productions in the Sun:

(6*b*)	$^{15}\mathrm{N} + \mathrm{p}$	$\rightarrow$	$^{16}\mathrm{O}$	12.13 MeV
(7*b*)	$^{16}\mathrm{O} + \mathrm{p}$	$\rightarrow$	$^{17}\mathrm{F}$	0.60 MeV
(8*b*)	$^{17}\mathrm{F}$	$\rightarrow$	$^{17}\mathrm{O} + \mathrm{e}^+ + \nu_\mathrm{e}$	2.76 MeV
(9*b*)	$^{17}\mathrm{O} + \mathrm{p}$	$\rightarrow$	$^{14}\mathrm{N} + ^4\mathrm{He}$	1.19 MeV
				26.72 MeV

which returns to step (4) in CNO-I

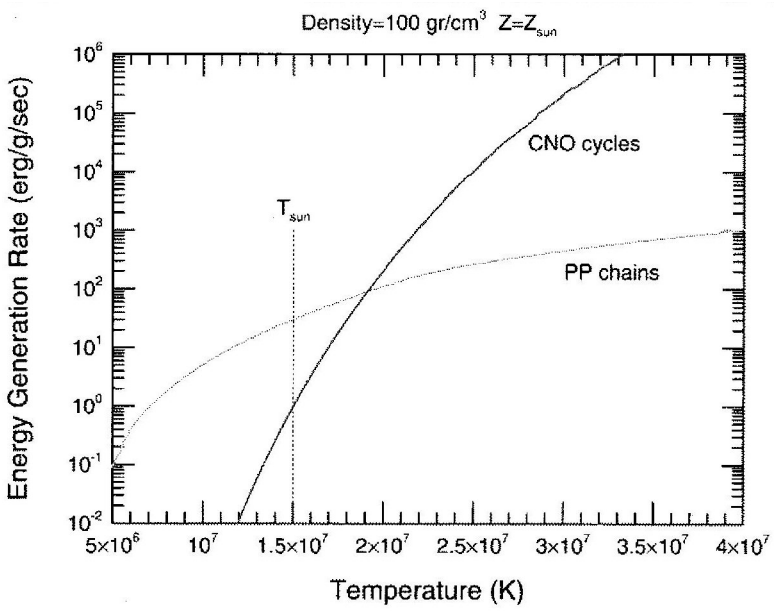


Figure G.3: Energy generation rates: CNO vs. PP processing

CNO-III, IV

The 'OF cycle' (which with CNO-I and CNO-II makes up the 'CNO tri-cycle') occurs in massive stars, and can be divided into CNO-III and CNO-IV; each branch starts from a ^{17}O produced in CNO-II:

(9*c*)	^{17}O + p	→	^{18}F + γ+	5.61 MeV
(10*c*)	^{18}F	→	^{18}O + e$^+$ + ν_e + γ	1.66 MeV
(11*c*)	^{18}O + p	→	^{15}N +^{4} He	

which returns to step (6b) in CNO-II; or, proceeding to CNO-IV:

(11*d*)	^{18}O + p	→	^{19}F + γ	7.99 MeV
(12*d*)	^{19}F + p	→	^{16}O +^{4} He	8.11 MeV

which returns to step (7b) in CNO-II

The only possible breakout from a closed cycle at temperatures relevant for quiescent hydrogen burning would be an alternative to step (12d),

(12*e*)	^{19}F + p	→	^{20}Ne + γ

but the rate is negligibly small, ensuring that the CNO cycles are completely closed.

We have seen that

$$\varepsilon \simeq \varepsilon_0 \rho T^{\alpha} \tag{G.9}$$

where $\alpha \simeq 4.5$ for proton-proton reactions in the Sun and $\alpha \simeq 18$ for CN processing. Because core temperature scales with mass (Section 11.5.3), PP dominates for lower-mass stars, while CN cycling dominates for higher-mass stars. The Sun lies just below the crossover point (fig. G.3), and although the PP chain dominates, the CN cycle is not negligible.

G.6 Helium burning

G.6.1 3α burning

Hydrogen burning dominates the stellar lifetime (the main-sequence phase), but the core pressure,

$$P = \frac{\rho k T}{\mu m(\mathrm{H})},$$

reduces as the mean molecular weight μ changes from 0.5 (for fully-ionized pure hydrogen) to 4/3 (for fully-ionized pure helium). As a consequence the core contracts, and heats. If the star is more massive than about $0.5\mathrm{M}_{\odot}$ the resulting core temperature and pressure are high enough to ignite helium burning ($\sim 10^8$K, 10^8 kg m^{-3}; lower-mass stars don't have enough gravitational potential energy); the reactions have a nett effect of

$$3 \times {}^4\mathrm{He} \quad \rightarrow \quad {}^{12}\mathrm{C}$$

However, the process is hindered by the absence of stable mass-5 (^{4}He + p) and mass-8 (^{4}He +^{4} He) nuclei; in particular, ^{8}Be is unstable, and decays back to a pair of alpha particles in only about $\sim 3{\times}10^{-16}$ s. Nonetheless, in equilibrium a small population of ^{8}Be particles exists (at a level of 1 for every $\sim 10^9$ α particles) and these can interact with ^{4}He under stellar-core conditions. Exceptionally therefore, because the lifetimes are so short, the production of ^{12}C is, essentially, a *3-body* process, with an energy-generation rate:

$$\varepsilon_{3\alpha} \simeq \varepsilon_0 \rho^2 T^{30}$$

(where $\varepsilon_0 \propto n(^4\mathrm{He})$ and the density-squared dependence is because of the three-body nature of the reaction).

(1)	^{4}He +^{4} He	$\leftrightarrow$	^{8}Be	−0.095 MeV
(2)	^{4}He +^{8} Be	$\leftrightarrow$	^{12}C*,	
	^{12}C*	$\rightarrow$	^{12}C + (2γ *or* e$^+$ + e$^-$)	7.37 MeV

The first stage is endothermic; ^{8}Be is more massive than two ^{4}He nuclei, so the relative binding energy is *negative*.

Reaction (2) is favoured by the existence of a resonance at 287 keV, which results in a ^{12}C nucleus excited 7.65 MeV above the ground state.[9] The lifetime of this excited state is very small (about 5×10^{-17} s!), and normally decays straight back to ^{4}He +^{8} Be, but 1 in ∼2400 decays is to a ground-state ^{12}C nucleus, with the emission of two photons. These decays are irreversible, and so a population of ^{12}C nuclei slowly builds up.

G.6.2 Further helium-burning stages

Once carbon has been created, still heavier nuclei can be built up:

$$\begin{aligned} ^{12}\mathrm{C} + ^{4}\mathrm{He} &\rightarrow {}^{16}\mathrm{O} && 7.16\ \mathrm{MeV} \\ ^{16}\mathrm{O} + ^{4}\mathrm{He} &\rightarrow {}^{20}\mathrm{Ne} && 4.73\ \mathrm{MeV} \end{aligned}$$

These processes therefore generate C, O, and Ne. ^{12}C and ^{16}O are the most abundant nuclei at the end of He burning (and the most cosmically abundant elements after H and He, with about 1 C or O for every 10^{3} hydrogens, or every 100 heliums) The situation is more complicated for ^{14}N, which is enhanced during CNO processing[10] but which is is destroyed during He burning by the reactions

$$\begin{aligned} ^{14}\mathrm{N} + ^{4}\mathrm{He} &\rightarrow {}^{18}\mathrm{O} + \mathrm{e}^{+} + \nu_{\mathrm{e}} \\ ^{18}\mathrm{O} + ^{4}\mathrm{He} &\rightarrow {}^{22}\mathrm{Ne} && 4.73\ \mathrm{MeV} \end{aligned}$$

G.7 Advanced burning

G.7.1 Carbon burning

After exhaustion of ^{4}He, the core of a high-mass star contracts further, and at $T \sim 10^{8}$–10^{9}K *carbon burning* can take place:

$$^{12}\mathrm{C} + ^{12}\mathrm{C} \rightarrow \begin{cases} ^{23}\mathrm{Na} + \mathrm{p} & 2.2\ \mathrm{MeV} \\ ^{20}\mathrm{Ne} + ^{4}\mathrm{He} & 4.6\ \mathrm{MeV} \\ ^{23}\mathrm{Mg} + \mathrm{n} & -2.6\ \mathrm{MeV} \\ ^{24}\mathrm{Mg} + \gamma & 13.9\ \mathrm{MeV} \end{cases}$$

[9] Hoyle (1954) deduced that such a resonance in a previously unknown excited state of carbon must exist to allow an α particle to combine with an ^{8}Be with sufficient probability for the triple-alpha process to proceed.

