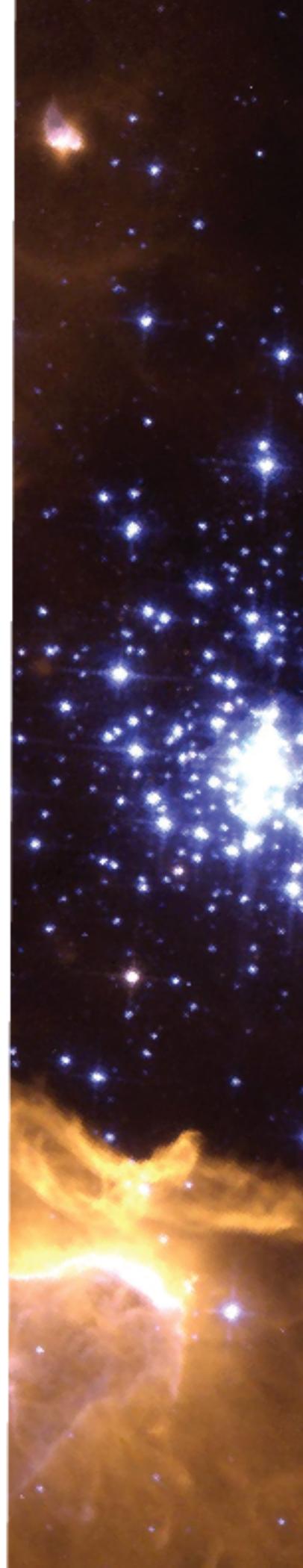


EDWARD BROWN

# TO BUILD A STAR

OPEN ASTROPHYSICS BOOKSHELF



*About the cover: This Hubble image of the nebula NGC 3603 shows stars at different life stages, from birth to main-sequence and beyond.*

Credit: Wolfgang Brandner (JPL/IPAC), Eva K. Grebel (U. Wash.), You-Hua Chu (UIUC), NASA

© 2018, 2020 Edward Brown  
git version 856bdf4 ...



Except where explicitly noted, this work is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.

## Preface

These notes grew out of a collection of handouts and exercises that I wrote while teaching the junior/senior undergraduate course on stars at Michigan State University in the autumn semesters of 2012, 2014, and 2016. In addition to deriving a basic physical description of how stars work, a secondary goal of the course was to train students to make simple physical models and order-of-magnitude estimates. This is a crucial skill that is not incorporated enough into the typical undergraduate physics courses. In keeping with this goal, many of the exercises asked the students to make estimates or to employ simple models, such as constant density throughout the star, rather than to perform elaborate calculations.

Concern about rising costs motivated me during the spring and summer of 2018 to assemble the handouts and exercises into a package that could be inexpensively distributed to students and eliminate the need for a required textbook. As I prepared for the transition to online teaching in the spring and summer of 2020, I added notes on three common numerical tasks: finding roots, solving ordinary differential equations, and interpolating tabulated data. These methods are used in the group computational project that is part of my course. Because there are many excellent references and numerical libraries available for these tasks, the goal of the appendix is just to introduce the techniques.

THERE ARE SEVERAL OPTIONS FOR THE ORDER IN WHICH TO PRESENT MATERIAL. One would be to start with chapter 2, which covers hydrostatic equilibrium and establishes estimates for the mean stellar density, pressure, and temperature. The material in chapter 1 on radiant intensity, flux, and thermal emission would then be introduced in chapters 3 and 4, which cover radiative heat transport in the stellar interior and the conditions at the photosphere. The remainder of chapter 1 on magnitudes would then come later, perhaps in chapter 6 where we discuss the main sequence.

Although this order is logical, after deliberation I decided on the layout used here for several reasons. First, radiative transfer is a difficult subject, and introducing the basic concepts early gives the students more time to become familiar with the topic. Second, finding Wien's law requires a numerical rootfind (exercise 1.5), and this is a good warmup for further numerical projects. Finally, the discussion of radiative intensity allows us to introduce magnitudes and color indices, thereby making contact with the subject's observational foundations at the start.

THE TEXT LAYOUT USES THE TUFTE-BOOK L<sup>A</sup>T<sub>E</sub>X CLASS<sup>1</sup>. The main features are a large outer margin in which the students can take notes and

<sup>1</sup><https://tufte-latex.github.io/tufte-latex/>

the tight integration of text, figures, and sidenotes. Sixty-six exercises, ranging from comprehension checks to longer, more challenging problems, are embedded throughout the text. A few of the exercises have a numerical component, denoted with a “” symbol. Because the exercises are spread throughout the text, there is a “List of Exercises” in the front. I’ve also added boxes containing more advanced material that I felt students should be exposed to, but were not essential to the main development of the course.

ONE EVENING I TRIED TO ENLIVEN THE CHAPTER TITLES. I noticed that the first two chapters had titles that were also titles for pop songs. I then decided to find song titles that would fit for the remaining chapters. When selecting titles, I imposed a rule that they all could plausibly go together on a playlist. This was challenging since the chapters originally had titles such as “The equation of state” and “The radiative opacity”. The credits for the chapter titles, in order, go to Muse, Queen/David Bowie, Greta van Fleet, Dio, Deep Purple, David Bowie, The Traveling Wilburys, and Muse.

PLEASE BE ADVISED THAT THESE NOTES ARE UNDER ACTIVE DEVELOPMENT. To refer to a specific version of the notes, use the eight-character stamp labeled “git version” on the copyright page.

## Contents

1	<i>Starlight</i>	1
2	<i>Under Pressure</i>	11
3	<i>Edge of Darkness</i>	25
4	<i>Rainbow in the Dark</i>	41
5	<i>Burn</i>	55
6	<i>Star</i>	69
7	<i>End of the Line</i>	87
A	<i>Algorithm</i>	99
	<i>Bibliography</i>	115



# Figures

1.1	The electric force in a light wave	2
1.2	Schematic of radiative intensity	2
1.3	Schematic of intensity being constant	3
1.4	Thermal spectra	4
1.5	Standard filters	8
1.6	Hipparcos color-magnitude diagram	9
2.1	A fluid element in hydrostatic equilibrium	11
2.2	The mass of a column of fluid	12
2.3	Fall to center	16
2.4	Star formation in Rho Ophiuchi	23
3.1	Schematic of mean free path	25
3.2	Mean free path of a hockey puck	26
3.3	Coordinates for radiative transport equation	29
3.4	The specific flux for a hypothetical opacity	30
3.5	Transport along a gradient	32
3.6	Schematic of a random walk	32
3.7	Distribution of positions in a random walk	33
4.1	Visible spectrum of the sun	41
4.2	Hertzsprung-Russell diagram of standard main-sequence stars	42
4.3	Spectral lines of neutral hydrogen	43
4.4	Standard stellar types	48
4.5	H $\gamma$ absorption line	48
4.6	Comparison of Lorentzian and Gaussian distributions	49
4.7	Spectra of two A1 stars	53
5.1	Schematic of the nuclear potential	56
5.2	Tunneling through the Coulomb potential barrier	59
5.3	Heat balance in a mass shell	67
6.1	Onset of convection	70
6.2	A boat with a weight	70
6.3	Illustration of criteria for convective instability	71

6.4	Solar convection cells	76	
6.5	Pre-main sequence stars	77	
6.6	Image of the massive star Eta Carinae	84	
7.1	Color-magnitude diagram for the globular cluster M55	88	
7.2	The planetary nebula NGC 2392	90	
7.3	Betelgeuse	91	
7.4	Schematic of a pulsar	96	
A.1	First bisection steps	101	
A.2	First five bisections	101	
A.3	Schematic of Newton's method	102	
A.4	Brent's method	103	
A.5	Second iteration, Brent's method	103	
A.6	Third iteration, Brent's method	104	
A.7	Solution for a sample set of ODE's	106	
A.8	Schematic of a single forward Euler step	106	
A.9	The second-order Runge-Kutta method	107	
A.10	The fourth-order Runge-Kutta method	108	
A.11	Cubic polynomial with varying right-hand slopes	110	
A.12	Two splines connected at an interior point	110	
A.13	Second derivatives of our spline functions.	110	
A.14	Example of a spline fit	111	

# *Tables*

2.1	Selected atomic masses	13	
2.2	Masses, radii, and luminosities for selected stellar types	22	
5.1	Liquid-drop coefficients	57	
5.2	Selected atomic mass excesses	58	
5.3	Parameters for non-resonant reactions	64	
6.1	Central densities and temperatures of zero-age main-sequence stars	78	
6.2	Characteristics of main-sequence stars	86	
7.1	Nuclear burning timescales for massive stars	91	



## *Boxes*

1.1	Solid angles	3	
1.2	Momentum transport and radiation pressure	6	
2.1	The sound speed	17	
2.2	Working with vectors	20	
3.1	Expansion in Legendre polynomials	36	
3.2	Decomposition of intensity into moments	39	
4.1	The partition function for neutral hydrogen	44	
4.2	The driven damped oscillator	50	
5.1	The thermally averaged reaction cross-section	61	
6.1	Along an adiabat	73	
6.2	The equations of stellar structure in Lagrangian form	77	
6.3	Identical particles	79	
7.1	Instability for a relativistic equation of state	94	
A.1	Interpolation	102	
A.2	Functions	105	
A.3	Solving a tridiagonal system	112	



# Exercises

1.1	Solar power	1
1.2	Flux from a distant star	2
1.3	Photon flux from the sun	2
1.4	Proof that $I_\lambda$ is conserved	3
1.5	Peak of thermal spectrum 	4
1.6	Frequency peak in thermal spectrum 	4
1.7	No net flux for thermal emission	5
1.8	Filter for observing sun-like star	8
1.9	Color index and temperature	9
1.10	An examination of the color-magnitude diagram	10
2.1	Pressure increase in the ocean	12
2.2	The isothermal atmosphere	13
2.3	Mean molecular weight for ionized helium	14
2.4	At the center	15
2.5	A star of constant density	15
2.6	Sound-crossing time	17
2.7	Applications of virial scalings	22
2.8	The relation between energy and temperature	22
2.9	The oscillation period of a star	23
2.10	Contraction of a constant density protostar	24
3.1	Mean free path of a hockey puck	26
3.2	Mean free path for electron scattering	26
3.3	Attenuation of light in an absorbing medium	27
3.4	Combined absorption and emission	27
3.5	Optical depth of the solar center	28
3.6	Formal solution of radiative transfer	28
3.7	Transport by frequency	30
3.8	Radiative transfer equation	31
3.9	Rosseland weighting	32
3.10	Radiative diffusion as a random walk	33
3.11	Gray emissivity?	34
3.12	Photospheric pressure	35

3.13	Odd-even powers of $\mu$	37	
3.14	The Eddington closure scheme	39	
3.15	Anisotropy of the intensity	40	
4.1	Partition function for neutral hydrogen	45	
4.2	Conditions for strong Balmer lines	47	
5.1	Depth of nuclear well	56	
5.2	If the strong force were long-range	57	
5.3	The nuclear landscape 	58	
5.4	Heat from hydrogen fusing to helium	58	
5.5	Energy released by various reactions	59	
5.6	Turning radius for proton-proton collision in solar plasma	60	60
5.7	Approximating a function as a power-law	64	
5.8	Conservation laws applied to reactions	65	
6.1	A boat with a weight	70	
6.2	Adiabatic relations	74	
6.3	Onset of convection	75	
6.4	Temperature and density within a star	76	
6.5	Central temperature and density during contraction	78	78
6.6	The mass-radius relation for a degenerate EOS	83	83
6.7	Minimum stellar mass	84	
6.8	Planetary masses and radii	84	
6.9	Radiation pressure	85	
6.10	Maximum stellar mass	85	
6.11	Nuclear burning timescale	85	
6.12	Mass-luminosity relation	86	
7.1	Horizontal branch lifetime	89	
7.2	Dynamical time of evolved stellar core	90	
7.3	Gravitational binding energy of a neutron star	95	95
7.4	Limiting spin frequency	97	
7.5	Accretion onto a neutron star	97	
A.1	Binary representation	100	
A.2	Spacing of model numbers	100	
A.3	Second-order Adams-Bashforth method	107	
A.4	Cubic polynomial	110	

# 1

## Starlight

### 1.1 Introduction: Our Sun

Let's start by considering the star we know best: the sun<sup>1</sup>. Measurements of the Earth-Sun distance and hence the size of the sun began in antiquity<sup>2</sup>. Observations of the planets' orbital periods, when combined with Kepler's laws, determine the planets' relative distances from the sun. The distance scale is then fixed using radar ranging of objects in the solar system, with the mean Earth-Sun distance, known as an ASTRONOMICAL UNIT, being defined<sup>3</sup> as

$$1 \text{ au} = 1.495\,978\,707\,00 \times 10^{11} \text{ m.}$$

From the sizes and periods of the planetary orbits, along with Newton's theory of gravity, we deduce the mass of the sun<sup>4</sup>:

$$M_{\odot} = 1.99 \times 10^{30} \text{ kg.}$$

The sun is roughly  $10^6$  times more massive than the Earth and 1000 times more massive than Jupiter. Knowing the Earth-Sun distance and the angular size of the sun, about  $0.5^\circ$  across its diameter, gives us its radius:

$$R_{\odot} = 6.96 \times 10^8 \text{ m.}$$

The total radiant power emitted by a star is its LUMINOSITY  $L$ . A detector with a collecting area  $\mathcal{A}$  located a distance  $d$  from a star intercepts a fraction  $\mathcal{A}/(4\pi d^2)$  of the star's light. We call  $F = L/(4\pi d^2)$  the FLUX, which has units  $\text{W m}^{-2}$ . From the mean solar flux<sup>5</sup> at 1 au,  $1360.8 \pm 0.5 \text{ W m}^{-2}$ , we infer the sun's luminosity:

$$L_{\odot} = 3.83 \times 10^{26} \text{ W.}$$

<sup>1</sup> The symbol  $\odot$  is used to denote the sun

<sup>2</sup> See Weinberg [2015] for a physicist's perspective on how this was done.

<sup>3</sup> Brian Luzum et al. The IAU 2009 system of astronomical constants. *Celestial Mechanics and Dynamical Astronomy*, 110(4):293–304, August 2011. DOI: 10.1007/s10569-011-9352-4

<sup>4</sup> Note that  $GM_{\odot}$  is known much more precisely than  $G$  or  $M_{\odot}$  separately.

<sup>5</sup> Greg Kopp and Judith L. Lean. A new, lower value of total solar irradiance: Evidence and climate significance. *Geophys. Res. Lett.*, 38(1):L01706, January 2011. DOI: 10.1029/2010GL045777

---

EXERCISE 1.1 — Suppose we wish to replace the MSU power plant—rated at 70 MW ( $70 \times 10^6 \text{ W}$ )—with a grid of solar panels. Under ideal conditions (direct light and 100% efficient panels), how many square meters of solar panels are needed?

---

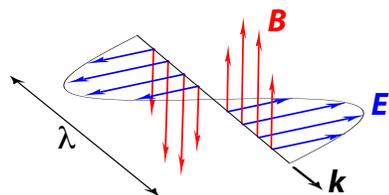


Figure 1.1: Schematic of the electric field (blue arrows) and magnetic field (red arrows) for a wave traveling along direction  $\mathbf{k}$  with wavelength  $\lambda$ .

<sup>6</sup> This velocity is exact; the meter is defined in terms of the speed of light.

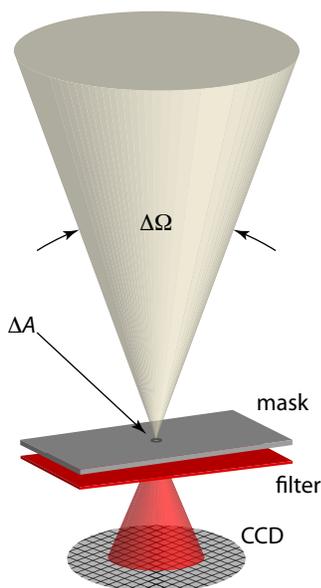


Figure 1.2: Schematic of radiative intensity

EXERCISE 1.2— What would the flux be from a star with  $L = 0.1L_{\odot}$  at a distance of 10 pc? Recall that a parsec (pc) is defined by the relation

$$\frac{1 \text{ au}}{1 \text{ pc}} = 1'' = \frac{1}{206265}.$$

Our information about the sun and the more distant stars is carried by light. Before going further, then, let's establish some basic properties of light and how observations are made.

## 1.2 Intensity and specific flux

When we detect light, what happens at the atomic level is that the charges in our detector (a CCD, an eye, a photographic emulsion) feel an electric (and magnetic) force that oscillates with frequency  $\nu$ . Imagine setting up a grid of detectors that measure the electric and magnetic forces per charge at each point in space and at each instant of time. We call these forces per charge the electric and magnetic fields,  $\mathbf{E}(\mathbf{x}, t)$  and  $\mathbf{B}(\mathbf{x}, t)$ . As light passes through this grid, we would notice a sinusoidal pattern, with amplitude proportional to  $|\mathbf{E}|^2 + |\mathbf{B}|^2$ , traveling at speed<sup>6</sup>  $c = 299792458 \text{ m/s}$  with a wavelength  $\lambda = c/\nu$ .

Take a detector and place (see Fig. 1.2) in front of the detector a filter that only lets through light with wavelengths in a range  $\Delta\lambda$ . Place a mask over the detector with a small pinhole of area  $\Delta A$  that restricts the light falling on the detector to fall in a narrow cone of solid angle  $\Delta\Omega$  about the normal to the detector. Then measure the energy  $\Delta E$  incident on the detector in a time  $\Delta t$ . The quantity

$$I_{\lambda} \equiv \frac{\Delta E}{\Delta t \Delta A \Delta \lambda \Delta \Omega} \quad (1.1)$$

is known as the *INTENSITY*, and is the basic quantity describing radiation. At sufficiently low intensity, we would find that energy is deposited into our detector in discrete quanta, known as *PHOTONS*, with each photon having an energy  $hc/\lambda = h\nu$ , with  $h = 6.63 \times 10^{-34} \text{ J s}$  being *PLANCK'S CONSTANT*. The light emitted by the sun (or any other source) consists of a huge number of photons; the distribution of photons over a range of wavelengths is known as a *SPECTRUM*.

EXERCISE 1.3— The peak of the sun's spectrum is at a wavelength of approximately 500 nm. Estimate the number of photons from the sun striking  $1 \text{ m}^2$  of Earth each second.

In situations in which the wavelength is small (relative to the system in question so we can neglect diffraction), light propagates along *RAYs*. By a ray of light, we mean the light emitted into a small cone of opening

solid angle  $d\Omega$  about a direction  $\hat{\mathbf{k}}$ . In the absence of any interactions with matter, the intensity is conserved along a ray if both source and receiver are stationary with respect to one another (Exercise 1.4).

### Box 1.1 Solid angles

Imagine that you are at the center of a great sphere of radius  $R$ , and you shine a light that emits rays into some solid angle. Orient your coordinates so that the rays are traveling along the  $z$ -axis. The light will illuminate an area

$$A = R^2 \int_0^{2\pi} \int_0^\theta \sin \theta \, d\theta \, d\phi.$$

Here  $\theta$  is the opening half-angle of the cone. The solid angle into which the light is emitted is  $\Omega = A/R^2$ . Astronomers often express the integral by changing variables to  $\mu = \cos \theta$ , so that the solid angle is

$$\Delta\Omega = \int_0^{2\pi} \int_{1-\Delta\mu}^1 d\mu \, d\phi.$$

If we integrate over all angles ( $0 \leq \theta \leq \pi$ , or  $-1 \leq \mu \leq 1$ ), then we get the area of a sphere,  $A = 4\pi R^2$ .

EXERCISE 1.4— Your friend (at top in Fig. 1.3) flashes a light: in a time  $\Delta t$  it emits energy  $\Delta E_{\text{emit}}$  in a waveband  $\Delta\lambda$ . The opening through which the light passes has area  $\Delta A_{\text{emit}}$ , and the light goes into a cone of opening solid angle  $\Delta\Omega_{\text{emit}}$ . Your friend therefore calculates her intensity as

$$I_{\lambda,\text{emit}} = \frac{\Delta E_{\text{emit}}}{\Delta t \Delta A_{\text{emit}} \Delta\lambda \Delta\Omega_{\text{emit}}}.$$

You stand (at bottom in Fig. 1.3) with a camera, the aperture of which has area  $\Delta A_{\text{obs}}$ , a distance  $d$  ( $d^2 \gg \Delta A_{\text{emit}}, \Delta A_{\text{obs}}$ ) from your friend. You measure an intensity  $I_{\lambda,\text{obs}}$ . Show that  $I_{\lambda,\text{obs}} = I_{\lambda,\text{emit}}$  by doing the following.

1. Calculate the incident energy that falls on your camera aperture  $\Delta E_{\text{obs}}$ .
2. Find the solid angle  $\Delta\Omega_{\text{obs}}$  subtended by the rays entering the aperture.
3. Now use parts 1 and 2 to compute your intensity

$$I_{\lambda,\text{obs}} = \frac{\Delta E_{\text{obs}}}{\Delta t \Delta A_{\text{obs}} \Delta\lambda \Delta\Omega_{\text{obs}}}.$$

and show that this is the same as  $I_{\lambda,\text{emit}}$ .

A related quantity is the SPECIFIC FLUX  $F_\lambda$ , which is the energy carried by light having wavelength in a range  $\Delta\lambda$  crossing an area  $\Delta A$  (from all angles) in a time  $\Delta t$ . We compute  $F_\lambda$  by multiplying the intensity by  $\cos \theta$ , where  $\theta$  is the angle between the ray and the normal of our area<sup>7</sup>

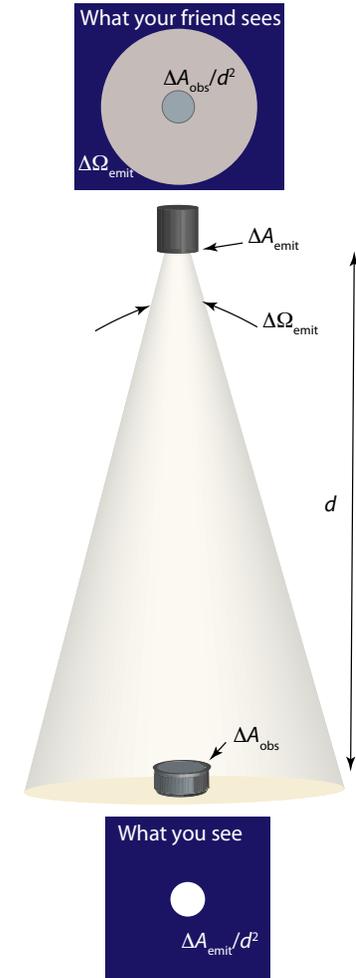


Figure 1.3: Schematic for exercise 1.4.

<sup>7</sup> this gives the projected area

and integrating over angle:

$$F_\lambda = \int I_\lambda \cos \theta \sin \theta d\theta d\phi. \quad (1.2)$$

### 1.3 Thermal emission

Imagine we had a material that emits and absorbs equally well at all wavelengths. We then made from this material a hollow box, and we heated this box to a temperature  $T$ . The hot atoms in the walls of the box would emit (and absorb) photons bouncing around in the cavity in this box, until the photons were in thermal equilibrium<sup>8</sup> with the walls of the box. If we then drilled a small hole in the side of the box, some photons would escape (but not so many as to disturb the thermal equilibrium). The intensity emerging from such a box is known as the **PLANCK SPECTRUM**:

$$I_\lambda(\text{Planck}) \equiv B_\lambda(T) = \frac{2hc^2}{\lambda^5} \left[ \exp\left(\frac{hc}{\lambda k_B T}\right) - 1 \right]^{-1}. \quad (1.3)$$

Here  $k_B = 1.381 \times 10^{-23} \text{ J K}^{-1}$  denotes **BOLTZMANN'S CONSTANT**. This spectrum is also known as a **BLACKBODY SPECTRUM**, because it is emitted from a material that absorbs (and therefore emits) equally well at all wavelengths. The emission is peaked at a wavelength  $\lambda_{\text{pk}} \sim hc/k_B T$ . Fig. 1.4 displays Planck's spectra for various temperatures. Note that  $B_\lambda$  increases at all wavelengths as the temperature increases.

<sup>8</sup> Meaning that the radiation field is on average neither gaining or losing energy from the walls of the box

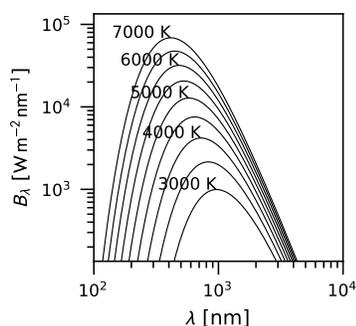


Figure 1.4: Thermal spectra for temperatures ranging from 3000 K to 7000 K.

---

**EXERCISE 1.5**— Show that the peak of the thermal spectrum, temperature  $T$ , occurs (i.e., where  $B_\lambda$  is maximum) at a wavelength

$$\lambda_{\text{pk}} = 290 \text{ nm} \left( \frac{10\,000 \text{ K}}{T} \right).$$

This result is known as **WIEN'S LAW**. Check this: what is the peak wavelength of the sun's emission? What is the peak wavelength for the cosmic microwave background ( $T_{\text{CMB}} = 2.73 \text{ K}$ )?  *Hint*: In finding the peak, you may need to find the root of a function numerically; that is, you will have an expression  $f(x)$  and you want to find  $x_r$  such that  $f(x = x_r) = 0$ . There are multiple ways to do this; see the notebooks `bisection` and `brent` in the `numerical-methods/notebooks` folder.

---

The Planck spectrum, expressed in terms of frequency, is

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \left[ \exp\left(\frac{h\nu}{k_B T}\right) - 1 \right]^{-1}. \quad (1.4)$$

---

**EXERCISE 1.6**— What is the frequency corresponding to  $\lambda_{\text{pk}}$  in Exercise 1.5?  Compute the frequency  $\nu_{\text{pk}}$  at which  $B_\nu$  is maximum. Is  $\nu_{\text{pk}}$  the same as the frequency corresponding to  $\lambda_{\text{pk}}$ ?

---

Using Eq. (1.2), we can compute the specific flux of this thermal radiation. Since  $B_\lambda$  doesn't depend on angle, the integral is easy:

$$F_\lambda = B_\lambda \int_0^{2\pi} \int_0^\pi \cos \theta \sin \theta \, d\theta \, d\phi = 0.$$

---

EXERCISE 1.7 — Explain, without using mathematical expressions, why there is no net flux for thermal emission.

---

Although the net flux is zero, if we just want the radiation escaping from our cavity, we should only integrate over the angles  $0 \leq \theta \leq \pi/2$ . If we do this, then our *outward-going* specific flux is

$$F_\lambda(\text{outward}) = B_\lambda \int_0^{2\pi} \int_0^{\pi/2} \cos \theta \sin \theta \, d\theta \, d\phi = \pi B_\lambda. \quad (1.5)$$

To find the total power emitted per area for thermal radiation, we need to integrate  $F_\lambda$  over wavelength:

$$F = \int_0^\infty F_\lambda(\text{outward}) \, d\lambda = \int_0^\infty \frac{2\pi hc^2}{\lambda^5} \frac{d\lambda}{\exp(hc/\lambda k_B T) - 1}. \quad (1.6)$$

By changing variables to  $x = hc/\lambda k_B T$ , we can write this integral as

$$F = \frac{2\pi k_B^4}{h^2 c} T^4 \times \underbrace{\int_0^\infty \frac{x^3}{e^x - 1} \, dx}_{=\pi^4/15} = \left[ \frac{2\pi^5}{15} \frac{k_B^4}{h^2 c} \right] T^4.$$

The integral over  $x$  can be converted into  $3! \times \zeta(4)$ , where  $\zeta$  is the Riemann zeta function; for further information, consult a text on mathematical methods in physics.

The quantity in  $[\cdot]$  is called the STEFAN-BOLTZMANN CONSTANT:

$$\sigma_{\text{SB}} = 5.7 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4};$$

The total energy radiated per second per area from a thermal emitter of temperature  $T$  is thus  $\sigma_{\text{SB}} T^4$ .

REAL STARS ARE NOT BLACKBODIES! That being said, their spectra are roughly thermal, so we can define an EFFECTIVE SURFACE TEMPERATURE

$$T_{\text{eff}} = \left[ \frac{F}{\sigma_{\text{SB}}} \right]^{1/4}.$$

The total power output, or *luminosity*, of a star of radius  $R$  is thus

$$L = 4\pi R^2 \sigma_{\text{SB}} T_{\text{eff}}^4. \quad (1.7)$$

For the sun,  $T_{\text{eff}} = 5780 \text{ K}$ .

### 1.4 The radiation energy density

We introduced

$$I_\lambda \equiv \frac{dE}{dt dA d\lambda d\Omega}$$

as the radiant energy  $dE$  crossing an area  $dA$  in a time  $dt$ , directed into a solid angle  $d\Omega$ , and carried by photons with wavelengths in a range  $d\lambda$ . Notice that in time  $dt$ , these photons will fill a volume  $dV = cdt \times dA$ . Hence we can write the intensity as

$$I_\lambda = c \frac{dE}{dV d\lambda d\Omega}.$$

Using this expressions, we define the radiant energy density per wavelength as

$$U_\lambda \equiv \frac{dE}{dV d\lambda} = \frac{1}{c} \int I_\lambda d\Omega. \quad (1.8)$$

If the radiation is thermal, that is, if  $I_\lambda = B_\lambda$ , then

$$U_\lambda = \frac{B_\lambda}{c} \int_0^{2\pi} \int_0^\pi \sin \theta d\theta d\phi = \frac{4\pi}{c} B_\lambda,$$

and the total radiant energy density is

$$U = \int_0^\infty U_\lambda d\lambda = \frac{4}{c} \pi \int_0^\infty B_\lambda d\lambda = \left[ \frac{4\sigma_{\text{SB}}}{c} \right] T^4.$$

In getting this result, we used equations (1.5) and (1.6). With

$$a \equiv \left[ \frac{4\sigma_{\text{SB}}}{c} \right] = 7.566 \times 10^{-16} \text{ J m}^{-3} \text{ K}^{-4},$$

the energy density of radiation in thermal equilibrium is  $U = aT^4$ .

It is common to denote the average (over angle) intensity as

$$J_\lambda = \frac{1}{4\pi} \int I_\lambda d\Omega; \quad (1.9)$$

the specific energy density is thus

$$U_\lambda = \frac{4\pi}{c} J_\lambda.$$

#### Box 1.2 Momentum transport and radiation pressure

In addition to transporting energy, photon also carry momentum. You will learn in your quantum mechanics course that the momentum of a photon of energy  $h\nu$  traveling along direction  $\hat{\mathbf{k}}$  is

$$\mathbf{p} = \frac{h\nu}{c} \hat{\mathbf{k}} = \frac{h}{\lambda} \hat{\mathbf{k}}.$$

## Box 1.2 continued

Here  $\nu$  and  $\lambda = c/\nu$  are the frequency and wavelength of the photon. Hence the momentum carried by photons of energy  $E_\nu$  along direction  $\hat{\mathbf{k}}$  is  $E/c$ . Since  $I_\nu$  is the amount of energy carried by photons per area per time along the direction  $\hat{\mathbf{k}}$ , the momentum transported by those photons per area per time along direction  $\hat{\mathbf{k}}$  must be  $(I_\nu/c)\hat{\mathbf{k}}$ .

TO RELATE THIS MOMENTUM TRANSPORT TO THE RADIATION PRESSURE, suppose we have a sheet of absorbing material with a normal  $\hat{\mathbf{n}}$  being impinged by a ray of photon traveling along  $\hat{\mathbf{k}}$ . As the photons are absorbed, they transfer momentum (along direction  $\hat{\mathbf{n}}$ ) of  $(E_\nu/c)\hat{\mathbf{n}} \cdot \hat{\mathbf{k}}$  to the matter. The projected area of the ray on the matter is  $dA \hat{\mathbf{n}} \cdot \hat{\mathbf{k}}$ . The rate of momentum transfer along  $\hat{\mathbf{n}}$  per area per frequency is therefore

$$\begin{aligned} P_\nu &= \frac{1}{c} \int I_\nu \underbrace{(\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})}_{\text{proj. area}} \underbrace{(\hat{\mathbf{n}} \cdot \hat{\mathbf{k}})}_{\text{comp. of } \mathbf{p} \text{ along } \hat{\mathbf{n}}} d\Omega \\ &= \frac{1}{c} \int_0^{2\phi} \int_{-1}^1 I_\nu \mu^2 d\mu d\phi. \end{aligned} \quad (1.10)$$

A change in momentum per time is a force; hence equation (1.10) represents the force per area, or PRESSURE, exerted by photons with frequencies in  $[\nu, \nu + d\nu]$ . The two factors of  $\mu = \cos \theta$  account for the projected area and the component of momentum along the normal to the surface  $\hat{\mathbf{n}}$ .

If the radiation is thermal, so that  $I_\nu = B_\nu$  and is independent of angle, then

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

We can integrate  $P_\nu$  over frequency to get the total radiation pressure,

$$P_{\text{rad}} = \frac{4\pi}{3c} \int_0^\infty B_\nu d\nu = \frac{4}{3c} \sigma_{\text{SB}} T^4 = \frac{1}{3} a T^4. \quad (1.11)$$

Note that the pressure is 1/3 of the energy density for thermal radiation. This is in general true for a gas of relativistic particles that have momentum proportional to energy.

## 1.5 Magnitudes

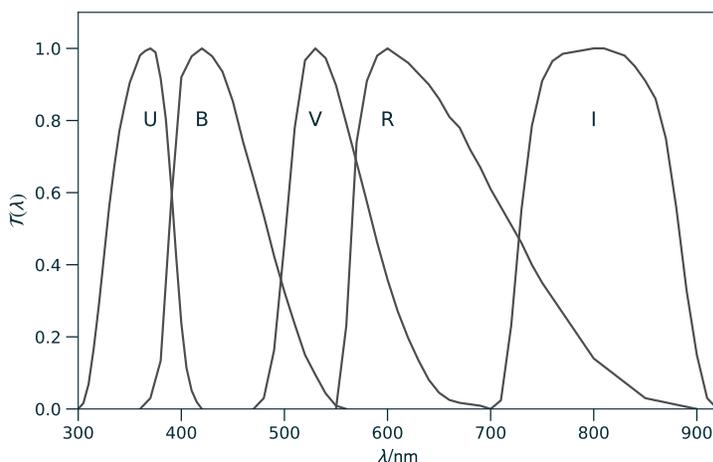
When observing a star, astronomers are collecting light over a range of frequencies. To compare observations, astronomers typically pass the light through standard filters and measure the transmitted flux. The flux

in a given band is then

$$F_{\text{band}} = \int F_{\lambda} \mathcal{T}(\lambda) d\lambda.$$

Here  $\mathcal{T}(\lambda)$  is the TRANSMISSION FUNCTION for that filter and specifies how much light is let through as a function of wavelength. The transmission functions for some common UV/optical/IR filters are shown in Figure 1.5. For example, the V-band filter is centered at  $\lambda = 551$  nm and has a width at half-max of 88 nm.

Figure 1.5: Some standard UV/optical/IR filters. The  $\mathcal{T}(\lambda)$  are normalized so that  $\max(\mathcal{T}) = 1$ .




---

EXERCISE 1.8— Suppose you wished to observe a sun-like star, and you wanted to observe wavelengths near the peak of the spectrum. Which filter would you choose, and why? What about for a star with a surface effective temperature  $T_{\text{eff}} = 8000$  K?

---

When making observations, a common practice is to compare fluxes in a particular band between two stars. Optical astronomers define the *apparent magnitude*  $m$  via

$$m(A) - m(B) = -2.5 \log \left[ \frac{F(A)}{F(B)} \right]. \quad (1.12)$$

Here  $F(A)$  and  $F(B)$  are two different measurements of flux (from two different stars, for example) in a particular waveband. In place of  $m$ , the usual practice, is to simply use the label of the waveband. For example, when an astronomer says, “The V-magnitude is 16.6”, what she means is that the apparent magnitude measured in the V-filter is 16.6.

Because equation (1.12) is a flux ratio, the absolute scale is not defined. Conversion from a magnitude to physical units, such as  $\text{W m}^{-2} \text{nm}^{-1}$ , is

NB. Throughout this text,  $\log \equiv \lg$  denotes  $\log_{10}$  and  $\ln$  denotes  $\log_e$ .

unfortunately difficult and not straightforward. Instead, the magnitude system is calibrated to specific stars. The zero-point is set so that Vega has  $U = B = V = \dots = 0^9$ .

Imagine comparing the flux from a star, at a distance  $d$ , with that from an imaginary identical star located at a distance of 10 pc. We'll call the magnitude of this imaginary star at 10 pc the ABSOLUTE MAGNITUDE  $M$  and define the DISTANCE MODULUS as

$$\begin{aligned} \text{DM} \equiv m - M &= m(d) - m(10 \text{ pc}) \\ &= -2.5 \log \left[ \frac{L/4\pi d^2}{L/4\pi(10 \text{ pc})^2} \right] \\ &= -2.5 \log \left[ \left( \frac{10 \text{ pc}}{d} \right)^2 \right] \\ &= 5 \log \left( \frac{d}{\text{pc}} \right) - 5. \end{aligned} \quad (1.13)$$

Since the absolute magnitude is a measure of the flux from the star *if it were* at a specified distance, the absolute magnitude is a proxy for the luminosity *in a given filter*. Because magnitudes do not have a straightforward conversion to physical units, the absolute magnitude is again a relative scale for comparing stellar luminosities.

WE CAN ALSO COMPARE THE FLUX FROM TWO DIFFERENT FILTERS FOR THE SAME STAR. The difference in magnitudes between two different filters defines a COLOR INDEX, which measures where the star's spectrum peaks and is therefore a rough proxy for surface effective temperature. For example, the index

$$B - V = -2.5 \log \left[ \frac{\int_{B\text{-band}} F_\lambda d\lambda}{\int_{V\text{-band}} F_\lambda d\lambda} \right]$$

measures the ratio of  $V$ -band flux to  $B$ -band flux for a particular star.

---

EXERCISE 1.9— How would the  $B - V$  index of the sun compare to that of a hotter star, e.g., one with  $T_{\text{eff}} = 8000 \text{ K}$ ?

---

Figure 1.6 shows the absolute  $M_V$  magnitudes and the  $B - V$  color indices for 20 853 nearby stars as measured by the *Hipparcos* satellite<sup>10</sup>. *Hipparcos* was launched by ESA in 1989 to measure parallaxes for over  $10^5$  nearby stars. Its successor, *Gaia*, is measuring positions and motions of about  $10^9$  stars in the Milky Way. Fig. 1.6 represents a wide range of stellar masses and evolutionary stages. In the remaining chapters of these notes, we'll explore how the groupings on this diagram come about and trace the life of a star from birth to death.

<sup>9</sup> The modern value is  $V(\text{Vega}) = 0.026 \pm 0.008$ ; Bohlin and Gilliland [2004].

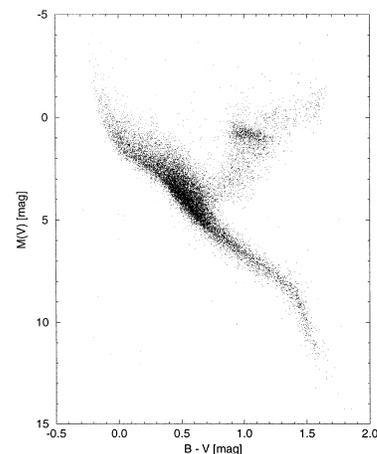


Figure 1.6: Distribution of stars cataloged by *Hipparcos* in  $M_V$  and  $B - V$ .

<sup>10</sup> M. A. C. Perryman et al. The *Hipparcos* Catalogue. *A&A*, 500:501–504, July 1997

---

EXERCISE 1.10 — Let's examine the Hipparcos color-magnitude diagram, Fig. 1.6, and understand the physical properties of the stars plotted on it.

1. On the axis  $B - V$ , indicate the direction "red→blue" in visual appearance; assuming thermal emission, indicate the direction of increasing surface temperature.
  2. On the axis  $M_V$ , indicate the direction of increasing luminosity. What is the dynamic range (ratio, largest to smallest) of stellar luminosities on this plot?
  3. Again assuming thermal emission, indicate the direction on the plot of increasing stellar size (radius).
  4. Look up  $M_V$  and  $B - V$  for the sun, and indicate it on the plot.
-

## 2

# Under Pressure

### 2.1 Hydrostatic equilibrium

Consider a fluid at rest in a gravitational field. By a FLUID, we mean that the pressure is isotropic<sup>1</sup> and directed perpendicular to any given surface. Let's now imagine a small fluid element, as depicted in Fig. 2.1. The gravitational acceleration is in the direction  $-\hat{r}$ ; the fluid element has thickness  $\Delta r$  along the direction of the gravitational force and cross-sectional area  $\Delta A$ .

Since our fluid is at rest, the forces must balance. This implies that the pressure only depends on  $r$ , so that there is no net sideways force on our fluid element. If the fluid has a density (mass per unit volume)  $\rho$ , then the mass of the fluid element is  $\rho\Delta A\Delta r$ , and the gravitational force on the fluid element is  $-(\rho\Delta A\Delta r)g(r)\hat{r}$ . This gravitational force is balanced by the *difference* in pressure  $P(r)$  between the upper and lower faces of the element.

The pressure force on the upper face is  $-\Delta A \times P(r + \Delta r)\hat{r}$ ; on the lower face,  $\Delta A \times P(r)\hat{r}$ . For the element to be in hydrostatic equilibrium the forces along  $\hat{r}$  must balance,

$$\Delta A [-P(r + \Delta r) + P(r) - \Delta r\rho g(r)] = 0.$$

Dividing by  $\Delta r$  and taking the limit  $\Delta r \rightarrow 0$  gives us the equation of hydrostatic equilibrium:

$$\frac{dP}{dr} = -\rho g(r). \quad (2.1)$$

This is a differential equation describing how the pressure varies in the star. We don't have enough information yet to solve it, however, because we haven't specified either the gravity  $g(r)$  or the density  $\rho$ .

#### Constant gravity

Let's start with a simple case: a thin fluid layer, over which gravity is approximately constant. This is a good approximation for Earth's atmosphere or ocean. Exercise 2.1 examines the pressure increase in an

<sup>1</sup> Meaning the pressure is the same in all directions.

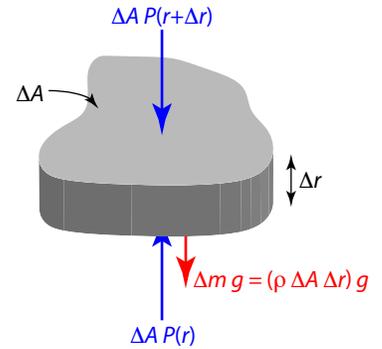


Figure 2.1: A fluid element in hydrostatic equilibrium.

The SI unit of pressure is the **Pascal**:  $1 \text{ Pa} = 1 \text{ N m}^{-2}$ . The mean pressure at terrestrial sea level is  $1 \text{ atm} = 1.013 \times 10^5 \text{ Pa}$ . Other common units of pressure are the **bar** ( $1 \text{ bar} = 10^5 \text{ Pa}$ ) and the **Torr** ( $760 \text{ Torr} = 1 \text{ atm}$ ).

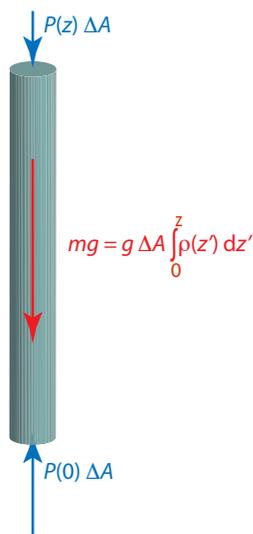


Figure 2.2: The mass of a column of fluid.

<sup>2</sup> By **IDEAL** gas, we mean that the particles are non-interacting; as a result, the energy of the gas only depends on the kinetic energy of the particles and in particular is independent of the volume.

<sup>3</sup> The constant  $N_A$  is known as **AVOGADRO'S NUMBER**.

incompressible (density is fixed) fluid. This is a bad approximation for a star, but a good one for Earth's oceans: the density of sea water increases by less than 5% between the surface and ocean floor.

---

**EXERCISE 2.1** — Seawater has a density  $\rho = 10^3 \text{ kg m}^{-3}$ . Solve eq. (2.1) to get an equation for the pressure as a function of depth in the ocean. How deep would you need to dive for the pressure to increase by  $1 \text{ atm} = 1.013 \times 10^5 \text{ Pa}$ ? Does this agree with your experience?

---

We can write equation (2.1) in the form of an integral:

$$\int_{P_0}^{P(z)} dP = -g \int_0^z \rho dz.$$

Consider a cylinder of cross-section  $\Delta A$  that extends from 0 to  $z$ . The mass of that cylinder is

$$m(z) = \Delta A \times \int_0^z \rho dz.$$

and its weight is  $m(z)g$ .

The difference in pressure between the bottom and top of the cylinder is just

$$P_0 - P(z) = gm(z)/\Delta A,$$

that is, the difference in pressure is the weight per unit area of our column. Let's apply this to our atmosphere: if we take the top of our column to infinity and the pressure at the top to zero, then the pressure at the bottom (sea level) is just the weight of a column of atmosphere with a cross-sectional area of  $1 \text{ m}^2$ .

### The isothermal ideal gas

In general the density  $\rho$  depends on the pressure  $P$  and temperature  $T$  via an **EQUATION OF STATE**. Let's relax our condition of constant density, but keep gravity and temperature constant and assume the fluid is an ideal gas<sup>2</sup>. For  $N$  particles in a volume  $V$  at pressure and temperature  $P$  and  $T$ , the ideal gas equation of state is

$$PV = Nk_B T. \quad (2.2)$$

In chemistry, it is convenient to count the number of particles by **MOLES**. One mole of a gas contains<sup>3</sup>  $N_A = 6.022 \times 10^{23}$  particles, and the number of moles in a sample is  $n = N/N_A$ . If we divide and multiply equation (2.2) by  $N_A$ , then our ideal gas equation becomes

$$PV = n [N_A k_B] T \equiv nRT,$$

where  $R = N_A k_B = 8.314 \text{ J K}^{-1} \text{ mol}^{-1}$  is the gas constant. This is perhaps the most familiar form of the ideal gas law—but it is not in a form useful to astronomers.

We astronomers don't care about little beakers of fluid—we have whole stars to model! Put another way, volume isn't a useful quantity since we are working in the middle of a large mass of fluid. Instead, define the NUMBER DENSITY as the number of particles per volume,  $N/V$ . The mass of each particle is  $\mathcal{A} \times m_u$ , where  $m_u$  has a mass of one ATOMIC MASS UNIT ( $\equiv 1 \text{ u}$ )<sup>4</sup>. Hence the mass per volume of our fluid is

$$\rho = \mathcal{A} m_u \times \frac{N}{V}.$$

We call  $\rho$  the MASS DENSITY, or density for short. This quantity appears in equation (2.1).

Starting with eq. (2.2), dividing by  $V$  and then multiplying and dividing by  $\mathcal{A} m_u$  gives

$$P = \left( \frac{\mathcal{A} m_u N}{V} \right) \frac{k_B}{\mathcal{A} m_u} T \equiv \rho \frac{k_B}{\mathcal{A} m_u} T. \quad (2.3)$$

Equation (2.3) is the form most convenient for fluid dynamics, because it is in terms of an intrinsic fluid property—the density  $\rho$ —rather than in terms of the volume.

<sup>4</sup> By definition, 1 mol of <sup>12</sup>C atoms in their ground state has a mass of 12 g. Thus,  $1 \text{ u} = 1.661 \times 10^{-27} \text{ kg}$  is  $1/(12N_A)$  of the mass of 1 mol of <sup>12</sup>C atoms in their ground state

---

EXERCISE 2.2 — Let's take a stab at modeling Earth's atmosphere with equation (2.1). Take Earth's atmosphere to be dry (no water, so we don't have to worry about condensation) and model it as an ideal gas. Also assume the temperature doesn't change with altitude. The average molecular mass of dry air is  $\mathcal{A} = 28.97$ . Integrate eq. (2.1) from  $z = 0$ , where  $P(z = 0) = P_0$ , to a height  $z$ . Show that the solution is  $P(z) = P_0 e^{-z/H_p}$ , where  $H_p$  is the PRESSURE SCALE HEIGHT—the height over which the pressure decreases by a factor  $1/e$ . Evaluate  $H_p$  for dry air at a temperature of 288 K (15 °C). Is the answer reasonable, based on your experience? Is the assumption of an isothermal atmosphere a good one? Explain why or why not.

---

The mass  $\mathcal{A}$  of an atom or nuclei, when expressed in atomic mass units, is approximately equal to the atomic number  $A$  (Table 2.1). The electron mass is  $m_e = 0.0005485 \text{ u}$ . Unless we need high accuracy, we can neglect the electron mass and take the mass of an atom or nuclide to be  $\mathcal{A} \times m_u$ .

## 2.2 Mass density and the mean molecular weight

For a mixture with different types of particles, it is useful to introduce the MEAN MOLECULAR WEIGHT  $\mu$ . This is computed by taking the total mass of a sample of particles and dividing by the total number of particles, so

We denote an atomic isotope or nuclide as <sup>A</sup>El, where  $A$  is the atomic number (total number of neutrons and protons in the nucleus) and  $El$  is the element abbreviation (corresponds to number of protons in the nucleus).

Table 2.1: Selected atomic masses

nuclide	$A$	$\mathcal{A}$	$( \mathcal{A} - A /A)$ (%)
n	1	1.00865	0.865
<sup>1</sup> H	1	1.00783	0.783
<sup>4</sup> He	4	4.00260	0.065
<sup>12</sup> C	12	12.00000	0.000
<sup>16</sup> O	16	15.99491	0.032
<sup>28</sup> Si	28	27.97693	0.082
<sup>56</sup> Fe	56	55.93494	0.116

that

$$\mu = \frac{\rho}{m_{\text{u}} n_{\text{tot}}} = \frac{1}{m_{\text{u}}} \frac{\sum_i m_i n_i}{\sum_i n_i}. \quad (2.4)$$

Some examples may make this clearer. Suppose we have  $n$  molecules of molecular hydrogen ( $\text{H}_2$ ). The mass of this sample is  $\approx 2 m_{\text{u}} \times n$ , since each molecule has 2 nucleons. The total number of particles in our sample is  $n$ , so

$$\mu(\text{H}_2) = \frac{2m_{\text{u}} \times n}{nm_{\text{u}}} = 2.$$

Now suppose we raise the temperature and dissociate those molecules into individual atoms. The mass of a sample of  $n$  atoms is  $1 \text{ u} \times n$ , so  $\mu = 1$ . Let's raise the temperature further, so that the gas ionizes into electrons and nuclei (protons). This is a bit trickier. The electrons contribute negligibly to the mass, so if we had  $n$  atoms, the mass is still  $1 \text{ u} \times n$ . The total number of particles has *doubled*, however, since for each atom there are now 2 particles (electron and nucleus). The mean molecular weight is therefore

$$\mu(^1\text{H} + e^-) = \frac{nm_{\text{u}}}{2nm_{\text{u}}} = \frac{1}{2}.$$

---

EXERCISE 2.3 — What is  $\mu$  for a fully ionized  $^4\text{He}$  gas ( $A = 4$ , with 2 electrons per atom)?

---

The ideal gas EOS (eq. [2.3]) as used in astronomy is thus

$$P = \rho \left( \frac{k_{\text{B}}}{\mu m_{\text{u}}} \right) T, \quad (2.5)$$

in which the composition enters via the mean molecular weight  $\mu$ .

### 2.3 The mass distribution

Now let's look at how gravity varies within a star. Suppose we are at a distance  $r$  from the stellar center. Newton discovered that the gravitational force inside a spherical shell vanished. This means that the net gravitational force arising from portions of the star exterior to our position vanishes. The gravitational force depends only on the amount of mass interior to our position,

$$m(r) = 4\pi \int_0^r \rho(r) r^2 dr,$$

or in differential form,

$$\frac{dm}{dr} = 4\pi r^2 \rho. \quad (2.6)$$

Furthermore, the gravitational force from a spherically symmetric mass is identical to that of a point particle of the same mass. Thus, the gravitational force at a radius  $r$  from the center is simply

$$g(r) = \frac{Gm(r)}{r^2}.$$

Using this expression for  $g$ , we can write eq. (2.1) as

$$\frac{dP}{dr} = -\rho \frac{Gm(r)}{r^2}. \quad (2.7)$$

EXERCISE 2.4 — What happens to  $m(r)$  and  $g(r)$  at the center ( $r \rightarrow 0$ )? Before doing any calculation, see if you can argue that  $g(r) \rightarrow 0$  at the center. If this is so, then what is  $dP/dr$  at the center? Assuming that  $\rho \rightarrow \rho_c \approx \text{const.}$  near the center, integrate eq. (2.6) over a small radius  $\Delta r$  to get  $m(\Delta r)$  and hence  $g(\Delta r)$ . Show explicitly that  $g(\Delta r) \rightarrow 0$  as  $\Delta r \rightarrow 0$ .

To recap, we now have two differential equations, (2.6) and (2.7), that describe the structure of a star. These equations are for the pressure  $P(r)$  and density  $\rho(r)$  in the star. We can't solve these equations, however, because we don't yet have a relation between  $P$  and  $\rho$ . For example, the ideal gas equation of state relates  $P$  and  $\rho$  via a temperature  $T$ , so at a minimum we need an equation for  $T(r)$ . We'll defer the derivation of additional stellar structures a bit, and explore what we can learn from just these two equations. To proceed, we need to prescribe  $\rho(r)$ ; let's use a constant density and see where that leads.

EXERCISE 2.5 — Let's suppose that  $\rho$  is constant throughout the star. In what follows, you should be able to express everything in terms of the star's mass  $M$  and radius  $R$ , along with physical constants such as  $G$  and  $k_B$ .

1. First, find  $\rho$  in terms of the total mass  $M$  and radius  $R$ .
2. Next, solve equation (2.6) to find  $m(r)$  in terms of  $M$  and  $r/R$ .
3. Use this expression for  $m(r)$  and your expression for  $\rho$  to integrate equation (2.7) and to find the pressure at the center,  $P_c = P(r = 0)$ .
4. Now that we have an expression for the central pressure in terms of  $M$  and  $R$ , let's try to understand what it means. Use your result from parts 1 and 3, as well as equation (2.5) to find the central temperature of the star, in terms of  $G$ ,  $M$ ,  $R$ , and the mean molecular weight of the gas  $\mu$ . Evaluate  $T_c$  for  $M = M_\odot$ ,  $R = R_\odot$ , and  $\mu = 0.6$ . Do you get a reasonable result?

## 2.4 A closer look at hydrostatic equilibrium

What would happen if the star were suddenly no longer in hydrostatic equilibrium? To explore this, imagine that we could turn off the pressure.

Once we did so, there would be no net outward force on a fluid element, just the downward pull of gravity (cf. Fig. 2.1). The gas in the star would therefore free-fall inwards once pressure were turned off.

HOW LONG WOULD THIS STELLAR IMPLOSION TAKE? We can certainly answer this by integrating the equation of motion for a particle free-falling from the surface of the star to the center. There is a simpler method, however, using Kepler's law. Start with a circular orbit and deform it into increasingly eccentric ellipses while keeping the sun at one focus, as shown in Fig. 2.3. The limit of these increasingly eccentric orbits is a fall into the center. The time to fall in is one-half of an orbital period, and the orbital period only depends on the semi-major axis, which in this limiting case is  $a = R/2$ :

$$\tau_{\text{ff}} = \frac{T}{2} = \frac{\pi}{\sqrt{GM}} \left(\frac{R}{2}\right)^{3/2}.$$

Notice that this expression contains the combination  $\sqrt{R^3/M} \propto 1/\sqrt{\bar{\rho}}$ , where  $\bar{\rho}$  is the mean density:

$$\tau_{\text{ff}} = \left(\frac{3}{32\pi}\right)^{1/2} \left(\frac{1}{G} \frac{4\pi R^3}{3M}\right)^{1/2} = \left(\frac{3}{32\pi}\right)^{1/2} \frac{1}{\sqrt{G\bar{\rho}}}. \quad (2.8)$$

The time to collapse is proportional to  $1/\sqrt{G\bar{\rho}}$  and depends on the average density of the star. We thus define the DYNAMICAL TIMESCALE of the star as  $t_{\text{dyn}} \equiv 1/\sqrt{G\bar{\rho}}$ .

Now imagine that after the collapse starts, we change our mind and turn the pressure back on. The infall of the star's outer layers compresses the gas within the star's interior and causes the pressure in the interior to rise. This will eventually restore hydrostatic equilibrium. How quickly can the star respond? A change in pressure is communicated to the rest of the star by sound waves, which travel at a speed (see Box 2.1)

$$c_s = \left(\gamma \frac{P}{\rho}\right)^{1/2} = \left(\gamma \frac{k_B T}{\mu m_u}\right)^{1/2}. \quad (2.9)$$

Here  $\gamma$  is the adiabatic index: for an ideal monatomic gas,  $\gamma = 5/3$ . The star can therefore respond to pressure changes on a SOUND-CROSSING TIME  $\tau_{\text{sc}} = R/\bar{c}_s$ . If we use the central temperature for a constant density star (exercise 2.5) to get an estimate for the mean  $c_s$ ,

$$\tau_{\text{sc}} \equiv \frac{R}{\bar{c}_s} \approx R \left(\frac{\mu m_u}{\gamma k_B T_c}\right)^{1/2} = R \left(\frac{2R}{\gamma GM}\right)^{1/2} = \left(\frac{3}{\sqrt{10\pi}}\right) \frac{1}{\sqrt{G\bar{\rho}}}.$$

Both the sound-crossing time,  $\tau_{\text{sc}}$ , and the free-fall time,  $\tau_{\text{ff}}$ , are of the order of the dynamical timescale  $1/\sqrt{G\bar{\rho}}$ .

We have another way of viewing hydrostatic equilibrium: the star is able to remain in balance because the time for pressure disturbances to

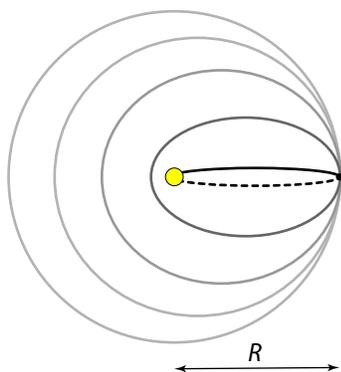


Figure 2.3: Deformation of an orbit until it becomes a fall to the center, denoted by the yellow dot.

propagate is comparable to the time for large-scale motions of the fluid

$$\tau_{\text{sc}} \sim \tau_{\text{ff}} \sim \tau_{\text{dyn}}.$$

For the sun,  $\bar{\rho} = 1400 \text{ kg m}^{-3}$ ; this is just a bit denser than you<sup>5</sup>. The dynamical timescale for the sun is thus about one hour.

<sup>5</sup> The density of an adult human body, which is about 60% water, is close to  $10^3 \text{ kg m}^{-3}$ ; most of us are close to neutral buoyancy in water.

---

EXERCISE 2.6 — The central temperature  $T_c$  is a measure of the average kinetic energy of a particle at the stellar center. Use the central temperature that you found for the constant density star in exercise 2.5 and estimate the time that such a particle would take to cross a distance  $R$ . How does this time compare to the orbital period of a satellite orbiting just outside the stellar surface?

---

### Box 2.1 The sound speed

Suppose we have a long tube filled with gas at pressure  $P(x, t) = P_0$ , density  $\rho(x, t) = \rho_0$ , and velocity  $u(x, t) = U_0 = 0$ . We then tap on one end of the tube; this causes a disturbance to propagate down the tube. Denote the cross-sectional area of the tube by  $A$ , and consider the volume  $A\Delta x$  located between  $x$  and  $x + \Delta x$ ; the mass in this small volume is  $\rho A \Delta x$ .

As a result of the disturbance, the pressure in the tube becomes  $P(x, t) = P_0 + \sigma P_1(x, t)$  and the fluid acquires a velocity  $u(x, t) = \sigma u_1(x, t)$ . This motion compresses or rarifies the gas:  $\rho(x, t) = \rho_0 + \sigma \rho_1(x, t)$ . In these expressions, terms with a subscript “1” are small perturbations compared with the equilibrium quantities; also,  $\sigma$  is a bookkeeping parameter used to keep track of the order of terms in the expansion. After expanding to terms linear in  $\sigma$ , we’ll set  $\sigma = 1$ .

The nonuniform perturbed pressure subjects our small fluid mass  $\rho A \Delta x$  to a net force

$$A [P(x) - P(x + \Delta x)] \approx -\sigma A [P_1(x + \Delta x) - P_1(x)].$$

This force causes the mass to accelerate:

$$\begin{aligned} \rho A \Delta x \frac{\partial u}{\partial t} &= A \Delta x (\rho_0 + \sigma \rho_1) \frac{\partial (0 + \sigma u_1)}{\partial t} \\ &\approx \sigma \left[ A \Delta x \rho_0 \frac{\partial u_1}{\partial t} \right] + \mathcal{O}(\sigma^2); \end{aligned}$$

equating this to the expression for the force, taking the limit  $\Delta x \rightarrow 0$ ,  $\sigma \rightarrow 1$ , and canceling common factors gives

$$\frac{\partial u_1}{\partial t} = -\frac{1}{\rho_0} \frac{\partial P_1}{\partial x}. \quad (2.10)$$

Because of the non-uniform velocity, the volume  $V$  and hence

## Box 2.1 continued

density of our little mass will also change:

$$\frac{\partial V}{\partial t} = A [u(x + \Delta x) - u(x)] = \sigma A \Delta x \left[ \frac{u_1(x + \Delta x) - u_1(x)}{\Delta x} \right]$$

or

$$\frac{1}{V} \frac{\partial V}{\partial t} = \sigma \frac{\partial u_1}{\partial x}. \quad (2.11)$$

This change in volume is related to the change in pressure. We are interested in fluctuations that are sufficiently quick that no heat is transferred into or out of our mass. This is an adiabatic process, for which  $PV^\gamma$  remains constant. Here  $\gamma$  is the **ADIABATIC INDEX**; for an ideal gas,  $\gamma = C_p/C_v$ . (We shall discuss adiabatic processes more thoroughly in chapter 6.)

As the pressure changes adiabatically from  $P_0$  to  $\sigma P_1$ , the volume changes as

$$\frac{dV}{V} = d \ln V = -\frac{1}{\gamma} \frac{dP}{P} = -\frac{1}{\gamma} d \ln P.$$

Hence

$$\frac{\partial \ln V}{\partial t} = -\frac{1}{\gamma} \frac{\partial \ln(P_0 + \sigma P_1)}{\partial t} \approx -\frac{\sigma}{\gamma P_0} \frac{\partial \ln P_1}{\partial t} = \sigma \frac{\partial u_1}{\partial x}. \quad (2.12)$$

The last equality comes from equation (2.11).

We therefore have two equations for the perturbed velocity to order  $\sigma$ :

$$\begin{aligned} \frac{\partial u_1}{\partial t} &= -\frac{1}{\rho_0} \frac{\partial P_1}{\partial x} \\ \frac{\partial u_1}{\partial x} &= -\frac{1}{\gamma P_0} \frac{\partial P_1}{\partial t}; \end{aligned}$$

differentiating the top equation with respect to  $x$  and the bottom with respect to  $t$ , and equating the expressions for  $\partial^2 u_1 / \partial t \partial x$  gives

$$\frac{\partial^2 P_1}{\partial t^2} = \left( \frac{\gamma P_0}{\rho_0} \right) \frac{\partial^2 P_1}{\partial x^2}. \quad (2.13)$$

This is the equation for a wave: the solutions are  $P_1(x, t) = P_1(x \pm c_s t)$ , where the sound speed is  $c_s = \sqrt{\gamma P / \rho}$ .

## 2.5 Virial Equilibrium

With the assumption that  $\rho = \text{constant}$ , we showed (exercise 2.5) that the central temperature and pressure depends on the total mass  $M$ , total

radius  $R$ , and the gravitational constant  $G$  as

$$T_c = \frac{1}{2} \left\{ \frac{GM}{R} \frac{\mu m_u}{k_B} \right\} \quad (2.14)$$

$$P_c = \frac{3}{8\pi} \left\{ \frac{GM^2}{R^4} \right\}. \quad (2.15)$$

Our task now is to show that the scalings of  $T_c$  and  $P_c$  with  $M$  and  $R$ —the quantities in  $\{ \}$ —hold in general for a star in mechanical equilibrium.

To show this, we shall employ a form of the *virial theorem*. Suppose we have a collection of  $N$  particles, all moving about and exerting forces on one another. If we let this system settle down into some kind of bound configuration, then we can add up the kinetic and potential energies of all the particles to get a total kinetic energy  $K$  and a total potential energy  $\Omega$ . The virial theorem asserts that  $K$  is proportional to, and comparable in magnitude to,  $\Omega$ ; indeed if the potential between a pair of particles scales as  $r^{-1}$ ,  $r$  being the distance between the particles, then  $K = -\Omega/2$ , as we'll now show.

Let us take the position and momentum of particle  $i$  to be  $\mathbf{r}_i = (x, y, z)_i$  and  $\mathbf{p}_i = (p_x, p_y, p_z)_i$ . Then the total kinetic energy is

$$\begin{aligned} K &= \frac{1}{2} \sum_{i=1}^N \mathbf{p}_i \cdot \frac{d\mathbf{r}_i}{dt} \\ &= \frac{1}{2} \left[ \frac{d}{dt} \left( \underbrace{\sum_{i=1}^N \mathbf{p}_i \cdot \mathbf{r}_i}_{\equiv \mathcal{G}} \right) - \sum_{i=1}^N \mathbf{r}_i \cdot \underbrace{\frac{d\mathbf{p}_i}{dt}}_{=\mathbf{F}_i} \right]. \end{aligned} \quad (2.16)$$

The quantity  $\mathcal{G} = \sum_i \mathbf{p}_i \cdot \mathbf{r}_i$  is called the “virial” of the system. By expressing the force  $\mathbf{F}_i = d\mathbf{p}_i/dt$  on particle  $i$  as the gradient of a potential  $\Omega$ ,  $\mathbf{F}_i = -\nabla_i \Omega$ , we can rewrite eq. (2.16) as

$$2K = \frac{d\mathcal{G}}{dt} + \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega. \quad (2.17)$$

$$\nabla_i \equiv \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)_i,$$

that is, it is the gradient with respect to the coordinates of particle  $i$ .

So far, we've just shuffled and relabeled terms. The crucial step comes in taking the time-average of the kinetic energy. The time average of a quantity  $f$  is defined as

$$\langle f \rangle \equiv \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^\tau f(t) dt.$$

Applying this to equation (2.17) gives

$$\begin{aligned} 2\langle K \rangle &= \left\langle \frac{d\mathcal{G}}{dt} \right\rangle + \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega \right\rangle \\ &= \lim_{\tau \rightarrow \infty} \left[ \frac{1}{\tau} \int_0^\tau \frac{d\mathcal{G}}{dt} dt \right] + \left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega \right\rangle \end{aligned}$$

$$= \underbrace{\lim_{\tau \rightarrow \infty} \left[ \frac{\mathcal{G}(\tau) - \mathcal{G}(0)}{\tau} \right]}_I + \underbrace{\left\langle \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega \right\rangle}_II$$

Now, if the system is bound and in mechanical equilibrium, then the positions and momenta of all particles are finite: none of the particles can escape, and the system doesn't violently collapse so that momenta are diverging. Hence both  $G(\tau)$  and  $G(0)$  are finite numbers, so as  $\tau \rightarrow \infty$ , term I vanishes.

As for term II, we can show that if the potential between pairs of particles depends on  $1/r$ , where  $r$  is the distance between those particles, then term II is just  $-\Omega$  (see Box 2.2). For now, I'll give a rough argument of why this is so: in a spherically symmetric system, then the potential just depends on the distance  $r$  from the origin; and since

$$r \frac{\partial}{\partial r} \left( \frac{1}{r} \right) = -\frac{1}{r},$$

the last term is just  $-\Omega$  and our equation is

$$2\langle K \rangle + \langle \Omega \rangle = 0. \quad (2.18)$$

This is the virial theorem, applied to a  $r^{-1}$  potential.

### Box 2.2 Working with vectors

In this sidebar we'll show that the second term in equation (2.17) is

$$\sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega = -\Omega. \quad (2.19)$$

First, we need an expression for  $\Omega$ . Suppose we pick a pair of particles,  $j$  and  $k$ . The potential between this pair is

$$-\frac{Gm_j m_k}{r_{jk}} = -\frac{Gm_j m_k}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}}.$$

Our total potential consists of a sum over the potentials between all  $N(N-1)/2$  unique pairs of particles,

$$\Omega = -\frac{Gm_1 m_2}{\sqrt{(\mathbf{r}_1 - \mathbf{r}_2)^2}} - \dots - \frac{Gm_j m_k}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} - \dots$$

When we take the derivative in eq. (2.19), we apply  $\mathbf{r}_i \cdot \nabla_i$  to each term in the potential and sum over all  $i$ . For the term with the pair  $j, k$ , this will give

$$\sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \left( -\frac{Gm_j m_k}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right)$$

## Box 2.2 continued

$$= -Gm_j m_k \left[ \mathbf{r}_j \cdot \nabla_j \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) + \mathbf{r}_k \cdot \nabla_k \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) \right].$$

Since many of you aren't yet comfortable with vector expressions, we'll do this in detail for the  $x$ -component:

$$\begin{aligned} & \left[ \mathbf{r}_j \cdot \nabla_j \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) + \mathbf{r}_k \cdot \nabla_k \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) \right]_x \\ &= x_j \frac{\partial}{\partial x_j} \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) + x_k \frac{\partial}{\partial x_k} \left( \frac{1}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) \\ &= -\frac{x_j(x_j - x_k)}{[(\mathbf{r}_j - \mathbf{r}_k)^2]^{3/2}} + \frac{x_k(x_j - x_k)}{[(\mathbf{r}_j - \mathbf{r}_k)^2]^{3/2}} \\ &= -\frac{(x_j - x_k)^2}{[(\mathbf{r}_j - \mathbf{r}_k)^2]^{3/2}} \end{aligned}$$

The  $y$ - and  $z$ -components are similar, giving

$$\begin{aligned} \sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \left( -\frac{Gm_j m_k}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right) &= Gm_j m_k \frac{(\mathbf{r}_j - \mathbf{r}_k)^2}{[(\mathbf{r}_j - \mathbf{r}_k)^2]^{3/2}} \\ &= -\left( -\frac{Gm_j m_k}{\sqrt{(\mathbf{r}_j - \mathbf{r}_k)^2}} \right). \end{aligned}$$

This can be done for every term in the sum, with the final result that

$$\sum_{i=1}^N \mathbf{r}_i \cdot \nabla_i \Omega = -\Omega.$$

For an ideal monatomic gas in thermal equilibrium, the mean kinetic energy of a particle in the gas is  $K = (3/2)k_B T$ , and we therefore may define an average temperature

$$2K = 3Nk_B \bar{T} = -\Omega. \quad (2.20)$$

The total number of particles is  $N = M/(\mu m_u)$ , and so

$$\bar{T} = -\frac{1}{3} \Omega \frac{\mu m_u}{M k_B}. \quad (2.21)$$

The total potential of the system depends on only three parameters:  $G$ ,  $M$ , and  $R$ . The only way to make a quantity having dimensions of energy is for

$$\Omega \propto -\frac{GM^2}{R},$$

and so

$$\bar{T} \propto \frac{GM}{R} \frac{\mu m_u}{k_B}.$$

The constant of proportionality depends on the distribution of gas in the system. By using the ideal gas law,  $\bar{P} = \bar{\rho}(k_B/\mu m_u)\bar{T}$ , we find the mean pressure

$$\bar{P} \propto \frac{GM^2}{R^4}$$

of the system.

As a concrete example, let's compute  $\Omega$  for a constant density sphere. If we bring a small amount of mass  $dm$  from infinity onto a sphere of mass  $m$  and radius  $r$ , then the change in potential is

$$d\Omega = -\frac{Gm}{r} dm.$$

For a constant density,  $r = R(m/M)^{1/3}$ ; upon substituting for  $r$  we have

$$\Omega_{\text{const. den.}} = -\int_0^M \frac{GM^{1/3}m^{2/3}}{R} dm = -\frac{3}{5} \frac{GM^2}{R}. \quad (2.22)$$

Using this in equation (2.21) gives us the mean temperature, and hence pressure, for a constant density sphere,

$$\bar{T} = \frac{1}{5} \frac{GM}{R} \frac{\mu m_u}{k_B}, \quad (2.23)$$

$$\bar{P} = \frac{3}{20\pi} \frac{GM^2}{R^4}. \quad (2.24)$$

These are comparable to the central values, eqn. (2.14) and (2.15).

---

EXERCISE 2.7— We can infer a great deal from our simple virial scalings. Table 2.2 provides masses, radii, and luminosities, in units of  $M_\odot$ ,  $R_\odot$ , and  $L_\odot$ , for stars from type B (hot blue stars) to type M (cool red stars). Using the constant density model, compute  $\rho/\rho_\odot$ ,  $T_c/T_{c,\odot}$ , and  $P_c/P_{c,\odot}$ . You should find that each quantity depends only on  $m = M/M_\odot$  and  $r = R/R_\odot$ . Describe your findings: do  $P_c/P_{c,\odot}$ ,  $\rho/\rho_\odot$ , and  $T_c/T_{c,\odot}$  vary in a similar fashion? If not, how do they change with stellar type?

---

Table 2.2: Masses, radii, and luminosities for selected stellar types. The type—B2, B8, F0, and so forth—indicates what features are present in the star's spectrum and indicates the star's surface effective temperature.

	B2	B8	F0	F5	G5	M0	M7
$M/M_\odot$	9.8	3.8	1.6	1.3	0.92	0.51	0.12
$R/R_\odot$	5.6	3.0	1.5	1.3	0.92	0.60	0.18
$L/L_\odot$	5800.0	180.0	6.5	3.2	0.79	0.08	0.003

---

EXERCISE 2.8— Using the constant density model, derive an expression for the total energy (kinetic plus potential) as a function of central temperature for a given mass. Plot this relation. What happens to the central temperature if additional heat is injected into the star?

---

EXERCISE 2.9— Let's examine the behavior of a thin layer at the surface of our constant density model. The pressure on the upper surface of our layer vanishes, and the pressure on the bottom surface of our layer is  $P(R)$ .

1. As a mathematical preliminary, suppose we have a function  $f(x) = Ax^\alpha$  and that we expand about a point  $x_0$  with  $f_0 = Ax_0^\alpha$ . Show that to lowest order in  $\delta x$ ,

$$f(x_0 + \delta x) \approx f_0 \left( 1 + \alpha \frac{\delta x}{x} \right).$$

2. Write  $dP/dr$  as  $\Delta P/\Delta r$ , where  $\Delta r$  is the thickness of the layer, and then integrate the equation of hydrostatic balance (2.7) over the surface to show that

$$4\pi R^2 P(R) - \frac{GMm}{R^2} = 0, \quad (2.25)$$

where  $m = 4\pi R^2 \rho \Delta r$  is the mass of the layer. For the rest of this exercise, we shall take  $m$  as fixed.

3. Now suppose of star expands by a small amount  $\delta R$ . Use the result of part 1 to find the new density  $\rho'$  in terms of the original density  $\rho$  and  $\delta R$ , to lowest order in  $\delta R/R$ .
4. If the expansion is adiabatic, then the new pressure obeys a relation  $P'V'^\gamma = PV^\gamma$  (Box 2.1). For our shell of mass  $m$ , show that this implies that  $P'\rho'^{-\gamma} = P\rho^{-\gamma}$ , and thus find  $P'$  in terms of  $P$ ,  $\delta R/R$ , and  $\gamma$ , to lowest order in  $\delta R/R$ .
5. Insert the expressions for  $P'$  and  $R' = R + \delta R$  into equation (2.25) and cancel any common factors; you should find that the pressure and gravitational forces no longer balance. Express the residual force in terms of  $GMm/R^2$ ,  $\gamma$  and  $\delta R/R$ .
6. Equate this residual force with the acceleration of the shell,  $m\ddot{\delta R}$ , and show that the shell oscillates. For  $\gamma = 5/3$  (appropriate for an ideal gas), find the period of oscillation in terms of  $\rho = 3M/4\pi R^3$ .

## 2.6 The Kelvin-Helmholtz timescale

Stars are born when a cold, dense<sup>6</sup> cloud of gas and dust becomes unstable to gravitational collapse. The details of this process is a topic of current research; for our purposes, however, after a period of time a PRE-MAIN SEQUENCE star forms. At this point, the star is in hydrostatic balance, but with a radius much larger than its main-sequence value and a cool temperature. We observe such pre-main sequence stars in star-forming regions, such as in the Rho Ophiuchi complex (Fig. 2.4). The average age of the newly formed stars in this complex is estimated to be  $\sim 30\,000$  yr.