[10] All the initial ^{12}C and ^{16}O ends up as ^{14}N.

with a temperature dependence of

$$\varepsilon_{\mathrm{C}} \simeq \varepsilon_0 \rho T^{32}$$

G.7.2 Neon burning

Neon burning takes place after carbon burning if the core temperature reaches $\sim 10^9$K, but at these temperatures photodisintegration also occurs:

$$\gamma + {}^{20}\mathrm{Ne} \quad \rightarrow \quad {}^{16}\mathrm{O} + {}^{4}\mathrm{He}$$

These 'new' alpha particles can then react with undissociated neons:

$${}^{20}\mathrm{Ne} + {}^{4}\mathrm{He} \quad \rightarrow \quad {}^{24}\mathrm{Mg} + \gamma$$

G.7.3 Oxygen burning

After neon burning the core consists mainly of ^{16}O and ^{24}Mg. Oxygen burning occurs at $\sim 2 \times 10^9$K:

$${}^{16}\mathrm{O} + {}^{16}\mathrm{O} \quad \rightarrow \quad \begin{cases} {}^{32}\mathrm{S} + \gamma & 16.5\ \mathrm{MeV} \\ {}^{31}\mathrm{P} + \mathrm{p} & 7.6\ \mathrm{MeV} \\ {}^{31}\mathrm{S} + \mathrm{n} & 1.4\ \mathrm{MeV} \\ {}^{28}\mathrm{Si} + {}^{4}\mathrm{He} & 9.6\ \mathrm{MeV} \\ {}^{24}\mathrm{Mg} + 2\,{}^{4}\mathrm{He} & -0.4\ \mathrm{MeV} \end{cases}$$

with silicon being the most important product.

G.7.4 Silicon burning

At $\sim 3 \times 10^9$K, silicon burning can occur; the Si is slowly photodisintegrated, releasing protons, neutrons, and alpha particles (a process sometimes called 'silicon melting' as opposed to 'silicon burning'). Of particular interest is the reaction

$$\gamma + {}^{28}\mathrm{Si} \quad \rightarrow \quad {}^{24}\mathrm{Mg} + {}^{4}\mathrm{He}$$

These alpha particles then combine with undissociated nuclei to build more massive nuclei; for example, by way of illustration,

$$
\begin{aligned}
^{28}\mathrm{Si} + ^{4}\mathrm{He} \quad &\leftrightarrow \quad ^{32}\mathrm{S} + \gamma \\
^{32}\mathrm{S} + ^{4}\mathrm{He} \quad &\leftrightarrow \quad ^{36}\mathrm{Ar} + \gamma \\
^{36}\mathrm{Ar} + ^{4}\mathrm{He} \quad &\leftrightarrow \quad ^{40}\mathrm{Ca} + \gamma \\
&\ldots \\
^{52}\mathrm{Fe} + ^{4}\mathrm{He} \quad &\leftrightarrow \quad ^{56}\mathrm{Ni} + \gamma \\
&(\rightarrow \quad ^{56}\mathrm{Fe})
\end{aligned}
$$

The overall timescale is set by the slowest step, which is the initial photodisintegration of Si.

G.7.5 The α process

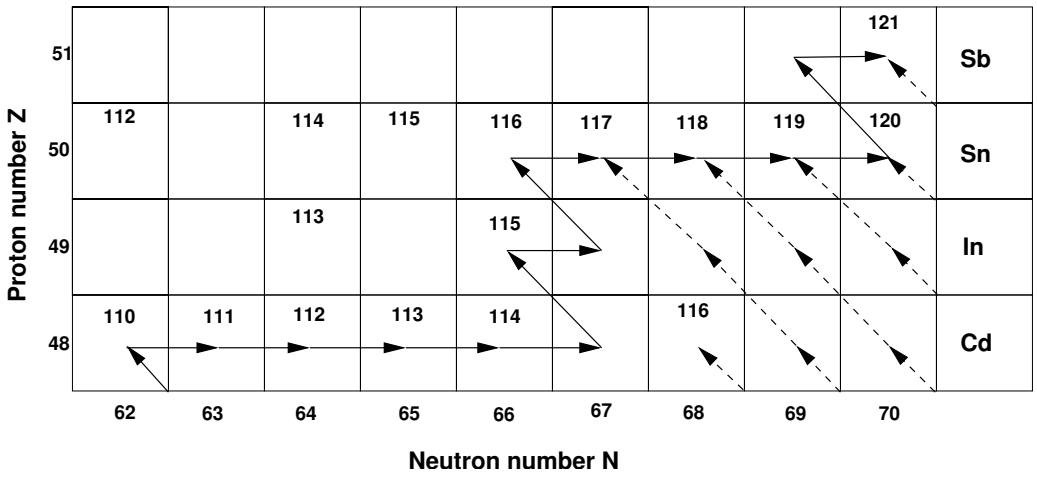
Throughout the preceding advanced processing stages, a recurrent mechanism is the capture of α particles (ie., helium nuclei) to form heavier elements. This 'α process' produces α-process elements whose most abundant isotopes have atomic masses that are integer multiples of four; the generally accepted list of α-process elements is:
^{16}O, ^{20}Ne, ^{24}Mg, ^{28}Si, ^{32}S, ^{36}Ar, ^{40}Ca, ^{48}Ti
(^{44}Sc is unstable; some authors consider C and N, produced by α capture, as α-process elements).

Because the binding energy per nucleon peaks around mass $A = 56$ (the 'iron-peak' elements V, Cr, Mn, Fe, Co, Ni) energy is absorbed to form heavier nuclei. Elements beyond the iron peak are therefore not formed during silicon burning.

G.8 Pre-main-sequence burning

Although not as important as energy-generating sources, some reactions involving light nuclei can occur at $\sim 10^6$K – i.e., lower temperatures than those discussed so far:

$^{2}\mathrm{D} + \mathrm{p}$	$\rightarrow$	$^{3}\mathrm{He} + \gamma$	5.4×10^5K	(step 2 of PP-I)
$^{6}\mathrm{Li} + \mathrm{p}$	$\rightarrow$	$^{3}\mathrm{He} + ^{4}\mathrm{He}$	2.0×10^6K	
$^{7}\mathrm{Li} + \mathrm{p}$	$\rightarrow$	$^{4}\mathrm{He} + ^{4}\mathrm{He}$	2.4×10^6K	
$^{9}\mathrm{Be} + ^{2}\mathrm{D}$	$\rightarrow$	$^{4}\mathrm{He} + ^{4}\mathrm{He} + ^{3}\mathrm{He}$	3.2×10^6K	
$^{10}\mathrm{B} + ^{2}\mathrm{D}$	$\rightarrow$	$^{4}\mathrm{He} + ^{4}\mathrm{He} + ^{4}\mathrm{He}$	4.7×10^6K	



These reactions generally *destroy* light elements such as lithium (produced, e.g., primordially) at relatively low temperatures.