What happens to this object? The pre-main sequence star is in hydrostatic balance, so it doesn't collapse. But the interior, and hence the surface, is warm, so it radiates energy. The only source of energy is gravitational, so the pre-main sequence star *must* contract. As we shall show,

<sup>6</sup> Dense is a relative term; here we mean  $\sim 100$  atoms per cubic centimeter

We'll discuss the pre-main sequence phase more thoroughly in § 6.2.



Figure 2.4: Star forming regions in the Rho Ophiuchi cloud. *Image credit:* NASA, JPL-Caltech, WISE.

this contraction is quite slow; so slow, in fact, that the pre-main sequence star is always in hydrostatic equilibrium to an excellent approximation. How long does the contraction take? To estimate this, we compute the total energy of the star and divide by its luminosity. For our sun, the total energy is

$$E_{\odot} = K + \Omega = \Omega/2 \approx -\frac{GM_{\odot}^2}{R_{\odot}};$$

the time to radiate this energy away is

$$t_{\text{KH}} = \frac{|E_{\odot}|}{L_{\odot}} \approx \frac{GM_{\odot}^2}{R_{\odot}L_{\odot}} \approx 3 \times 10^7 \text{ yr.} \quad (2.26)$$

This timescale is called the **KELVIN-HELMHOLTZ TIMESCALE**. Since  $t_{\text{KH}} \gg t_{\text{dyn}} = (G\bar{\rho})^{-1/2}$  the star is indeed in hydrostatic equilibrium, to an excellent approximation, throughout the whole contraction.

---

**EXERCISE 2.10**— Using the constant density model (constant here means “constant throughout the star at any given time”) of exercise 2.5 and the virial relations, give a qualitative sketch for how the pressure, density, temperature, radius, and total energy change with time as the protostar contracts.

---

Following the development of thermodynamics in the mid- to latter-half of the nineteenth century, gravitational contraction was proposed to explain the source of the sun’s luminosity. It was quickly realized, however, that the short Kelvin-Helmholtz timescale was in conflict with geological estimates of the age of the earth. This tension was finally resolved with the discovery of nuclear reactions in the 1930s.

# 3

## Edge of Darkness

We saw in chapter 2 that the equilibrium central temperature of a self-gravitating object—such as a star—with an ideal gas EOS depends *solely* on the mass, radius, and composition of that star. For the sun, this temperature is  $\approx 15$  MK and is much higher than the surface effective temperature  $T_{\text{eff},\odot} = 5780$  K. We don't see X-rays coming from the interior of the sun; the photons emitted from the sun are all coming just from the cooler surface layers.

PHOTONS IN A PLASMA, SUCH AS IN THE INTERIOR OF THE SUN, TRANSPORT ENERGY. Were the sun transparent, these photons would immediately stream out, and the sun would release its stored energy in a fiery blast. This doesn't happen: a photon can only travel a short distance before being scattered or absorbed. The net effect is that photons generated in the core must travel a tortuous path, rather like a pinball, before reaching the surface and escaping.

### 3.1 Interaction of radiation and matter

How far does a photon—or any particle, for that matter—travel, on average, in the interior of the sun? Imagine a particle traveling with speed  $v$ . Draw a cylinder, of length  $\ell$  and cross-sectional area  $\mathcal{A}$ , around its path, as shown in Fig. 3.1. What the particle “sees” is that the cylinder is partly blocked by obstacles—other particles in its path.

What is the probability of our particle making it through the cylinder unscathed? The probability of the particle hitting an obstacle is the ratio

$$\mathcal{P} = \frac{\text{total area covered by obstacles}}{\text{area of cylinder}}$$

Denote the cross-sectional area of each particle by  $\sigma$ . If the density of particles is  $n$ , then the number of obstacles in the cylinder is  $n \times (\mathcal{A}\ell)$ , and therefore the fraction of the area blocked by the obstacles is

$$\mathcal{P} = \frac{n \times (\mathcal{A}\ell) \times \sigma}{\mathcal{A}} = n\sigma\ell. \quad (3.1)$$

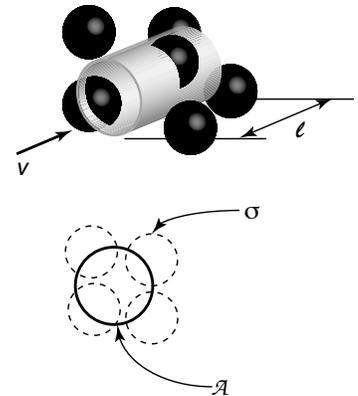


Figure 3.1: Schematic of a particle incident on a group of scattering or absorbing particles.

We are taking  $\ell$  and  $\mathcal{A}$  sufficiently small that we don't have to worry about particles overlapping.

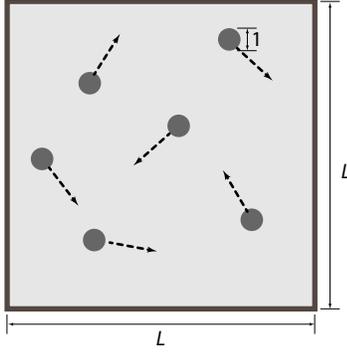


Figure 3.2: Schematic for Exercise 3.1

The particle will suffer a collision when  $\mathcal{P} \rightarrow 1$ , or when

$$\ell = \frac{1}{n\sigma}. \quad (3.2)$$

We call  $\ell$  the “mean free path”: it is the mean distance the particle travels freely before colliding.

---

EXERCISE 3.1— Suppose we have a flat, slippery surface on which hockey pucks are sliding around, as shown in Fig. 3.2. The pucks bounce off the walls as they slide around. Suppose there are  $N$  pucks, each with unit diameter, and the table is square with sides of length  $L$ . Estimate the mean free path of a puck.

---

Although we have motivated this derivation with a classical picture, since the cross-section  $\sigma$  is just related to the probability of an interaction we can define it for quantum mechanical systems as well.

---

EXERCISE 3.2— In the sun, free electrons scatter photons; the cross-section for this is

$$\sigma_{\text{Th}} = \left(\frac{8\pi}{3}\right) \left(\frac{e^2}{4\pi\epsilon_0 m_e c^2}\right)^2 = 6.65 \times 10^{-29} \text{ m}^2.$$

What is the mean free path against this process for a photon at the average density of the solar interior?

---

As the ray of light traverses a small distance  $\Delta s$  through some matter, the probability of a photon being absorbed is  $\mathcal{P} = n\sigma\Delta s$ . Thus, out of every  $N$  photons,  $\Delta N = N \times \mathcal{P} = N \times n\sigma\Delta s$  are absorbed. Since the intensity  $I_\nu$  is proportional to the number of photons, the change in intensity across  $\Delta s$  is just

$$\Delta I_\nu = -n\sigma I_\nu \Delta s.$$

Dividing by  $\Delta s$  and taking the limit  $\Delta s \rightarrow 0$ , we obtain an equation for the absorption of light,

$$\left. \frac{dI_\nu}{ds} \right|_{\text{absorption}} = -n\sigma I_\nu. \quad (3.3)$$

Rather than work with the microscopic cross-section, it is convenient to define the ABSORPTION OPACITY,

$$\kappa_\nu^{\text{abs}} = \frac{n\sigma}{\rho},$$

The units of opacity are  $\text{m}^2/\text{kg}$ .

so that  $dI_\nu/ds = -\rho\kappa_\nu^{\text{abs}}I_\nu$ . We use a subscript  $\nu$  to indicate that the opacity is a function of frequency. In terms of the opacity, the photon mean free path is  $\ell = (\rho\kappa_\nu^{\text{abs}})^{-1}$ .

---

EXERCISE 3.3— A ray of light crosses a slab of absorbent material. Calculate the intensity  $I_\nu$  as a function of distance traveled. Your expression should be in terms of  $\rho$  and  $\kappa_\nu^{\text{abs}}$ . How far does the ray go before its intensity has dropped to  $1/e$  of its original value?

---

IN ADDITION TO ABSORBING PHOTONS, MATTER CAN ALSO SPONTANEOUSLY EMIT THEM. Denote the power emitted per wavelength per volume per steradian by  $\rho j_\nu$ . After traveling a distance  $\Delta s$  through matter with this EMISSIVITY, the ray will increase in intensity by  $\rho j_\nu \Delta s$ ; dividing by  $\Delta s$  and taking the limit  $\Delta s \rightarrow 0$ ,

The units of  $j_\nu$  are  $\text{W kg}^{-1} \text{Hz}^{-1}$ .

$$\left. \frac{dI_\nu}{ds} \right|_{\text{emission}} = \rho j_\nu. \quad (3.4)$$

---

EXERCISE 3.4— Suppose a ray traverses matter that both absorbs (opacity  $\kappa_\nu^{\text{abs}}$ ) and emits (emissivity  $j_\nu$ ), so that

$$\frac{dI_\nu}{ds} = \rho j_\nu - \rho \kappa_\nu^{\text{abs}} I_\nu.$$

Solve for  $I_\nu(s)$  assuming  $\rho$ ,  $j_\nu$ , and  $\kappa_\nu$  are constant, and show that  $I_\nu \rightarrow j_\nu / \kappa_\nu^{\text{abs}}$  as  $s \rightarrow \infty$ .

---

FINALLY, MATTER CAN ALSO SCATTER LIGHT. This removes photons from a ray, similar to absorption, but the photons are redirected into a ray propagating in a different direction. If we assume that the direction into which the photon is scattered is random and isotropic (as is most often the case), then if the intensity in our ray  $I_\nu$  is greater than the angle-average  $J_\nu = (4\pi)^{-1} \int I_\nu d\Omega$ , scattering will cause a net reduction in intensity as more photons are scattered out of the ray than are scattered into it. Conversely, if  $I_\nu < J_\nu$ , then more photons will be scattered into the ray than out of it. Thus, the effect of scattering can be described via

$$\left. \frac{dI_\nu}{ds} \right|_{\text{scattering}} = -\rho \kappa_\nu^{\text{sca}} (I_\nu - J_\nu). \quad (3.5)$$

The effect of scattering is to drive the intensity towards its angle-averaged value.

### 3.2 The equation of radiative transfer

Combining our expressions for absorption, emission, and scattering gives the full expression for how the intensity changes along a ray,

$$\frac{dI_\nu}{ds} = -\rho (\kappa_\nu^{\text{abs}} + \kappa_\nu^{\text{sca}}) I_\nu + \rho j_\nu + \rho \kappa_\nu^{\text{sca}} J_\nu. \quad (3.6)$$

This is a complicated INTEGRODIFFERENTIAL equation: it contains both the derivative  $dI_\nu/ds$  of the intensity as well as its integral  $J_\nu$ .

In general, eq. (3.6) must be solved numerically; but conditions in the deep interior of the star and near the surface allow us to make simplifying approximations and to obtain a solution that gives some insight into the physics. Before doing that, let's clean up eq. (3.6): define a new quantity, the OPTICAL DEPTH  $\tau_\nu$ , via

$$\frac{d\tau_\nu}{ds} = \rho\kappa_\nu \equiv \rho(\kappa_\nu^{\text{abs}} + \kappa_\nu^{\text{sca}}).$$

Next, divide through by  $\rho\kappa_\nu = \rho(\kappa_\nu^{\text{abs}} + \kappa_\nu^{\text{sca}})$ ,

$$\frac{1}{\rho\kappa_\nu} \frac{dI_\nu}{ds} = \frac{1}{d\tau_\nu/ds} \frac{dI_\nu}{ds} - I_\nu + \left[ \frac{j_\nu + \kappa_\nu^{\text{sca}} J_\nu}{\kappa_\nu} \right],$$

and change variables,  $dI_\nu/ds = (dI_\nu/d\tau_\nu) \cdot (d\tau_\nu/ds)$  so the left-hand side is just  $dI_\nu/d\tau_\nu$ . Finally, define the SOURCE FUNCTION

$$S_\nu \equiv \left[ \frac{j_\nu + \kappa_\nu^{\text{sca}} J_\nu}{\kappa_\nu} \right].$$

Doing all that gives us the deceptively simple-looking equation,

$$\frac{dI_\nu}{d\tau_\nu} = -I_\nu + S_\nu. \quad (3.7)$$

This prettifying doesn't advance us any closer to the solution, of course, but notice! The optical depth has a simple meaning:

$$\tau_\nu = \int_0^s \rho\kappa_\nu ds = \int_0^s n\sigma_\nu ds = \int_0^s \frac{ds}{\ell}.$$

That is, the optical depth measures distance along the ray in units of mean free path. Said differently, to travel one optical depth is to travel one mean free path.

EXERCISE 3.5 — For the electron scattering cross-section (Exercise 3.2), estimate the optical depth between the solar center and the solar photosphere.

The other advantage of organizing eq. (3.6) is that this can help develop our intuition about how the solutions should behave, and this will guide our analysis in § 3.3. Although  $S_\nu$  depends on the integral of  $I_\nu$ , we can get a feel for the general behavior by considering a simple model where  $S_\nu$  is a known function of  $\tau_\nu$ .

EXERCISE 3.6 — Suppose that  $S_\nu$  were constant, and solve eq. (3.7) for  $I_\nu(\tau_\nu)$  with boundary condition  $I_\nu(\tau_\nu = 0) = I_{\nu,0}$ . How does  $I_\nu$  behave in the limiting cases  $\tau_\nu \ll 1$  and  $\tau_\nu \gg 1$ ?

The solution to the simple case of exercise 3.6 should make sense. If we have a ray of intensity  $I_{\nu,0}$  incident on an object with  $\tau_{\nu} \ll 1$ , then photons will hardly be absorbed or scattered (cf. exercise 3.3). If we are deep inside an object with  $\tau_{\nu} \gg 1$ , then we shouldn't expect any incident light to matter, so  $I_{\nu}$  shouldn't depend on  $I_{\nu,0}$ ; further if  $\tau_{\nu} \gg 1$ , then going one mean free path shouldn't affect our solution, meaning  $dI_{\nu}/d\tau_{\nu} \rightarrow 0$ , or  $I_{\nu} \rightarrow S_{\nu}$  (cf. exercise 3.4).

SUPPOSE WE ARE IN A CAVITY IN WHICH THE RADIATION AND MATTER ARE IN A STEADY-STATE. That is, the matter is neither gaining nor losing energy to the radiation. Maintaining a steady-state requires balancing the

$$\text{energy emitted per unit volume} = \rho \int j_{\nu} d\nu d\Omega$$

with the

$$\text{energy absorbed per unit volume} = \rho \int \kappa_{\nu}^{\text{abs}} I_{\nu} d\nu d\Omega,$$

so that

$$\int_0^{\infty} (j_{\nu} - \kappa_{\nu}^{\text{abs}} J_{\nu}) d\nu = 0. \quad (3.8)$$

We don't include scattering in this expression because scattering doesn't transfer energy between the radiation and the gas.

If in addition to being in steady-state, the matter and radiation are also in thermal equilibrium,<sup>1</sup> so that  $J_{\nu} = B_{\nu}$ , then eq. (3.8) implies that

$$\frac{j_{\nu}}{\kappa_{\nu}^{\text{abs}}} = B_{\nu}(T). \quad (3.9)$$

Now  $j_{\nu}$  and  $\kappa_{\nu}^{\text{abs}}$  are properties of the matter, and do not depend on the state of the radiation field. Hence, equation (3.9), known as DETAILED BALANCE, must hold whenever the matter is in equilibrium, *regardless of the state of the ambient radiation.*

### 3.3 Radiative diffusion

We can now examine how heat transport works in the deep interior of a star. First, we need to establish our coordinate system. In equation (3.6), the coordinate  $s$  is distance along a ray; but it is more convenient to use coordinates that are tied to the star. We therefore use radial distance  $r$  as a coordinate, and measure the optical depth along it:  $d\tau_{\nu} = \rho\kappa_{\nu} dr$ . Since  $dr = \mu ds$ , where  $\mu = \cos \theta$  is the cosine between  $ds$  and  $dr$  (Fig. 3.3), the equation of transfer becomes

$$\mu \frac{dI_{\nu}}{dr} = -\rho\kappa_{\nu} (I_{\nu} - S_{\nu}). \quad (3.10)$$

<sup>1</sup> meaning they are at the same temperature with each having a thermal distribution of states

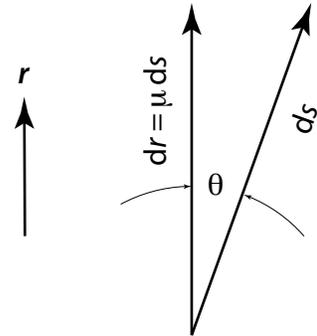


Figure 3.3: Schematic of the coordinate system used for solving the radiative transport equation.

LET'S EXAMINE THE TYPICAL SCALES OF TERMS IN THE RADIATIVE TRANSFER EQUATION, FOR CONDITIONS IN THE DEEP SOLAR INTERIOR. We'll start by indicating some expected scales for eq. (3.6):

$$\underbrace{\mu \frac{dI_\nu}{dr}}_{\sim I_\nu/R_\odot} = - \underbrace{\rho \kappa_\nu I_\nu}_{\sim I_\nu/\ell} + \rho \kappa_\nu S_\nu.$$

If we are far from the surface of the star, then we should expect the intensity to change over lengthscales comparable to  $R_\odot$ . Of course, it won't be exactly this, but—as we'll show—the exact value doesn't matter so long as  $|dI_\nu/dr|$  is in the ballpark of  $I_\nu/R_\odot$ . Notice the enormous disparity in scales:

$$\frac{|dI_\nu/dr|}{\rho \kappa_\nu I_\nu} \sim \frac{\ell}{R_\odot}.$$

The left-hand side of eq. (3.6) is smaller than the terms on the right by the ratio of the mean free path to the solar radius. This implies that conditions in the deep solar interior are nearly homogeneous. They are also isotropic, so that  $I_\nu = J_\nu$ . We expect that collisions are fast enough so that the matter is in thermal equilibrium and  $j_\nu = \kappa_\nu^{\text{abs}} B_\nu$ . We also know from exercise 3.4 that  $I_\nu \rightarrow j_\nu/\kappa_\nu^{\text{abs}} = B_\nu$ . From eq. (3.7) it follows that  $S_\nu = B_\nu$  as well.

We can't have  $I_\nu = B_\nu$  exactly, however, since in that case there is no net flux! We'll therefore treat the intensity as being thermal plus a perturbation:

$$I_\nu = B_\nu + I_\nu^{(1)},$$

where the superscript "(1)" indicates that this is a small correction. Inserting this expansion into eq. (3.6) and keeping only the lowest-order terms on each side gives

$$I_\nu^{(1)} = -\mu \frac{dB_\nu}{d\tau_\nu}. \quad (3.11)$$

$B_\nu$  is a function of the temperature  $T$ , so  $dB_\nu/d\tau_\nu = dB_\nu/dT \cdot dT/d\tau_\nu$ . To get the flux, multiply eq. (3.11) by  $\mu$  and integrate over angles:

$$F_\nu = \int \mu I_\nu^{(1)} d\Omega = - \int \mu^2 \frac{dB_\nu}{dT} \frac{dT}{d\tau_\nu} d\Omega = - \frac{4\pi}{3} \frac{dB_\nu}{dT} \frac{dT}{d\tau_\nu}.$$

Switching variables from  $\tau_\nu$  back to  $r$  gives

$$F_\nu = - \frac{4\pi}{3} \left[ \frac{1}{\rho \kappa_\nu} \frac{\partial B_\nu}{\partial T} \right] \frac{dT}{dr}. \quad (3.12)$$

The flux  $F_\nu$  is therefore proportional to the temperature gradient  $dT/dr$ , and the term in  $[\cdot]$  controls which frequencies have the largest flux and are therefore most responsible for energy transport.

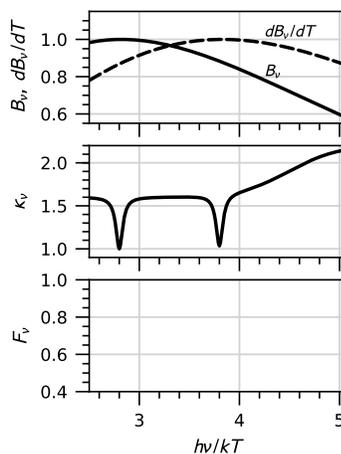


Figure 3.4: The specific flux for a hypothetical opacity

---

EXERCISE 3.7 — Let's examine the term  $[\cdot]$  in eq. (3.12) more closely.

Fig. 3.4 shows  $B_\nu$  and  $dB_\nu/dT$  (top panel) and a hypothetical  $\kappa_\nu$  (middle). Sketch  $F_\nu$  on the bottom panel. For which frequencies is it maximum?

---

To get the total flux, we integrate  $F_\nu$  over all frequencies.

$$\begin{aligned} F &= \int_0^\infty F_\nu \, d\nu = -\frac{4\pi}{3} \left[ \int_0^\infty \frac{1}{\rho\kappa_\nu} \frac{\partial B_\nu}{\partial T} \, d\nu \right] \frac{dT}{dr} \\ &\equiv -\frac{4\pi}{3} \frac{1}{\rho\kappa_R} \frac{\partial}{\partial T} \left[ \int_0^\infty B_\nu \, d\nu \right] \frac{dT}{dr}. \end{aligned}$$

Here we've defined the ROSSELLAND MEAN of the opacity:

$$\frac{1}{\kappa_R} = \left( \int_0^\infty \frac{\partial B_\nu}{\partial T} \, d\nu \right)^{-1} \int_0^\infty \frac{1}{\kappa_\nu} \frac{\partial B_\nu}{\partial T} \, d\nu. \quad (3.13)$$

This is a weighted average of  $\kappa_\nu^{-1}$  with weight  $\partial B_\nu/\partial T$ . Since (eq. [1.6])  $\int B_\nu \, d\nu = \sigma_{\text{SB}} T^4/\pi = caT^4/4\pi$ , we can write the equation for the flux as

$$F = -\frac{1}{3} \frac{c}{\rho\kappa_R} \frac{d}{dr} aT^4. \quad (3.14)$$

Equation (3.14) is known as the equation for radiative diffusion, for reasons that will become apparent in the next section.

If we multiply the flux by the surface area of a shell in the star we obtain the luminosity  $L = 4\pi r^2 F$ ; we can therefore recast eq. (3.14) into an equation for the thermal gradient:

$$\frac{dT}{dr} = -\frac{3\rho\kappa_R}{4acT^3} \frac{L(r)}{4\pi r^2}. \quad (3.15)$$

EXERCISE 3.8— Let's dissect eq. (3.14) to see how it sets the luminosity.

1. To keep the algebra simple, assume that  $F$  is constant throughout the star and that  $aT^4$  is linear in  $r$ —that is,  $aT^4 = aT_c^4(1 - r/R)$ . Since  $F$  is constant, you can express it in terms of the luminosity at the surface  $L$ . Use this to transform eq. (3.14) into an expression for  $L$  in terms of  $R$  and  $T_c$  (along with  $\rho$ ,  $\kappa_R$ , and  $c$ ).
2. Write the luminosity as  $L = E_\gamma/\tau$ , where  $E_\gamma$  is the total radiative energy of the star, and  $\tau$  some as-yet-undetermined *diffusion timescale*. Give an estimate of  $E_\gamma$  in terms of the mean temperature  $T$  and the radius  $R$  of the star.
3. Finally, assume that the photon mean free path  $\ell = (\rho\kappa_R)^{-1}$  is constant. Substitute the results from parts 1 and 2 into equation (3.14). After simplifying, you should end up with a simple expression for  $\tau$  in terms of  $c$ ,  $R$ , and  $\ell$ . For Thomson scattering, what is  $\tau$  (express in years)?

### 3.4 Diffusion

In the presence of scattering or absorption, photons take short hops averaging one mean free path  $\ell$  in length. Imagine a small cube with sides of length  $\ell$  and filled with photons. The total radiant energy in the cube is  $\Delta E$ . In a time  $\Delta t = \ell/c$ , all of the photons will leave this cube.

The total luminosity of the cube is thus  $\Delta E/\Delta t = c\Delta E/\ell$ . If everything is isotropic, then the flux out of any one face is 1/6 of the luminosity, divided by the area of that face:

$$F = \frac{1}{6\ell^2} \frac{c\Delta E}{\ell} = \frac{1}{6}cU,$$

where  $U = E/\ell^3$  is the radiative energy density.

Now place two of these cubes against one another, with their common face located at position  $x$ . The energy density of the two cubes need not be the same; the energy density of the left cube is  $U(x-\ell)$  and of the right cube is  $U(x+\ell)$  (see Fig. 3.5). The net flux traveling in the  $x$ -direction through the common face is then

$$F = \frac{1}{6}cU(x-\ell) - \frac{1}{6}cU(x+\ell) \approx -\frac{1}{3}c\ell \frac{dU}{dx}.$$

This is an expression for a **DIFFUSIVE FLUX**. Although we gave a heuristic explanation, the formula is in general true:

$$\begin{aligned} (\text{flux of something}) &= -\frac{1}{3} \times (\text{speed of carriers}) \times (\text{MFP of carriers}) \\ &\quad \times \nabla(\text{density of something}) \end{aligned} \quad (3.16)$$

For radiation, the “something” is “radiative energy” and the carriers are photons.

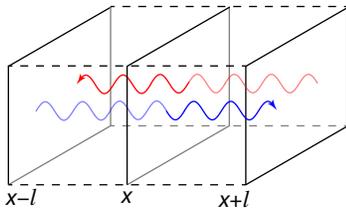


Figure 3.5: Illustration of net flux crossing a face between regions with slightly different energy densities.

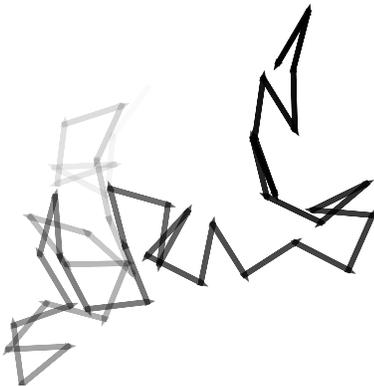


Figure 3.6: Schematic of a random walk of 50 steps.

To keep things simple, we’ll imagine that after absorption the atom immediately emits an identical photon in a random direction.

---

EXERCISE 3.9— Compare this diffusion equation,

$$F = -\frac{1}{3}c\ell \frac{dU}{dr},$$

with eq. (3.14) and use eq. (3.13) to write an expression for the *average* mean free path of a photon (the “mean mean free path”?). Can you give a physical interpretation for the weighting function used in computing the average (cf. exercise 3.7)?

---

FOR AN ALTERNATE VIEW ON PHOTON DIFFUSION, imagine a photon random-walking throughout the stellar interior. The photon moves at speed  $c$ , but it can only go one mean free path  $\ell$  before being absorbed or scattered, at which point it is sent off in a random direction. The path of the photon will therefore look something like that in Fig. 3.6.

We will just do our calculation for motion along a diameter, with the photon starting at the center. On each hop, the photon either goes left or right with equal probability. On average, the photon doesn’t go anywhere; but after enough hops, there is some probability for the photon to reach the edge of the star and escape. Figure 3.7 shows the distribution of positions for walks of length  $n = 10, 30, 100, 300$  steps, with each step having length 1.0. Suppose the edge of the star is at  $x = \pm 10$  (red dotted

lines). Although the average position is at  $x = 0$ , for  $n \gtrsim 100$  steps, there is a reasonable probability of the photon eventually escaping.

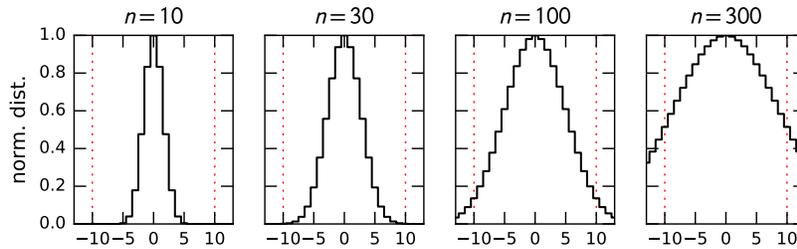


Figure 3.7: Distribution of positions after  $n$  steps in a random walk.

Recall that a random walk is described by a binomial distribution: after  $n$  steps, the probability that  $m$  of them were to the right is

$$\mathcal{P}_n(m; p) = \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m}. \quad (3.17)$$

Here  $p$  is the probability of any single step being to the right. The mean and root variance of  $m$  are

$$\langle m \rangle = np \quad (3.18)$$

$$\left[ \langle (m - \langle m \rangle)^2 \rangle \right]^{1/2} = [np(1-p)]^{1/2}. \quad (3.19)$$

In exercise 3.10, you will use these quantities to estimate the diffusion timescale  $\tau$ .

#### EXERCISE 3.10—

1. Show from equation (3.18) that the mean distance traveled by the photon after  $n$  steps is  $\langle d \rangle = \ell(2np - n)$ , so for  $p = 1/2$ ,  $\langle d \rangle = 0$ .
2. If all the steps were in the same direction, how many steps would be needed to reach the edge, at a distance  $R$  from the center? Assume all steps have the same length  $\ell$ .
3. We want the distribution of steps (Fig. 3.7) to be wide enough to reach the edge. Set the root variance—a measure of the width of the probability distribution—equal to the number of steps found in part 2 and use equation (3.19),

$$\left[ \langle (m - \langle m \rangle)^2 \rangle \right]^{1/2} = [n_{\text{edge}} p(1-p)]^{1/2},$$

to find  $n_{\text{edge}}$  in terms of  $R$  and  $\ell$ .

4. What is the *total* distance traveled by the photon after  $n_{\text{edge}}$  steps? If the photon traveled at speed  $c$ , how long did it take? Compare your answer with that for part 3 of exercise 3.8.

### 3.5 The photosphere

We are now ready to investigate heat transport near the star's edge, where the optical depth  $\tau_\nu \lesssim 1$  and photons begin to freely escape. Near the edge, we cannot use the approximation of radiative diffusion, because conditions are changing over distances of order one mean free path. We therefore return to equation (3.6) for radiative transport:

$$\frac{dI_\nu}{ds} = -\rho (\kappa_\nu^{\text{abs}} + \kappa_\nu^{\text{sca}}) I_\nu + \rho j_\nu + \rho \kappa_\nu^{\text{sca}} J_\nu.$$

This is difficult to solve: for some frequencies, the atmosphere is nearly transparent, while for other frequencies it is quite opaque. Rather than develop the numerical machinery to solve this equation, we shall make a few simplifying assumptions (indicated by **highlighted bold text** in the margins) to obtain an approximate solution for the temperature of the stellar atmosphere.

Opacities are gray

First, we assume that the opacity is gray—that is, independent of frequency. Although unphysical, the solutions for temperature and pressure near the stellar photosphere will still have the correct qualitative behavior. Because the opacity is gray, we shall drop the “ $\nu$ ” subscript in  $\kappa$  and  $\tau$ .

---

EXERCISE 3.11 — Does matter with a gray opacity in thermal equilibrium also have a gray emissivity  $j_\nu = j$ ?

---

We next define a coordinate system. Since we are in a thin layer near the edge of the star, we will adopt planar coordinates, with  $z$  being the altitude above some point. We'll pick  $z = 0$  to be a point deep enough in the star that  $I_\nu \approx B_\nu$ . Then we define the optical depth as

$$\tau(z) = \int_z^\infty \rho (\kappa^{\text{abs}} + \kappa^{\text{sca}}) dz; \quad (3.20)$$

differentiating this expression gives

$$\frac{d\tau}{dz} = -\rho (\kappa^{\text{abs}} + \kappa^{\text{sca}}) \equiv -\rho\kappa.$$

Note the “ $-$ ”: in these coordinates, as  $z$  gets larger,  $\tau$  gets smaller. Alternatively, you can view  $\tau$  as being the optical depth for a photon traveling *into* the star.

Using eq. (3.20), we then rewrite the equation of hydrostatic balance (2.1) as

$$\begin{aligned} -\rho g = \frac{dP}{dz} &= \frac{dP}{d\tau} \frac{d\tau}{dz} = -\rho\kappa \frac{dP}{d\tau}, \\ \frac{dP}{d\tau} &= \frac{g}{\kappa}. \end{aligned} \quad (3.21)$$

Since we are in a thin layer, we can take the gravitational acceleration  $g$  to be approximately constant. By integrating hydrostatic equilibrium from where  $\tau = 0, P = 0$  to where  $\tau = 1$ , we obtain an approximate value of the photospheric pressure,

$$P_{\text{ph}} = \int_0^{P_{\text{ph}}} dP = \int_0^1 \frac{g}{\kappa} d\tau \approx \frac{g}{\kappa}.$$

The surface gravity sets the pressure at the **photosphere**, the location where the optical depth is of order unity and where photons can escape from the star.

EXERCISE 3.12 — Suppose you observe a star that has a 10% larger mass and 10% larger radius than the sun. All else being equal, how does the pressure at the photosphere of this star compare to that of the sun?

For our second approximation, we assume that the matter is in LOCAL THERMAL EQUILIBRIUM (LTE). This means there is a well-defined temperature at each depth. Furthermore, the emissivity is related to the absorption opacity,

$$j_\nu = \kappa^{\text{abs}} B_\nu.$$

Note that this does *not* imply the radiation field is actually Planckian.

We then take the radiative transfer equation (3.6) and substitute our definition of optical depth (eq. [3.20]) to obtain

$$\mu \frac{dI_\nu}{d\tau} = I_\nu - S_\nu. \quad (3.22)$$

Here

$$S_\nu = \frac{j_\nu + \kappa^{\text{sca}} J_\nu}{\kappa} = \frac{\kappa^{\text{abs}} B_\nu + \kappa^{\text{sca}} J_\nu}{\kappa}.$$

If, in addition, the matter is in steady-state, then the rate at which energy is absorbed from the radiation field,  $\int \kappa^{\text{abs}} I_\nu d\nu d\Omega$ , must equal the rate at which energy is emitted,  $\int j_\nu d\nu d\Omega$ :

$$\begin{aligned} \int (j_\nu - \kappa^{\text{abs}} I_\nu) d\nu d\Omega &= \kappa^{\text{abs}} \int (B_\nu - I_\nu) d\nu d\Omega \\ &= 4\pi \kappa^{\text{abs}} \int (B_\nu - J_\nu) d\nu = 0. \end{aligned}$$

In this expression we use the LTE expression for  $j_\nu$  and we pull  $\kappa^{\text{abs}}$  from the integral because it is independent of frequency.

Since  $J = \int J_\nu d\nu = \int B_\nu d\nu = B$ , it follows that  $S = \int S_\nu d\nu = B$  as well.

*For a gray atmosphere in steady-state, local thermal equilibrium, the integrated source function and mean intensity equal the Planck value:*

$$S(\tau) = J(\tau) = B(\tau),$$

Atmosphere is in steady-state LTE

Note that this does *not* imply that  $I_\nu = B_\nu$  or  $J_\nu = B_\nu$ : it only means their frequency-integrated averages are equal.

If we integrate eq. (3.22) over all angles,

$$\begin{aligned} \int \mu \frac{dI_\nu}{d\tau} d\Omega &= \int I_\nu d\nu - \int S_\nu d\Omega \\ \frac{dF_\nu}{d\tau} &= 4\pi (J_\nu - S_\nu), \end{aligned}$$

and then integrate over all frequencies and use the steady-state, LTE relation,

$$\frac{dF}{d\tau} = \frac{d}{d\tau} \int F_\nu d\nu = 4\pi \int (J_\nu - S_\nu) d\nu = 0.$$

We thus have the remarkable result:

*For a steady-state gray atmosphere in local thermal equilibrium, the total flux  $F = \int F_\nu d\nu$  is constant.*

That is, the atmosphere does not add to, or detract from, the radiation flowing through it.

We still have the problem that eq. (3.22) includes both the derivative and integral of  $I_\nu$ . To get around this, we expand  $I_\nu$  in Legendre polynomials<sup>2</sup>,

$$I_\nu(\tau, \mu) = I_{\nu,0}(\tau)\mathcal{P}_0(\mu) + I_{\nu,1}(\tau)\mathcal{P}_1(\mu) + I_{\nu,2}(\tau)\mathcal{P}_2(\mu) + \dots$$

and only retain the first two terms,  $\mathcal{P}_0(\mu) = 1$ ,  $\mathcal{P}_1 = \mu$ . That is, we assume  $I_\nu$  is *linear* in  $\mu$ :  $I_\nu = I_{\nu,0}(\tau) + I_{\nu,1}(\tau)\mu$ .

In terms of this expansion, the angle-averaged specific intensity is

$$J_\nu(\tau) = \frac{1}{4\pi} \int I_\nu d\mu d\phi = I_{\nu,0}(\tau),$$

and hence the specific energy density is  $U_\nu = 4\pi/c \cdot J_\nu = 4\pi/c \cdot I_{\nu,0}$ . The specific flux is

$$F_\nu(\tau) = \int \mu I_\nu d\mu d\phi = \frac{4\pi}{3} I_{\nu,1}(\tau).$$

We can therefore use these relations for  $I_{\nu,0}$  and  $I_{\nu,1}$  to express the intensity as

$$I_\nu(\tau) = \frac{c}{4\pi} U_\nu(\tau) + \frac{3\mu}{4\pi} F_\nu(\tau). \quad (3.23)$$

<sup>2</sup> see Box 3.1

**Intensity is linear in  $\mu$**

### Box 3.1 Expansion in Legendre polynomials

You may recall from electrostatics that we can decompose the field from a set of charges into a sum of moments: dipole, quadrupole, and so on. The basis functions for this expansion are

## Box 3.1 continued

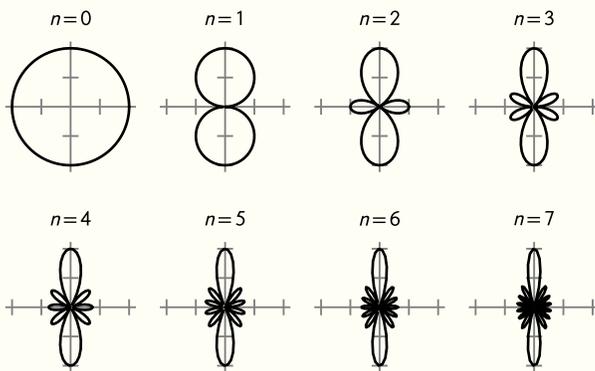
the Legendre polynomials  $\mathcal{P}_n(\cos \theta)$ , defined via

$$\frac{1}{\sqrt{1-2\mu z+z^2}} \equiv \sum_{n=0}^{\infty} \mathcal{P}_n(\mu) z^n,$$

for  $-1 < \mu < 1$ ,  $|z| < 1$ . The first four polynomials are

$$\begin{aligned} \mathcal{P}_0(\mu) &= 1 & \mathcal{P}_2(\mu) &= \frac{1}{2}(3\mu^2 - 1) \\ \mathcal{P}_1(\mu) &= \mu & \mathcal{P}_3(\mu) &= \frac{1}{2}(5\mu^3 - 3\mu), \end{aligned}$$

and the first eight Legendre polynomials are plotted below.



As  $n$  increases, the angular variations become finer.

The Legendre polynomials are *orthogonal* in the following sense:

$$\int_{-1}^1 \mathcal{P}_n(\mu) \mathcal{P}_m(\mu) d\mu = \begin{cases} 0 & m \neq n \\ \frac{2}{2n+1} & m = n \end{cases}. \quad (3.24)$$

As a result of this orthogonality, we can decompose the radiative intensity into multipoles:

$$I = \sum_{n=0}^{\infty} I_n \mathcal{P}_n(\mu). \quad (3.25)$$

---

EXERCISE 3.13— Use eq. (3.24) to show that  $(4\pi)^{-1} \int I d\Omega = I_0$  and  $\int \mu I d\Omega = (4\pi/3)I_1$ , for  $I = I_0 + I_1\mu$ .

---

Since the flux in the atmosphere is constant, it must be equal to its value far from the star where  $\tau \rightarrow 0$ :  $F(\tau = 0) = \sigma_{\text{SB}} T_{\text{eff}}^4$ . We can therefore integrate eq. (3.23) over frequency and substitute for  $F$ :

$$I(\mu, \tau) = \frac{c}{4\pi} U(\tau) + \frac{3\mu}{4\pi} \sigma_{\text{SB}} T_{\text{eff}}^4. \quad (3.26)$$

To solve for  $U(\tau)$ , integrate eq. (3.22) over frequency,  $\mu \, dI/d\tau = I - S$ , substitute for  $I$  on the left-hand side using eq. (3.26), then multiply by  $\mu$  and integrate over all angles:

$$\begin{aligned} \frac{d}{d\tau} \int \mu^2 I \, d\Omega &= \int \mu I \, d\Omega - \int \mu S \, d\Omega \\ \frac{c}{4\pi} \frac{d}{d\tau} \int \mu^2 U \, d\Omega + \frac{3}{4\pi} \sigma_{\text{SB}} T_{\text{eff}}^4 \int \mu^3 \, d\Omega &= F - \int \mu S \, d\Omega \\ \frac{c}{3} \frac{dU}{d\tau} &= F = \sigma_{\text{SB}} T_{\text{eff}}^4 \end{aligned} \quad (3.27)$$

The integrals over  $\mu^3$  and  $\mu S$  vanish because  $S$  is independent of angle and  $\int_{-1}^1 \mu \, d\mu = \int_{-1}^1 \mu^3 \, d\mu = 0$ . Also,  $U$  is independent of  $\mu$ , and  $\int \mu^2 \, d\Omega = \int_0^{2\pi} \int_{-1}^1 \mu^2 \, d\mu \, d\phi = (4\pi/3)$ .

Equation (3.27) is a first-order ODE, which upon integration yields

$$U(\tau) = \frac{3}{c} F(\tau + \tau_0) = \frac{3}{c} \sigma_{\text{SB}} T_{\text{eff}}^4 (\tau + \tau_0), \quad (3.28)$$

where  $\tau_0$  is an integration constant. Substituting this back into the expression for the intensity, eq. (3.26), gives

$$I(\mu, \tau) = \frac{3}{4\pi} \sigma_{\text{SB}} T_{\text{eff}}^4 (\tau + \tau_0 + \mu).$$

To fix the integration constant  $\tau_0$ , evaluate this expression at  $\tau = 0$ . Far outside the star, all of the radiation must be outward-bound. Hence if we integrate  $\mu I(\mu, \tau = 0)$  over  $0 \leq \mu \leq 1$ , we should recover the flux:

$$\sigma_{\text{SB}} T_{\text{eff}}^4 = \int_0^{2\pi} \int_0^1 \mu I(\mu, \tau = 0) \, d\mu \, d\phi = \frac{3}{4} \sigma_{\text{SB}} T_{\text{eff}}^4 \left( \tau_0 + \frac{2}{3} \right),$$

which fixes  $\tau_0 = 2/3$ .

We are almost finished! To recap, we now have the expressions for the intensity, flux, and radiative energy density under the assumption of a gray atmosphere in steady-state, local thermal equilibrium and under the approximation of the intensity being linear in  $\mu$ :

$$\begin{aligned} I &= \frac{3}{4\pi} \sigma_{\text{SB}} T_{\text{eff}}^4 \left( \tau + \mu + \frac{2}{3} \right) \\ F &= \sigma_{\text{SB}} T_{\text{eff}}^4 \\ U &= \frac{3}{c} \sigma_{\text{SB}} T_{\text{eff}}^4 \left( \tau + \frac{2}{3} \right). \end{aligned}$$

To finish this, we note that in the atmosphere  $J = B$  since we are in steady-state local thermal equilibrium. The radiative energy density can thus be written as  $U(\tau) = (4\pi/c)J(\tau) = (4\pi/c)B(\tau) = (4\sigma_{\text{SB}}/c)T^4(\tau)$ . Substituting this into eq. (3.28) gives us an expression for the temperature in terms of optical depth,

$$T^4(\tau) = \frac{3}{4} T_{\text{eff}}^4 \left( \tau + \frac{2}{3} \right). \quad (3.29)$$

This equation, along with eq. (3.21), determines the structure of the stellar atmosphere.

### Box 3.2 Decomposition of intensity into moments

Our integration of the radiative-transfer equation (3.22) is known as taking a **MOMENT** of the equation. A moment is simply a weighted average, where the weight is a power of  $\mu$ . For example, to take the zeroth-order moment of the radiative intensity, we multiply  $I_\nu$  by  $\mu^0 = 1$ , integrate over all angles, and divide by  $4\pi$ :

$$J_\nu = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^1 I_\nu \, d\mu \, d\phi.$$

To take the first-order moment  $H_\nu$ , we use a weight  $\mu^1$ :

$$H_\nu = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^1 \mu I_\nu \, d\mu \, d\phi.$$

To take the second-order moment  $K_\nu$ , we use a weight  $\mu^2$ :

$$K_\nu = \frac{1}{4\pi} \int_0^{2\pi} \int_{-1}^1 \mu^2 I_\nu \, d\mu \, d\phi.$$

The first three moments have physically interpretable meanings: the specific radiative energy density, flux, and pressure are  $U_\nu = (4\pi/c)J_\nu$ ,  $F_\nu = 4\pi H_\nu$ , and  $P_\nu = (4\pi/c)K_\nu$ , respectively.

By taking moments of the radiative-transfer equation (3.22), we reduce the complicated integro-differential equation into a simpler ordinary differential equation. This comes at a cost, however; because the left-hand side contains  $\mu d/d\tau$ , the left hand side will have a higher-order moment than the right-hand side. By multiplying eq. (3.22) by successively higher powers of  $\mu$  and integrating, we generate an infinite series of ODE's for successively higher moments of  $I_\nu$ . The trick to this procedure is to adopt a **CLOSURE RELATION** that truncates this series. The classic scheme, attributed to Eddington, is to take  $K = J/3$ . The Eddington closure scheme is equivalent to expanding the radiative intensity to terms linear in  $\mu$ .

---

**EXERCISE 3.14** — Show that if we approximate the intensity as  $I_\nu(\mu, \tau) = I_{\nu,0}(\tau) + \mu I_{\nu,1}(\tau)$  (eq. [3.23]), then  $K_\nu = J_\nu/3$  identically.

---

---

EXERCISE 3.15— Deep in the star, we expect the radiation to be nearly isotropic, while it becomes outward-bound as  $\tau \rightarrow 0$ . Let's investigate this. We'll measure the anisotropy of the radiation field using the first two moments of the intensity (Box 3.2).

1. Demonstrate that  $H_\nu/J_\nu = 0$  if the radiation is isotropic.
2. Next, suppose the radiation is completely anisotropic: all the photons are headed in a narrow cone about the direction  $\mu = 1$ . To make this precise, let

$$I_\nu(\mu) = \begin{cases} I_{\nu,0} & 1 - \varepsilon \leq \mu \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

where  $I_{\nu,0}$  is a constant. Show that  $H_\nu/J_\nu \rightarrow 1$  as  $\varepsilon \rightarrow 0$ .

3. Now compute  $H(\tau)/J(\tau)$  for our gray atmosphere. What is the degree of anisotropy at  $\tau = 0$ ? at  $\tau = 2/3$ ? at  $\tau = 10$ ?
-

## 4

# Rainbow in the Dark

Now that we've discussed radiative transport in the star, we'll explore how the emergent spectrum of a star serves as a diagnostic of ambient conditions in the photosphere.

### 4.1 Overview

If light from the sun is passed through a grating (a piece of glass with finely etched lines), the light is dispersed in wavelength and creates a spectrum, such as the highly detailed one shown in Fig. 4.1. Superposed on the slow variation from red to violet are dark ABSORPTION LINES. The ions, atoms, and molecules in the solar atmosphere absorb light at specific frequencies and create these lines.

Beginning in the late 1800's, astronomers began classifying stars by the observed absorption lines in the spectra. At this time, Edward Pickering and Williamina Fleming of the Harvard College Observatory began amassing a vast catalog of stellar spectra. They classified these spectra according to the strength of observed hydrogen Balmer lines (the first four are  $H\alpha$ : 657 nm;  $H\beta$ : 486 nm;  $H\gamma$ : 434 nm;  $H\delta$ : 410 nm). Stars, such as Vega, with the strongest Balmer lines were classified as type "A", those with the next strongest were type "B", and so forth. Annie Jump Cannon, who would later succeed Fleming as curator of astronomical photography at the observatory, simplified and reorganized the scheme, and added decimal subdivisions (0 . . . 9) for each type<sup>1</sup>. When stellar color is taken into account, the ordering of stars, from blue to red, is "OBAFGKM".

The classification of large numbers of stars allowed for comparison of brightness and spectral type for stars in clusters (and hence at the same distance). Hertzsprung and Russell independently noticed<sup>2</sup> that most stars tended to lie along a band, termed the MAIN SEQUENCE, in a plot of absolute magnitude (or luminosity) against stellar type (now known as a HERTZSPRUNG-RUSSELL DIAGRAM). Figure 4.2 shows some standard main-sequence stars, along with their stellar type and approximate color.

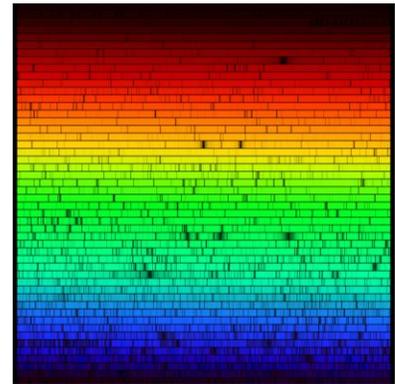


Figure 4.1: Visible spectrum of the sun. Frequency increases along a row from left to right, and by rows from top to bottom. Image credit & copyright: N.A.Sharp (NSF), FTS, NSO, KPNO, NAO/AURA/NSF.

<sup>1</sup> For example, the sun's type is G2

<sup>2</sup> Ejnar Hertzsprung. Über die Sterne der Unterabteilungen c und ac nach der Spektralklassifikation von Antonia C. Maury. *Astronomische Nachrichten*, 179(24):373, January 1909. DOI: 10.1002/asna.19081792402; and Henry Norris Russell. Relations Between the Spectra and Other Characteristics of the Stars. *Popular Astronomy*, 22:275–294, May 1914

<sup>3</sup> C. H. Payne. *Stellar Atmospheres; a Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars*. PhD thesis, RADCLIFFE COLLEGE., 1925

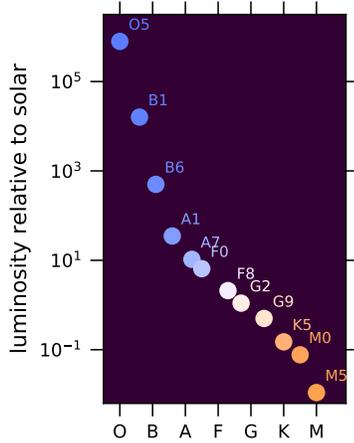


Figure 4.2: Hertzsprung-Russell diagram showing standard main-sequence stars. Colors are approximate translations of the spectra [Harre and Heller, 2021].

<sup>4</sup> J. D. Kirkpatrick, I. N. Reid, J. Liebert, et al. Dwarfs Cooler than “M”: The Definition of Spectral Type “L” Using Discoveries from the 2 Micron All-Sky Survey (2MASS). *ApJ*, 519:802–833, July 1999

<sup>5</sup> Michael C. Cushing, J. Davy Kirkpatrick, Christopher R. Gelino, et al. The Discovery of Y Dwarfs using Data from the Wide-field Infrared Survey Explorer (WISE). *ApJ*, 743:50, December 2011. DOI: 10.1088/0004-637X/743/1/50

The classification of stars based on spectra was put on a firm physical foundation with the influential PhD thesis of Cecilia Payne-Gaposhkin<sup>3</sup>, who applied the Boltzmann and Saha equations to show that different stellar spectra were consistent with changes in temperature, rather than composition, of the stellar photosphere. The sequence of stellar types is therefore a temperature sequence, with “O” stars being the hottest.

In the 1990’s the “L” and “T” classes were added<sup>4</sup> for cool stars and brown dwarfs (stellar-like objects that do not reach central temperature sufficient for fusion of hydrogen into helium). With the introduction of the “Y” stellar type<sup>5</sup>, this classification was further extended to even cooler objects having  $T_{\text{eff}} \lesssim 500$  K.

## 4.2 The hydrogen atom

To understand why the Balmer lines are strongest in a certain range of temperatures, we first need to review the workings of a hydrogen atom.

The electrons bound to an atom or molecule can only occupy states having a discrete set of energies. For example, the electron in a hydrogen atom only has energies

$$E_n = -13.6 \text{ eV} \times \frac{1}{n^2}, \quad (4.1)$$

where  $n > 0$  is an integer known as the PRINCIPAL QUANTUM NUMBER. These energies are negative, relative to a free electron. For example, the ground state ( $n = 1$ ) has energy  $-E_{\text{Ry}} = -13.6$  eV, meaning that 13.6 eV is required to remove an electron in its ground state from the atom.

Because the electrons in an atom can only have certain energies, the atom can only absorb or emit light at specific wavelengths, such that the energy of the photon matches the difference in energy between two levels. For example, a hydrogen atom in its ground state can absorb a photon of energy

$$E_{1 \rightarrow 2} = -E_{\text{Ry}} \left( \frac{1}{2^2} - \frac{1}{1^2} \right) = 10.2 \text{ eV}$$

corresponding to the energy required to excite the electron from level  $n = 1$  to  $n = 2$ . The wavelengths that can be absorbed by a hydrogen atom at rest can be found by substituting  $E = hc/\lambda$  into equation (4.1):

$$\lambda_{m \rightarrow n} = \lambda_0 \left( \frac{1}{m^2} - \frac{1}{n^2} \right)^{-1}, \quad (4.2)$$

where  $\lambda_0 = 91.2$  nm and  $n > m$ . The transitions from the lowest levels are named after their discoverers: Lyman for  $1 \rightarrow n$ , Balmer for  $2 \rightarrow n$ , Paschen for  $3 \rightarrow n$ . A greek letter is used to denote the higher state: for example Lyman  $\alpha$  (abbr. Ly $\alpha$ ) means  $1 \rightarrow 2$ , with  $\lambda_{\text{Ly}\alpha} = 121.6$  nm. Note that  $\lambda_{m \rightarrow n} > \lambda_0$ ; photons with wavelengths  $\lambda < 91.2$  nm have sufficient

energy to knock the electron out of the atom, thereby producing a hydrogen ion and a free electron. The first line transition in the Balmer series is  $2 \rightarrow 3$ , and is designated  $H\alpha$ :  $\lambda_{H\alpha} = 656.3 \text{ nm}$ . The first 20 lines for each of the Lyman, Balmer, and Paschen series are shown in Fig. 4.3; note the  $3 \rightarrow 4$  transition is outside the plot range. The Balmer lines lie in the visible range.

### 4.3 The Boltzmann Equation

In order to produce a Balmer absorption line, we must have some hydrogen atoms in the photosphere with electrons in the energy level  $n = 2$ . The more atoms in a state  $n = 2$ , the more absorption and the stronger the line. To find the number of atoms with energy level  $n = 2$ , we make use of a fundamental result, due to Boltzmann, from statistical (thermal) physics; namely, that if our sample of atoms is in thermal equilibrium, then the ratio of the number of atoms with energy  $E_i$  to the number of atoms with energy  $E_j$  is

$$\frac{N_i}{N_j} = \frac{g_i}{g_j} \exp\left(-\frac{E_i - E_j}{k_B T}\right). \quad (4.3)$$

Here the number  $g_n$  gives the number<sup>6</sup> of quantum mechanical states having energy  $E_n = -E_{Ry}/n^2$ . For an energy level  $n$ , there are  $n^2$  possible states, each having a different angular momentum. For each of these  $n^2$  states, both the electron and proton may each have 2 possible spins. The total number of states for energy  $E_n$  is therefore  $g_n = 2 \times 2 \times n^2$ .

Suppose we wish to know the fraction of atoms in a given state  $i$ : that is, we wish to know

$$x_i = \frac{N_i}{N_1 + N_2 + \dots + N_i + \dots}$$

Using equation (4.3), we can express  $x_i$  as

$$\begin{aligned} x_i &= \frac{g_i e^{-E_i/k_B T}}{g_1 e^{-E_1/k_B T} + g_2 e^{-E_2/k_B T} + \dots + g_i e^{-E_i/k_B T} + \dots} \\ &\equiv \frac{g_i e^{-E_i/k_B T}}{Q}, \end{aligned} \quad (4.4)$$

where the quantity

$$Q = \sum_n g_n \exp\left(-\frac{E_n}{k_B T}\right) \quad (4.5)$$

is the *partition function*. Loosely speaking, the partition function indicates the number of ways the sample of atoms can be partitioned among the different energy levels.

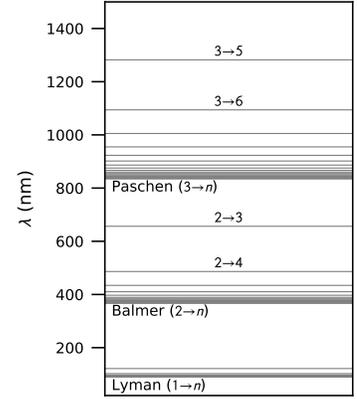


Figure 4.3: Spectral lines of neutral hydrogen

<sup>6</sup>  $g_n$  is known as the *degeneracy* of a given level  $n$

### Box 4.1 The partition function for neutral hydrogen

The partition function for neutral hydrogen, eq. (4.5), diverges as  $n \rightarrow \infty$ . To see this, substitute  $g_n = 4n^2$  and  $E_n = -E_{\text{Ry}}/n^2$  and factor out common terms to obtain

$$Q = 4e^{\beta E_{\text{Ry}}} \sum_n n^2 e^{-\beta E_{\text{Ry}}(1-1/n^2)},$$

with  $\beta = (k_{\text{B}}T)^{-1}$ . The sum evidently diverges, since for  $n \gg 1$  the individual terms approach  $n^2 e^{-\beta E_{\text{Ry}}}$ . In practice, this divergence isn't a problem, as there is an upper limit on  $n$  set by ambient conditions. For example, the mean distance of the electron from the nucleus is  $\approx a_{\text{B}} n^2$ , where  $a_{\text{B}} = 5.29 \times 10^{-11}$  cm is the Bohr radius. As a result, each atom takes up a volume  $\approx a_{\text{B}}^3 n^6$ ; if the atoms are not to overlap, then the volume per atom,  $V/N \equiv 1/\xi$ , must be larger than this by some factor. To make this concrete, set the volume of an atom to be less than half of that available in our gas:

$$a_{\text{B}}^3 n^6 \lesssim \frac{1}{2} \frac{V}{N} = \frac{1}{2\xi}.$$

Thus the maximum level is  $n < (2a_{\text{B}}^3 \xi)^{-1/6}$ . For a typical A-star photospheric density  $\xi \sim 10^{15} \text{ cm}^{-3}$ , the energy level cutoff is  $n \approx 35$ . In practice the cutoff will be even lower because of collisions.

You might worry that this estimate for the maximum value of  $n$  is a bit sloppy. Fortunately, the precise maximum value of  $n$  is unimportant for most applications. The reason is that the terms in the partition function increase only slowly. As an example, the terms and cumulative sum in the partition function at a temperature  $T = 10^4$  K are as follows.

$n$	$n^2 e^{-\beta E_{\text{Ry}}(1-1/n^2)}$	$4 \sum_{i=1}^n i^2 e^{-\beta E_{\text{Ry}}(1-1/i^2)}$
1	1.00e+00	4.0000
2	2.88e-05	4.0001
3	7.23e-06	4.0001
	⋮	
26	9.62e-05	4.0038
	⋮	
52	3.78e-04	4.0274
	⋮	
268	9.99e-03	7.5901

As we can see from the cumulative sum (rightmost column), the value of the partition function is insensitive to cutoff until  $n$  is quite large; indeed, for many applications it is reasonably accurate to just use  $Q \approx 4e^{\beta E_{\text{Ry}}}$ .

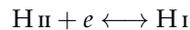
---

EXERCISE 4.1 — Assuming that the first term  $g_1 e^{-E_1/k_B T}$  dominates the sum in the partition function (see Box 4.1), plot the fraction of neutral hydrogen in its  $n = 2$  state as a function of temperature, for  $5\,000\text{ K} < T < 20\,000\text{ K}$ .

---

#### 4.4 Ionization: The Saha equation

As the temperature in the gas rises, there are more photons with sufficient energy to eject electrons from an atom. In addition, collisions between atoms also become sufficiently energetic to ionize the atom. In astronomical nomenclature, the ionization state is denoted by a small Roman numeral: Fe I denotes neutral iron, Fe II denotes singly-ionized iron (charge +1), Fe III denotes doubly-ionized iron (charge +2), and so on. In thermal equilibrium, the rate at which atoms are ionized must equal the rate at which ions and electrons recombine: for example, in a gas consisting of hydrogen atoms, hydrogen ions (i.e., protons), and electrons the reaction



is in equilibrium. We'd like to extend equation (4.3) to find the ratio of two ionization states  $N_{i+1}/N_i$ . Although deriving this equation, termed the SAHA EQUATION<sup>7</sup>, is beyond the scope of the course, what we shall do is take the equation apart and try to understand how it works. The Saha equation for the ratio of the populations of two ionization states<sup>8</sup>  $N_{i+1}$  and  $N_i$  is

$$\frac{N_{i+1}}{N_i} = \left[ \frac{2}{n_e} \left( \frac{m_e k_B T}{2\pi \hbar^2} \right)^{3/2} \right] \frac{Q_{i+1}}{Q_i}. \quad (4.6)$$

In this equation,  $n_e$  denotes the electron density—the number of free electrons per unit volume—and  $m_e$  is the electron mass. The terms  $Q_{i+1}$  and  $Q_i$  are the partition functions for the two states, *both measured with respect to the same zero-point for energy*. To understand better what this equations means, let's consider each of the color-coded factors in turn.

Start with the term  $Q_{i+1}/Q_i$ . If both partition functions are dominated by the ground state term<sup>9</sup> then

$$\begin{aligned} \frac{Q_{i+1}}{Q_i} &= \frac{g_{i+1,1}}{g_{i,1}} e^{-\beta(E_{i+1,1} - E_{i,1})} \\ &= \frac{g_{i+1,1}}{g_{i,1}} e^{-\beta E_{\text{ion}}} \end{aligned}$$

Here  $E_{\text{ion}} = E_{i+1,1} - E_{i,1}$  is the energy needed to remove an electron from an ion in state  $i$  and we use the common shorthand  $\beta = (k_B T)^{-1}$ . Thus  $Q_{i+1}/Q_i$  resembles the Boltzmann equation (4.3).

The additional factor in [·] in eq. (4.6) arises because we also need to allow for the number of free electron states. When the atom is ionized, each electron quickly acquires an average kinetic energy  $(3/2)k_B T$ . There

<sup>7</sup> Derived by Meghnad Saha in 1920

<sup>8</sup> In this context,  $N_{i+1}/N_i$  refers to ratios such as  $N_{\text{Fe II}}/N_{\text{Fe I}}$ . Each population  $N$  can be divided into sub-populations based on the different electron energy levels. For example,  $N_{\text{H I}} = N_{\text{H I},1} + N_{\text{H I},2} + \dots$ , where  $N_{\text{H I},2}$  are the number of hydrogen in ionization state I with an electron in the second energy level.

<sup>9</sup> see Box 4.1

are many different states having this energy: the electron can be in different locations and moving in different directions, for example.

You might think that there would be an infinitude of possible electron states. Quantum mechanics, however, sets limitations on this number. First, we have the Pauli exclusion principle: no two electrons can be in the same location with the same momentum and same spin. What do we mean by same location and momentum? Recall the Heisenberg uncertainty principle: the electrons  $x$ -position and  $x$ -momentum are spread about a range of values  $\Delta x$  and  $\Delta p_x$ , and these uncertainties are related via

$$\Delta x \Delta p_x \gtrsim h.$$

Thus, if we imagine dividing our volume into little boxes of volume

$$\Delta V = \Delta x \Delta y \Delta z \approx \frac{h^3}{\Delta p_x \Delta p_y \Delta p_z},$$

<sup>10</sup> Because electrons have spin 1/2, we can put two electrons into the same position and momentum state if their spins are oppositely directed.

each box can hold two electrons.<sup>10</sup> Suppose we have a volume  $V$ ; how many boxes are there? The number of available boxes is

$$\frac{V}{\Delta V} \approx \frac{V \Delta p_x \Delta p_y \Delta p_z}{h^3}.$$

To estimate the size of  $\Delta p_x \Delta p_y \Delta p_z$ , let's take  $\Delta p_x \sim p_x$  and similarly for  $\Delta p_y$  and  $\Delta p_z$ ; further, if everything is isotropic then  $p_x \approx p_y \approx p_z$  on average, so  $\Delta p_x \Delta p_y \Delta p_z \sim p_x^3$ . Now the kinetic energy of the electron is  $p^2/2m_e$ , and  $p^2 = p_x^2 + p_y^2 + p_z^2 \approx 3p_x^2$ . Hence the kinetic energy is  $(3/2)p_x^2/m_e$ ; in thermal equilibrium, however, the kinetic energy has an average value of  $(3/2)k_B T$ . The value of  $p_x^2$  is therefore

$$p_x^2 \approx m_e k_B T,$$

and the number of boxes is

$$\frac{V}{\Delta V} \sim V \frac{p_x^3}{h^3} \sim V \frac{(m_e k_B T)^{3/2}}{h^3}.$$

If our volume  $V$  contains  $N_e$  electrons, then the number of states electron is

$$\frac{2V}{N_e \Delta V} \sim \frac{2V}{N_e} \frac{(m_e k_B T)^{3/2}}{h^3}.$$

The factor of 2 appears because each box can hold 2 electrons. Recognizing that  $N_e/V = n_e$ , we see that this number of states per free electrons corresponds to the factor in [·] in equation (4.6). When the numerical calculation is done correctly, the additional factor of  $2\pi$  arises.

The number of states per free electron plays an important role in setting the temperature at which a species ionizes. You might expect, since a term  $e^{-E_{\text{ion}}/k_B T}$  appears in the ratio  $N_{i+1}/N_i$ , that a species would

ionize at a temperature  $E_{\text{ion}}/k_{\text{B}}$ . In fact the ionization temperature is much lower. To see how this works, define

$$\zeta = \ln \left[ \frac{1}{n_e} \left( \frac{m_e k_{\text{B}} T}{2\pi \hbar^2} \right)^{3/2} \right].$$

We can then write eq. (4.6)—with the approximation that the partition functions are dominated by the ground state—as

$$\frac{N_{i+1}}{N_i} = \frac{2g_{i+1,1}}{g_{i,1}} \exp(\zeta - \beta E_{\text{ion}}).$$

Now the factor  $g_{i+1,1}/g_{i,1}$  is of order unity. Hence, when the gas ionizes and  $N_{i+1,1} \approx N_{i,1}$ , we must have that  $\zeta \approx \beta E_{\text{ion}}$ ; put differently, the ionization temperature will not be  $E_{\text{ion}}/k_{\text{B}}$  but rather  $E_{\text{ion}}/k_{\text{B}}\zeta$ . Under conditions in the photosphere of an A star ( $T \approx 10^4$  K,  $n \sim 10^{15}$  cm $^{-3}$ ),  $\zeta \approx 15$ .

In more intuitive terms, when an electron is ejected from an atom, it has an enormously large number  $\sim e^{15}$  number of different states available. To rejoin with an ion requires being in the right place at the right time with the right energy. The large number of available states makes this unlikely, so the electron must wander lonely through a vast and desolate phase space until at long last it reunites with an ion. In a sense, the large number of available states per electron makes ionization easier than recombination; as a result the temperature at which ionization occurs is considerably lower than  $E_{\text{ion}}/k_{\text{B}}$ .

---

**EXERCISE 4.2** — Let  $n_i$  be the density of H I and  $n_{\text{II}}$  be the density of H II. Denote the fraction of neutral hydrogen as  $x = n_i/(n_i + n_{\text{II}})$ , so that  $1 - x = n_{\text{II}}/(n_i + n_{\text{II}})$  is the fraction of ionized hydrogen. Take  $n_i + n_{\text{II}} = 10^{15}$  cm $^{-3}$ , and assume that all free electrons come from the ionization of hydrogen, so that  $n_e = n_{\text{II}}$ . Plot  $x$  as a function of temperature for  $7\,500$  K  $\leq T \leq 15\,000$  K, and find the temperature at which  $x = 1/2$ . Then multiply  $x$  by the fraction  $n_2/n_1$ , as set by the Boltzmann equation, to find the fraction of hydrogen in the  $n = 2$  level.

---

As shown in exercise 4.2, the Balmer lines, which correspond to transitions  $2 \rightarrow 3$ ,  $2 \rightarrow 4$ , ..., are most prominent in A stars. These stars have  $T_{\text{eff}} = (7\,500\text{--}9\,500)$  K. At lower temperatures, the population of hydrogen atoms in the level  $n = 2$  decreases as  $e^{-E_2/k_{\text{B}}T}$  and the lines become weak. At higher temperatures, the number of neutral hydrogen atoms decreases; most of the hydrogen is ionized, and the Balmer lines again become weaker.

These arguments apply to other species present in the stellar photosphere. Figure 4.4 displays spectra for selected stellar types at optical wavelengths. In the hottest stars (type O:  $T_{\text{eff}} > 30\,000$  K), hydrogen is mostly ionized and the lines are from He II and multiply-ionized metals. As the temperature cools into the B and A series, the hydrogen lines

increase in strength. Going from F into G ( $T_{\text{eff}} = (5\,000\text{--}6\,000\text{ K})$ ), the hydrogen lines decrease, while lines from singly-ionized and neutral metals such as Ca II, Ca I, and Fe I become strong. At still lower temperatures in the K and M ( $T_{\text{eff}} < 3\,500\text{ K}$ ) types, absorption from molecules such as TiO becomes prominent. An example is the broad trough seen in the K spectrum near  $\lambda = 500\text{ nm}$ .

Figure 4.4: Spectra from main-sequence stars of spectral types O–K. Data from Jacoby et al. [1984].

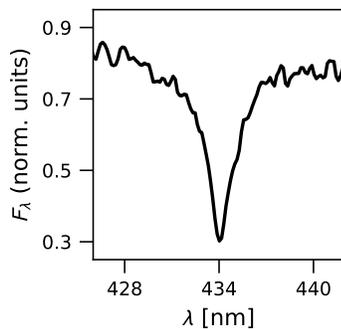
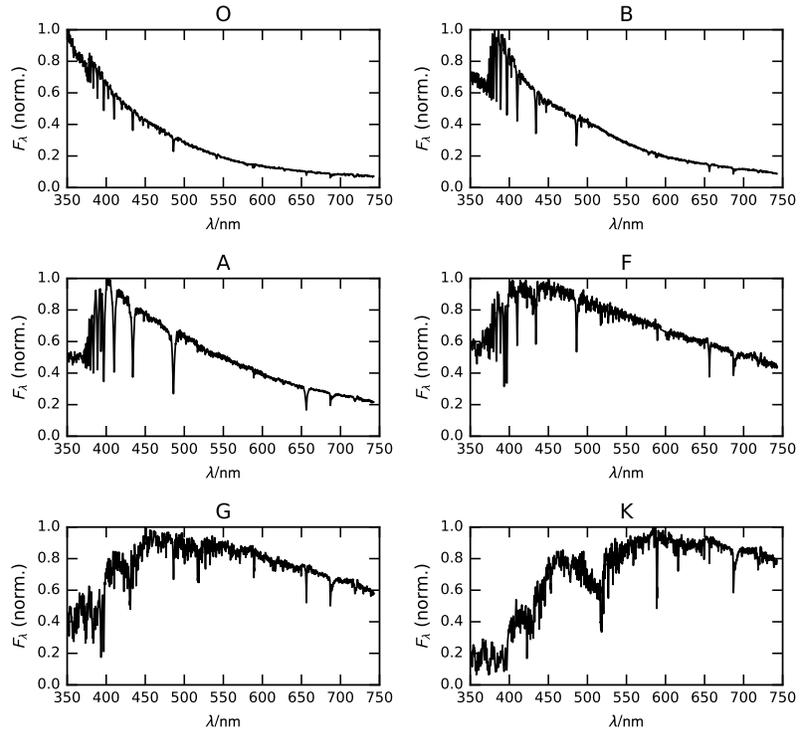


Figure 4.5:  $H\gamma$  absorption line observed from the main-sequence A1 star HD16608. Spectrum from Jacoby et al. [1984].

#### 4.5 Pressure broadening of lines

We've now demonstrated how stars may be classified by the absorption lines in their spectra, and how this classification gives us the photosphere effective temperature. We can also obtain information about the pressure at the photosphere, and hence the surface gravity of the star, by looking at the shape of the absorption lines. A zoomed-in view of the  $H\gamma$  line ( $2 \rightarrow 5$  transition in H I) from a main-sequence A1 star is shown in Fig. 4.5. The line is spread over a few nanometers, compared against a central value of 434 nm.

TO UNDERSTAND WHAT SETS THE SHAPE, AND WIDTH, OF THE ABSORPTION LINE, we need to model our atomic transition. Consider an electronic

transition in an atom between two energy levels,  $E_m$  and  $E_n$ . The natural frequency of this transition is  $\nu_0 = |E_n - E_m|/h$ . Light incident on the atom with frequency  $\nu \neq \nu_0$  drives the electron at frequency  $\nu$ .

Since the transition between two states has a definite frequency associated with it, let's start with a simple harmonic oscillator, which is described by an equation

$$\frac{d^2x}{dt^2} + \omega_0^2x = 0.$$

Here  $\omega_0 = 2\pi\nu_0$ . Light is described as an electromagnetic wave, so classically the electron feels a force  $eE \cos(\omega t)$ , where  $\omega = 2\pi\nu$ . An accelerating electron radiates, which damps the acceleration of the electron. The damping can be modeled as a force that is proportional to the velocity,  $-m\Gamma dx/dt$ . Classically, the transition in an atom can therefore be modeled as an electromagnetic oscillator with damping and driving terms,

$$\frac{d^2x}{dt^2} + \Gamma \frac{dx}{dt} + \omega_0^2x = \frac{eE}{m} \cos(\omega t).$$

This has a well known solution (see Box 4.2). The amplitude of oscillation is proportional to the energy removed from the incident light, which is proportional to the cross-section. The classical cross-section for absorption of radiant energy by an electromagnetic oscillator is thus

$$\sigma = \left( \frac{\pi e^2}{m_e c} \right) \left\{ \frac{\Gamma/4\pi}{(\nu_0 - \nu)^2 + (\Gamma/4\pi)^2} \right\}. \quad (4.7)$$

The function

$$\mathcal{L}(\nu; \Gamma) = \frac{1}{\pi} \frac{\Gamma/4\pi}{(\nu_0 - \nu)^2 + (\Gamma/4\pi)^2}$$

is known as a **LORENTZIAN**. In contrast to a Gaussian, a Lorentzian is characterized by broad "wings" (Fig. 4.6) away from the central frequency  $\nu_0$ . The actual value of the cross-section must be calculated using quantum mechanics. The overall shape of the cross-section is still in the form of equation (4.7), however, so the opacity is just

$$\rho\kappa_\nu = n_{\text{ion},m} \left( \frac{\pi e^2}{m_e c} \right) f_{mn} \left\{ \frac{\Gamma/4\pi}{(\nu_0 - \nu)^2 + (\Gamma/4\pi)^2} \right\}. \quad (4.8)$$

In this equation,  $f_{mn}$  is a number, called the *oscillator strength*, that results from the calculation of the transition probability from state  $m$  to state  $n$ , and  $n_{\text{ion},m}$  is the density of atoms in state  $m$ . The key point is that  $f_{mn}$  depends only on the details of the transition: the energies, spins, and parities of the atomic states. It does not depend on environmental parameters such as temperature and pressure. As a result,  $f_{mn}$  can be measured or computed once and then tabulated.

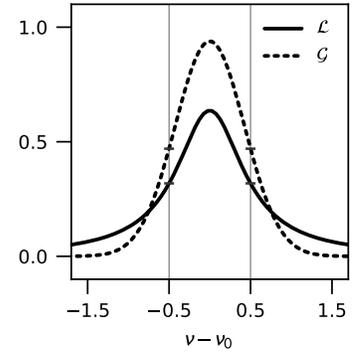
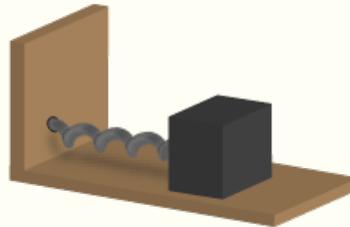


Figure 4.6: Comparison of a Lorentzian ( $\mathcal{L}$ , solid line) and a Gaussian ( $\mathcal{G}$ , dotted line), both with  $\text{FWHM} = 1$ . The area under each curve is unity.

### Box 4.2 The driven damped oscillator

Let's begin with a simple system: a mass  $m$  attached to a spring with force  $F = -kx$ .



If we put the origin of our coordinate system where the mass is at rest with the spring relaxed, then the equation of motion of the mass is

$$\frac{d^2x}{dt^2} + \frac{k}{m}x = 0. \quad (4.9)$$

You have solved this equation before: the most general solution is

$$x(t) = x_0 \cos(\omega_0 t) + \frac{v_0}{\omega_0} \sin(\omega_0 t) \quad (4.10)$$

with  $\omega_0^2 = k/m$  and with  $x_0$  and  $v_0$  being the initial position and velocity of the mass. The angular frequency  $\omega_0$  is related to the period of oscillation  $T$  as  $\omega_0 = 2\pi/T = 2\pi\nu$ .