Note that the first step, burning of *pre-existing* deuterium, defines brown dwarfs – objects with cores too cool to produce deuterium by proton-proton reactions.

G.9 Synthesis of heavy elements

G.9.1 Neutron capture: *r* and *s* processes

Carbon burning, oxygen burning etc. can generate heavy elements in the cores of very massive stars, but only as far as the iron peak. However, a quite different set of reactions can occur at lower temperatures ($\sim 10^8$ K, comparable to that need for 3α burning).

Since neutrons are electrically neutral, they see no Coulomb barrier, and can be absorbed into nuclei even at quite low energies (in fact, heavy nuclei have relatively large neutron-capture cross-sections). Neutron absorption produces a heavier isotope (increases A but not Z); a change in element may then result if the nucleus is unstable to β decay ($\mathrm{n} \rightarrow \mathrm{p} + \mathrm{e}^- + \overline{\nu}_\mathrm{e}$).

Following the pioneering work of Burbidge, Burbidge, Fowler & Hoyle (Rev. Mod. Phys., 29, 547, 1955), it is conventional to distinguish between *r* and *s* processes, depending on whether neutron capture is *r*apid or *s*low compared to the β-decay timescale. If it is rapid, then more and more massive isotopes accumulate; if it is slow, then decay to a higher-Z element takes place. Suppose we start off with a neutron capture to produce some new isotope:

$$(Z, A-1) + \mathrm{n} \rightarrow (Z, A).$$

Then if neutron capture happens *s*lowly compared to decay for this new isotope, β decay precedes any further neutron capture, and a new element is formed:

$$(Z, A) \rightarrow (Z + 1, A) + \mathrm{e}^- + \overline{\nu}_\mathrm{e}.$$

However, if neutron capture is *r*apid then a further isotope is produced,

$$(Z, A) + \mathrm{n} \rightarrow (Z, A + 1),$$

which will in turn β-decay,

$$(Z, A + 1) \rightarrow (Z + 1, A + 1) + \mathrm{e}^- + \overline{\nu}_\mathrm{e},$$

or assimilate a further neutron.

The timescales involved for the *r* and *s* processes are largely set by the relevant nuclear timescales.[11] The *s* process occurs during non-catastrophic evolutionary phases (principally the AGB phase); we *know* this from the observation that technetium occurs in S-type stars (moderately carbon rich M stars). Even the longest-lived technetium isotope, ^{99}Tc, has a half-life only of order 10^4 yr, and so it must be produced within stars during normal evolutionary processes.

Where do the free neutrons come from? For the *s* process, the CNO cycle establishes an appreciable abundance of ^{13}C (step 2 in the sequence set out in Section G.5), which can react with ^{4}He:

$$^{13}\mathrm{C} + ^4\mathrm{He} \quad \rightarrow \quad ^{16}\mathrm{O} + \mathrm{n} \qquad\qquad (-0.91\ \mathrm{MeV})$$

This is the main source of neutrons in AGB stars; at higher temperatures,

$$^{22}\mathrm{Ne} + ^4\mathrm{He} \quad \rightarrow \quad ^{25}\mathrm{Mg} + \mathrm{n} \qquad\qquad (-0.48\ \mathrm{MeV}).$$

is significant.

Neutron-capture cross-sections are exceptionally small for certain nuclear neutron numbers. Because it's harder for the corresponding isotopes to increase in mass through neutron capture, they build up in abundance. We see this effect as peaks in the element-abundance distribution for elements such as $^{88}_{38}$Sr, $^{138}_{56}$Ba, and $^{208}_{82}$Pb.

Elements beyond bismuth ($Z = 83$) cannot be produced through the *s* process, the terminating cycle

[11] Just to have some sense of the numbers, the *s* process typically operates on timescales of $\sim 10^4$ yr at neutron densities of $\sim 10^{11}$ m^{-3}; corresponding numbers for the *r* process are a few seconds at $\sim 10^{25}$ m^{-3}.

	115	116	117	118	119	120	121	122	123	124	125	126	127	128
$_{83}$Bi									ε	ε	ε	**100**	β	α
$_{82}$Pb					ε	ε	ε	**1.42**	ε	**24.1**	**22.1**	**52.3**	β	β
$_{81}$Ti					ε	ε	ε	**29.5**	β	**70.5**	β	β		
$_{80}$Hg	ε	**0.15**	ε	**10.0**	**16.8**	**23.1**	**13.3**	**29.8**	β	**6.9**	β			
$_{79}$Au	ε	ε	ε	**100**	β	β								

Figure G.4: Isotopes of gold–bismuth. The top row lists the number of neutrons in the isotope, while the atomic number (number of protons) is given by the element name. Unstable isotopes decay by conversion of a proton to a neutron (electron capture, ε), conversion of a neutron to a proton (β decay), or emission of a helium nucleus (α decay, Bi^{211} only). Numbers give natural percentage abundances of stable isotopes (blanks are for isotopes that do not occur in nature).
The dashed line shows the *s*-process path from the only stable isotope of gold (Au^{197}) to the only stable isotope of bismuth (Bi^{209}). Hg^{204} is an example of an isotope that can be made only by the *r* process.

being

$$^{209}\mathrm{Bi} + \mathrm{n} \rightarrow ^{210}\mathrm{Bi}$$
$$^{210}\mathrm{Bi} \rightarrow ^{210}\mathrm{Po}(+\mathrm{e}^- + \overline{\nu}_\mathrm{e})$$
$$^{210}\mathrm{Po} \rightarrow ^{206}\mathrm{Pb} + ^{4}\mathrm{He}$$

$$^{206}\mathrm{Pb} + 3\mathrm{n} \rightarrow ^{209}\mathrm{Pb}$$
$$^{209}\mathrm{Pb} \rightarrow ^{209}\mathrm{Bi}(+\mathrm{e}^- + \overline{\nu}_\mathrm{e})$$

(involving $Z = 84$ polonium and $Z = 82$ lead in addition to bismuth).

Many, but not all, elements at lower atomic masses can be produced by both *r* and *s* processes; *s*-only products include $^{87}_{38}$Sr and $^{187}_{76}$Os.

The *r* process requires very high neutron fluxes, so that neutron capture rates exceed or compete with β-decay rates. These conditions can only occur during catastrophic, short-timescale phases – supernova explosions. Although some isotopes can be produced by both processes, in general there are significant differences between their products.

G.9.2 The *p* process (for reference only)

In their seminal paper, Burbidge, Burbidge, Fowler & Hoyle (B^2FH) identified the need for a process to create certain relatively proton-rich nuclei, heavier than iron, that cannot be produced by either of the *r* or *s* processes (e.g., ^{190}Pt, ^{168}Yb).

They originally envisaged a proton-capture process, but we now believe that these proton-rich nuclei are not produced by addition of protons, but by removal of neutrons by photodisintegration (i.e., impact by

high-energy photons).[12] This occurs through neutron photodisintegration (ejection of a neutron) or α photodisintegration (emission of an α particle). These processes require high temperatures (i.e., high-energy photons), and is believed to occur during core collapse of supernovae.

G.10 Summary

Hydrogen and helium were produced primordially. After these, CNO are the most abundant elements, with CO produced through helium burning,[13] with nitrogen generated in CNO processing.