NOW LET'S PUSH ON OUR MASS WITH AN OSCILLATING FORCE,  $F \cos(\omega t)$  WITH  $\omega \neq \omega_0$ . A real world example would be holding a vibrating tuning fork near another fork tuned to a different frequency. The equation of motion is now

$$\frac{d^2x}{dt^2} + \omega_0^2 x = \frac{F}{m} \cos(\omega t). \quad (4.11)$$

You can verify by substitution that a general solution is

$$x(t) = \frac{F/m}{(\omega_0^2 - \omega^2)} \cos(\omega t) + A \cos(\omega_0 t) + B \sin(\omega_0 t).$$

Let's start with our harmonic oscillator at rest ( $v_0 = dx/dt|_{t=0} = 0$ ) and at  $x|_{t=0} = 0$ . With these conditions, we can determine the constants  $A$  and  $B$ ; the solution is

$$x(t) = \frac{F/m}{(\omega_0^2 - \omega^2)} [\cos(\omega t) - \cos(\omega_0 t)].$$

Let's recast this by defining  $\Delta = \omega_0 - \omega$  and  $\omega_m = (\omega_0 + \omega)/2$ . Then

$$\begin{aligned} \omega_0^2 - \omega^2 &= (\omega_0 - \omega)(\omega_0 + \omega) = 2\Delta\omega_m, \\ \cos(\omega_0 t) &= \cos(\omega_m t + \Delta t/2), \\ \cos(\omega t) &= \cos(\omega_m t - \Delta t/2); \end{aligned}$$

## Box 4.2 continued

using the cosine addition rules and combining terms, we can write the solution as

$$x(t) = \left[ \frac{F/m}{\Delta\omega_m} \sin(\Delta t/2) \right] \sin(\omega_m t). \quad (4.12)$$

This illustrates the phenomena of **BEATS**: the oscillation consists of a **carrier signal** at frequency  $\omega_m$  with the **amplitude** modulated at the slower frequency  $\Delta/2$ . Notice that the amplitude increases as  $\Delta \rightarrow 0$ , i.e.,  $\omega \rightarrow \omega_0$ .

NOW LET'S MAKE OUR MODEL EVEN MORE REALISTIC. We add a frictional force that is proportional to velocity,  $F_{\text{friction}} = -m\Gamma dx/dt$ . Our complete equation of motion is then

$$\frac{d^2x}{dt^2} + \Gamma \frac{dx}{dt} + \omega_0^2 x = \frac{F}{m} \cos(\omega t). \quad (4.13)$$

The solution to this is straightforward to find, although the algebra is tedious (trust me on this). The general solution for initial conditions  $x|_{t=0} = x_0$  and  $dx/dt|_{t=0} = v_0$  is

$$\begin{aligned} x(t) = & \frac{F(\omega_0^2 - \omega^2)/m}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \cos(\omega t) \\ & + \frac{\Gamma\omega F/m}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \sin(\omega t) \\ & + \left[ x_0 - \frac{F(\omega_0^2 - \omega^2)/m}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \right] e^{-\Gamma t/2} \cos(\omega_\Gamma t) \\ & + \left[ \frac{v_0}{\omega_\Gamma} - \frac{\Gamma\omega F/m}{(\omega_0^2 - \omega^2)^2 + \Gamma^2\omega^2} \frac{\omega}{\omega_\Gamma} \right] e^{-\Gamma t/2} \sin(\omega_\Gamma t), \end{aligned} \quad (4.14)$$

with

$$\omega_\Gamma = \omega_0 \left( 1 - \frac{\Gamma^2}{4\omega_0^2} \right)^{1/2}.$$

Let's simplify this a bit. First, the last two terms decay as  $e^{-\Gamma t/2}$ : these are transients set by the initial conditions. After a time  $t \gg 2/\Gamma$ , therefore, only the first two terms, which oscillate at the driving frequency  $\omega$ , will remain.

We can simplify these first two terms even further: if we write

$$\cos(\omega t) = \frac{e^{i\omega t} + e^{-i\omega t}}{2}, \quad \sin(\omega t) = \frac{e^{i\omega t} - e^{-i\omega t}}{2i},$$

we can combine them and obtain

$$x(t) = \frac{F}{2m} \left[ \frac{1}{(\omega_0^2 - \omega^2) + i\Gamma\omega} \right] e^{i\omega t}$$

## Box 4.2 continued

$$\begin{aligned}
& + \frac{F}{2m} \left[ \frac{1}{(\omega_0^2 - \omega^2) - i\Gamma\omega} \right] e^{-i\omega t} \\
= & \Re \left\{ \frac{F}{m} \left[ \frac{1}{(\omega_0^2 - \omega^2) + i\Gamma\omega} \right] e^{i\omega t} \right\}. \quad (4.15)
\end{aligned}$$

We use the symbol “ $\Re$ ” to denote taking the real part of a complex quantity. The oscillation is thus described as the real part of a complex quantity  $Ae^{i\omega t}$ , with

$$A = \frac{F}{m} \left[ \frac{1}{(\omega_0^2 - \omega^2) + i\Gamma\omega} \right]$$

being the (complex) amplitude.

For  $\omega \approx \omega_0$ , we approximate  $(\omega_0^2 - \omega^2) \approx 2\omega_0(\omega_0 - \omega)$  and take the square of the amplitude to find,

$$\begin{aligned}
|A|^2 &= \left( \frac{F}{2m\omega_0} \right)^2 \frac{1}{(\omega_0 - \omega)^2 + (\Gamma/2)^2} \\
&= \frac{\pi}{2\Gamma} \left( \frac{F}{m\omega_0} \right)^2 \left\{ \frac{1}{\pi} \frac{\Gamma/2}{(\omega_0 - \omega)^2 + (\Gamma/2)^2} \right\} \quad (4.16)
\end{aligned}$$

We rewrote the amplitude in the second line so that the term in  $\{\cdot\}$  is normalized; in fact, it is the Lorentzian function  $\mathcal{L}(\omega; \Gamma)$  in terms of the driving frequency  $\omega$ .

<sup>11</sup> This is an application of the *Stark* effect that you learn about in quantum mechanics.

IN A STELLAR ATMOSPHERE, THE WIDTH  $\Gamma$  IS SET BY COLLISIONS. For example, when an electron passes close by our atom, the electric field shifts the energy levels of the atom<sup>11</sup>. The greater the collision rate, the larger the width. If we have two stars of the same photospheric temperature (so that both stars have the same lines), then a way to increase the collision rate is to increase the pressure. Recall, however, that in the stellar atmosphere  $P = (g/\kappa)\tau$ ; as a result, stars with a higher surface gravity will have broader lines. The inset in Figure 4.7 illustrates the broadening of the Balmer H $\gamma$  line ( $2 \rightarrow 5$ ) in the spectrum of a main-sequence A1 star compared with that of a supergiant A1 star.

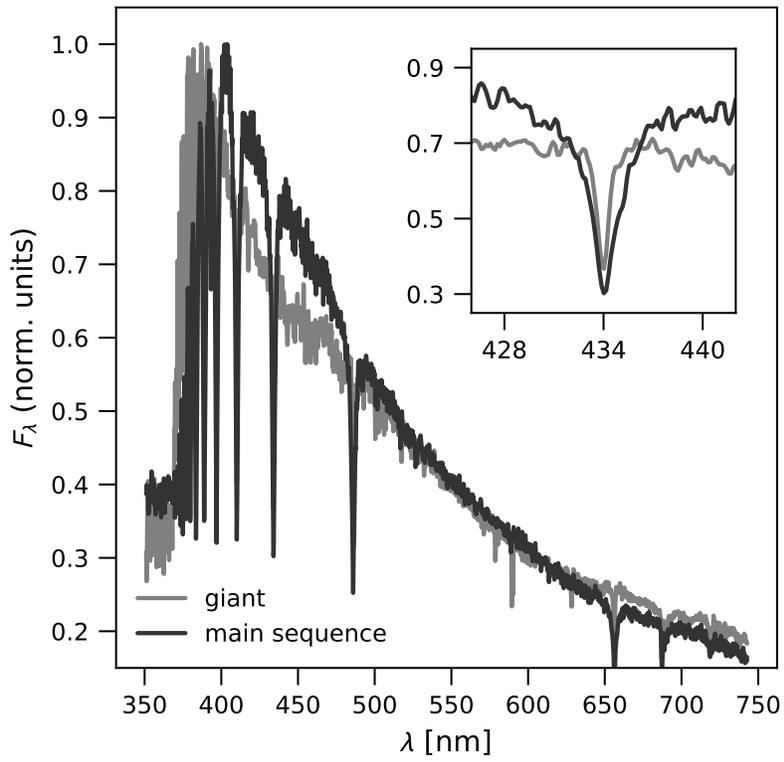


Figure 4.7: Spectra of two A1 stars, HD 16608 (a main sequence star) and SAO 12149 (a supergiant star). Spectra are from Jacoby et al. [1984].



# 5

## Burn

To recap, we have established a description for the basic features of a self-gravitating fluid:

1. For a set mass and radius, hydrostatic equilibrium (balance of pressure and gravity) is established on the time needed for a sound wave to cross the star. Once this equilibrium is established, the central pressure, density, and temperature are established.
2. The gradient in temperature from center to surface drives a luminosity, which is controlled by the opacity of material in the stellar interior.
3. The ambient pressure and temperature near the stellar photosphere (where  $\tau \sim 1$ ) are set by the surface gravity and opacity.

In this chapter we now discuss how the luminosity is generated by nuclear reactions in the core of a star, and the conditions needed to generate that luminosity.

### 5.1 The nucleus

Experimentally, nuclei are on the order of femtometers<sup>1</sup> in size. Like an atom, the nucleus also has excited states; typical energies for these states<sup>2</sup> are on the order of MeV. It therefore makes sense to use fm and MeV as our units of length and energy. In these units, the combination

$$\hbar c = 197 \text{ MeV fm}$$

to three significant digits. In quantum field theory, the strength of the electromagnetic interaction is characterized by the dimensionless FINE STRUCTURE CONSTANT

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137},$$

again to three significant digits. From these two quantities, we find the electron (or proton) charge in these units,

$$\frac{e^2}{4\pi\epsilon_0} = \alpha\hbar c = 1.44 \text{ MeV fm}.$$

<sup>1</sup> 1 fm =  $10^{-15}$  m. This unit is sometimes called a FERMI.

<sup>2</sup> 1 MeV =  $10^6$  eV; an ELECTRON VOLT (eV) is the energy acquired by an electron being accelerated through a potential difference of 1 volt.

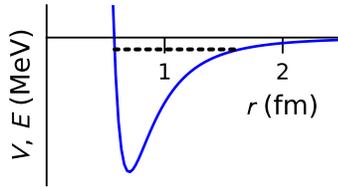


Figure 5.1: Schematic of the nuclear potential for a deuteron ( ${}^2\text{H}$ ). The binding energy of the deuteron is shown as a black dotted line.

In our units of MeV and fm, some relevant masses are

$$\begin{aligned} m_n &= 939.6 \text{ MeV}/c^2 \\ m_p &= 938.3 \text{ MeV}/c^2 \\ m_u &= 931.5 \text{ MeV}/c^2 \\ m_e &= 0.5110 \text{ MeV}/c^2 \end{aligned}$$

Put another way, the Coulomb potential energy between two protons separated by 1 fm is 1.44 MeV.

The strong nuclear force differs from electromagnetism and gravity in several ways. First, the strong nuclear force is short-range: the interaction vanishes for distances  $\gtrsim 2$  fm. It is weakly attractive for distances  $1 \text{ fm} \lesssim r \lesssim 2 \text{ fm}$  and becomes strongly repulsive at distances  $\ll 1$  fm. The potential between the neutron and proton in a deuterium ( ${}^2\text{H}$ ) nucleus (called a deuteron) therefore looks something like that sketched in Fig. 5.1. The deuteron’s ground state (black dotted line) is at  $E_d = -2.2 \text{ MeV}$ , so the nucleus is weakly bound ( $|E_d| \ll |V|$ , where  $V$  is the depth of the potential well).

---

EXERCISE 5.1 — We can estimate the depth of the well in Fig. 5.1. Since this is a two-body problem, transfer to center-of-mass coordinates and solve for a single particle with a reduced mass  $m_p m_n / (m_p + m_n) \approx m_n / 2$ . Use the uncertainty principle, with  $\Delta x$  being the width of the well, to get an estimate of  $p \sim \Delta p$  and from this estimate the kinetic energy of the particle. Finally, use the small value of the binding energy (sum of potential and kinetic energies) to estimate the depth of the potential well.

---

Also unlike electromagnetism and gravity, the strong nuclear force does not obey superposition: we cannot write the energy of the nucleus as a sum over the potential between all pairs of nucleons. Further, the strong nuclear force is not a central force, meaning that it depends on more than just the distance between any two nucleons. The atomic nucleus is thus much more complicated to describe than the electronic structure of the atom.

DESPITE THESE COMPLICATIONS, WE CAN CONSTRUCT A PHENOMENOLOGICAL FORMULA FOR THE NUCLEAR MASS THAT IS REASONABLY ACCURATE. Let us write the mass of a nucleus with  $A$  nucleons— $Z$  protons and  $N = A - Z$  neutrons—as

$$M(Z, N) = Zm_p + Nm_n - B(Z, N)/c^2,$$

where  $B(Z, N)$  is the BINDING ENERGY—the amount of energy that must be supplied to the nucleus in order to break it into its constituent protons and neutrons. Because the nuclear force is weakly attractive for separations  $1 \text{ fm} \lesssim r \lesssim 2 \text{ fm}$  and repulsive at shorter distances (Fig. 5.1), there is a characteristic spacing between nucleons that is a bit larger than 1 fm. In a large nucleus, we therefore expect the nucleons to have a roughly constant density, so that the volume of the nucleus is proportional to  $A$ ; experimentally, the radius of the nucleus is roughly<sup>3</sup>

$$r_A = (1.1 \text{ to } 1.8) \text{ fm} \times A^{1/3}.$$

<sup>3</sup> The value of the radius depends on how it is measured; scattering with various light particles (protons, neutrons, alpha, electrons) agree, however, that  $r_A \propto A^{1/3}$ .

Notice that because the nucleon-nucleon potential is short-ranged, nucleons in a large nucleus only interact with their nearest neighbors. Indeed the nucleon-nucleon interaction is similar in form to the potential between molecules in a fluid, such as a water drop. This motivates developing a simple formula that gives a decent approximation for the binding energy. For the first term, we estimate the binding energy of a large nucleus as just the (constant) binding energy of a single nucleon multiplied by the number of nucleons. Experimentally, it is found that for large nuclei this is the case: the binding energy per nucleon is roughly constant. We say that the nuclear interaction SATURATES, so that  $B(Z, N) \propto A = (Z + N)$ .

---

EXERCISE 5.2 — To see how the nuclear force differs from the long-range Coulomb and gravitational forces, suppose instead that the nuclear force acted like a super-gravity: that is, the potential was  $\propto 1/r$ . Use the results from our constant-density model of a star (eq. [2.22]) to derive how the binding energy would scale with  $A$  in this case.

---

It is energetically favorable to have equal numbers of neutrons and protons. We therefore define an asymmetry parameter  $\eta \equiv (N - Z)/(N + Z) = 1 - 2Z/A$ , so that  $-1 \leq \eta \leq 1$ . The nuclear contribution to the binding energy is maximized for  $\eta = 0$  (equal numbers of protons and neutrons). Because the nuclear force does not distinguish between neutrons and protons, the binding energy is quadratic in  $\eta$ , so that  $B$  doesn't depend on the sign of  $\eta$ . Thus our first approximation for the binding energy is  $B \approx (a_V - a_A \eta^2)A$ . Here  $a_V$  and  $a_A$  are as-yet-undetermined coefficients.

In a fluid drop there is a correction for the surface tension. Heuristically, we imagine that nuclei in the surface have fewer neighbors and are therefore not as bound. We therefore subtract from our formula a term proportional to the surface area,  $\propto r_A^2 \propto A^{2/3}$ . The next iteration of our liquid-drop approximation is thus  $B \approx (a_V - a_A \eta^2)A - a_S A^{2/3}$ .

Finally, the protons in the nucleus are charged and therefore repel one another. This Coulomb repulsion also reduces the binding energy. We therefore subtract a term  $\propto Z^2/r_A \propto Z^2/A^{1/3}$  from our mass formula to obtain

$$B = (a_V - a_A \eta^2)A - a_S A^{2/3} - a_C \frac{Z^2}{A^{1/3}}. \quad (5.1)$$

This is a version of the SEMI-EMPIRICAL MASS FORMULA, also known as the BETHE-WEIZSÄCKER MASS FORMULA. The coefficients  $a_V, a_A, a_S, a_C$  are found by fitting the formula to measured nuclear masses (Table 5.1).

This fit should have another term to account for the pairing of neutrons and protons, so that the binding energy is increased for even  $Z$  and  $N$ .

We omit that term here for simplicity.

Table 5.1: Coefficients for the fit to nuclear masses, (5.1), in units of MeV.

$a_V$	$a_A$	$a_S$	$a_C$
15.5	22.7	16.6	0.71

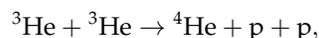
---

EXERCISE 5.3 —  For a given nuclear mass number  $A$ , derive an expression for the charge number  $Z_*(A)$  that maximizes the binding energy (eq. [5.1] with coefficients from Table 5.1).

1. Plot the ratio  $Z_*/A$  for  $4 \leq A \leq 128$ . Give a physical explanation for the behavior of  $Z_*/A$ .
  2. Plot the binding energy per nucleon  $B/A$  as a function of  $Z_*$  and  $A$ , for  $4 \leq A \leq 128$ .
  3. Find the atomic number  $Z$  and atomic mass  $A$  of the nucleus with the maximum  $B/A$ .
- 

## 5.2 Nuclear reactions

From mass-energy conservation, the heat evolved during a nuclear reaction equals the change in mass of the reacting system. For example, in the reaction



the binding energy of  ${}^3\text{He}$  is 7.718 MeV and that of  ${}^4\text{He}$  is 28.296 MeV; the heat evolved by this reaction is therefore

$$\begin{aligned} & 2 [2m_p + m_n - B({}^3\text{He})] - [2m_p + 2m_n - B({}^4\text{He})] - 2m_p \\ &= B({}^4\text{He}) - 2B({}^3\text{He}) \\ &= 28.296 \text{ MeV} - 15.437 \text{ MeV} \\ &= 12.859 \text{ MeV}. \end{aligned}$$

---

EXERCISE 5.4 — Fusion of hydrogen into helium entails converting 4 hydrogen atoms (including the 4 electrons) into 1 helium atom (2 protons, 2 neutrons, 2 electrons) with  $B = 28.296$  MeV. What is the heat evolved *per hydrogen atom*? Assume that the sun has been shining with its current luminosity over its life. What mass of hydrogen atoms would need to undergo fusion to supply this energy? How large is this mass relative to the total mass of the sun?

---

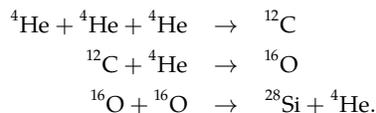
Table 5.2: Selected atomic mass excesses, taken from Tuli [2011].

isotope	$\Delta/\text{MeV}$
n	8.071
${}^1\text{H}$	7.289
${}^4\text{He}$	2.425
${}^{12}\text{C}$	0.000
${}^{16}\text{O}$	-4.737
${}^{28}\text{Si}$	-21.493
${}^{56}\text{Fe}$	-60.606

From the definition of the atomic mass unit  $m_u$ ,  $\Delta({}^{12}\text{C}) \equiv 0$ .

Because stellar reactions often involve electrons, it is convenient to define the ATOMIC MASS EXCESS  $\Delta(Z, A) = \mathcal{M}(Z, A) - Am_u$ , where  $\mathcal{M}$  is the atomic mass, including electrons. Some common mass excesses are listed in Table 5.2. Neglecting the electron binding energy ( $\sim$  eV), we can relate the atomic mass excess to the binding energy via  $\mathcal{M}(Z, A) = Am_u + \Delta(Z, A) = M(Z, N = A - Z) + Zm_e$ .

EXERCISE 5.5— Compute the energy released by the following reactions:



YOU MIGHT THINK THAT BECAUSE THE NUCLEAR INTERACTION IS SHORT-RANGE, THE CROSS-SECTION IS SOMETHING LIKE  $\pi r_n^2$ , WHERE  $r_n \approx (1 \text{ TO } 2) \text{ fm} \times A^{1/3}$ . Things are a bit more subtle, however, and in this section we shall explore how the reaction rate works. First, the “size” of a particle is in general proportional to the “size” of the wavefunction. From the uncertainty principle,

$$\pi \Delta x^2 \approx \pi \left( \frac{\hbar}{\Delta p} \right)^2 = \pi \frac{\hbar^2}{2mE}.$$

where we’ve taken  $\Delta p \sim p$ . Notice that if we multiply and divide by  $c^2$ , then we can estimate the area of the wavepacket as

$$\frac{(\hbar c)^2}{m_p c^2} \frac{1}{E} \sim 4 \times 10^4 \text{ fm}^2 \times \left( \frac{\text{keV}}{E} \right) = 400 \text{ b} \left( \frac{\text{keV}}{E} \right).$$

Here we’ve introduced a convenient unit for nuclear cross-sections, the BARN<sup>4</sup> (b), with  $1 \text{ b} = 10^{-28} \text{ m}^2 = 100 \text{ fm}^2$ .

The key point is that the size of the wave packet is  $\propto 1/E$ , which is in general true. This geometrical size of the wave packet is then multiplied by the probability of the nucleons forming a bound state, so we write the nuclear portion of the cross-section as

$$\sigma_{\text{nuclear}}(E) = \frac{S(E)}{E}. \quad (5.2)$$

The function  $S(E)$  contains the details of the nuclear interaction and is in general measured experimentally.

The final part of the cross-section concerns the Coulomb potential. Because protons repel one another, at large separations the nuclei interact *only* via the Coulomb potential. Consider the case of two nuclei with masses<sup>5</sup>  $A_1 m_u$  and  $A_2 m_u$ . Transform to the center-of-mass frame; the problem then reduces to that of one particle, mass  $A m_u = A_1 A_2 / (A_1 + A_2) \times m_u$ , moving in a potential (Fig. 5.2) that at large separations is purely Coulomb,

$$\frac{Z_1 Z_2 e^2}{4\pi\epsilon_0 r} = \frac{Z_1 Z_2 \alpha \hbar c}{r} = 1.44 \text{ MeV} \times Z_1 Z_2 \left( \frac{1 \text{ fm}}{r} \right).$$

Although the nuclear interaction forms a deep potential well (blue  $V_n$ , Fig. 5.2) at short distances, outside the nucleus the Coulomb potential (red  $V_C$ , Fig. 5.2) dominates.

<sup>4</sup> as in hitting the broad side of

<sup>5</sup> When doing kinematics, we shall make the approximation  $m \approx A m_u$ .

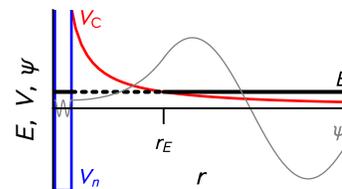


Figure 5.2: Tunneling through the Coulomb potential barrier. Not to scale.

---

EXERCISE 5.6— For the sun, typical center-of-mass energies are  $E \sim 1$  keV (horizontal black line in Fig. 5.2). Suppose we have two protons heading towards one another with this kinetic energy. What is their distance of closest approach?

---

As shown in Exercise 5.6, the turning radius  $r_E$  at typical stellar energies is much larger than the nuclear radius. Classically the particle can't penetrate the region  $r_n < r < r_E$  where  $E < V$  (dotted black line, Fig. 5.2); under classical physics, there would be no nuclear reactions at typical stellar temperatures because two particles would never find themselves close enough to be bound by the nuclear force.

The world is quantum, however, and the uncertainty in a particle's position means there is a small probability for the nucleons to be close enough for the nuclear force to come into play. In the classically forbidden region  $r_n < r < r_E$ , the particle wavefunction (thin gray line, Fig. 5.2) decreases exponentially, and the probability to reach  $r \sim 1$  fm is

$$\mathcal{P} \approx \exp \left[ -2\pi^2 \frac{r_E}{\lambda} \right]$$

where  $\lambda = h/p$  is the particle's wavelength and  $p$  is the momentum. It is convenient to rewrite the argument of the exponential in terms of the particle's energy,

$$\frac{2\pi^2 r_E}{\lambda} = 2\pi^2 \left( \frac{Z_1 Z_2 e^2}{E} \right) \left( \frac{p}{h} \right) = \left[ \pi \frac{Z_1 Z_2 e^2 \sqrt{2m}}{4\pi\epsilon_0 \hbar} \right] \left( \frac{1}{E} \right)^{1/2},$$

so that the probability of "tunneling" through the Coulomb barrier is

$$\mathcal{P} \approx \exp \left[ - \left( \frac{E_G}{E} \right)^{1/2} \right], \quad (5.3)$$

with

$$E_G \equiv \text{"Gamow Energy"} = \left[ \frac{2\pi^2 Z_1^2 Z_2^2 e^4 m}{(4\pi\epsilon_0)^2 \hbar^2} \right] = Z_1^2 Z_2^2 A \times 979 \text{ keV}.$$

Combining eqs. (5.2) and (5.3), we write the reaction cross-section as the nuclear cross-section multiplied by the probability of tunneling:

$$\sigma(E) = \frac{S(E)}{E} \exp \left[ - \left( \frac{E_G}{E} \right)^{1/2} \right]. \quad (5.4)$$

For many reactions  $S(E)$  is nearly constant over the range of typical energies in a stellar plasma. This is useful, as the reaction cross-section can be measured in the lab at higher energies and then extrapolated to the much lower stellar energies using eq. (5.4).

TO GET THE REACTION RATE FROM THE CROSS-SECTION, recall that the mean-free path of a particle is  $\ell = (n\sigma)^{-1}$ , where  $n$  is the density of targets. For definiteness, let us consider a plasma with two species present,

type 1 and type 2. The mean-free path of any given nucleus of type 1 against reactions with nuclei of type 2 is  $\ell = (n_2\sigma)^{-1}$ . If the nuclei are traveling with relative speed  $v = |\mathbf{v}_1 - \mathbf{v}_2|$ , then the mean time between collisions is  $\ell/v$ . Thus in a large ensemble of nuclei, the mean rate of reactions is

$$r_{12} = \frac{n_1 v}{\ell} = n_2 n_1 \langle \sigma v \rangle. \quad (5.5)$$

Here  $\langle \sigma v \rangle$  is the mean value of  $\sigma v$  for all pairs of nuclei in the plasma. For reactions between nuclei of the same type, we replace  $n_1 n_2$  with  $n^2/2$ ; the factor of 1/2 is to avoid double-counting.

A detailed calculation of the thermally averaged cross-section  $\langle \sigma v \rangle$  is presented in Box 5.1; here we'll just give a brief physical explanation for its value. There are two competing terms. First, the cross-section (eq. 5.4) has an exponential term  $\exp[-(E_G/E)^{1/2}]$  that increases rapidly with energy: more energetic particles have a much higher probability of tunneling through the Coulomb barrier. On the other hand, in thermal equilibrium the number of particles with energy  $E$  decreases as  $\exp(-E/k_B T)$ . As a result, reactions predominately occur in a narrow window of energies about a sort of geometric mean between  $E_G$  and  $k_B T$ :

$$E_{\text{pk}} = \frac{E_G^{1/3} (k_B T)^{2/3}}{4^{1/3}}.$$

The reaction rate is suppressed for  $E \ll E_{\text{pk}}$  because the probability of penetrating the Coulomb barrier is so small; the reaction rate is suppressed for  $E \gg E_{\text{pk}}$  because there are so few nuclei with those energies. Using this approximation the rate can be evaluated, with  $\langle \sigma v \rangle$  being given by eq. (5.10).

#### Box 5.1 The thermally averaged reaction cross-section

Since the cross-section depends on energy, the rate at which any given nucleus of type 1, traveling with velocity  $\mathbf{v}_1$ , will react with nuclei of type 2 having velocities  $\mathbf{v}_2$  in a range  $d^3 v_2$  is

$$n_2 \sigma |\mathbf{v}_1 - \mathbf{v}_2| \left( \frac{m_2}{2\pi k_B T} \right)^{3/2} \exp\left(-\frac{m_2 v_2^2}{2k_B T}\right) d^3 v_2.$$

The extra terms are because the nuclei have a Maxwell-Boltzmann distribution of velocities. To get the total rate per unit volume, we then have to multiply by the number of nuclei of type 1 having velocities  $\mathbf{v}_1$  in a range  $d^3 v_1$  and integrate over  $d^3 v_1 d^3 v_2$ :

$$r_{12} = n_1 n_2 \left[ \frac{m_1 m_2}{(2\pi k_B T)^2} \right]^{3/2} \times \int \sigma(E) v \exp\left(-\frac{m_1 v_1^2}{2k_B T} - \frac{m_2 v_2^2}{2k_B T}\right) d^3 v_1 d^3 v_2. \quad (5.6)$$

## Box 5.1 continued

Now  $E$  and  $v$  are the relative energies and velocity in the center-of-mass frame. We can change variable using the relations

$$\begin{aligned} \mathbf{v}_1 &= \mathbf{V} - \frac{m_2}{m_1 + m_2} \mathbf{v} \\ \mathbf{v}_2 &= \mathbf{V} + \frac{m_1}{m_1 + m_2} \mathbf{v}. \end{aligned}$$

where  $V$  is the center-of-mass velocity. It is straightforward to show that  $dv_{1,x} dv_{2,x} = dV_x dv_x$ , and likewise for the  $y, z$  directions. Furthermore,  $m_1 v_1^2 + m_2 v_2^2 = (m_1 + m_2)V^2 + mv^2$ , and multiplying and dividing the integral in equation (5.6) by  $m_1 + m_2$  allows us to write

$$\begin{aligned} r_{12} &= n_1 n_2 \left( \frac{m_1 + m_2}{2k_B T} \right)^{3/2} \left( \frac{m}{2k_B T} \right)^{3/2} \\ &\quad \times \int d^3 V \int d^3 v \sigma(E) v \exp \left[ -\frac{mv^2}{2k_B T} \right] \exp \left[ -\frac{(m_1 + m_2)V^2}{2k_B T} \right]. \end{aligned}$$

The integral over  $d^3 V$  can be factored out and is normalized to unity. Hence we have for the reaction rate between a pair of nuclei of types 1 and 2,

$$\begin{aligned} r_{12} &= n_1 n_2 \left\{ \left( \frac{m}{2\pi k_B T} \right)^{3/2} \int_0^\infty \sigma(E) v \exp \left( -\frac{mv^2}{2k_B T} \right) 4\pi v^2 dv \right\}. \\ &\equiv n_1 n_2 \langle \sigma v \rangle. \end{aligned} \quad (5.7)$$

The term in  $\{\}$  is the averaging over the joint distribution of the cross-section times the velocity, and is usually denoted as  $\langle \sigma v \rangle$ . Note that if nuclei 1 and 2 were identical, then we would need to divide  $r_{12}$  by 2.

Changing variables to  $E = mv^2/2$  in equation (5.7) and inserting the formula for the cross-section, equation (5.4), gives

$$\langle \sigma v \rangle = \left( \frac{8}{\pi m} \right)^{1/2} \left( \frac{1}{k_B T} \right)^{3/2} \int_0^\infty S(E) \exp \left[ -\left( \frac{E_G}{E} \right)^{1/2} - \frac{E}{k_B T} \right] dE. \quad (5.8)$$

Now, we've assumed that  $S(E)$  varies slowly; but look at the argument of the exponential. This is a competition between a rapidly rising term  $\exp[-(E_G/E)^{1/2}]$  and a rapidly falling term  $\exp(-E/k_B T)$ . As a result, the exponential will have a strong peak, and we can expand the integrand in a Taylor series about the maximum. Let

$$f(E) = -\left( \frac{E_G}{E} \right)^{1/2} - \frac{E}{k_B T}.$$

## Box 5.1 continued

Then we can write

$$\int_0^{\infty} S(E) \exp \left[ - \left( \frac{E_G}{E} \right)^{1/2} - \frac{E}{k_B T} \right] dE$$

$$\approx \int_0^{\infty} S(E_{pk}) \exp \left[ f(E_{pk}) + \frac{1}{2} \frac{d^2 f}{dE^2} \Big|_{E=E_{pk}} (E - E_{pk})^2 \right] dE.$$

Here  $E_{pk}$  is found by solving  $(df/dE)|_{E=E_{pk}} = 0$ . By expanding the argument of the exponential, we have approximated the integrand by a Gaussian,

$$\exp \left[ - \frac{(E - E_{pk})^2}{2\zeta^2} \right]$$

where

$$\frac{1}{\zeta^2} = - \frac{d^2 f}{dE^2} \Big|_{E=E_{pk}}.$$

This trick of approximating a steeply peaked function as a Gaussian is known as the METHOD OF STEEPEST DESCENT.

Solving for  $E_{pk}$ , we get

$$E_{pk} = \frac{E_G^{1/3} (k_B T)^{2/3}}{2^{2/3}},$$

and

$$\exp [f(E_{pk})] = \exp \left[ -3 \left( \frac{E_G}{4k_B T} \right)^{1/3} \right].$$

Further,

$$\frac{1}{2} \frac{d^2 f}{dE^2} \Big|_{E=E_{pk}} = - \frac{3}{2(2E_G)^{1/3} (k_B T)^{5/3}} = - \frac{3}{4E_{pk} k_B T}.$$

Defining a variable  $\Delta = 4(E_{pk} k_B T/3)^{1/2}$ , our integral becomes

$$\langle \sigma v \rangle = \left( \frac{8}{\pi m} \right)^{1/2} \left( \frac{1}{k_B T} \right)^{3/2} S(E_{pk})$$

$$\times \exp \left[ -3 \left( \frac{E_G}{4k_B T} \right)^{1/3} \right] \int_0^{\infty} \exp \left[ - \frac{(E - E_{pk})^2}{(\Delta/2)^2} \right] dE. \quad (5.9)$$

Another simplification can be made because both the Gaussian and the original integrand go to zero as  $E \rightarrow 0$ . As a result, we can extend the lower bound of our integral (eq. [5.9]) to  $-\infty$ , which allows us to evaluate the integral analytically and obtain

$$\langle \sigma v \rangle \approx \left( \frac{8}{m} \right)^{1/2} \left( \frac{1}{k_B T} \right)^{3/2} S(E_{pk}) \exp \left[ -3 \left( \frac{E_G}{4k_B T} \right)^{1/3} \right] \frac{\Delta}{2}$$

$$= \frac{2^{13/6} E_G^{1/6} S(E_{pk})}{\sqrt{3m} (k_B T)^{2/3}} \exp \left[ -3 \left( \frac{E_G}{4k_B T} \right)^{1/3} \right]. \quad (5.10)$$

The rate has the temperature dependence

$$r \propto T^{-2/3} \exp \left[ -3 \left( \frac{E_G}{4k_B T} \right)^{1/3} \right];$$

since  $E_G \propto Z_1^2 Z_2^2 A$ , at any given temperature lighter nuclei typically have much faster reaction rates. Also note that at stellar energies, reaction rates are incredibly sensitive to temperature. To quantify this, approximate the rate at a given temperature as a power-law,  $r(T) \propto T^n$ . Then the exponent is

$$n(T) = \frac{\partial \ln r}{\partial \ln T} = -\frac{2}{3} + \left( \frac{E_G}{4k_B T} \right)^{1/3}, \quad (5.11)$$

as you can verify for yourself (Exercise 5.7). Table 5.3 lists  $E_G$ ,  $E_{pk}$ , and  $n$  for some common reactions. In the table, the peak reaction energy  $E_{pk}$  and exponent  $n(T)$  are evaluated at  $T = 10^7$  K ( $k_B T = 0.86$  keV). Note the large value of  $n(T)$  at stellar temperatures—this is a consequence of the largeness of  $E_G/k_B T$ .

Table 5.3: Parameters for non-resonant reactions

	p + p	p + $^3\text{He}$	$^3\text{He} + ^3\text{He}$	p + $^7\text{Li}$	p + $^{12}\text{C}$
$E_G$ (MeV)	0.489	2.94	23.5	7.70	32.5
$E_{pk} _{T=10^7 \text{ K}}$ (keV)	4.5	8.2	16.3	11.3	18.2
$n(T = 10^7 \text{ K})$	4.6	8.8	18.3	12.4	20.5

EXERCISE 5.7 — Suppose we wish to approximate a function  $f(x)$  at a point  $x_0$  with a power-law,  $p(x; A, n) = Ax^n$ . Impose the condition  $p(x_0; A, n) = f(x_0)$  and  $dp/dx|_{x=x_0} = df/dx|_{x=x_0}$  to find the parameters  $A$  and  $n$ , and show that

$$n = \frac{d \ln f}{d \ln x}.$$

Apply this to the reaction rate, eq. (5.10), and thus derive eq. (5.11).

### 5.3 Stellar nuclear reactions

#### *Hydrogen burning via pp reactions: the weak nuclear interaction*

In the previous section, we established that lighter nuclei, because of their lower Coulomb repulsion, will tend to fuse at lower temperatures. Thus we expect that the first reaction that can occur is p + p and therein lies a problem: there is no bound state of  $^2\text{He}$ . The only possible way for two protons to fuse is for one of the protons to transmute into a neutron, giving the reaction



This reaction is possible because there are two nuclear forces: the strong and the weak. The strong is what binds nuclei together; the weak medi-

More precisely, what the strong force mainly does is bind quarks into neutrons or protons.

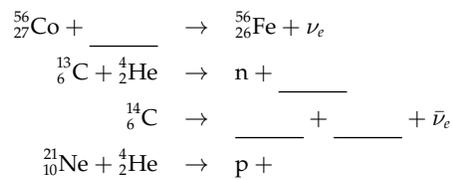
ates the conversion of a neutron into a proton (and vice versa). Two leptons are also involved (either emitted or absorbed) in this type of weak reaction: an electron (or its anti-particle, the positron) and an electron neutrino (or anti-neutrino). Three conservation laws determine which particles are involved:

1. the number of nucleons is conserved;
2. the charge is conserved; and
3. the number of leptons is conserved.

With regard to item 3, electrons ( $e^-$ ) and electron neutrinos ( $\nu_e$ ) have lepton number +1 while positrons ( $e^+$ ) and anti-electron neutrinos ( $\bar{\nu}_e$ ) have lepton number -1. Neutrinos, as the name implies, do not carry charge.

Applying these rules to the reaction (5.12), the number of nucleons on both sides of this reaction is the same, so rule 1 is satisfied. The positron on the right hand side balances charge to satisfy rule 2. Finally, the emission of an electron neutrino ensures that the lepton number on the right-hand side is zero to satisfy rule 3.

EXERCISE 5.8 — Complete the following reactions. Note that the symbol  ${}^A_Z\text{El}$  indicates that nuclide “El” has atomic charge number  $Z$  and atomic mass number  $A$ .



The weak cross section goes roughly as  $\sigma_{\text{weak}} \sim 10^{-20} \text{ b} (E/\text{keV})$ , so that

$$\frac{\sigma_{\text{weak}}}{\sigma_{\text{nuc}}} \sim 10^{-23} \left( \frac{E}{\text{keV}} \right).$$

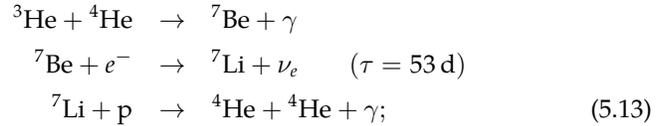
As a result the characteristic temperature for reaction (5.12) to occur is  $\approx 1.5 \times 10^7 \text{ K}$ , much higher than the temperature at which  $\text{p} + {}^2\text{H}$  occurs; at this temperature, the lifetime of a proton to forming deuterium via capture of another proton is about 6 Gyr. Once a deuterium nucleus is formed, it is immediately destroyed via  ${}^2\text{H} + \text{p} \rightarrow {}^3\text{He}$ . The reaction  $\text{p} + {}^3\text{He} \rightarrow {}^4\text{Li}$  cannot occur because  ${}^4\text{Li}$  is unbound and decays back into  $\text{p} + {}^3\text{He}$  with a lifetime of  $10^{-22} \text{ s}$ . The nucleus  ${}^6\text{Be}$  is likewise unbound ( $\tau \sim 5 \times 10^{-21} \text{ s}$ ), but it decays into two p and a  ${}^4\text{He}$  nucleus. As a result, the next reaction that occurs is  ${}^3\text{He} + {}^3\text{He} \rightarrow \text{p} + \text{p} + {}^4\text{He}$ . Despite having a

Because the weak cross section is so small, the first reaction that occurs in a contracting pre-main sequence star is  ${}^2\text{H} + \text{p} \rightarrow {}^3\text{He}$ ; in fact, this reaction can occur in objects as small as  $\approx 12 M_{\text{Jupiter}}$ . The small primordial abundance of deuterium prevents this reaction from doing anything more than slowing contraction slightly.

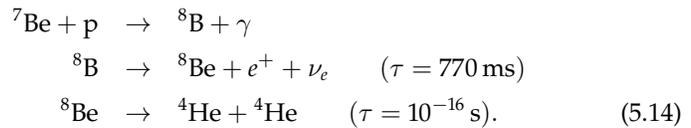
much greater Gamow energy than  $p + p$ , this reaction is still much faster than  $p + p$  owing to the small weak cross-section.

In addition to capturing another  ${}^3\text{He}$ , it is also possible for  ${}^3\text{He}$  to react with  ${}^4\text{He}$  and trigger the reactions

In eq. (5.13) and (5.14),  $\tau$  refers to the half-life for the nucleus on the left.



furthermore, at slightly higher temperatures  ${}^7\text{Be}$  can capture a proton instead of an electron to yield



The end result of these chains is the conversion of hydrogen to helium.

Although the conversion of hydrogen to helium involves many different reactions, the slowest step by far is the  $p + p$  reaction, so we can use this as a good approximation to the net reaction rate  $r$ . From eq. (5.5),  $r = n_p^2 \langle \sigma v \rangle_p / 2$ , where the factor of 2 avoids double counting. Since each  $pp$  reaction destroys 2 hydrogen, however, the rate at which hydrogen is depleted is  $2r = n_p^2 \langle \sigma v \rangle_{pp}$ . To get the net heating rate  $\varepsilon_{pp}$ , we multiply the rate at which hydrogen is destroyed by the heat released per hydrogen (see exercise 5.4).

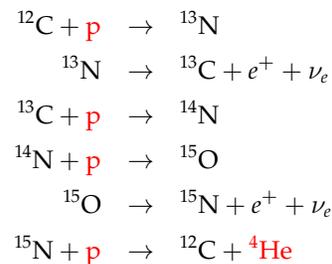
Because the solar interior is transparent to neutrinos,  $\varepsilon_{pp}$  is reduced slightly because the neutrinos carry away some energy. The amount of energy carried away by neutrinos differs from one chain to the next, so the net energy release varies slightly with temperature.

### *Hydrogen burning via the CNO cycle*

As we saw in the previous section, the smallness of the  $p + p$  cross-section means that proton captures onto heavier nuclei can occur at similar, or even faster rates, than  $p + p$  despite the larger Coulomb barrier. At  $T_6 = 10$ ,  $p + {}^{12}\text{C}$  has a comparable cross-section to  $p + p$ ; at  $T_6 = 20$ ,  $p + {}^{16}\text{O}$  has a comparable cross-section. Thus at temperatures slightly greater than that in the solar center, the following catalytic cycle<sup>6</sup>

$$T_6 \equiv T/10^6 \text{ K}$$

<sup>6</sup> known as the CNO CYCLE



becomes possible. The net result of this cycle (indicated by nuclei in red) is the ingestion of 4 protons and release of 1 helium. The reaction  $^{14}\text{N} + \text{p} \rightarrow ^{15}\text{O}$  is by far the slowest step in the cycle; as a result, all of the CNO elements in a sufficiently hot stellar core are quickly converted into  $^{14}\text{N}$ , and this reaction controls the net heating rate  $\varepsilon_{\text{CNO}}$ . At  $T = 2 \times 10^7$  K,  $\partial \ln \varepsilon_{\text{CNO}} / \partial \ln T = 18$ ; in contrast the  $\text{p} + \text{p}$  reaction has a temperature exponent of only 4.5. This rapid increase with temperature ensures that consumption of hydrogen via the CNO cycle is the dominant source of energy for stars more massive, and therefore having hotter cores, than the sun.

#### 5.4 The luminosity equation

Suppose we have a shell of mass  $\Delta m = 4\pi r^2 \rho \Delta r$  lying between surfaces  $r$  and  $r + \Delta r$  (Fig. 5.3). Nuclear reactions in the shell heat it at a rate  $\Delta m \times \varepsilon$ , where  $\varepsilon$  is the heating rate per unit mass. In addition, heat enters the shell from the bottom at a rate  $L(r)$  and leaves from the top at a rate  $-L(r + \Delta r)$ . If the shell is neither gaining or losing heat, then all these terms must balance:

$$4\pi r^2 \rho \varepsilon \Delta r + L(r) - L(r + \Delta r) = 0.$$

Dividing by  $\Delta r$  and taking the limit  $\Delta r \rightarrow 0$  produces our fourth equation of stellar structure,

$$\frac{dL}{dr} = 4\pi r^2 \rho \varepsilon.$$

At the center,  $L(r)_{r \rightarrow 0} \rightarrow 0$ , while at the surface  $L(r)_{r \rightarrow R} \rightarrow 4\pi R^2 \sigma_{\text{SB}} T_{\text{eff}}^4$ .

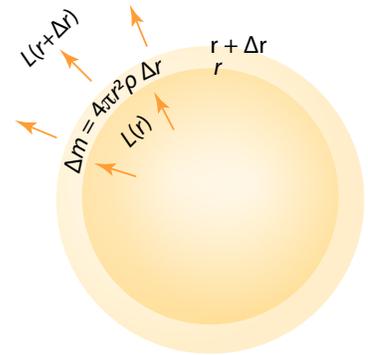


Figure 5.3: Heat balance in a shell  $\Delta m$ .



# 6

## Star

We now have almost all of the physics necessary to describe the structure of a star. We only need two additional items: we must consider whether the fluid can slowly circulate and thereby transport heat, and we must discuss how the equation of state deviates from that of a classical ideal gas at high densities and at high temperatures. The changes in the equation of state at high densities are important for low-mass stars and set the minimum stellar mass; the changes at high temperatures are important for high-mass stars and set the maximum stellar mass.

### 6.1 Convection

We've established that in the interior of the star a temperature gradient,

$$\frac{dT}{dr} = -\frac{3\rho\kappa_R}{4acT^3} \frac{L(r)}{4\pi r^2},$$

arises to transport heat outward (cf. eq. [3.15]). This gradient becomes steeper as we increase either the flux  $L/4\pi r^2$  or the mean opacity  $\kappa_R$ . There is a limit, however, to the magnitude of  $|dT/dr|$ : if the gradient is too steep, the warm fluid becomes buoyant relative to the cooler fluid above it and begins to rise. You are familiar with this phenomenon: picture a hot summer day. As the ground absorbs sunlight, it warms the air just above the ground. The warm air rises and forms updrafts. You have perhaps seen hawks circling as they are carried aloft by these updrafts. This circulation of fluid induced by a temperature gradient is known as CONVECTION.

You can do a home demonstration of convection. Brew tea, and pour the hot tea into a saucepan that is on an unlit burner. Use a straw to inject a layer of cold milk under the hot tea in the saucepan. The temperature difference between the tea and milk inhibits premature mixing. Light the burner, and watch for the development of convection—you will know it when you see it (Fig. 6.1).

Convection can also occur in stars, in regions of high flux and/or high opacity. During convection, the fluid velocities in question are typically

As a bonus, you can drink the results of this demonstration.

Figure 6.1: Onset of convection in a tea-milk mixture.



quite subsonic, so hydrostatic equilibrium abides. But the fluid motions do make an enormous difference to heat transport! Warm fluid is carried upward and cool fluid sinks. The net result is that heat is transported upward much faster than it would have been if only diffusion had been operating. This upward transport of heat modifies the temperature gradient. In this chapter, we shall derive the condition for the onset of convection, and the value of the temperature gradient in the presence of subsonic, efficient convection.

#### *The onset of convection*

To understand when convection starts, it helps to recall why a parcel of warm air rises. Recall Archimedes' law:

*The buoyant force on an object, either wholly or partially immersed in a fluid under a constant gravitational acceleration, equals the weight of the fluid it displaces.*

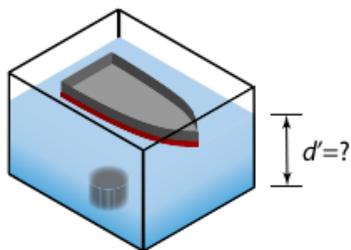
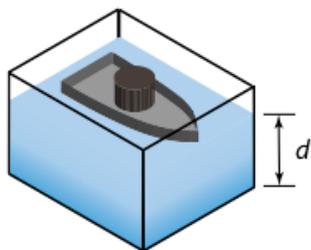


Figure 6.2: A boat with a weight in a tank.

What does this mean? When floating (buoyant force equals weight), a boat of mass  $m$  displaces (pushes aside) a volume  $v$  of water of density  $\rho_w$ . The weight of this displaced water,  $\rho_w v g$ , must equal the weight of the boat  $mg$ , so that  $v = m/\rho_w$ .

---

EXERCISE 6.1 — Suppose we have a toy boat carrying a weight and floating in a tank as shown in the top panel of Fig. 6.2. The depth of the water in the tank is  $d$ . The weight is then removed from the boat and allowed to sink to the bottom of the tank (bottom panel, Fig. 6.2). Does the depth of water in the tank increase, decrease, or stay the same? Explain your reasoning.

---

We can use Archimedes' law—which is just an application of hydrostatic equilibrium—to determine whether a fluid in planar geometry and hydrostatic equilibrium,

$$\frac{dP}{dr} = -\rho g, \quad (6.1)$$

and with a temperature gradient is unstable to convection. We start with a fluid at rest in hydrostatic equilibrium in spherical symmetry,

so  $P$ ,  $\rho$ , and  $T$  depend on radial position  $r$ . Imagine moving a blob of fluid upwards from  $r$  to  $r + h$ . We raise the blob slowly enough that it is in hydrostatic equilibrium with its new surroundings,<sup>1</sup>  $P_b(r + h) = P(r + h)$ . We do move the blob quickly enough, however, that it doesn't exchange heat with its surroundings and therefore doesn't remain in *thermal* equilibrium with its new environment.

As a result of this lack of heat exchange, the upward motion of the blob is **ADIABATIC**. To understand what this means, recall the first law of thermodynamics, which relates the change in internal energy  $dU$  and in volume  $dV$  to the heat transferred  $dQ = T dS$ :

$$dQ = T dS = dU + P dV, \tag{6.2}$$

where  $P$  is the pressure,  $T$  the temperature, and  $S$  the entropy. During an adiabatic process,  $dQ = T dS = 0$ . The entropy of the blob is therefore constant,  $S_b(r + h) = S_b(r) = S(r)$ , and is therefore not equal, in general, to the entropy of the surrounding gas at  $r + h$ :  $S_b(r + h) \neq S(r + h)$ .

As in our discussion of the equation of state (cf. eq. [2.2]), it isn't really convenient to write things in terms of volume. To put eq. (6.2) into a more convenient form, divide both sides by the mass of the blob  $m_b = \rho V$ :

$$\begin{aligned} d\left(\frac{Q}{m_b}\right) &= T d\left(\frac{S}{m_b}\right) = d\left(\frac{U}{m_b}\right) + P d\left(\frac{V}{m_b}\right) \\ T ds &= du + P d\left(\frac{1}{\rho}\right) \\ T ds &= du - \frac{P}{\rho^2} d\rho \end{aligned} \tag{6.3}$$

Here we denote the *entropy per mass* and the *energy per mass* by  $s$  and  $u$  respectively; and we identify the volume per mass with  $1/\rho$ , where  $\rho$  is the mass density. Equation (6.3) is the first law of thermodynamics as written for fluid dynamics.

The blob expands as it moves upward from  $r$  to  $r + h$ :

$$\rho_b(r + h) = \rho[P_b(r + h), s_b(r + h)] = \rho[P(r + h), s(r)].$$

Here we've written the density as a function of pressure and entropy,  $\rho(P, s)$ . Now we can apply Archimedes' law: if the density of the blob at  $r + h$  is greater than that of the surrounding fluid, then the buoyant force will be less than the weight of the blob; as a consequence, the blob will sink back to its original location. The fluid is thus **stable**. In contrast, if the density of the blob at  $r + h$  is less than that of the surrounding fluid, then the buoyant force is greater than the weight of the blob; as a result, the blob will continue to rise. In this case, the fluid is **unstable**: a small perturbation leads to the acceleration of the blob upwards. Figure 6.3 has a schematic of this criterion.

<sup>1</sup> We'll use the subscript  $b$  to denote properties of the blob; quantities without a subscript refer to the background fluid.

Pressure equilibrium in the blob is established over the time a sound wave takes to cross the blob. Thus, moving the blob slowly enough to maintain pressure equilibrium means that the motion is quite subsonic. Moving the blob quickly enough to prevent heat transport means (cf. exercise 3.10) that the blob is much larger than a mean free path so the time for photons to random walk across the blob is longer than time taken to raise the blob.

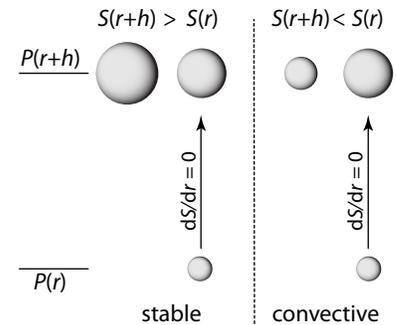


Figure 6.3: Illustration of criteria for convective instability. On the left, raising a blob a distance  $h$  adiabatically and in pressure balance with its surrounding results in a higher density  $\rho_b < \rho$ , or  $\rho_b > \rho$ . This is stable: the blob will sink back. On the right, the blob is less dense and hence buoyant: it will continue to rise.

Thus, for the fluid to be stable, we require that the density of the displaced blob be greater than that of the surrounding fluid:

$$\begin{aligned}\rho_b(r+h) &> \rho(r+h) \\ \rho[P(r+h), s(r)] &> \rho[P(r+h), s(r+h)].\end{aligned}\quad (6.4)$$

If condition (6.4) is satisfied, then the blob will be restored to its original location after a perturbation, and the system is stable. If condition (6.4) is not satisfied, then the blob will continue to rise following a perturbation, and the system is unstable.

We take  $h$  to be an infinitesimal displacement, so we can expand the right-hand side of eq. (6.4) about  $\rho[P(r+h), s(r)]$ :

$$\rho[P(r+h), s(r+h)] \approx \rho[P(r+h), s(r)] + \left(\frac{\partial\rho}{\partial s}\right)_P \frac{ds}{dr} h.$$

Here the notation  $(\partial\rho/\partial s)_P$  means taking the derivative of  $\rho$  with respect to  $s$  while holding  $P$  fixed. After canceling common factors, equation (6.4) therefore reduces to

$$\left(\frac{\partial\rho}{\partial s}\right)_P \frac{ds}{dr} h < 0. \quad (6.5)$$

We can put eq. (6.5) into a more useful form by changing variables from entropy  $\rho$  to temperature  $T$  via

$$\left(\frac{\partial\rho}{\partial T}\right)_P = \left(\frac{\partial\rho}{\partial s}\right)_P \left(\frac{\partial s}{\partial T}\right)_P = \left(\frac{\partial\rho}{\partial s}\right)_P \frac{C_P}{T},$$

where we introduce the specific heat at constant pressure,  $C_P \equiv T(ds/dT)_P$ . Inserting this expression into eq. (6.5) and dropping  $h$  (since it is positive by construction) gives

$$\frac{T}{C_P} \left(\frac{\partial\rho}{\partial T}\right)_P \frac{ds}{dr} < 0,$$

Now,  $(\partial\rho/\partial T)_P$  is negative (gas expands on being heated), while  $C_P$  is positive; hence eq. (6.5) will be satisfied wherever

$$\frac{ds}{dr} > 0. \quad (6.6)$$

*In a convectively stable star, the entropy must increase with radius.*

If this condition is not satisfied—if  $ds/dr < 0$ —then convection ensues: high-entropy material is buoyant and rises upward, while lower-entropy material sinks inward. Eventually this circulation mixes the rising high-entropy fluid with the sinking low-entropy fluid and drives the entropy gradient toward the marginally stable configuration  $ds/dr = 0$ .

### *The adiabatic thermal gradient*

Although succinct, the condition (6.6) for convective stability is not directly useful since our equations of stellar structure do not directly involve the entropy. We'd instead like to express the criterion for the onset

of convection in terms of pressure and temperature, since those quantities appear in our stellar structure equations.

### Box 6.1 Along an adiabat

To describe adiabatic processes, let's start from the first law expressed in terms of mass-specific quantities (eq. [6.3]):

$$dq = T ds = du - \frac{P}{\rho^2} d\rho.$$

We write the energy  $u$  as a function of temperature  $T$  and density  $\rho$ ,  $u = u(\rho, T)$ , and take the differential,

$$du = \left( \frac{\partial u}{\partial T} \right)_{\rho} dT + \left( \frac{\partial u}{\partial \rho} \right)_{T} d\rho.$$

We insert this expression for  $du$  into eq. (6.3) and find

$$dq = T ds = \left( \frac{\partial u}{\partial T} \right)_{\rho} dT + \left[ \left( \frac{\partial u}{\partial \rho} \right)_{T} - \frac{P}{\rho^2} \right] d\rho.$$

Hence the heat needed to raise the temperature of one kilogram of fluid while holding density fixed is

$$C_{\rho} \equiv T \left( \frac{\partial s}{\partial T} \right)_{\rho} = \left( \frac{\partial u}{\partial T} \right)_{\rho}. \quad (6.7)$$

For an ideal gas  $C_{\rho}$  is a constant and the internal energy  $u = u(T)$  is independent of density. In what follows, we'll specialize to the case of an ideal gas.

To express the first law in terms of  $T$  and  $P$ , rather than  $T$  and  $\rho$ , write eq. (6.3) as

$$T ds = C_{\rho} dT - \frac{P}{\rho} \frac{d\rho}{\rho}. \quad (6.8)$$

Then take the logarithm of the ideal gas equation of state,  $\ln(P) = \ln(\rho) + \ln(T) + \ln(k_B/\mu m_u)$ , and differentiate to find

$$\frac{dP}{P} = \frac{d\rho}{\rho} + \frac{dT}{T}.$$

Use this to eliminate  $d \ln \rho$  in eq. (6.8) to obtain an expression for the heat transferred as a function of temperature and pressure,

$$T ds = \left[ C_{\rho} + \frac{P}{\rho T} \right] dT - \frac{1}{\rho} dP = \left[ C_{\rho} + \frac{k_B}{\mu m_u} \right] dT - \frac{1}{\rho} dP. \quad (6.9)$$

The heat needed to raise the temperature of a mass of fluid while holding pressure fixed is therefore

$$C_P \equiv T \left( \frac{\partial s}{\partial T} \right)_P = C_{\rho} + \frac{k_B}{\mu m_u}. \quad (6.10)$$

## Box 6.1 continued

For a plasma of ions and electrons, the specific heats are

$$C_\rho = \frac{3}{2} \frac{k_B}{\mu m_u}, \quad (6.11)$$

$$C_P = \frac{5}{2} \frac{k_B}{\mu m_u}; \quad (6.12)$$

the ratio of these specific heats is

$$\gamma = \frac{C_P}{C_\rho} = \frac{5/2}{3/2} = \frac{5}{3}. \quad (6.13)$$

This value of  $\gamma$  is for an ideal, monatomic gas and does not hold universally.

DURING ADIABATIC MOTION, THERE IS NO HEAT EXCHANGE: hence, the entropy is constant and eq. (6.9) relates temperature and pressure:

$$T ds = dq = 0 = C_P dT - \frac{1}{\rho} dP. \quad (6.14)$$

Rearranging terms and using the ideal gas equation of state to eliminate  $1/\rho$ , we obtain

$$\begin{aligned} C_P dT &= \frac{k_B}{\mu m_u} \frac{T}{P} dP; \\ \frac{dT}{T} &= \frac{C_P - C_\rho}{C_P} \frac{dP}{P} \\ &= \frac{\gamma - 1}{\gamma} \frac{dP}{P}. \end{aligned} \quad (6.15)$$

Here we've used the ratio of specific heats  $\gamma$ , eq. (6.13), and the equation of  $C_P$ , eq. (6.10) for an ideal gas. Integrating eq. (6.15) gives us  $T(P)$  along an adiabat<sup>2</sup>:

$$T = T_0 \left( \frac{P}{P_0} \right)^{(\gamma-1)/\gamma}. \quad (6.16)$$

Here  $T_0$  and  $P_0$  are the temperature and pressure at the beginning of the adiabatic process.

---

EXERCISE 6.2— Use equation (6.16) and the equation of state of an ideal gas to derive a relation between temperature and density, as well as a relation between density and pressure, during an adiabatic process.

---

In a convectively stable star,  $ds/dr > 0$  (eq. 6.6)), and so

$$C_P \frac{dT}{dr} = \frac{1}{\rho} \frac{dP}{dr} + T \frac{ds}{dr} > \frac{1}{\rho} \frac{dP}{dr}. \quad (6.17)$$

<sup>2</sup> We used a similar relation in determining the sound speed (Box 2.1).

Now,  $dP/dr = -\rho g$  because the fluid is in hydrostatic equilibrium; and this is true even if convection is occurring, so long as the convective motions are quite subsonic. Since the pressure  $P$  decreases monotonically with radius, we can use the equation of hydrostatic balance to write  $dT/dr = (dT/dP)(dP/dr)$ . Inserting this into equation (6.17), changing the direction of the inequality since  $dP/dr < 0$ , and rearranging slightly gives

$$\frac{dT}{dP} < \frac{1}{\rho C_P}$$

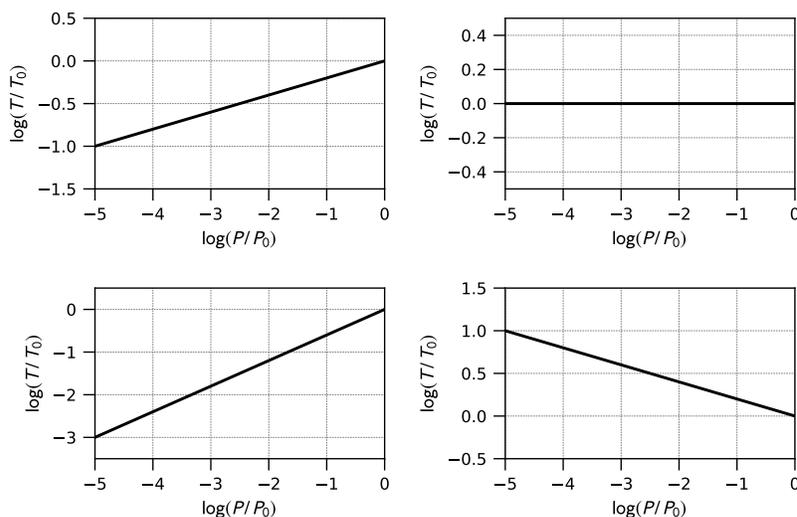
$$\frac{P}{T} \frac{dT}{dP} < \frac{P}{C_P \rho T} = \frac{\gamma - 1}{\gamma} = \frac{P}{T} \left( \frac{\partial T}{\partial P} \right)_s. \quad (6.18)$$

Although we derived eq. (6.18) for the case of an ideal gas, it is in general true. The left-hand side is the relation between temperature and pressure in the star, as set by radiative conduction; the right-hand side is the relation between temperature and pressure along an adiabat. The condition for stability, eq. (6.18), can be more succinctly written as

$$\nabla_{\text{rad}} \equiv \frac{d \ln T}{d \ln P} < \left( \frac{\partial \ln T}{\partial \ln P} \right)_s \equiv \nabla_{\text{ad}}. \quad (6.19)$$

for the fluid to be stable against convection.

EXERCISE 6.3 — The figure shows some hypothetical runs of temperature with respect to pressure in a gas in hydrostatic equilibrium. Indicate which of these situations is convectively unstable, and explain why. Draw on that plot the pressure-temperature relation that would ensue once convection sets in.



When convection is absent, the temperature gradient in the star is (eq. [3.15])

$$\frac{dT}{dr} = -\frac{3\rho\kappa}{4acT^3} \frac{L(r)}{4\pi r^2}.$$

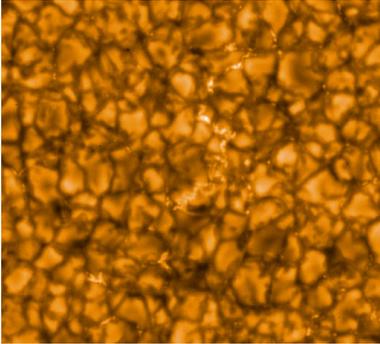
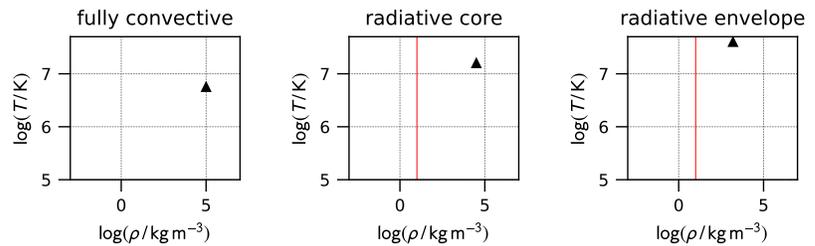


Figure 6.4: Solar convection cells, imaged with the Hinode Solar Optical Telescope. *Image credit:* Hinode JAXA/NASA/PPARC.

Here  $\kappa$  is the opacity and  $L(r)$  is the luminosity at radius  $r$ :  $L/4\pi r^2$  is the flux. If this thermal gradient,  $|dT/dr|$ , becomes too steep, however, then—just as in the pot of tea and milk on the stove, Fig. 6.1—the fluid becomes unstable: warm fluid begins to rise while cold fluid sinks. Over a wide range of stellar conditions this mixing drives the temperature gradient in the convectively unstable region to the adiabatic relations (eqs. [6.15]–[6.16]). As an example, the sun has a convective region in its outer region, Fig. 6.4.

**EXERCISE 6.4**— The figure below indicates the central density and temperature (*triangle*) for 3 hypothetical stars: (*left*) a star that is fully convective; (*center*) a star with a radiative (i.e., stable against convection) core (densities greater than  $10 \text{ kg m}^{-3}$ ) and a convective envelope; (*right*) a star with a convective core and a radiative envelope. For each star, sketch a plausible run of temperature with density within the star. In the center and right panels, the boundary between radiative and convective regions is marked with a vertical solid line.



WE CAN NOW COLLECT THE EQUATIONS DESCRIBING THE STRUCTURE OF A STAR IN STEADY-STATE. Previously, we established the relations for the enclosed mass,

$$\frac{dm}{dr} = 4\pi r^2 \rho, \quad (6.20)$$

the pressure,

$$\frac{dP}{dr} = -\rho \frac{Gm}{r^2}, \quad (6.21)$$

and the luminosity,

$$\frac{dL}{dr} = 4\pi r^2 \rho \varepsilon. \quad (6.22)$$

To these we add the equations for the temperature,

$$\frac{dT}{dr} = -\frac{L}{4\pi r^2} \frac{3\rho\kappa}{4acT^3} \quad \text{where radiative; and} \quad (6.23)$$

$$\frac{dT}{dr} \simeq \frac{T}{P} \left( \frac{\partial \ln T}{\partial \ln P} \right)_s \frac{dP}{dr} \quad \text{where convective.} \quad (6.24)$$

Equations (6.20)–(6.24), or equivalently (6.25)–(6.29), are supplemented by an equation of state  $P = P(\rho, T, \{X\})$ , opacity  $\kappa = \kappa(\rho, T, \{X\})$ , and heating rate  $\varepsilon = \varepsilon(\rho, T, \{X\})$ . Here  $\{X\}$  refers to the mass fractions of the various isotopes. Note that to have a full description we would also

need to include equations for the change in composition ( $dX/dt$ ) due to nuclear burning, as well as terms containing  $dr/dt$  to describe expansion or contraction.

### Box 6.2 The equations of stellar structure in Lagrangian form

In general, the equations (6.20), (6.21), (6.23)-(6.24), and (6.22) must be solved numerically. In practice, the radius  $r$  is not the most convenient variable to use as a coordinate. In one dimension, the mass in each shell remains distinct, so the enclosed mass

$$m(r) = \int_0^r 4\pi r^2 \rho \, dr$$

makes a useful coordinate. Using the enclosed mass as a coordinate is called a LAGRANGIAN description of the star. Upon changing variables from  $r$  to  $m$ , the structure equations become

$$\frac{dr}{dm} = \frac{1}{4\pi r^2 \rho} \quad (6.25)$$

$$\frac{dP}{dm} = \frac{dP}{dr} \frac{dr}{dm} = -\frac{Gm}{4\pi r^4} \quad (6.26)$$

$$\frac{dL}{dm} = \epsilon \quad (6.27)$$

$$\frac{dT}{dm} = -\frac{3\kappa L}{64\pi^2 r^4 a c T^3} \quad \text{where radiative} \quad (6.28)$$

$$\frac{dT}{dm} \simeq -\frac{T}{P} \left( \frac{\partial \ln T}{\partial \ln P} \right)_s \frac{Gm}{4\pi r^4} \quad \text{where convective.} \quad (6.29)$$

## 6.2 Contraction to the main sequence

Stars are formed when clouds of gas and dust fall out of pressure balance and become unstable to gravitational collapse. Often, the cloud fragments into a myriad of small collapsing regions, such as in the Soul and Taurus Nebulae pictured in Fig. 6.5. In the center of these dense knots, a core comes into hydrostatic equilibrium and grows in mass as matter continues to infall. Much of this process is obscured from view by the surrounding clouds of gas and dust.

While the nebula thins out, the pre-main sequence star contracts slowly on a Kelvin-Helmholtz timescale, eq. (2.26), since the core is too cool for nuclear reactions to generate sufficient power. As the central temperature rises, the nuclear reaction rate increases rapidly until the heat released by reactions balances that radiated from the surface. At that point the star joins the ZERO-AGE MAIN SEQUENCE (ZAMS). Of course, not all collapsing stellar-like objects reach the ZAMS—objects that are too low in mass will not ignite hydrogen fusion, while objects that are too



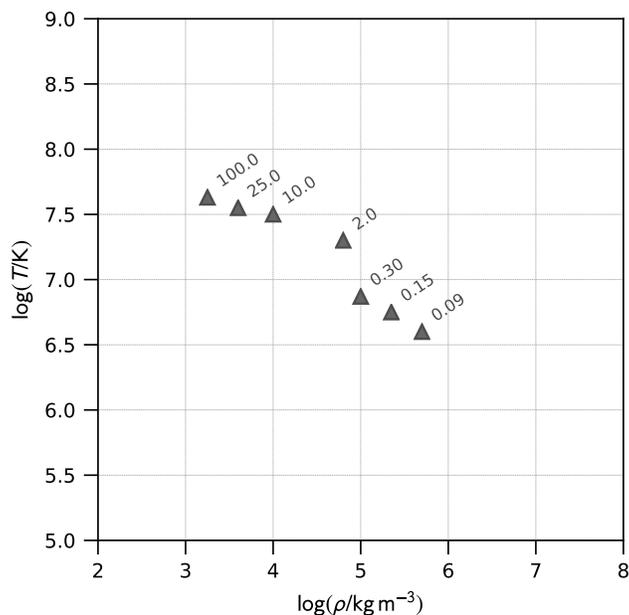
Figure 6.5: (Top) Image of the Soul Nebula (IC 1848) in the constellation Cassiopeia. Image credit: José Jiménez Priego (Astromet). (Bottom) The Taurus molecular cloud. At right are million-year old pre-main sequence stars still embedded in dust clouds. Image credit & copyright: Lloyd L. Smith, Deep Sky West.

high in mass tend to be unstable and eject mass. We'll now explore these limits in turn.

**EXERCISE 6.5**— This exercise revisits problem 2.10. In that exercise you modeled how the density and temperature changed as a pre-main-sequence star contracted. Table 6.1 gives central densities and temperatures of stars at the onset of hydrogen fusion (known as the ZERO-AGE MAIN SEQUENCE). These central temperatures and densities are plotted below and labeled by stellar mass. Assume an ideal-gas equation of state and use the virial relations for the temperature and central density to plot the tracks in this plane each star followed during its contraction.

Table 6.1: Selected central densities and temperatures of zero-age main-sequence stars, computed with the MESA stellar evolution code [Paxton et al., 2011].

$M/M_{\odot}$	$\log(\rho_c/\text{kg m}^{-3})$	$\log(T_c/\text{K})$
0.09	5.70	6.60
0.15	5.35	6.75
0.30	5.00	6.87
2.0	4.80	7.30
10.0	4.00	7.50
25.0	3.60	7.55
100.0	3.25	7.63



You will use this plot for exercises 6.7 and 6.10 as well.

### Degeneracy

As a star contracts, the particles within it are packed ever closer together. As we saw from our discussion of ionization, quantum mechanics enters the description of particle behavior when the separation between particles is of the order of the uncertainty in their positions. Said differently, our classical description breaks down when the particle density exceeds roughly

$$n > \frac{1 \text{ particle}}{(\Delta x)^3} = \left(\frac{\Delta p}{h}\right)^3 \approx \left(\frac{mk_B T}{h^2}\right)^{3/2}. \quad (6.30)$$

Another way to put this is that quantum effects become important when there is more than roughly 1 particle in a normalized phase space volume

$$d^3x d^3p/h^3.$$

If the particles are this densely packed, their wavefunctions overlap and we must describe the system as a single quantum state. Suppose we have two identical particles in a quantum state. Since the particles are identical, if we exchange them the wavefunction can only change by a phase factor<sup>3</sup>  $e^{i\delta}$ . If we exchange the particles again, we are back to our original state; as a result,  $e^{2i\delta} = 1$ , and therefore  $\delta = 0$  or  $\pi$ . Hence upon the exchange of particles, the wavefunction either is unchanged ( $\delta = 0, e^{i\delta} = 1$ ) or it changes sign ( $\delta = \pi, e^{i\pi} = -1$ ).

<sup>3</sup> See Box 6.3

*There are two types of wavefunctions in this world: those that change sign under exchange; and those that don't.*

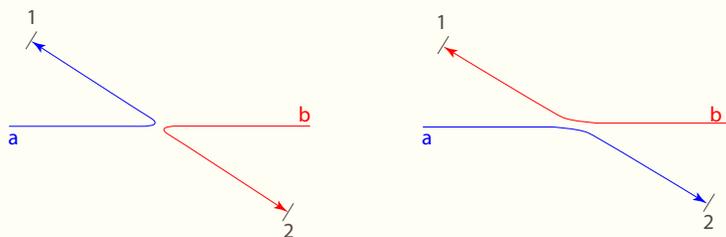
Particles whose wavefunctions don't change sign under exchange are called **BOSONS** and have integer spin. Photons (spin = 1) are bosons. Particles whose wavefunctions change sign under exchange are called **FERMIONS** and have half-integer spin. Electrons, neutrinos, protons, and neutrons (spin = 1/2) are all fermions.

A consequence of the fermion wavefunction changing sign when any two particles are exchanged is that the wavefunction vanishes if any two particles are in the same state—that is, they have the same position, momentum, and spin. For spin-half particles like electrons, this means we can put at most two such electrons in the same position and momentum state; we do this by having their spins antiparallel.

**Box 6.3 Identical particles**

To understand how the interchange of identical particles works in more detail, let's start by recalling some features of quantum mechanics. This discussion is based on Feynman et al. [1989]. We denote a particle's state as  $|a\rangle$ , where  $a$  is just a label. For example,  $a$  could be "electron with such-and-such properties". The probability of finding the electron in some other state  $|\phi\rangle$  is given by  $|\langle\phi|a\rangle|^2$ , where  $\langle\phi|a\rangle$  is a complex number known as the probability amplitude formed via an inner product of  $|\phi\rangle$  and  $|a\rangle$ .

Now suppose we have two particles,  $a$  and  $b$ , and we scatter them so that one particle ends up in detector 1 and the other ends up in detector 2. There are two ways this can go, as shown here.