Stars more massive than $\sim 8\mathrm{M}_\odot$ go on to produce elements such as neon, sodium, and magnesium, with stars more massive than $\sim 11\mathrm{M}_\odot$ proceding to silicon burning, thereby generating nuclei all the way up to the iron peak.

Subsequent processing primarily involves neutron capture (although other processes, such as spallation and proton capture, have a small role).

The timescales for various burning stages are progressively shorter, as energy production rates increase to compensate increasing energy losses (e.g., by increasing neutrino losses). Only massive stars have enough gravitational potential energy to power the most advanced burning stages, so we review the timescales for a 25-$\mathrm{M}_\odot$ star:

Burning stage	Timescale	$T_\mathrm{c}/10^9\mathrm{K}$	ρ_c (kg m^{-3})	Products
H	7×10^6 yr	0.06	5×10^4	He; N (CNO process)
He	5×10^5 yr	0.1	7×10^5	C, O
C	6×10^2 yr	0.6	2×10^8	Ne, Na, Mg, etc.
Ne	1×10^0 yr	1	4×10^9	O, Na, Mg, etc.
O	5×10^{-1} yr	2	1×10^{10}	Si, S, P, etc.
Si	1 d	3	3×10^{10}	Mn, Cr, Fe, Co, Ni etc.

[12] Luckily, 'photodisintegration' fits the description '*p* process' as well as 'proton capture' does! There is a proton-capture mechanism, now called the *rp* process, but it is generally less important than the *p* process.

[13] The balance between C and O is determined by the balance between the rate of production of C and the rate of destruction (in O formation). If the ratio favoured O only a little more, then we wouldn't be here.

Appendix H

Composition: abundances in astrophysics

Astronomers routinely use a number of different terminologies to describe abundances in stars (galaxies, interstellar medium. . .). Because densities vary over an enormous range in astrophysical environments, abundances aren't always best characterized in absolute terms (e.g., numbers of atoms per unit volume); instead, standard nomenclatures utilise *relative* abundances.

- Mass fractions:

  $X =$ mass fraction of hydrogen

  $Y =$ mass fraction of helium

  $Z =$ mass fraction of metals.

  Here 'metals' are all elements heavier than helium, grouped together. Solar photospheric values given by Asplund et al. (ARA&A 47, 481, 2009) are

  $X_\odot = 0.7381$

  $Y_\odot = 0.2485$

  $Z_\odot = 0.0134$

- Finer descriptions of individual element abundances are usually based on relative abundances by number of nuclei. This may be done by reference to total numbers of atoms; e.g.,

  $x = n(\mathrm{H})/[n(\mathrm{H}) + n(\mathrm{He}) + n(\mathrm{metals})]$,

  $y = n(\mathrm{He})/[n(\mathrm{H}) + n(\mathrm{He}) + n(\mathrm{metals})]$,

  $z = n(\mathrm{metals})/[n(\mathrm{H}) + n(\mathrm{He}) + n(\mathrm{metals})]$

or relative to the most abundant element. Because hydrogen is normally the most abundant element by orders of magnitude, it's convenient to put the relative abundance on a $\log_{10}$ scale. One might reasonably suppose that the obvious figures to employ would then simply be, e.g., $\log[n(\mathrm{X})/n(\mathrm{H})]$ for element X, and indeed this is often done; but more usually a rescaling by 12 dex is used, with results denoted either by square brackets, or sometimes by 'ε'. That is,

$$[\mathrm{X}] \equiv \log \varepsilon(\mathrm{X}) = \log[n(\mathrm{X})/n(\mathrm{H})] + 12$$

which ensures that the numerical results are almost always positive. For example, the solar carbon abundance can be expressed as

$$[\mathrm{C}] = \log \varepsilon(\mathrm{C})_{\odot} = 8.39$$

meaning there is one carbon nucleus for every $\sim 10^{3.61}$ hydrogens.

- Finally, it's often useful to express abundances relative to solar values, in order to have an immediate impression of how 'normal' the results are. This is again done on a relative logarithmic scale, and the differential aspect denoted by the use of square brackets; so for two elements A, B we have

  $$[A/B] \equiv \log\left[\frac{n(A)}{n(B)}\right] - \log\left[\frac{n(A)}{n(B)}\right]_{\odot}$$

  For example, a star with an iron abundance with respect to hydrogen that is one-tenth that of the Sun would have [Fe/H] = −1.0. One minor disadvantage of this approach for precise work is that if the accepted value of the solar abundance changes for some element, any other published values that use the original solar values require adjustment (unless originally determined differentially with respect to the Sun).

  This nomenclature is commonly used for *global* scaling of metal abundances, such that (e.g.) $[M/\mathrm{H}] = +1.0$ would mean all metals enhanced by a factor 10 over solar; in practice iron is often used as a surrogate for 'all metals', in which case the same result may be expressed as [Fe/H] = +1.0.

'Cosmic' Abundances *by number* (Asplund et al., 2005ASPC..336...25A)

Elem.	*Photosphere*	*Meteorites*	*Elem.*	*Photosphere*	*Meteorites*
H	≡ 12.00	8.25 ± 0.05	Ru	1.84 ± 0.07	1.77 ± 0.08
He	[10.93 ± 0.01]	1.29	Rh	1.12 ± 0.12	1.07 ± 0.02
Li	1.05 ± 0.10	3.25 ± 0.06	Pd	1.69 ± 0.04	1.67 ± 0.02
Be	1.38 ± 0.09	1.38 ± 0.08	Ag	0.94 ± 0.24	1.20 ± 0.06
B	2.70 ± 0.20	2.75 ± 0.04	Cd	1.77 ± 0.11	1.71 ± 0.03
C	8.39 ± 0.05	7.40 ± 0.06	In	1.60 ± 0.20	0.80 ± 0.03
N	7.78 ± 0.06	6.25 ± 0.07	Sn	2.00 ± 0.30	2.08 ± 0.04
O	8.66 ± 0.05	8.39 ± 0.02	Sb	1.00 ± 0.30	1.03 ± 0.07
F	4.56 ± 0.30	4.43 ± 0.06	Te	2.19 ± 0.04	
Ne	[7.84 ± 0.06]	−1.06	I	1.51 ± 0.12	
Na	6.17 ± 0.04	6.27 ± 0.03	Xe	[2.27 ± 0.02]	−1.97
Mg	7.53 ± 0.09	7.53 ± 0.03	Cs	1.07 ± 0.03	
Al	6.37 ± 0.06	6.43 ± 0.02	Ba	2.17 ± 0.07	2.16 ± 0.03
Si	7.51 ± 0.04	7.51 ± 0.02	La	1.13 ± 0.05	1.15 ± 0.06
P	5.36 ± 0.04	5.40 ± 0.04	Ce	1.58 ± 0.09	1.58 ± 0.02
S	7.14 ± 0.05	7.16 ± 0.04	Pr	0.71 ± 0.08	0.75 ± 0.03
Cl	5.50 ± 0.30	5.23 ± 0.06	Nd	1.45 ± 0.05	1.43 ± 0.03
Ar	[6.18 ± 0.08]	−0.45	Sm	1.01 ± 0.06	0.92 ± 0.04
K	5.08 ± 0.07	5.06 ± 0.05	Eu	0.52 ± 0.06	0.49 ± 0.04
Ca	6.31 ± 0.04	6.29 ± 0.03	Gd	1.12 ± 0.04	1.03 ± 0.02
Sc	3.05 ± 0.08	3.04 ± 0.04	Tb	0.28 ± 0.30	0.28 ± 0.03
Ti	4.90 ± 0.06	4.89 ± 0.03	Dy	1.14 ± 0.08	1.10 ± 0.04
V	4.00 ± 0.02	3.97 ± 0.03	Ho	0.51 ± 0.10	0.46 ± 0.02
Cr	5.64 ± 0.10	5.63 ± 0.05	Er	0.93 ± 0.06	0.92 ± 0.03
Mn	5.39 ± 0.03	5.47 ± 0.03	Tm	0.00 ± 0.15	0.08 ± 0.06
Fe	7.45 ± 0.05	7.45 ± 0.03	Yb	1.08 ± 0.15	0.91 ± 0.03
Co	4.92 ± 0.08	4.86 ± 0.03	Lu	0.06 ± 0.10	0.06 ± 0.06
Ni	6.23 ± 0.04	6.19 ± 0.03	Hf	0.88 ± 0.08	0.74 ± 0.04
Cu	4.21 ± 0.04	4.23 ± 0.06	Ta	−0.17 ± 0.03	
Zn	4.60 ± 0.03	4.61 ± 0.04	W	1.11 ± 0.15	0.62 ± 0.03
Ga	2.88 ± 0.10	3.07 ± 0.06	Re	0.23 ± 0.04	
Ge	3.58 ± 0.05	3.59 ± 0.05	Os	1.45 ± 0.10	1.34 ± 0.03
As	2.29 ± 0.05		Ir	1.38 ± 0.05	1.32 ± 0.03
Se	3.33 ± 0.04		Pt	1.64 ± 0.03	
Br	2.56 ± 0.09		Au	1.01 ± 0.15	0.80 ± 0.06
Kr	[3.28 ± 0.08]	−2.27	Hg	1.13 ± 0.18	
Rb	2.60 ± 0.15	2.33 ± 0.06	Tl	0.90 ± 0.20	0.78 ± 0.04
Sr	2.92 ± 0.05	2.88 ± 0.04	Pb	2.00 ± 0.06	2.02 ± 0.04
Y	2.21 ± 0.02	2.17 ± 0.04	Bi	0.65 ± 0.03	
Zr	2.59 ± 0.04	2.57 ± 0.02	Th	0.06 ± 0.04	
Nb	1.42 ± 0.06	1.39 ± 0.03	U	< −0.47	−0.52 ± 0.04
Mo	1.92 ± 0.05	1.96 ± 0.04			