## Box 6.3 continued

Classically, we would argue that the probability of getting either particle in detector 1 is just

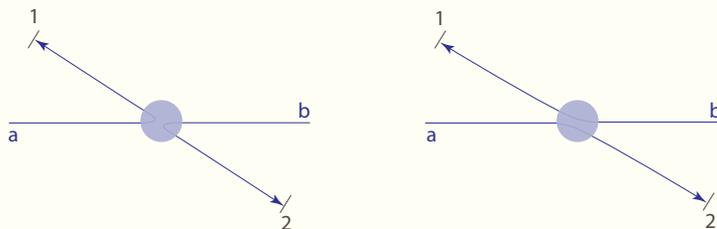
$$\mathcal{P}(\text{a or b in 1}) = \mathcal{P}(\text{a in 1}) + \mathcal{P}(\text{b in 1}). \quad (6.31)$$

If particles a and b are different—e.g., one is a  $^{12}\text{C}$  nucleus and the other is an  $^{16}\text{O}$  nucleus—then this holds in quantum mechanics as well. Quantum mechanically, we write

$$\mathcal{P}(\text{a or b in 1}) = |\langle 1|a\rangle\langle 2|b\rangle|^2 + |\langle 2|a\rangle\langle 1|b\rangle|^2. \quad (6.32)$$

If the particles are identical, however—for example, if a and b are two electrons with identical spin—then this picture is wrong.

Because of the uncertainty principle, we cannot follow the trajectories of a and b with infinite precision to see which is which; instead, the situation is more analogous to the depiction shown here.



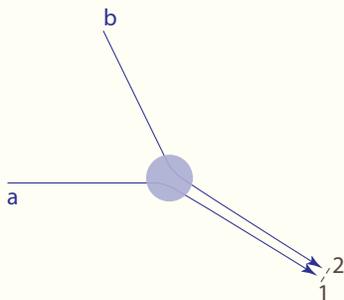
There are now two indistinguishable ways of arriving at the final state—in this case, an electron in detector 1 and an electron in detector 2. According to quantum mechanics, we must therefore sum the amplitudes for getting to the final state, *before taking the square*. That is, the probability for one particle to end up in detector 1 and the other to end up in detector 2 is

$$\begin{aligned} \mathcal{P}(\text{a or b in 1}) &= |\langle 1|a\rangle\langle 2|b\rangle + \langle 2|a\rangle\langle 1|b\rangle|^2 & (6.33) \\ &= |\langle 1|a\rangle\langle 2|b\rangle|^2 + |\langle 2|a\rangle\langle 1|b\rangle|^2 \\ &\quad + [\langle 1|a\rangle^* \langle 2|b\rangle^* \langle 2|a\rangle\langle 1|b\rangle + \langle 2|a\rangle^* \langle 1|b\rangle^* \langle 1|a\rangle\langle 2|b\rangle]. \end{aligned}$$

The probability of scattering one electron into detector 1 and the other into detector 2 is the value for distinguishable particles, eq. (6.32), *plus* an interference term in [·].

TO SEE THE EFFECT OF THIS INTERFERENCE TERM ON THE THERMAL PROPERTIES OF THE SYSTEM, let's imagine putting two particles into the same small volume. To do this, we imagine the detectors 1 and 2 sliding together until they overlap, as shown here.

## Box 6.3 continued



Since detectors 1 and 2 are approaching one another, we must have

$$|\langle 1|a\rangle\langle 2|b\rangle|^2 = |\langle 2|a\rangle\langle 1|b\rangle|^2. \quad (6.34)$$

For short, let's call  $|\langle 1|a\rangle\langle 2|b\rangle|^2 \equiv |ab|^2$ . For distinguishable particles, we add probabilities, so from eq. (6.32),

$$\mathcal{P}(\text{a or b in 1}) = |\langle 1|a\rangle\langle 2|b\rangle|^2 + |\langle 2|a\rangle\langle 1|b\rangle|^2 = 2|ab|^2.$$

For indistinguishable particles, however, the amplitudes can differ by a phase factor,  $\langle 2|a\rangle\langle 1|b\rangle = e^{i\delta}\langle 1|a\rangle\langle 2|b\rangle$ . We argued that since interchanging the particles twice brings us back to the original state, we must have  $e^{i\delta} = \pm 1$ .

If  $e^{i\delta} = 1$ , so that  $\langle 2|a\rangle\langle 1|b\rangle = \langle 1|a\rangle\langle 2|b\rangle$ , then from equation (6.33)

$$\mathcal{P}(\text{a or b in 1}) = |\langle 1|a\rangle\langle 2|b\rangle + \langle 2|a\rangle\langle 1|b\rangle|^2 = 4|ab|^2. \quad (6.35)$$

This is *twice* the classical value: the probability of the particles entering the same state is *enhanced*.

In contrast, if  $e^{i\delta} = -1$  and  $\langle 2|a\rangle\langle 1|b\rangle = -\langle 1|a\rangle\langle 2|b\rangle$ , then equation (6.33) implies that

$$\mathcal{P}(\text{a or b in 1}) = |\langle 1|a\rangle\langle 2|b\rangle + \langle 2|a\rangle\langle 1|b\rangle|^2 = 0 \quad (6.36)$$

*We cannot put 2 identical particles into a state with with the same momentum, position, and spin if their wavefunction changes sign when the particles are exchanged.*

Particles with integer spin (i.e., their angular momentum is an integer multiple of  $\hbar$ ) have wavefunctions that do not change sign under exchange; these particles are said to obey BOSE-EINSTEIN STATISTICS and are called **BOSONS**. Particles with half-integer spin have wavefunctions that do change sign under exchange; these particles are said to obey FERMI-DIRAC STATISTICS and are called **FERMIONS**. Photons are bosons; electrons, protons, neutrons, and neutrinos are fermions.

The denominator  $h^3$  can be derived from the density of states for a free particle in a box with periodic boundary conditions.

TO ACCOUNT FOR FERMI-DIRAC STATISTICS WITHIN THE EQUATION OF STATE, imagine a small volume containing  $N$  electrons. Motivated by eq. (6.30), divide the phase space into cells,

$$\frac{d^3x d^3p}{h^3},$$

and into each cell place 2 electrons with opposing spins. We always add the electrons to the lowest open energy level, and repeat the process until we have added all  $N$  electrons. This procedure is represented by the equation

$$N = \frac{2}{h^3} \int_V d^3x \int_0^{E_F} d^3p \quad (6.37)$$

In this equation  $E_F$ , the *Fermi energy*, is the energy of the last electron added and is the largest filled energy level.

If our volume is isotropic, then we can change variables: first, to spherical momentum coordinates,  $d^3p = 4\pi p^2 dp$ ; second, from  $dp$  to  $d\varepsilon$ . Since  $p = \sqrt{2m\varepsilon}$ , where  $\varepsilon$  is the energy of a single electron,

$$dp = \sqrt{\frac{m}{2\varepsilon}} d\varepsilon;$$

upon changing variables and integrating over  $\varepsilon$  from 0 to  $E_F$  we obtain

$$N = \frac{8\pi}{h^3} V \int_0^{E_F} \sqrt{2m}^{3/2} \varepsilon^{1/2} d\varepsilon = \frac{8\pi}{3h^3} V (2m)^{3/2} E_F^{3/2}.$$

Solving for the Fermi energy gives

$$E_F = \frac{h^2}{2m} \left( \frac{3N}{8\pi V} \right)^{2/3}. \quad (6.38)$$

What is the total energy of our system? We again integrate over phase space, with each electron multiplied by its energy  $\varepsilon$ :

$$E = \frac{8\pi}{h^3} V \int_0^{E_F} \sqrt{2m}^{3/2} \varepsilon^{3/2} d\varepsilon = \frac{8\pi}{5h^3} V (2m)^{3/2} E_F^{5/2}. \quad (6.39)$$

Using eq. (6.38) to substitute for  $E_F$  in eq. (6.39), we can find the energy per unit volume,

$$\frac{E}{V} = \frac{3}{5} \left( \frac{3}{8\pi} \right)^{2/3} \frac{h^2}{2m} n^{5/3} = \frac{3}{5} n E_F,$$

where  $n = N/V$  is the density of electrons.

For a non-relativistic gas the pressure is  $P = (2/3)(E/V)$ . Hence the pressure of our electron gas is

$$P = \frac{2}{3} \frac{E}{V} = \frac{2}{5} n E_F = \frac{2}{5} \left( \frac{3}{8\pi} \right)^{2/3} \frac{h^2}{2m} n^{5/3}. \quad (6.40)$$

Notice that the pressure is independent of the temperature.

In a stellar plasma, there is a mix of electrons and nuclei. The density at which a gas becomes degenerate, eq. (6.30), depends on the mass of the particles in question. Since electrons are more than 1000 times lighter than nuclei, they become degenerate at a lower density than nuclei. For most cases, we can treat the nuclei in the stellar plasma as a classical gas. As the electrons become degenerate, their pressure increases over the classical value, and once the star is degenerate, we can treat the pressure as being provided solely by the electrons to a good approximation. It's useful to express eq. (6.40) in terms of the mass density  $\rho$ . To do this, suppose our composition consists of isotopes with charge  $Z_i$  and mass number  $A_i$ . Then the number of electrons per unit volume<sup>4</sup> is

<sup>4</sup> assuming complete ionization

$$n_e = \sum_i n_i Z_i = \frac{\rho}{m_u} \sum_i X_i \frac{Z_i}{A_i}.$$

By analogy with the mean molecular weight, we define an electron mean weight

$$\mu_e \equiv \left( \sum_i X_i \frac{Z_i}{A_i} \right)^{-1} \quad (6.41)$$

so that  $n_e = \rho / (m_u \mu_e)$ . Equation (6.40) then provides a relationship between pressure and density for a given  $\mu_e$ .

EXERCISE 6.6 — Express the pressure as a function of mass density  $\rho$  and  $\mu_e$ . Then use the virial scalings for  $P(M, R)$  and  $\rho(M, R)$  to obtain a relation  $R(M)$  for a degenerate object.

As you found in exercise 6.6, when the star becomes degenerate, there is a unique radius for a given mass and composition. This is in contrast to the non-degenerate case, for which a star of a given mass can have a wide range of possible radii depending on the internal temperature.

Consider a contracting pre-main-sequence star. Initially, the star has a low density and the equation of state is that of an ideal non-degenerate gas. According to the virial theorem, as the radius decreases, both the central temperature and density increase. The radius decreases because the star is radiating away energy, and a star with an ideal, non-degenerate equation of state has a total energy that depends on its radius.

At some density, the equation of state will become degenerate. At this point, contraction comes to a halt. The star continues to radiate energy, but instead of contracting, the star simply cools while remaining at constant radius. If the contracting pre-main-sequence star is to become a main-sequence star, then, it must reach temperatures sufficient for hydrogen fusion to occur *before* becoming degenerate.

<sup>5</sup>J. D. Kirkpatrick, I. N. Reid, J. Liebert, et al. Dwarfs Cooler than “M”: The Definition of Spectral Type “L” Using Discoveries from the 2 Micron All-Sky Survey (2MASS). *ApJ*, 519:802–833, July 1999; and Michael C. Cushing, J. Davy Kirkpatrick, Christopher R. Gelino, et al. The Discovery of Y Dwarfs using Data from the Wide-field Infrared Survey Explorer (WISE). *ApJ*, 743:50, December 2011. DOI: 10.1088/0004-637X/743/1/50

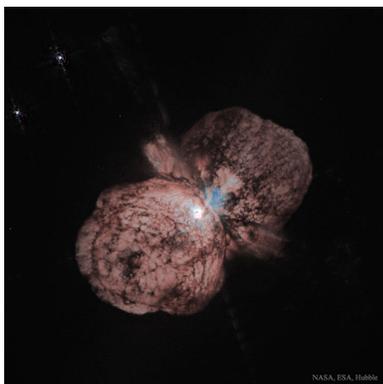


Figure 6.6: Image of the massive star Eta Carinae. Image credit: J. Morse (Arizona State U.), K. Davidson (U. Minnesota) et al., WFPC2, HST, NASA.

---

EXERCISE 6.7— The equation of state becomes degenerate roughly where  $k_B T = E_F$ , with  $E_F$  given by eq. (6.38). From this and eq. (6.41), assuming a H-He composition with  $X_H = 0.7$  and  $X_{He} = 0.3$ , derive a relation between  $\log(T)$  and  $\log(\rho)$ . Plot this relation on the phase diagram in exercise 6.5, and on the plot indicate which side of the relation is degenerate. Given that contraction halts when the equation of state becomes degenerate, what does this plot imply for the minimum mass required to initiate hydrogen fusion?

---

AS SHOWN IN EXERCISE 6.7, THERE IS A MINIMUM MASS NEEDED TO INITIATE HYDROGEN FUSION. Contracting star-like objects of lower mass are known as BROWN DWARFS. Although dim, they are observable with spectral types “L”, “T” or “Y”<sup>5</sup>.

---

EXERCISE 6.8— You might notice that the degenerate mass-radius relation you found in exercise 6.6 can’t hold for very light objects (or very heavy ones, for that matter). Earth, for example has a much larger mass than Mars, and also has a larger radius, contrary to what the degenerate relation predicts. What happens is that at low pressures, the Coulomb force comes into play—the atomic and molecular bonds that add variety to life. These bonds set the size and spacing of atoms, and therefore fix the density of matter. Let’s model this. The typical size of an atom is the Bohr radius,

$$a_B = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = 5.29 \times 10^{-11} \text{ m.}$$

1. Assume a solar composition,  $X_H = 0.7$ ,  $X_{He} = 0.3$  with one average nuclear mass per volume  $a_B^3$ . What is the density? Is it plausible?
  2. For an object with this density, what is  $R(M)$ ?
  3. You should find from part 2 that when the Coulomb force is important, more massive objects have *larger* radii, unlike the case from exercise 6.6 when degeneracy dominates. At what mass do these two relations meet? This sets the mass scale at which degeneracy becomes important for a cold object. Compare the planetary masses in our solar system with this mass scale; are there any that are near or above it?
- 

### Radiation pressure

Radiation in thermal equilibrium exerts a pressure (eq. 1.11):  $P_{\text{rad}} = aT^4/3$ . Because of its strong dependence on temperature, radiation exerts an increasingly large fraction of the total pressure for massive stars. Stars that are radiation-pressure dominated tend to be unstable: they have strong winds and violent fits of mass ejection (see the image of Eta Carinae, Fig. 6.6). As a result, they lose copious amounts of mass while on the main sequence. This mass-shedding effectively sets a rough upper limit on the range of main-sequence stellar masses.

---

EXERCISE 6.9— Use the virial relations for density and temperature to estimate how the ratio  $P_{\text{rad}}/P_{\text{gas}}$  depends on the mass of the star.

---



---

EXERCISE 6.10— The equation of state becomes dominated by radiation roughly where  $P(\text{ideal gas}) \approx P(\text{radiation})$ . Derive from this criterion a relation between  $\log(T)$  and  $\log(\rho)$ , and plot this relation on the figure for exercise 6.5. Indicate which side of this relation is radiation-pressure dominated. What do your findings in this exercise imply for the mass range of main-sequence stars?

---

### 6.3 Life on the main-sequence

With the initiation of hydrogen fusion, the star settles into thermal and mechanical equilibrium, with its structure described by the solution of equations<sup>6</sup> (6.20)–(6.22), along with the equation of state and prescriptions for the opacity  $\kappa$  and heating rate  $\varepsilon$ .

The reason for star’s stability on the main sequence is a consequence of the relation, derived in exercise 2.8, between the star’s total energy and temperature. If the reaction rate were to increase and deposit more energy into the star, then since the total energy is  $\propto -GM^2/R$ , the star would expand. This expansion would cause the central temperature to decrease, thereby reducing the reaction rate.

The star is not in complete equilibrium, however, as hydrogen in the core is gradually being converted to helium. The timescale over which the composition changes is much longer than the dynamical timescale (sets hydrostatic equilibrium), the radiative diffusion timescale (sets thermal gradient), and the Kelvin-Helmholtz timescale (sets core temperature via growth or contraction of stellar radii). The gradual build-up of a helium-rich core does not, therefore, affect the stability of the star, but it does lead to a slow brightening of the star over its main sequence lifetime. For our sun, the gradual enrichment of the core in helium causes a slow increase in luminosity of  $\approx 10\%$  for each billion years.

---

EXERCISE 6.11— You computed in exercise 5.4 the energy released from the conversion of 4 hydrogen atoms into helium. Express this number in terms of the energy released per mass of hydrogen burned; this number should be in units of J/kg. Now assume that the sun’s luminosity comes from the fusion of hydrogen into helium in the innermost 10% of the sun’s mass. For a composition that is 70% hydrogen by mass, how long would it take to deplete the hydrogen in the solar core? This sets the main-sequence lifetime of the sun.

---

The cool outer layers of low-mass stars have large opacities: for example many elements are not ionized, so there are many potential lines for absorption. As a result, stars with  $M \lesssim M_{\odot}$  have convective regions in

<sup>6</sup> Or, in Lagrangian form, (6.25)–(6.27).

Although this slow increase in luminosity is not a drastic change, it has significant implications for life on Earth. The expected warming is sufficient to make Earth uninhabitable within about a billion years from now.

their outer parts. The fraction of the star that is convective is larger for low-mass, cool stars; and stars with  $M \lesssim 0.3 M_\odot$  are fully convective, so that the whole interior lies along an adiabat. For more massive, hotter, stars, the opacities are lower, and as a result, the outer convective region vanishes for stars with  $M \gtrsim M_\odot$ .

---

EXERCISE 6.12— We can estimate how the luminosity depends on stellar mass for stars that have a mostly radiative structure. Start with equation (6.23) for the temperature gradient and approximate  $dT/dr \approx -T_c/R$ ,  $\rho \approx \bar{\rho}$ ,  $L/4\pi r^2 \approx L/4\pi R^2$ , and  $T \approx T_c$ . Take the opacity  $\kappa$  to be constant, use the virial estimate for the central temperature  $T_c$  and express the mean density  $\bar{\rho}$  in terms of stellar mass  $M$  and radius  $R$ . After some algebra, you should find that the luminosity  $L$  depends on  $M$  to some power. Compare this scaling against the data in Table 2.2. Obtain an expression for the stellar lifetime as a function of mass, and calibrate it to the sun's main-sequence lifetime,  $\tau_\odot \sim 10$  Gyr.

---

STARS MORE MASSIVE THAN THE SUN HAVE SUFFICIENTLY HIGH CORE TEMPERATURES FOR HYDROGEN TO BE CONSUMED VIA THE CNO CYCLE. The strong temperature dependence of the CNO burning has two effects on the structure of the star. First, it makes the central temperature nearly constant over a wide range of stellar masses for  $M > 1 M_\odot$ —a small rise in temperature is sufficient to raise the heat production  $\varepsilon$  to match the rise in luminosity. A nearly constant central temperature implies, via the virial theorem, that  $R \propto M$  on the upper main sequence. The second consequence is that nearly all of the star's luminosity is generated in a small region about the stellar center. The flux,  $L/4\pi r^2$ , in this small region is enormous, and this makes the core of the star convective. The convection can mix hydrogen fuel into the core, which makes the lifetime somewhat longer than the estimate from exercise 6.12. A summary of the structure of main sequence stars is contained in Table 6.2.

Table 6.2: Characteristics of main-sequence stars

	$M \lesssim M_\odot$	$M \gtrsim M_\odot$
$4p \rightarrow {}^4\text{He} \dots$	pp	CNO
core is...	radiative	convective
envelope is...	convective	radiative

# 7

## *End of the Line*

The depletion of hydrogen in the core heralds the end of the star's placid main-sequence life. We shall first give an overview of the changes that ensue. Fusion of helium requires a temperature  $\gtrsim 10^8$  K, substantially higher than that required for the fusion of hydrogen. As a result, when the hydrogen is used up helium burning cannot immediately begin, and the core contracts, similar to what happened before the star was born, with one crucial difference. As the helium core contracts, hydrogen continues to fuse in a shell surrounding the core. This shell burning intensifies as the core contracts and causes drastic changes to the star's radius, surface temperature, and luminosity.

Once the core becomes sufficiently hot, helium fuses into carbon, and the core again reaches a state of thermal and mechanical equilibrium. After a brief helium-burning phase, the core becomes depleted in helium and must again contract. As with pre-main sequence stars, the critical question is whether the core becomes degenerate before the next fusion reaction can ignite. For stars with main-sequence masses  $\lesssim (8-10) M_{\odot}$ , the core becomes degenerate before the onset of  $^{12}\text{C}$  fusion, which requires temperatures  $\approx 8 \times 10^8$  K. Indeed, for stars around a solar mass, the fusion of  $^4\text{He}$  occurs under moderately degenerate conditions.<sup>1</sup> As a result, the cores of low-mass stars end up composed of carbon and oxygen (or perhaps oxygen and neon) and supported by degenerate electrons; such objects are known as WHITE DWARFS.

For stars with masses  $\gtrsim (8-10) M_{\odot}$ , the core is hot enough to avoid degeneracy until reactions in the core have made heavier isotopes up to  $^{56}\text{Fe}$ . At this point the matter reaches its maximum binding energy<sup>2</sup>, so further heating from nuclear reactions is curtailed. A degenerate core forms and grows in mass due to reactions in shells surrounding the core. There is a maximum mass, known as the CHANDRASEKHAR MASS, that can be supported by electron degeneracy pressure. When the core exceeds this mass, it violently implodes. The implosion halts when matter reaches nuclear density and the repulsive strong nuclear force provides pressure support. In this implosion, most of the electrons and protons

<sup>1</sup> Stars with masses  $\lesssim 0.5 M_{\odot}$  will become degenerate before reaching temperatures sufficient for helium to fuse; the main-sequence lifetime of such stars is much greater than the age of the universe, so making a helium white dwarf requires some kind of mass loss, such as in a binary.

<sup>2</sup> see exercise 5.3

<sup>3</sup> At densities substantially above that of an atomic nucleus other constituents, such as hyperons, may appear.

combine,  $e^- + p \rightarrow n + \nu_e$ . The resulting torrent of neutrinos injects energy into the outer layers of the star; in many cases this is sufficient to eject the outer layers of the star and produce a SUPERNOVA. Left behind will be the core, now composed mostly of neutrons<sup>3</sup> and known as a NEUTRON STAR.

If the envelope is not ejected, matter will fall back onto the neutron star. The maximum mass that can be supported by the nuclear force is uncertain, but is somewhere between  $(2\text{--}3) M_\odot$ ; when this maximum mass is exceeded, the neutron star collapses into a black hole. Having sketched the various end-of-life scenarios, we shall now explore them in more detail.

## 7.1 Low-mass stars

### Ascent of the red-giant branch

With the depletion of hydrogen in the core, the core contracts. During this contraction, hydrogen fusion continues in a shell surrounding the core. The shell hydrogen fusion produces helium, which adds to the core mass. As the core contracts its temperature rises. The rising temperature and pressure at the base of the hydrogen-burning shell causes the reactions in the shell to go at an ever-increasing rate. The resulting increase in luminosity inflates the envelope, now fully convective, to large radii and hence to a low surface temperature: the star becomes a red giant. The high luminosity, combined with the low surface gravity of the distended envelope, drives a strong wind so that the star loses a substantial amount of mass during the giant phase.

Figure 7.1 shows a color-magnitude diagram for the globular cluster M55. The figure plots the absolute  $V$ -band magnitude against the  $(B - V)$  index—brighter and bluer stars at top left, dimmer and redder stars at lower right. The surface effective temperatures is indicated along the top axis, and the luminosity in solar units is indicated along the right axis.

Each dot on the plot represents a star, and the colors indicate how the star would appear. The main sequence forms a band running from the lower right to the center of the figure. At the time of the cluster's birth, the main-sequence would have continued on to the upper left of the plot. Stellar mass increases as one moves upwards and leftwards along the main-sequence, and since more massive stars evolve faster, those bright, blue stars originally on the upper left of the main sequence have ended their hydrogen-burning tenure and moved on. From the location of the main-sequence turn-off, the age of the cluster is estimated<sup>4</sup> to be  $12.9 \pm 0.8$  Gyr. The red giant branch arcs from the center of the plot towards the upper right. As stars turn away from the main sequence and their helium core mass grows, the stars move up the red giant branch

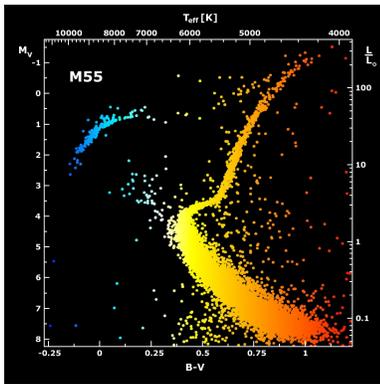


Figure 7.1: Color-magnitude diagram for the globular cluster M55. *Image credit:* B.J. Mochejska, J. Kaluzny (CAMK), 1m Swope Telescope.

<sup>4</sup> Don A. Vandenberg and P. A. Denisenkov. Constraints on the Distance Moduli, Helium and Metal Abundances, and Ages of Globular Clusters from their RR Lyrae and Non-variable Horizontal-branch Stars. III. M55 and NGC 6362. *ApJ*, 862(1):72, July 2018. DOI: 10.3847/1538-4357/aaca9b

becoming redder and more luminous.

### *Helium burning: the horizontal branch*

There are no stable isotopes with mass number  $A = 5$  or  $A = 8$ , which makes the fusion of  ${}^4\text{He}$  somewhat tricky. Although unstable, the isotope  ${}^8\text{Be}$  is relatively long-lived ( $10^{-16}$  s) compared to a nuclear timescale<sup>5</sup>. As a result, when the core temperature reaches  $\approx 10^8$  K, the reaction



builds up a minute abundance of  ${}^8\text{Be}$ . This abundance is sufficient for the reaction



to make a small abundance of  ${}^{12}\text{C}$  in an excited state (denoted by the  $*$ ). While most of the  ${}^{12}\text{C}^*$  decays back into  ${}^8\text{Be} + {}^4\text{He}$ , a small fraction decays instead to the ground state,  ${}^{12}\text{C}^* \rightarrow {}^{12}\text{C} + \gamma$ . The net result is  $3 {}^4\text{He} \rightarrow {}^{12}\text{C}$ , known as the **TRIPLE-ALPHA REACTION**.

Once core  ${}^4\text{He}$  has ignited, the star settles onto a “helium main sequence;” observationally this is the **HORIZONTAL BRANCH**, so called because these stars lie in a clump on a Hertzsprung-Russell diagram. The luminosity on the horizontal branch is about  $(30\text{--}100) L_{\odot}$ . The higher luminosity and the much lower energy release from the triple-alpha reaction make the horizontal branch lifetime much shorter than that of the main-sequence (e.g., the horizontal branch lifetime is  $\sim 10^8$  yr for a solar-mass star). The horizontal branch is clearly visible as the blue arc in the upper-left quadrant of Fig. 7.1.

---

**EXERCISE 7.1** — Use the result of exercise 5.5 to find the heat released per kilogram from fusing 3  ${}^4\text{He}$  nuclei into  ${}^{12}\text{C}$ . Take the core mass to be  $0.45 M_{\odot}$  (the minimum core mass needed for the ignition of helium), and find the lifetime for core helium burning for a horizontal branch luminosity of  $30 L_{\odot}$ .

---

### *The asymptotic giant branch and emergence of a white dwarf*

As the mass of  ${}^{12}\text{C}$  builds up in the core, the reaction  ${}^{12}\text{C} + {}^4\text{He} \rightarrow {}^{16}\text{O}$  begins to compete with the triple alpha reaction. As a result, the core becomes composed of a  ${}^{12}\text{C}/{}^{16}\text{O}$  mixture. With the depletion of  ${}^4\text{He}$ , the core—now composed of  ${}^{12}\text{C}$  and  ${}^{16}\text{O}$ —again contracts, while the growing luminosity from the H- and He-burning shells again inflate the envelope to large radii. Observationally, this phase is the **ASYMPTOTIC GIANT BRANCH**: the stars move away from the horizontal branch and become redder and more luminous. This branch can be observed in Fig. 7.1 arcing from the horizontal branch and asymptotically approaching the red giant branch at upper right.

<sup>5</sup> Roughly the time for a pion to cross a nucleus,  $\sim 10^{-22}$  s.

The mass of a  ${}^8\text{Be}$  nucleus is 91 keV greater than the mass of two  ${}^4\text{He}$  nuclei; at a temperature  $\approx 10^8$  K, the kinetic energy of the  ${}^4\text{He}$  nuclei is just enough to make up the difference.

The triple-alpha reaction is incredibly temperature-sensitive:  $\partial \ln \epsilon_{3\alpha} / \partial \ln T \approx 40$  at  $T = 10^8$  K. This sensitivity, combined with the mildly degenerate conditions of the core, makes the ignition of  ${}^4\text{He}$  somewhat unstable for solar-mass stars.



Figure 7.2: The planetary nebula NGC 2392. *Image credit:* NASA, ESA, Hubble, Chandra; *Processing &*   *License:* Judy Schmidt.

During the ascent of the asymptotic giant branch, the star’s hydrogen-rich envelope is consumed at its base by the H- and He-burning shells and expelled at the surface by an increasingly strong wind. The expelled envelope resembles a nebula and is termed a **PLANETARY NEBULA** (Fig. 7.2). After the envelope has dispersed, the hot core—observed as a white dwarf—slowly cools. For a solar-mass star, the expected final mass of the core, and hence of the white dwarf, is  $\approx 0.6 M_{\odot}$ .

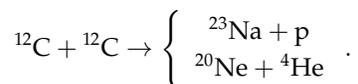
## 7.2 Massive stars

For stars with main-sequence masses  $\gtrsim (8\text{--}10) M_{\odot}$ , the fusion of  $^{12}\text{C}$  commences while the core is non-degenerate and at a temperature  $\approx 8 \times 10^8$  K. At this temperature, electron-positron pairs form and annihilate ( $e^- + e^+ \longleftrightarrow \gamma\gamma$ ); occasionally instead of decaying into photons, the reaction

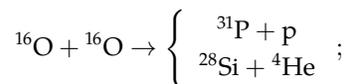
$$e^- + e^+ \longrightarrow \nu_e + \bar{\nu}_e$$

occurs instead and generates a neutrino-antineutrino pair. The mean free path for the neutrinos is larger than the radius of the star; as a result, the neutrinos stream out and take energy from the core. Because the neutrinos can easily leave the star, they end up carrying away the bulk of the heat from the core at these high core temperatures.

Within the core,  $^{12}\text{C}$  is consumed by the reactions



The p and  $^4\text{He}$  capture onto other nuclei that are present. At slightly higher temperatures,  $^{20}\text{Ne} + \gamma \rightarrow ^{16}\text{O} + ^4\text{He}$  releases  $^4\text{He}$  nuclei that subsequently capture onto other  $^{16}\text{O}$ ,  $^{20}\text{Ne}$ , and  $^{24}\text{Mg}$ . As the temperature increases, the next significant burning stage is



as with  $^{12}\text{C} + ^{12}\text{C}$ , the p and  $^4\text{He}$  combine with ambient nuclei with the end result being a distribution of isotopes about  $^{28}\text{Si}$ .

---

**EXERCISE 7.2**— At the onset of  $^{16}\text{O}$  burning in a  $25 M_{\odot}$  star, the central density (Table 7.1) is  $3.6 \times 10^9 \text{ kg m}^{-3}$ . What is the dynamical time of the core?

---

The strong Coulomb barrier inhibits the fusion of nuclei beyond  $^{16}\text{O}$ ; instead, photodissociation reactions such as  $^{28}\text{Si} + \gamma \rightarrow ^{24}\text{Mg} + ^4\text{He}$  liberate n, p, and  $^4\text{He}$ . These light nuclei then capture onto heavier nuclei, and the composition gradually becomes composed of isotopes about  $^{56}\text{Fe}$ . This is **NUCLEAR STATISTICAL EQUILIBRIUM**: the composition is in the

lowest energy state (most bound) for the ambient density and temperature. As a result, there is no further release of nuclear energy possible. The (mostly  $^{56}\text{Fe}$ ) core contracts and becomes degenerate; its mass gradually increases from the burning of surrounding material.

The amount of energy available from the reactions with heavy nuclei is low; as a consequence, the time required for the core to deplete the available fuel grows shorter and shorter, with the final stages occurring in a day (column labeled  $\tau$  in Table 7.1). After the ignition of carbon, the core evolves too quickly for the envelope to keep up. Thus the external appearance of the star provides no window into the final days of burning.

hydrogen						
$M_{\text{ZAMS}}$	$T_c$	$\rho_c$	$L$	$L_\nu$	$\tau$	
$M_\odot$	$10^7 \text{ K}$	$10^3 \text{ kg m}^{-3}$	$10^3 L_\odot$	$L_\odot$	Myr	
15	3.53	5.81	28	—	11.1	
25	3.81	3.81	110	—	6.7	
helium						
$M_{\text{ZAMS}}$	$T_c$	$\rho_c$	$L$	$L_\nu$	$\tau$	
$M_\odot$	$10^8 \text{ K}$	$10^6 \text{ kg m}^{-3}$	$10^3 L_\odot$	$L_\odot$	Myr	
15	1.78	1.39	41	1	1.97	
25	1.96	0.76	182	20	0.84	
carbon						
$M_{\text{ZAMS}}$	$T_c$	$\rho_c$	$L$	$L_\nu$	$\tau$	
$M_\odot$	$10^8 \text{ K}$	$10^9 \text{ kg m}^{-3}$	$10^3 L_\odot$	$10^3 L_\odot$	kyr	
15	8.34	2.39	83	90	2.03	
25	8.41	1.29	245	2600	0.52	
oxygen						
$M_{\text{ZAMS}}$	$T_c$	$\rho_c$	$L$	$L_\nu$	$\tau$	
$M_\odot$	$10^9 \text{ K}$	$10^9 \text{ kg m}^{-3}$	$10^3 L_\odot$	$10^6 L_\odot$	yr	
15	1.94	6.66	87	2	2.58	
25	2.09	3.60	246	6000	0.40	
silicon						
$M_{\text{ZAMS}}$	$T_c$	$\rho_c$	$L$	$L_\nu$	$\tau$	
$M_\odot$	$10^9 \text{ K}$	$10^{10} \text{ kg m}^{-3}$	$10^3 L_\odot$	$10^6 L_\odot$	d	
15	3.34	4.26	87	$10^5$	18.3	
25	3.65	3.01	246	$10^6$	0.7	

The best-known example of an evolved massive star is Betelgeuse, which is large enough and close enough to be resolved (Fig. 7.3 shows a reconstructed image made with interferometry). Betelgeuse probably started as a blue main-sequence star of approximately  $20 M_\odot$  and is currently burning helium in its core. The extended envelope,  $\approx 5 \text{ au}$  in radius, has large convective cells (bright spots in image) and pulsates violently. As can be inferred from Table 7.1, in less than 1 Myr Betelgeuse's core will reach nuclear statistical equilibrium; no more nuclear energy will be available and Betelgeuse will transform into either a neutron star

Table 7.1: Nuclear burning timescales for massive stars. Values taken from Woosley et al. [2002]; neutrino luminosities are taken from Weaver et al. [1978] and do not exactly correspond to the same stellar models for the other parameters.

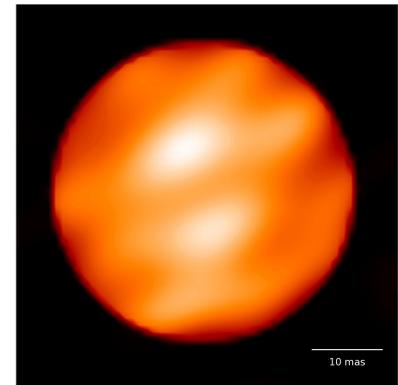


Figure 7.3: A reconstructed image of Betelgeuse made using interferometry. *Image credit:* Xavier Haubois et al. (Observatoire de Paris)

or black hole, as we describe next.

### Core collapse

When the core of a massive star reaches nuclear statistical equilibrium, there are no further sources of energy available. Fusion reactions in the shells surrounding the core add mass to it, causing it to contract. The increasing density raises the electron Fermi energy. When the Fermi energy approaches the rest mass of the electrons— $m_e c^2 = 0.511 \text{ MeV}$ —the electrons move relativistically. This dramatically alters the equation of state.

A particle's energy, including rest mass, is

$$E = \sqrt{p^2 c^2 + m^2 c^4} = mc^2 \left[ 1 + \left( \frac{p}{mc} \right)^2 \right]^{1/2};$$

when  $p \ll mc$ , we can expand this as  $E \approx mc^2 + p^2/2m$ —that is, as the sum of the rest mass and the Newtonian kinetic energy. In the opposite limit, when  $p \gg mc$ ,  $E \approx pc$ . Let's see how this relativistic limit affects the degenerate equation of state. Recall that we fill energy states, starting with the lowest open levels until we have added all  $N$  electrons (eq. [6.37]):

$$N = \frac{2}{h^3} \int_V d^3x \int_0^{E_F} d^3p.$$

Change variables,  $d^3p = 4\pi p^2 dp = 4\pi c^{-3} \varepsilon^2 d\varepsilon$ , where  $\varepsilon = pc$  is the energy of a single, relativistic electron:

$$N = \frac{8\pi}{h^3 c^3} V \int_0^{E_F} \varepsilon^2 d\varepsilon = \frac{8\pi}{3h^3 c^3} V E_F^3.$$

This gives the Fermi energy,

$$E_F = hc \left( \frac{3N}{8\pi V} \right)^{1/3}.$$

To get the total energy, multiply each electron by its energy  $\varepsilon$  and integrate over phase space:

$$E = \frac{8\pi}{h^3 c^3} V \int_0^{E_F} \varepsilon^3 d\varepsilon = \frac{1}{4} \frac{8\pi}{h^3 c^3} V E_F^4 = \frac{3}{4} N E_F.$$

For a relativistic gas, the pressure is  $P = (1/3)(E/V)$  (cf. Box 1.2), so that

$$P = \frac{1}{4} n E_F = \frac{1}{4} \left( \frac{3}{8\pi} \right)^{1/3} h c n^{4/3}, \quad (7.1)$$

with  $n = \rho/\mu_e m_u$ . Instead of  $P \propto \rho^{5/3}$ , as for a non-relativistic gas,  $P \propto \rho^{4/3}$ .

### The Chandrasekhar mass

In exercise 6.6, we constructed a mass-radius relation for white dwarfs by combining the virial relations,

$$P \propto \frac{GM^2}{R^4}$$

$$\rho \propto \frac{M}{R^3}$$

and the equation of state for a non-relativistic, degenerate, ideal gas. We found that  $R \propto M^{-1/3}$ . If we try that with our relativistic equation of state, eq. (7.1), we get

$$\frac{GM^2}{R^4} \propto P = \frac{1}{4} \left( \frac{3}{8\pi} \right)^{1/3} hc \left( \frac{\rho}{m_u \mu_e} \right)^{4/3} \propto \frac{hc}{m_u^{4/3}} \frac{M^{4/3}}{R^4}.$$

The radius  $R$  cancels, and what we have is a relation  $M \propto (hc/G)^{3/2}/m_u^2$ . This is rather odd: a gas with a relativistic equation of state in hydrostatic balance has a characteristic mass defined in terms of fundamental constants.