Appendix I

Abbreviations

ABBREVIATIONS:		
IR	infra-red	
ISM	InterStellar Medium	
UV	ultraviolet	
c	Speed of light (in vacuo)	
D	Depletion factor	**??**
e	Electron charge	
f	Filling factor	**??**
	Oscillator strength	
h	Planck's constant	
k	Boltzmann's constant	
m_e	Electron mass	
W	Work function	**??**
γ	classical damping constant	**??**
ν	Photon frequency	

Appendix J

Einstein (radiative) coefficients

Einstein (1916) proposed that there are three purely radiative processes which may be involved in the formation of a spectral line: induced emission, induced absorption, and spontaneous emission, each characterized by a coefficient reflecting the probability of a particular process.

[1] A_{ji} (s^{-1}) – the Einstein coefficient, or transition probability, for spontaneous decay from an upper state j to a lower state i, with the emission of a photon (radiative decay). The probability that the emission will occur in time dt and in solid angle $d\omega$ is $A_{ji}dtd\omega$, and the average time taken for an electron in state j to spontaneously decay to state i is $1/A_{ji}$.

If n_j is the number density of atoms in state j then the change in the number density of atoms in that state per unit time due to spontaneous emission will be

$$\frac{dn_j}{dt} = -\sum_{i<j} A_{ji} n_j$$

while level i is populated according to

$$\frac{dn_i}{dt} = +\sum_{j>i} A_{ji} n_j \,.$$

The contribution to the volume emissivity of spontaneous emission is just

$$j_\nu = n_j A_{ji} h\nu \,.$$

[2] B_{ji} (s^{-1} J^{-1} m^2 sr): the Einstein coefficient for radiatively induced de-excitation from an upper state to a lower state (stimulated emission).

$$\frac{dn_j}{dt} = -\sum_{i<j} B_{ji} n_j I_\nu,$$

$$\frac{dn_i}{dt} = +\sum_{j>i} B_{ji} n_j I_\nu$$

[3] B_{ij} (s^{-1} J^{-1} m^2 sr): the Einstein coefficient for radiative excitation from a lower state i to an upper state j, with the absorption of a photon.

$$\frac{dn_i}{dt} = -\sum_{j>i} B_{ij} n_i I_\nu,$$
$$\frac{dn_j}{dt} = +\sum_{i<j} B_{ij} n_i I_\nu$$

where I_ν is the specific intensity at the frequency ν corresponding to E_{ij}, the energy difference between excitation states.

The volume opacity arising from this bound–bound transition is

$$k_\nu = n_i B_{ij} h\nu - n_j B_{ji} h\nu$$

For reference, we state, without proof, the relationships between these coefficients:

$$A_{ji} = \frac{2h\nu^3}{c^2} B_{ji};$$
$$B_{ij} g_i = B_{ji} g_j$$

where g_i is the statistical weight of level i.

In astronomy, it is common to work not with the Einstein A coefficient, but with the absorption oscillator strength f_{ij}, where

$$A_{ji} = \frac{8\pi^2 e^2 \nu^2}{m_e c^3} \frac{g_i}{g_j} f_{ij}$$

and f_{ij} is related to the absorption cross-section by

$$a_{ij} \equiv \int a_\nu \, d\nu = \frac{\pi e^2}{m_e c} f_{ij}.$$

Because of the relationships between the Einstein coefficients, we also have

$$B_{ij} = \frac{4\pi^2 e^2}{m_e h\nu c} f_{ij},$$
$$B_{ji} = \frac{4\pi^2 e^2}{m_e h\nu c} \frac{g_i}{g_j} f_{ij}$$

J.0.1 Collisional coefficients

For collisional processes we have analogous coefficients:

[4] C_{ji} ($m^3\ s^{-1}$): the coefficient for collisional de-excitation from an upper state to a lower state.

$$\frac{dn_j}{dt} = -\sum_{j>i} C_{ji} n_j n_e,$$
$$\frac{dn_i}{dt} = +\sum_{i<j} C_{ji} n_j n_e$$

(for excitation by electron collisions)

[5] C_{ij} ($m^3\ s^{-1}$): the coefficient for collisional excitation from a lower state to an upper state.

$$\frac{dn_i}{dt} = -\sum_{j>i} C_{ij} n_i n_e,$$
$$\frac{dn_j}{dt} = +\sum_{i<j} C_{ij} n_i n_e$$

These coefficients are related through

$$\frac{C_{ij}}{C_{ji}} = \frac{g_j}{g_i} \exp\left\{-\frac{h\nu}{kT_{ex}}\right\}$$

for excitation temperature T_{ex}.

The rate coefficient has a Boltzmann-like dependence on the kinetic temperature

$$C_{ij}(T_k) = \left(\frac{2\pi}{T_k}\right)^{1/2} \frac{h^2}{4\pi^2 m_e^{3/2}} \frac{\Omega(ij)}{g_i} \exp\left\{\frac{-\Delta E_{ij}}{kT_k}\right\}$$
$$\propto \frac{1}{\sqrt{T_e}} \exp\left\{\frac{-\Delta E_{ij}}{kT_k}\right\} \qquad [m^3\ s^{-1}] \tag{J.1}$$

where $\Omega(1,2)$ is the so-called 'collision strength'.

J.0.2 Statistical Equilibrium

Overall, for any ensemble of atoms in equilibrium, the number of de-excitations from any given excitation state must equal the number of excitations into that state – the principle of statistical equilibrium. That is,

$$\sum_{j>i} B_{ij} n_i I_\nu + \sum_{j\neq i} C_{ij} n_i n_e = \sum_{j>i} A_{ji} n_j + \sum_{j>i} B_{ji} n_j I_\nu + \sum_{j\neq i} C_{ji} n_j n_e \tag{J.2}$$

The solution of this equation is at the heart of the calculation of theoretical spectra for stellar atmospheres (and other applications). Although superficially innocuous, it is salutory to realise that the summations may be over hundreds or thousands of energy levels, with the complication that the photorates depend on the specific intensity – which depends on the statistical equlibrium (and not the statistical equilibrium of a single ionic species, but of all species simultaneously).

J.0.3 Detailed Balance

Appendix K

Atomic spectra

K.1 Notation

For any given atom (or ion) there exist a number of discrete energy levels, corresponding to bound orbits in the Bohr atom. From the innermost orbits outwards, each orbit is usually occupied by the maximum number of electrons allowed. Electrons in the outermost occupied energy levels are those that normally participate in transitions from one bound state to another (bound-bound transitions, giving rise to absorption or emission lines). The same outermost, or valence, levels are usually involved when an electron goes from a bound state to a free state (ionization), or from a free state to a bound state (recombination).

In the Bohr model – the 'old' quantum mechanics – a bound electron could be described by just one number, which in effect represented the orbital radius: the principal quantum number, n. The more complete view of quantum mechanics developed by Schrödinger and others has electrons occupying three-dimensional space, requiring three three quantum numbers (in effect, three co-ordinates), to describe the orbitals in which electrons can be found. These are the principal, angular, and magnetic quantum numbers, n, ℓ, and m. These quantum numbers describe the size, shape, and spatial orientation of the orbitals; a fourth quantum number, s, describes the electron spin.

Bound electrons are described by four quantum numbers:

1. XXXXXThe principal quantum number (n) describes the size of the orbital. Orbitals for which n = 2 are larger than those for which n = 1, for example. Because they have opposite electrical charges, electrons are attracted to the nucleus of the atom. Energy must therefore be absorbed to excite an electron from an orbital in which the electron is close to the nucleus (n = 1) into an orbital in which it is further from the nucleus (n = 2). The principal quantum number therefore indirectly describes the energy of an orbital.

2. XXXXThe angular quantum number (l) describes the shape of the orbital. Orbitals have shapes that are best described as spherical (l = 0), polar (l = 1), or cloverleaf (l = 2). They can even take on more complex shapes as the value of the angular quantum number becomes larger.

3. XXXa third quantum number, known as the magnetic quantum number (m), is needed to describe the orientation in space of a particular orbital. (It is called the magnetic quantum number because the effect of different orientations of orbitals was first observed in the presence of a magnetic field.)

Orbitals that have the same value of the principal quantum number form a shell. Orbitals within a shell are divided into subshells that have the same value of the angular quantum number. Chemists describe the shell and subshell in which an orbital belongs with a two-character code such as 2p or 4f. The first character indicates the shell (n = 2 or n = 4). The second character identifies the subshell. By convention, the following lowercase letters are used to indicate different subshells.

s: l = 0 p: l = 1 d: l = 2 f: l = 3

Although there is no pattern in the first four letters (s, p, d, f), the letters progress alphabetically from that point (g, h, and so on). Some of the allowed combinations of the n and l quantum numbers are shown in the figure below.

The electron configuration of an atom describes the orbitals occupied by electrons on the atom. The basis of this prediction is a rule known as the aufbau principle, which assumes that electrons are added to an atom, one at a time, starting with the lowest energy orbital, until all of the electrons have been placed in an appropriate orbital.

A hydrogen atom (Z = 1) has only one electron, which goes into the lowest energy orbital, the 1s orbital. This is indicated by writing a superscript "1" after the symbol for the orbital.

H (Z = 1): 1s1

The next element has two electrons and the second electron fills the 1s orbital because there are only two possible values for the spin quantum number used to distinguish between the electrons in an orbital.

He (Z = 2): 1s2

The third electron goes into the next orbital in the energy diagram, the 2s orbital.

Li (Z = 3): 1s2 2s1

The fourth electron fills this orbital.

Be (Z = 4): 1s2 2s2

After the 1s and 2s orbitals have been filled, the next lowest energy orbitals are the three 2p orbitals. The fifth electron therefore goes into one of these orbitals.

B (Z = 5): 1s2 2s2 2p1

When the time comes to add a sixth electron, the electron configuration is obvious.

C (Z = 6): 1s2 2s2 2p2

However, there are three orbitals in the 2p subshell. Does the second electron go into the same orbital as the first, or does it go into one of the other orbitals in this subshell?

To answer this, we need to understand the concept of degenerate orbitals. By definition, orbitals are degenerate when they have the same energy. The energy of an orbital depends on both its size and its shape because the electron spends more of its time further from the nucleus of the atom as the orbital becomes larger or the shape becomes more complex. In an isolated atom, however, the energy of an orbital doesn't depend on the direction in which it points in space. Orbitals that differ only in their orientation in space, such as the 2px, 2py, and 2pz orbitals, are therefore degenerate.

In atoms with several electrons, each energy level is characterized by a total orbital angular momentum number J (the vector sum of orbital and spin angular momenta for all electrons).

Each J level contains g_J separate states with the same energy (where the statistical weight $g_J \simeq 2J + 1$).

Several 'fine structure levels', with different J values and slightly different energies, constitute a 'spectroscopic term', designated $S, P, D, F \ldots$ (with corresponding total orbital angular momenta $L = 0, 1, 2, 3 \ldots$).

The number of different J levels is the multiplicity ($= 2S + 1$).

Appendix L

Structure

Atomic structure is described by:

1. Configuration – the description of an atom (or ion) by quantum numbers of individual electrons only. For example, the ground state of the sodium atom is described by

$$1s^2\ 2s^2\ 2p^6\ 3s$$

   The notation here, $n\ell^m$, refers to the principal quantum number, n; the electron orbital angular momentum, ℓ ($\ell = 0, 1, 2, 3, 4, \ldots \equiv s, p, d, f, g, \ldots$); and the number of electrons m in that condition. Often the leading terms are omitted (as 'obvious'). (Letters come from old sharp, principal, diffuse, fundamental spectroscopic notation.)

2. Term – describes a pair of L, S values. In LS coupling, L is the vector sum of the orbital angular momenta ℓ of individual electrons; S is the vector sum of the spins s of individual electrons. The *multiplicity* of terms is normally given as $= (2S + 1)$.

The total angular momentum, J, is the vector sum of $L + S$.

$$J = L + S, \ldots, |L - S|$$

Fine structure transitions involve a change in J, i.e. a change in the electron spin-orbit interaction.

(Magnetic quantum numbers M_L, M_S, and M express the components of L, S, and J in the direction of the magnetic field; but interstellar mag fields are generally weak, and Zeeman splitting undetectable.)

Nuclei with odd numbers of (neutrons plus protons) have a nuclear spin, I. For non-zero J there is *hyperfine* structure, F :

$$F = I + J, \ldots, |I - J|$$

Thus *hyperfine* transitions involve a change in F

1. Example 1: The ground state of neutral H. The ground state of hydrogen is described by

$$1s \quad {}^2S_{1/2}$$

This shows that there is one electron with $s = 1$. The multiplicity is 2, so the total vector sum of (electron) spins is $\frac{1}{2}$ (since mult. = 2S+1) - which is sensible! The orbital angular momentum L is described by 'S', i.e. L=0. The single nuclear proton has spin $I = \frac{1}{2}$ so $F = 1, 0$. The energy separation of the hyperfine levels is 21.1cm.

2. Example 2: The Na D lines. The ground state of neutral sodium is described by

$$1s^2 2s^2 2p^6 3s \quad {}^2S_{1/2}$$

The 'D' lines are the result of a change in L :

$$3p\ {}^2P_{1/2} \rightarrow 3s\ {}^2S_{1/2},$$

$$3p\ {}^2P_{3/2} \rightarrow 3s\ {}^2S_{1/2}.$$

(Hyperfine splitting is also present within the D lines.)

3. Example 3: The 'nebular' lines of O^{2+}. The ground state configuration is

$$1s^2 2s^2 2p^2$$

The nebular lines arise from transitions between the three lowest terms (^3P, ^1D, and ^1S).

^{16}O has $I = 0$ so there is *no* hyperfine structure.

L.1 Transitions

L.1.1 Electric dipole transitions.

So-called 'allowed' transitions give lines of appreciable strength, and correspond to *electric dipole transitions* for classical oscillators. The selection rules for electric dipoole transitions are

$$\Delta \ell = \pm 1$$

$$\Delta S = 0$$

$$\Delta L = 0, \pm 1 \quad (L + L' \geq 1)$$

$$\Delta J = 0, \pm 1 \quad (J + J' \geq 1)$$

The Na D lines (for example) obey these rules; but the 21cm and nebular lines do *not*. They are *forbidden lines* in terms of electric dipole transitions.

L.1.2 Magnetic dipole transitions.

For *magnetic dipole transitions* ('semi-forbidden' lines)

$$\Delta\ell = 0$$

$$\Delta S = 0$$

$$\Delta L = 0$$

$$\Delta J = 0, \pm 1 \quad (J + J' \geq 1)$$

and also

$$\Delta F = 0, \pm 1$$

Thus the 21cm line is allowed as a magnetic dipole transition. The oxygen lines are not.

L.1.3 Electric quadrupole transitions.

For *electric quadrupole transitions* ('forbidden' lines)

$$\Delta\ell = 0, \pm 2$$

$$\Delta S = 0$$

$$\Delta L = 0, \pm 1, \pm 2 \quad (L + L' \geq 2)$$

$$\Delta J = 0, \pm 1, \pm 2 \quad (J + J' \geq 2)$$

The ${}^1D_2 \rightarrow {}^1S_0$ transition is electric quadrupole. The ${}^3P \rightarrow {}^1D$ lines violate $\Delta S = 0$; but that selection rule is not rigorous (it holds strictly only in pure LS coupling, and not at all in jj coupling.)

L.2 Transition probabilities.

Transition probabilities are expressed using the Einstein A coefficient (the probability of spontaneous emission). For optical lines, typically

$$\text{Electric dipole} \qquad A = 10^8 s^{-1}$$

$$\text{Magnetic dipole} \qquad 10^3$$

$$\text{Electric quadrupole} \qquad 10^0$$

Thus for the Na lines, $A \sim 10^8$/s. For O^{2+}

$${}^1D_2 \rightarrow {}^1S_0 \quad 4363 \quad 1.6/s$$

$$^3P_1 \rightarrow ^1 D_2 \quad 4958 \quad 0.007$$

$$^3P_2 \rightarrow ^1 D_2 \quad 5007 \quad 0.021$$

The $^3P_0 \rightarrow ^1 D_2$ has $\Delta J = 2$ and has a much smaller probability ($\sim 10^{-6}$)

What transition probability do we expect for 21cm? A figure $\sim 10^3$ is given above for optical lines; but *(inter alia)* $A \sim \lambda^{-3}$. Since $21cm = 2100 \times 10^6$Å, we expect

$$A \sim 10^3/(10^6)^3 \sim 10^{-15}/s$$

Accurate calculations give

$$A = 2.876 \times 10^{-15}/s$$

The lifetime in the upper state is therefore

$$1/A = 3.5 \times 10^{14} s \sim 10^7 \text{yr}$$

The (de-excitation) timescale in interstellar clouds is such that this spontaneous de-excitation can occur; and the amount of hydrogen is enough to give a strong signal.

Once a 'forbidden'-line photon is emitted the chance of being re-absorbed in the same transition is small. Thus the nebular lines easily escape a nebula, and the 21cm lines can traverse large volumes of cold H gas.

Appendix M

Another go...

M.1 Quantum Numbers of Atomic States

- Principal quantum number n defines the energy level
- Azimuthal quantum number ℓ
  - states with $\ell = 0$ are called s states
  - states with $\ell = 1$ are called p states
  - states with $\ell = 2$ are called d states
  - states with $\ell = 3$ are called f states
- 'orbits' of s states become more distorted as n increases
- Electron transitions take place between adjacent angular momentum states (i.e., $\Delta\ell = 1$)
  - 'sharp series' lines from p to higher s states
  - 'principal series' lines from s to higher p states
  - 'diffuse series' lines from p to higher d states
  - 'fundamental series' lines from d to higher f states
- The first line(s) of the principal series (s to p) are called resonance lines since they involve the ground level
- In alkali metals, the p, d, and f levels are doubled (e.g., the Na D lines) due to the coupling between the magnetic moment of the orbital motion and the spin of the electron (the spin quantum number s, which takes values $+1/2$ or $-1/2$).

M.2 Spectroscopic Notation

- The total angular momentum quantum number is $j = \ell + S$
  - For s states, $j = 1/2$
  - For p states, $j = 1/2$ or $3/2$.
- Electron levels are designated by the notation '$n^2(L)_J$'
- n is the total quantum number
- The superscript 2 indicates the levels are doubled
- L is the azimuthal quantum number (S, P, D, F)
- J denotes the angular-momentum quantum number
- e.g., for sodium the ground level is $3s\,^2S_{1/2}$
- The two lowest p levels are $3p\,^2P_{1/2}$ and $3p\,^2P_{3/2}$
- The Na D lines are described as
- $3s\,^2S_{1/2}$–$3p\,^2P_{3/2}$ λ5889.953 and $3s\,^2S_{1/2}$–$3p\,^2P_{1/2}$ λ5895.923Å

M.3 More Spectroscopic Vocabulary

- The Pauli exclusion principle requires that two s electrons in the same state must have opposite spin
- Therefore $S = 0$, and these are called 'singlet' states
- The ground state of He is a singlet state, $^1S_0$
  - The superscript 1 means singlet
  - The subscript 0 means $J = 0$
- In the first excited state of He, one electron is in the $1s$ state, and the second can be in either the $2s$ or the $2p$ state.
- Depending on how the electron spins are aligned, these states can either be singlets or triplets
- Electrons can only jump between singlet states or between triplet states
- The state of the electrons is described with a 'term' for each electron above the closed shell.

- For carbon atoms, '$1s^2 2s^2 2p^2$' means there are
  - 2 electrons in the $1s$ state
  - 2 electrons in the $2s$ state
  - 2 electrons in the $2p$ state

M.4 Allowed and Forbidden Transitions

- Transitions with $\Delta\ell = 1$ and $\Delta J = 1$ and 0 are 'allowed' (except $J = 0 \rightarrow 0$)
- Other transitions are 'forbidden'
- For some electron states there are no allowed transitions to lower energy states. Such levels are called 'metastable'
- The situation is more complex in atoms with more electrons!
- A multiplet is the whole group of transitions between two states, say ^3P–^3D

M.5 Spectral Line Formation

M.5.1 Spectral Line Formation-Line Absorption Coefficient

- Radiation damping (atomic absorptions and emissions aren't perfectly monochromatic – uncertainty principle)
- Thermal broadening from random kinetic motion
- Collisional broadening – perturbations from neighbouring atoms/ions/electrons)
- Hyperfine structure
- Zeeman effect

M.6 Classical Picture of Radiation

- Photons are sinusoidal variations of electromagnetic fields
- When a photon passes by an electron in an atom, the changing fields cause the electron to oscillate
- We treat the electron as a classical harmonic oscillator:
  mass × acceleration = external force − restoring force − dissipative

M.7 Atomic Absorption Coefficient

- n_0 is the number of bound electrons per unit volume
- the quantity $\nu - \nu_0$ is the frequency separation from line centre
- ε is the dielectric constant ($\varepsilon = 1$ in free space)
- and $\gamma = g/m$ is the classical damping constant

The atomic absorption coefficient includes atomic data (f, ε, γ) and the state of the gas (n_0), and is a function of frequency. The equation expresses the natural broadening of a spectral line.

M.8 The Classical Damping Constant

- For a classical harmonic oscillator,
- The shape of the spectral line depends on the size of the classical damping constant
- For $\nu - \nu_0 >> \gamma/4\pi$ the line falls off as $(\nu - \nu_0)^{-2}$
- Accelerating electric charges radiate.
- The Na D lines have a wavelength of 590 nm, $\gamma = 6.4 \times 10^7$ s^{-1}

M.9 Line Absorption with QM

- Replace γ with Γ
- Broadening depends on lifetime of level
- Levels with long lifetimes are sharp
- Levels with short lifetimes are fuzzy
- QM damping constants for resonance lines may be close to the classical damping constant
- QM damping constants for other Fraunhofer lines may be 5,10, or even 50 times bigger than the classical damping constant

Appendix N

Temperatures

Lots of possibilities, e.g.,

- Effective temperature, $T_{\rm eff}$.
  The temperature of a blackbody which would produce the same amount of radiant energy per unit area as the star (averaged over its surface if necessary).
- Colour temperature, $T_{\rm col}$.
  Measure the colour, normally as the difference of magnitudes in two filters (i.e., the flux ratio between two wavelengths). The colour temperature is the temperature of a blackbody which would produce the flux ratio.
- Brightness temperature, $T_{\rm b}$.
  The temperature of a black body that would reproduce the observed surface brightness at some specific wavelength (or passband).
- Radiation temperature, $T_{\rm rad}$.
  In bolometric radiative equilibrium. In LTE frequency equlibrium
- Excitation temperature, $T_{\rm ex}$
  The temperature for which the Boltzmann equation reproduces an observed ratio of two energy levels in a given species.
- Ionization temperature, $T_{\rm I}$
  The temperature for which the Saha equation reproduces an observed ratio of two ionization stages for a given element.
- Kinetic temperature, $T_{\rm k}$.
  If (as is usually the case in stellar astrophysics) the velocities of particles in a gas follow a Maxwellian distribution, the kinetic temperature is what characterizes the magnitude of the

motions. In astrophysics, we can't measure the velocity distribution directly, and so it is customary to *assume* a Maxwell-Boltzmann distribution (unless there are reasons not to). A typical application would then be to use this assumption in estimating the kinetic temperature of electrons that is required to match the level populations for a collisionally excited emission line, giving the

- Electron temperature, $T_{\rm e}$.
  Other species-specific temperatures may be used as surrogates for the kinetic temperature; e.g., the ionic temperature, $T_{\rm ion}$ (understood to reflect the motion of nuclei in a gas).

In thermodynamic equilibrium (TE) $T_{\rm ex} = T_{\rm I} = T_{\rm k}$ (and the same kinetic temperature holds for all types of particle); we might collectively refer to all these descriptors as 'the gas temperature', $T_{\rm gas}$. In this case, it should be clear that the velocity distributions are Maxwellian, and the ionization and excitation are described by Saha-Boltzmann. Moreover, in TE, Kirchhoff's law

Real systems are very rarely in true thermal equlib

Jo:

LTE refers to the occupation numbers, i.e., Saha Boltzmann, and that the velocity distribution of ALL particles follows a Maxwellian with the same temperature (i.e., $T_{\rm e} = T_{\rm ion}$)

THUS: 'true' opacities and emissivities follow the Kirchhoff–Planck law, i.e.,
$\eta_{\rm true} = \kappa_{\rm true} * B_\nu \rightarrow S_{\rm true} = B_\nu$

But: there is (almost) always scattering, which also in LTE makes life complicated.

Thus $S_{\rm total} = (\kappa_{\rm true} * S_{\rm true} + \kappa_{\rm scat} * J)/(\kappa_{\rm true} + \kappa_{\rm scat})$

(in the above line, it could be also some integrals over J or I, if one accounts for partial redistribution).

Moreover, $J \neq B$, because at the outer boundary (and such a thing is defined in LTE, though not in TE), the photons can escape, and then $J < B$. Of course as well, $I \neq B$.

b) What you are after is thermalization. For large optical depths, $J \rightarrow B$, and thus $S_{\rm total} \rightarrow B$, since the scattering part becomes thermalized.

In this case, $J \simeq S \simeq B$. But I is still angle dependent, because otherwise you would have no flux (isotropic I gives flux = 0)

All this can be also described in terms of the diffusion approximation, where the same is valid.

Conclusion: In LTE, $J \simeq S \simeq B$, but I has an angle-dependent term, which can be approximated by $I \simeq B + \mu {\rm d}B/{\rm d}\tau$, BUT ONLY FOR LARGE OPTICAL DEPTH.

In that part where the light is emitted, i.e., for optical depths below unity, one has only that $S_{\text{true}} = B$, and nothing else.

Webby stuff:

In astrophysics the distinction between 'thermal equilibrium' and 'thermodynamic equilibrium' is not carefully made, because there is rarely if ever a situation in this context in which thermal equilibrium might hold without thermodynamic equilibrium.

The most common situation in which the presence or absence of LTE is considered is for a star. There is a flow of heat from the interior of a star to its atmosphere at large radii, and the temperature varies as a function of radius. So clearly gas near the atmosphere is not in thermodynamic (or thermal) equilibrium with gas in the core.

However, models of stellar interiors can be vastly simplified if one recognizes that there is still a local thermodynamic (thermal) equilibrium in the sense that the kinetic distribution of the free electrons, the plasma ionization state, and the radiation field at each radius can all be very well described by a single number, a local temperature TT. The electron velocities obey a Maxwell-Boltzmann distribution, the ionization state follows from the Saha equation, and the radiation field is described by a Planck function (blackbody), all evaluated for some common TT.

As the link you provide explains, this works as long as the mean free path of any particles that might transport heat (e.g. photons, electrons) is very small compared to the length scale over which the temperature is changing. In the atmospheres of stars, where the photon mean free path grows large, LTE in the above sense can break down. However, if the electron mean free path is still small enough, it can helpful to apply an even more limited notion of LTE in which one takes the electron velocities to be in LTE, while acknowledging that the radiation field may depart from a Planck function.

We can't treat the entirety of a star as being in TE – obviously, the core is hotter than the photospheric layers, and no one temperature describes the entire object. However, if conditions don't change 'significantly' over some 'small' distance – where 'significantly' and 'small' will be discussed shortly – then we can say that a system is in *local* thermodynamic equilibrium, or LTE.

Notice that here LTE is described entirely by the properties of the particles; radiation has not made an appearance. But of course, radiation almost always has some role to play; the question is, how important is that role?

For example: on a sunny day in temperate latitudes, the gas temperature...

Index (experimental – work in progress!!)