Let's investigate this further. Suppose we have a box with adjustable sides, which we pack with  $N$  degenerate electrons. We add some nuclei for mass, so that the total mass in the box is  $\mu_e m_u N$ . The volume of the box  $V \sim R^3$ , and since the electrons are degenerate, the volume per electron is roughly  $\lambda^3$ , where  $\lambda \sim h/p$  is the wavelength of the electrons. As a result,  $N = (R/\lambda)^3$ ; further, the momentum of an electron is

$$p \sim \frac{h}{\lambda} \sim h \frac{N^{1/3}}{R}.$$

If our electrons were non-relativistic, the total, kinetic plus gravitational, energy of our box would be

$$E_{\text{total}} = N \frac{p^2}{2m_e} - \frac{GM^2}{R} \sim N^{5/3} \frac{h^2}{R^2 m_e} - GN^2 \mu_e^2 m_u^2 \frac{1}{R}.$$

For a given  $N$ , we can adjust  $R$  to make  $E_{\text{total}} < 0$ , and indeed, if we satisfy the virial theorem, we will recover the  $R \propto M^{-1/3}$  scaling.

If, however, the electrons are relativistic then the total energy is

$$E_{\text{total}} = Npc - \frac{GM^2}{R} = \frac{1}{R} \left[ hcN^{4/3} - GN^2(\mu_e m_u)^2 \right]$$

$$= G(\mu_e m_u)^2 \frac{N^{4/3}}{R} \left[ \frac{hc}{G(\mu_e m_u)^2} - N^{2/3} \right].$$

Look at the term in  $[\cdot]$ . If  $N < [hc/G/(\mu_e m_u)^2]^{3/2}$ , then  $E_{\text{total}} > 0$ ; by making  $R$  larger, however, we can lower the energy until the electrons are no longer relativistic, and then we can again recover the virial scaling. If  $N > [hc/G/(\mu_e m_u)^2]^{3/2}$  then  $E_{\text{total}} < 0$ ; by making  $R$  smaller, however, we can keep reducing  $E_{\text{total}}$  indefinitely.

There is no bound state with finite  $R$  for  $M > (hc/G)^{3/2}(\mu_e m_u)^{-2}$ .

### Box 7.1 Instability for a relativistic equation of state

There is another way of looking at the onset of instability which is instructive (this treatment follows that in Cox [1980]). In exercise 2.9 you found that during a contraction or expansion, the equation of motion for a thin layer at the star's surface was

$$\delta\ddot{R} = \frac{GM}{R^2} [4 - 3\gamma] \frac{\delta R}{R}.$$

Here  $M$  and  $R$  are the total stellar mass and radius, and the adiabatic pressure-density relation is  $P \propto \rho^\gamma$ .

For a non-relativistic gas with  $\gamma = 5/3$ , we have  $\delta\ddot{R} \propto -\delta R$ : the star oscillates with a period that is comparable to the dynamical timescale of the star. If, however,  $\gamma < 4/3$  the equation of motion is  $\delta\ddot{R} \propto \delta R$ , which has an exponential solution: squeeze the star slightly, and it will implode!

Let's work out a more physical explanation for what is happening. Suppose we have a star in virial equilibrium, with the central pressure and density

$$P \propto \frac{GM^2}{R^4}$$

$$\rho \propto \frac{M}{R^3}.$$

Now if the star contracts by a small amount, say  $\delta R/R = -1\%$ , then the density increases by an amount  $\delta\rho/\rho = -3\delta R/R = 3\%$ . How does the pressure respond? If the star contracts slowly, on a Kelvin-Helmholtz timescale, then there is time for heat to radiate away, so that the internal pressure can increase by the amount needed to maintain virial equilibrium: in this case  $\delta P/P = -4\delta R/R = 4\%$ . Under an *adiabatic* contraction, however, there is not enough time for the star to radiate away excess heat; as a consequence, the pressure and density are linked, so that  $\delta P/P = \gamma\delta\rho/\rho = -3\gamma\delta R/R$ .

If the adiabatic index is  $\gamma = 4/3$ , then during an adiabatic compression of  $\delta R/R = -1\%$ , the density increases by  $3|\delta R/R| = 3\%$  and the pressure increases by  $3\gamma|\delta R/R| = 4\%$ , which is precisely the increase needed to maintain mechanical equilibrium. As a result, the star remains in hydrostatic balance at its new, smaller radius. This is why there was no mass-radius relation for  $\gamma = 4/3$ ; it takes no energy to contract (or expand) the star.

For  $\gamma > 4/3$ , the central pressure increases during contraction by  $3\gamma|\delta R/R| > 4|\delta R/R|$ . As a result, the pressure becomes greater

## Box 7.1 continued

than the amount needed for hydrostatic balance. This excess pressure pushes the star outward and acts as a restoring source. During an expansion, the pressure falls below the amount needed for hydrostatic equilibrium, so gravity halts the expansion and forces the star to contract. Hence, for  $\gamma > 4/3$ , the star responds to a radial perturbation by oscillating with a period comparable to the dynamical timescale (cf. exercise 2.9).

In contrast, if  $\gamma < 4/3$  the increase in pressure during contraction is  $3\gamma|\delta R/R| < 4|\delta R/R|$ . The gas pressure does not increase enough to maintain hydrostatic equilibrium, and so the star's contraction accelerates. A small perturbation inwards leads to implosion.

Thus, there is a limit to the total mass that can be supported in hydrostatic equilibrium by degenerate electrons. An exact calculation for the maximum mass of a cold, degenerate star yields

$$M_{\text{Ch}} = 1.456 \left( \frac{2}{\mu_e} \right)^2 M_{\odot}. \quad (7.2)$$

When the mass reaches this limiting value, known as the CHANDRASEKHAR MASS<sup>6</sup>, the electrons become relativistic and  $\partial P/\partial\rho \rightarrow 4/3$ ; the star becomes unstable and collapses.

When the core of a massive star begins its collapse, the electron Fermi is  $\sim \text{MeV}$ , which is sufficient to induce electron captures on iron-group nuclei. These captures increase  $\mu_e$  and reduce  $M_{\text{Ch}}$ . As the core begins the final plunge, the rapidly rising temperature induces the photodissociation of iron-group nuclei into neutrons, protons, and helium nuclei. This process is endothermic, which further robs the core of pressure support and accelerates the collapse. The effective  $\gamma = \partial P/\partial\rho < 4/3$  on account of the photodissociation and electron captures, and the core implodes.

As the core density approaches  $0.16 \text{ fm}^{-3}$ , the nucleons begin to repel one another on account of the strong nuclear force. This abruptly halts the collapse and launches a shockwave outwards. The core now consists mostly of neutrons and is termed a NEUTRON STAR.

<sup>6</sup> Derived by S. Chandrasekhar at age 20(!) while traveling from India to England in 1930

the density of an atomic nucleus

---

EXERCISE 7.3 — What is the mass density if the number density of nucleons is  $0.16 \text{ fm}^{-3}$ ? What is the gravitational binding energy for an object with a mass  $1.4 M_{\odot}$  at this density?

---

The outward traveling shockwave soon stalls as the outer layers of the star fall inward. The energy needed to blow the envelope off is about 1% of the gravitational binding energy of the core, so there is plenty of en-

<sup>7</sup> W. David Arnett, John N. Bahcall, Robert P. Kirshner, and Stanford E. Woosley. Supernova 1987a. *Annual Review of Astronomy and Astrophysics*, 27 (1):629–700, 1989. doi: 10.1146/annurev.aa.27.090189.003213. URL <https://doi.org/10.1146/annurev.aa.27.090189.003213>

<sup>8</sup> By timing pulsars (next section) in a binary system, the orbital parameters and hence the mass of the neutron star can be deduced; the largest measured mass is  $2.14^{+0.10}_{-0.09} M_{\odot}$  [Cromartie et al., 2020].

<sup>9</sup> S. M. Adams, C. S. Kochanek, J. R. Gerke, K. Z. Stanek, and X. Dai. The search for failed supernovae with the Large Binocular Telescope: confirmation of a disappearing star. *MNRAS*, 468:4968–4981, July 2017. doi: 10.1093/mnras/stx816

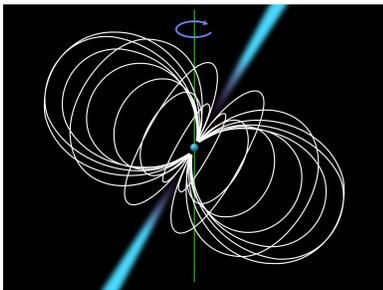


Figure 7.4: Schematic of a pulsar. The white lines indicate the dipole magnetic field, the green line is the rotation axis, and the light blue beams are the radiation. *Image credit:* Made by Mysid in Inkscape, based on `Pulsar_schematic.jpg` by Roy Smits. 

ergy available to disperse the envelope if this energy can be tapped. Most of the gravitational binding energy released by the imploding core is carried outwards by neutrinos. This has been confirmed observationally<sup>7</sup>. In February 1987 the star Sk -69 202, a B3 supergiant in the Large Magellanic Cloud, became supernova 1987A. Just before the optical brightening, a burst of neutrinos were detected in the Kamiokande II (Japan) and IMB (Ohio) water Cherenkov detectors.

During the collapse, the neutrino mean free path becomes smaller than the core radius for two reasons: the weak interaction cross-section increases as the nucleons reach temperatures  $\gtrsim 10^{10}$  K, and the mean free path  $\ell = (n\sigma)^{-1}$  decreases as the density rises. As a result, the neutrinos become trapped and must diffuse out of the collapsing core. As the neutrinos diffuse out, they transfer a small fraction of their energy to the material, which heats it. This heating tends to push the shock outward, and a competition arises between the ram pressure of infalling matter and the heating from the neutrinos. If the neutrinos can transfer enough energy to the envelope, then the envelope will be blown off in a supernova. If not, then matter will continue to accumulate onto the neutron star. The maximum mass of a neutron star is uncertain<sup>8</sup>, but on physical grounds is likely  $< 3 M_{\odot}$ . If the shock is not re-energized, then conceivably the entire star could implode into a black hole—the star would disappear without a corresponding luminous supernovae. There is evidence that this has happened<sup>9</sup>.

### 7.3 Stellar resurrection

In the previous section, we learned that stars with  $M \lesssim (8\text{--}10) M_{\odot}$  eventually become white dwarfs composed of carbon and oxygen and supported by electron degeneracy pressure; and that more massive stars have cores that collapse, either to form neutron stars supported by the strong nuclear interaction or to collapse fully into black holes.

Both the white dwarfs and neutron stars that emerge from the ashes of isolated stars slowly cool and dim. The cooling of white dwarfs can be modeled accurately enough that observations of white dwarfs in clusters can be used to infer the ages of and distances to their host clusters. No such capability is possible with isolated neutron stars: most are too dim to be observed, and there are vast uncertainties about the composition of the deep interior, where the density is several times higher than that of an atomic nucleus. Rather, efforts have been on using observations of the handful of isolated neutron stars with measured surface temperatures to constrain models of nuclear matter.

Many observed neutron stars are endowed with strong magnetic fields, with a surface dipole field strength  $\gtrsim 10^8$  T. If the neutron star spins rapidly enough and the dipole is misaligned with the spin axis, a

tremendous voltage is generated at the surface that accelerates charges above the polar caps. These accelerated charges emit photons that fan outward from the poles, as illustrated in Fig. 7.4. As the neutron star spins, the beams of radiation are swept around; a distant observer therefore observes light pulsing at the rotation frequency of the star. These systems, known as PULSARS, were discovered by Jocelyn Bell<sup>10</sup> and Anthony Hewish in 1967<sup>11</sup>.

---

EXERCISE 7.4 — The Crab pulsar spins at 33 Hz. For a star of  $1 M_{\odot}$ , find the maximum radius such that material at the equator remains bound to the star when spinning at that rate. Based on these results, argue that the Crab pulsar cannot be a white dwarf.

---

MANY STARS ARE IN BINARY SYSTEMS. If the binary happens to survive the evolution off the main sequence, it can often happen that the orbit is close enough for matter to be tidally stripped from the companion and ACCRETED onto the compact star (i.e., white dwarf, neutron star, or black hole). As matter falls into the gravitational potential, it liberates a considerable amount of energy. This makes the system bright.

---

EXERCISE 7.5 — Let's estimate the luminosity and surface temperature of an accreting neutron star. Assume a mass of  $1.4 M_{\odot}$  and a radius of 10 km.

1. Compute the gravitational energy (in MeV) released when a proton falls onto the surface (use a Newtonian approximation for the gravitational potential). How does this compare to the energy released (per proton) from the fusion of hydrogen into helium?
  2. Now suppose the neutron star is accreting  $10^{14} \text{ kg s}^{-1}$ , which is a typical rate for many observed systems. What would be the luminosity generated by this accretion?
  3. Suppose the luminosity were emitted thermally from the surface of the neutron star. What would be the surface effective temperature? In what band (e.g., visible, IR, UV, X-ray) would you want to observe this system?
- 

When sufficient material<sup>12</sup> has accumulated on the surface of a white dwarf or neutron star, thermonuclear reactions can ignite in the accreted layer. This ignition is typically thermally unstable and leads to an explosion. On a white dwarf, this explosion presents as a NOVA<sup>13</sup> as the white dwarf abruptly brightens and then dims over several weeks to months. The mass of the burning layer is typically ( $10^{-5}$  to  $10^{-4}$ )  $M_{\odot}$ ; at typical accretion rates  $\lesssim 10^{-9} M_{\odot} \text{ yr}^{-1}$  the time between the explosions is thousands of years or longer. The amount of mass necessary for ignition decreases strongly with the mass of the white dwarf, however, so that the time between explosions can be years to decades. In these systems

<sup>10</sup> Jocelyn Bell was a 24 year old graduate student at Cambridge at the time

<sup>11</sup> A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins. Observation of a Rapidly Pulsating Radio Source. *Nature*, 217(5130):709–713, February 1968. doi: 10.1038/217709a0

The radio emission from several pulsars, including the Crab, was independently detected by Airman C. Schisler at the Ballistic Missile Early Warning Site, Clear Air Force Station, Alaska.

<sup>12</sup> The accreted matter is usually mostly hydrogen, but if the companion star is evolved it could be enriched in helium or even, if the companion star is itself a white dwarf, carbon and oxygen.

<sup>13</sup> from the Latin *novus* meaning “new”

the novae are observed to reoccur and they are called—appropriately enough—RECURRENT NOVAE. On a neutron star, the explosion is observed as an X-RAY BURST that lasts (10–100) s. The strong gravity makes the amount of material needed for ignition much less than on a white dwarf: roughly  $10^{-12} M_{\odot}$ . As a consequence, the time between bursts can be as short as hours to days.

SOME NEUTRON STARS AND BLACK HOLES ARE IN TIGHT (SHORT ORBITAL PERIOD) BINARIES with another neutron star or black hole. In this case, the system has an oscillating mass quadrupole; further, just as an oscillating electrical dipole radiates electromagnetic waves (light), the orbiting stars will radiate GRAVITATIONAL RADIATION. The gravitational radiation carries energy away from the binary and forces the orbit to shrink. As the orbit shrinks, the emission of gravitational radiation intensifies, and the rate of orbital shrinkage increases. Monitoring of the orbital period of the binary pulsar PSR 1913+16<sup>14</sup> found that the orbital period, and hence the semi-major axis, were indeed decreasing at a rate consistent with predictions from General Relativity<sup>15</sup>.

Direct detection of gravitational radiation was finally achieved in 2015 by the LIGO<sup>16</sup> and Virgo gravitational wave observatories. The event GW150914 was the merger of two black holes with masses  $36_{-4}^{+5} M_{\odot}$  and  $29_{-4}^{+4} M_{\odot}$ <sup>17</sup>. Two years later LIGO and Virgo observed the merger of two neutron stars, GW170817<sup>18</sup>. The event was also detected in the  $\gamma$ -ray, X-ray, optical and infrared bands<sup>19</sup>. The fading afterglow of the merger is consistent with the ejecta containing large amounts of high-opacity lanthanides. This suggests that the copious amounts of heavy elements were formed in the merger, and that perhaps such mergers are the origin of elements such as gold, platinum, and lead. Your jewelry may be a souvenir of the violent merger of two neutron stars long ago.

<sup>14</sup> R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *ApJ*, 195:L51–L53, January 1975. doi: 10.1086/181708

<sup>15</sup> J. H. Taylor and J. M. Weisberg. A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16. *ApJ*, 253:908–920, February 1982. doi: 10.1086/159690

<sup>16</sup> Laser Interferometer Gravitational-Wave Observatory

<sup>17</sup> B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116: 061102, Feb 2016. doi: 10.1103/PhysRevLett.116.061102

<sup>18</sup> B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, October 2017. doi: 10.1103/PhysRevLett.119.161101

<sup>19</sup> B. P. Abbott et al. Multi-messenger observations of a binary neutron star merger. *The Astrophysical Journal*, 848 (2):L12, oct 2017. doi: 10.3847/2041-8213/aa91c9

# A

## Algorithm

Computing is now ubiquitous in science and technology, and forms a triad with theory and experiment. Stellar modeling has a long and storied history in this area. Of course, libraries of numerical routines are now widely available, and for research it is far better to use a well-written and well-tested routine than trying to build your own. One still needs to have a basic understanding, however, of how a computational technique works! In this appendix, we wish to give a flavor of a few common numerical techniques.

### A.1 Numerical precision

Before diving into the sea of computational techniques, we need to wade a bit in the shallows of floating-point arithmetic. Numbers are stored in base-2 (binary) format: a sequence of 1's and 0's known as `bits`. The number of bits in this sequence is fixed<sup>1</sup>, and the processor and compiler define a particular model<sup>2</sup> to represent numbers.

In symbols, an integer  $d$  can be written, using  $N$  bits, as

$$d = (-1)^s \times \sum_{k=1}^{N-1} d_k 2^{k-1}.$$

Here  $s$  is the `SIGN BIT` and  $d_k = 0, 1$ . The largest integer in this representation is  $2^{N-1} - 1$ . For example, suppose we are using  $N = 4$  bits; one bit is reserved to indicate the sign, and with the remaining 3 bits we represent the positive integers from 0 to  $2^{4-1} - 1 = 7$  as

0	1	2	3	4	5	6	7
000	001	010	011	100	101	110	111

Real numbers can be modeled as follows: for  $x \neq 0$ ,

$$x = (-1)^s \times 2^\varepsilon \times \sum_{k=1}^p f_k \times 2^{-k}. \quad (\text{A.1})$$

Here the exponent  $\varepsilon_{\min} \leq \varepsilon \leq \varepsilon_{\max}$  and  $f_k = 0, 1$  with  $f_1 = 1$ . For example, if  $\varepsilon = -6$  and  $f_k = 11010000 \dots$  (i.e.,  $f_k = 1$  for  $k = \{1, 2, 4\}$  and is 0 otherwise), then  $x = 2^{-6} \times (2^{-1} + 2^{-2} + 2^{-4}) = 0.0126953125$ .

<sup>1</sup> On most modern systems the default is 64 bits.

<sup>2</sup> Notation used here follows Metcalf et al. [2018].

The IEEE 754 64-bit model uses 11 bits for representing the exponent  $\varepsilon$  and the remaining 53 for the fractional part, known as the `MANTISSA`. Because  $f_1 = 1$  always, it is not stored to make room for the sign bit.

---

EXERCISE A.1— What is the floating point number 10.0 in the model representation, eq. (A.1)? Specify  $\varepsilon$  and the  $f_k$ .

---

Using the numerical inquiry functions in Fortran 2018, I wrote a small program to report on the arithmetic on my MacBook Pro (Intel Core i5 processor) for 64-bit floating-point arithmetic. The results are as follows.

```

exponent = [-1021, 1024]
digits = 53
[tiny, huge] = [ 2.2251E-308, 1.7977E+308]
precision = 15
epsilon = 2.2204E-16

```

Let's understand what these numbers mean. First, the exponent,  $\varepsilon$  in eq. (A.1), ranges over  $-1021 \leq \varepsilon \leq 1024$ ; given this, what is the smallest positive number, *tiny*, that can be represented by the model (eq. [A.1])? Since in the model  $f_1 = 1$ , the smallest representable number has  $\varepsilon = \varepsilon_{\min}$  and  $f_{k \neq 1} = 0$ :

$$\text{tiny} = 2^{\varepsilon_{\min}} \times 2^{-1} = 2^{-1022} = 2.2251 \times 10^{-308}.$$

What is the largest positive number, *huge*, that can be represented by eq. (A.1)? This number has  $\varepsilon = \varepsilon_{\max}$  and  $f_k = 1, \forall k = 1, \dots, p$ , and  $p = 53$  (digits):

$$\text{huge} = 2^{\varepsilon_{\max}} \times \sum_{k=1}^{53} 2^{-k} = 2^{\varepsilon_{\max}} \times (1 - 2^{-53}) = 1.7977 \times 10^{308}.$$

Finally, we can ask: what is the smallest number that, when added to one, gives a number different than one? The number one is represented with  $\varepsilon = 1$  and  $f_k = 10000 \dots$ , i.e.,  $f_1 = 1$  and  $f_{k \neq 1} = 0$ . The next number larger than one that can be represented in the model (eq. [A.1]) has  $f_k = 10000 \dots 00001$ ; that is,  $f_1 = f_p = 1$  and all other  $f_k = 0$ . Hence, the closest number to 1.0 that we can model with eq. (A.1) differs from 1.0 by  $\text{epsilon} = 2 \times 2^{-p} = 2^{-52} = 2.2204 \times 10^{-16}$ ; our precision in decimal format is therefore about 15 digits.

---

EXERCISE A.2— Using the representation of exercise A.1, what is the nearest number to 10.0 that can be represented in the model of eq. (A.1)?

---

## A.2 Finding the root of a function

A common task in computation is to find the root of a function: that is, given a function  $f(x)$  defined on an interval  $a \leq x \leq b$  with  $f(a)$  and  $f(b)$  having opposite signs ( $f(a) \times f(b) < 0$ ), find  $r \in [a, b]$  such that  $f(r) = 0$ .

If you are counting, it may seem that  $\varepsilon_{\min}$  should be -1023; however, a couple special values of  $\varepsilon$  are reserved to indicate exceptions: NaN (not a number), e.g.,  $\sqrt{-1}$ ;  $\pm\text{Infinity}$ , e.g.,  $1/0$ ; or underflow—the number is too small to be represented in this model with  $f_1 = 1$ .

In this expression use the formula for the sum of the geometric series,

$$\sum_{k=1}^p r^{-k} = \frac{1 - r^{-p}}{r - 1}$$

with  $r = 2$ .

### Bisection

A robust method for finding  $r$  is **BISECTION**. We find the midpoint  $m = (a + b)/2$  and compute  $f(m)$  (red dot, left, Fig. A.1). We then determine the interval— $[a, m]$  or  $[m, b]$  in which the root lies. For example, in Fig. A.1, the root lies in  $[m, b]$  since  $f(m)$  and  $f(b)$  have opposite signs. We thus reset the bounds of our interval—in this case, setting  $a = m$ —and repeat the process (right, Fig. A.1).

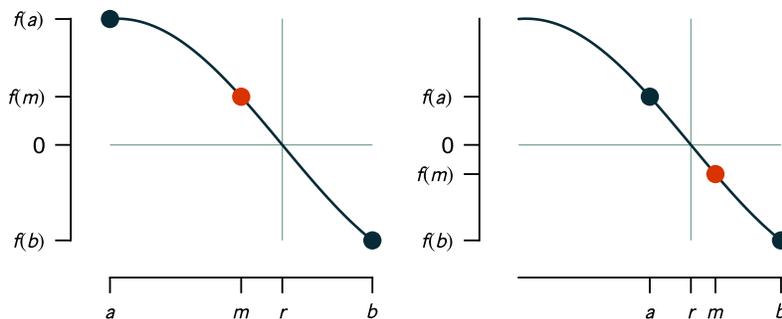


Figure A.1: During the first iteration (left), we bisect the interval and determine whether the root of the function lies in  $[a, m]$  or  $[m, b]$ . For the second iteration, we move the left-boundary  $a$  to  $m$  and compute the next midpoint.

On the second iteration, the root is found to lie in  $[a, m]$ , so for the third iteration, we set  $b = m$ . The midpoints gradually converge toward the root, as shown in Fig. A.2. The midpoint represents the current best estimate for the root; the uncertainty in this estimate is given by the width of the interval  $\Delta = |b - a|$ . On each iteration this uncertainty in our root is halved: if the uncertainty at the start is  $\Delta_0$ , then after one iteration the uncertainty is  $\Delta_0/2$ ; after two iterations,  $\Delta_0/2^2$ ; after  $n$  iterations,  $\Delta_0/2^n$ . If our desired **TOLERANCE** is  $\epsilon$ —that is, the root is known to lie in an interval of width  $\epsilon$ —then we should stop iterating when  $2^{-n}\Delta_0 < \epsilon$ , or after

$$n > \log_2 \left( \frac{\Delta_0}{\epsilon} \right) \text{ iterations.}$$

Note the  $\log_2$ : on each iteration, we gain another bit of precision on the root. Since our precision is limited to roughly 53 bits (§ A.1), this sets the upper limit on how many iterations we need, depending on the initial width of our bracket.

### Newton's method

Bisection is robust: it is guaranteed to converge to a root that is bracketed on some interval  $a \leq x \leq b$ . It converges to the root at a rather plodding pace, however, and you might wonder: can we speed up convergence a bit? If we can evaluate the derivative of our function,  $f'(x)$ , then a classic method for rapidly converging to a root from a good initial guess is **NEWTON'S METHOD**. In this method, on each iteration  $n$  with a trial root

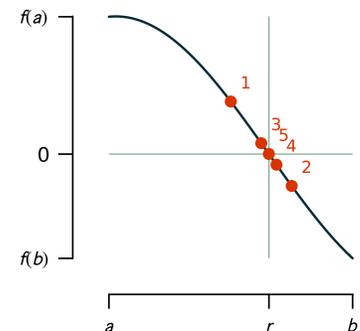


Figure A.2: The first five bisections illustrating the convergence to the root.

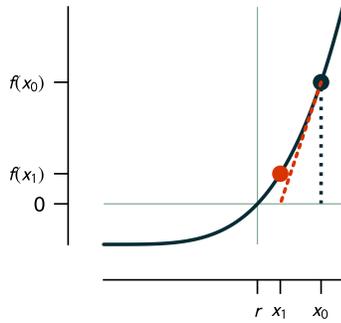


Figure A.3: Starting from an initial guess  $x_0$ , we extend a line (red dotted line) of slope  $f'(x_0)$  to where it crosses  $y = 0$ , thus giving our next guess  $x_1$ .

<sup>3</sup> that is,  $\approx \log_2 p$  iterations, where  $p = 53$  is the number of bits of precision, see eq. (A.1)

Newton's method is useful for quickly estimating roots of numbers: for example, given a guess  $x_n$  for a square root of a number  $r$ , an improved estimate is  $x_{n+1} = (x_n^2 + r)/(2x_n)$ . Often one can do one or two iterations mentally and thus estimate the root within a percent or so. For example, to estimate  $\sqrt{2}$ :  $x_0 = 1$ ;  $x_1 = 3/2$ ; and  $x_2 = 17/12$ , which is accurate to 0.2%.

<sup>4</sup> Richard P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, 1973

$x_n$ , evaluate  $f(x_n)$  and  $f'(x_n)$ . To get the next guess,  $x_{n+1}$ , for the root, construct a line through  $(x_n, f(x_n))$  and having a slope  $f'(x_n)$ ; and solve for where this line crosses the  $x$ -axis:

$$\frac{f(x_n) - 0}{x_n - x_{n+1}} = f'(x_n),$$

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Fig. A.3 illustrates the first iteration to determine  $x_1$  from an initial guess  $x_0$ . We then repeat this process to get  $x_2, x_3$ , and so on, with each one hopefully ever closer to the root.

Compared with bisection, Newton's method converges quite rapidly:  $|x_{n+1} - r| \sim |x_n - r|^2$ —that is, the number of decimal places of precision of the guess roughly *doubles* on each iteration. Thus, for  $f(x) = x^4 - 4$  with  $x_0 = 2$ , only 6 iterations<sup>3</sup> are needed to find the root to a tolerance  $< 10^{-15}$ . Nothing comes for free, however; if the initial guess is too far from the root, then Newton's method may converge much slower than bisection, or perhaps not even converge at all (what happens if  $x_0$  is near the left end of the curve in Fig. A.3?). For this reason, Newton's method is generally not preferred.

#### Brent's method

BRENT'S METHOD<sup>4</sup> is a rapidly converging, robust routine for finding roots. Like bisection, it requires that the root be bracketed on an interval  $x \in [a, b]$ . Rather than use the midpoint as a guess for the root, however, Brent's method instead uses, when possible, either linear or quadratic interpolation (see Box A.1) to construct the next guess for the root.

#### Box A.1 Interpolation

Through any two points  $(x_0, y_0), (x_1, y_1)$  with  $x_1 \neq x_0$ , we can fit a unique line,  $y = ax + b$ , with

$$a = \frac{y_1 - y_0}{x_1 - x_0}$$

$$b = \frac{y_0 x_1 - y_1 x_0}{x_1 - x_0}.$$

Through any three points  $(x_0, y_0), (x_1, y_1), (x_2, y_2)$  with  $x_0, x_1$ , and  $x_2$  distinct, we can fit a unique quadratic  $q = ax^2 + bx + c$ , with  $a, b, c$  determined by the equations

$$ax_0^2 + bx_0 + c = y_0$$

$$ax_1^2 + bx_1 + c = y_1$$

$$ax_2^2 + bx_2 + c = y_2.$$

Box A.1 continued

Continuing, through any 4 distinct points we can fit a polynomial of degree 3:  $p_3(x) = ax^3 + bx^2 + cx + d$ ; and so on. The formula for a polynomial of degree  $N$  passing through  $N + 1$  distinct points is known as the LAGRANGE POLYNOMIAL,

$$p_N(x) = \sum_{i=0}^N y_i \ell_i(x), \tag{A.2}$$

$$\ell_i(x) = \prod_{k=0, k \neq i}^N \frac{x - x_k}{x_i - x_k}. \tag{A.3}$$

For example,

$$p_1(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}; \tag{A.4}$$

$$p_2(x) = y_0 \frac{x - x_1}{x_0 - x_1} \cdot \frac{x - x_2}{x_0 - x_2} + y_1 \frac{x - x_0}{x_1 - x_0} \cdot \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_0}{x_2 - x_0} \cdot \frac{x - x_1}{x_2 - x_1}; \tag{A.5}$$

$$p_3(x) = y_0 \frac{x - x_1}{x_0 - x_1} \cdot \frac{x - x_2}{x_0 - x_2} \cdot \frac{x - x_3}{x_0 - x_3} + y_1 \frac{x - x_0}{x_1 - x_0} \cdot \frac{x - x_2}{x_1 - x_2} \cdot \frac{x - x_3}{x_1 - x_3} + y_2 \frac{x - x_0}{x_2 - x_0} \cdot \frac{x - x_1}{x_2 - x_1} \cdot \frac{x - x_3}{x_2 - x_3} + y_3 \frac{x - x_0}{x_3 - x_0} \cdot \frac{x - x_1}{x_3 - x_1} \cdot \frac{x - x_2}{x_3 - x_2}. \tag{A.6}$$

Fig. A.4 shows the first iteration of Brent’s method. Our test problem is to find the root of  $\sin(x)$  on the interval  $[1.3, 4.5]$ . We orient the interval so that  $|f(b)| < |f(a)|$ , so that  $b$  is the “best guess” for the root. We then use eq. (A.4) to construct a line, known as a SECANT, connecting  $a$  and  $b$  and determine the point  $s$  where that line crosses the  $x$ -axis. In this example,  $s$  becomes the new best guess for the root, so we set  $b = s$  and store the previous best guess as  $c$ .

For the next iteration, we have three points:  $a$ ,  $b$  (the current best guess), and  $c$  (the previous best guess). We can therefore fit a parabola (Fig. A.5) through the three points  $(a, f(a))$ ,  $(b, f(b))$ , and  $(c, f(c))$ , using eq. (A.5). Instead of fitting a parabola  $y = p_2(x)$ , however, we fit  $x = p_2(y)$ ; the reason is that we can simply set  $y = 0$  in eq. (A.5)—with  $x, y$  exchanged—to find the next guess.

The new guess for the root  $s$  (Fig. A.6) is on the same side of the root as  $a$ ; hence the previous best guess becomes  $a$ , and we set  $b = s$ . We are then ready to perform a linear interpolation to determine the next guess, Fig. A.6. After this third iteration, our new guess  $s$  is quite close to the

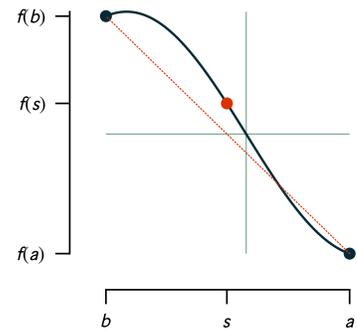


Figure A.4: A secant (thin red line) is constructed through the endpoints of an interval containing the root; where this secant crosses the  $x$ -axis is used as the next guess (red dot) for the root.

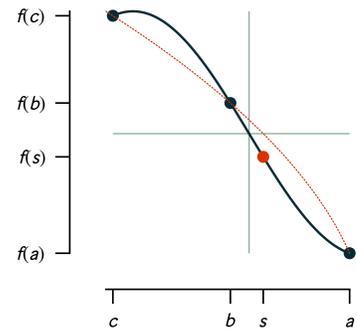


Figure A.5: A parabola  $x = p_2(y)$  (thin red curve) is fit through the previous guesses that bracket a root, and where this parabola intersects  $y = 0$  is used as the next guess  $s$  (red dot) for the root.

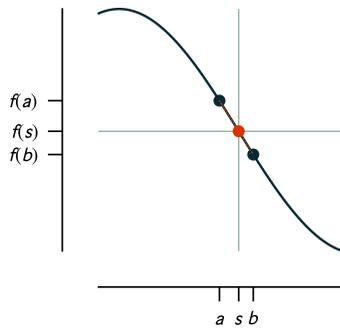


Figure A.6: Brent's method is repeated on the restricted interval  $[m, b]$  using the points  $m, g, b$  to fit a new parabola and determining the next guess (red dot) for the root.

root,  $|s - r| = 5 \times 10^{-4}$ . To reach machine precision for this problem took only 6 iterations.

Although the use of linear and quadratic interpolation can converge rapidly, for some cases the interpolation can fail. By keeping track of the previous best guesses for the root, Brent's method can test whether the best guesses for the root are converging as rapidly as bisection. If the guesses aren't improving quickly enough, or if the guess is out of bounds, the method falls back to taking a bisection step. This combination of rapid convergence and robustness makes Brent's method a workhorse routine for finding roots.

### A.3 Numerically solving a system of ordinary differential equations

Another common numerical task is to integrate a system of first-order ordinary differential equations (ODEs). That is, given a system of equations

$$\frac{dz}{dt} = f(t, z) \quad (\text{A.7})$$

with specified initial conditions  $z(t = t_0)$ , find  $z(t)$ . Here  $z$  is shorthand for an *array* of variables:  $z(t) = \{z_0(t), z_1(t), z_2(t), \dots\}$ . Likewise,  $f(t, z)$  is an array of known, specified functions:  $f(t, z) = \{f_0(t, z), f_1(t, z), f_2(t, z), \dots\}$ .

A prominent example is Newton's equation of motion,

$$\frac{d^2 \mathbf{r}}{dt^2} = \frac{\mathbf{F}}{m}. \quad (\text{A.8})$$

You may object that this is a second-order differential equation; but notice that if we define

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}$$

then we can recast this single second-order differential equation into a system of two<sup>5</sup> first-order differential equations of the form (A.7):

$$\frac{d\mathbf{r}}{dt} = \mathbf{v} \quad (\text{A.9})$$

$$\frac{d\mathbf{v}}{dt} = \frac{\mathbf{F}}{m}. \quad (\text{A.10})$$

As a worked example, we'll now show how to obtain an approximate numerical solution for the following system of ODEs,

$$\frac{dz_0}{dt} = 2\pi z_1 \quad (\text{A.11})$$

$$\frac{dz_1}{dt} = -2\pi z_0, \quad (\text{A.12})$$

with boundary conditions

$$z_0(t = 0) = 0, \quad z_1(t = 0) = 1. \quad (\text{A.13})$$

<sup>5</sup> Strictly speaking, this is a system of *six* ODEs: three components of  $\mathbf{r}$  and three components of  $\mathbf{v}$ .

The solution to equations (A.11)–(A.12) is

$$z_0 = \sin(2\pi t), \quad (\text{A.14})$$

$$z_1 = \cos(2\pi t).. \quad (\text{A.15})$$

as you can easily verify.

SUPPOSE WE KNOW  $z(t)$  AT SOME POINT  $t$  and we wish to make a numerical estimate of  $z$  at a nearby point  $t + h$ . We have the values of  $z$  and we therefore wish to implement eq. (A.7) for equations (A.11)–(A.12) into a single function.

### Box A.2 Functions

What is a function in the context of a program? Basically, a function is a self-contained group of statements that you name. For example, to implement  $dz/dt = f(t, z)$  for equations (A.11)–(A.12), we might write (in Python)

```
def f(t, z):
    """
    RHS of equation dz/dt = f(t, z)
    """
    # this makes an array of length 2,
    # each element of which is zero
    dzdt = np.zeros(2)

    # equation (A.5)
    dzdt[0] = 2.0*np.pi*z[1]
    # equation (A.6)
    dzdt[1] = -2.0*np.pi*z[0]
    return dzdt
```

In this listing we give our function the uninspired name `f`. A function may receive information, which is listed in the parentheses after the function name: `(t, z)`. Thus, the first line

```
def f(t, z):
```

says “bundle the following list of statements together and call it `f`. The statements expect as input two variables, called `ARGUMENTS`, which will be referred to in the function as `t` and `z`.” This function then does three things: it creates an array `dzdt` of length 2, sets the first element of this array to `2.0*np.pi*z[1]`, and sets the second element to `-2.0*np.pi*z[0]`. The final statement

```
    return dzdt
```

**Box A.2 continued**

says “finish, and provide the value of  $dz/dt$ ” to whatever CALLED the function. Thus, for example, after defining the function, you could write

$$k = f(x, y)$$

where  $x$  and  $y$  are variables you had already defined. Python would interpret this as: “Execute the statements in the function  $f$ . In those statements, set the value of  $t$  to be that of  $x$  and the value of  $z$  to be that of  $y$ . After executing those statements, store the value of the function variable  $dz/dt$  in  $k$  and carry on.”

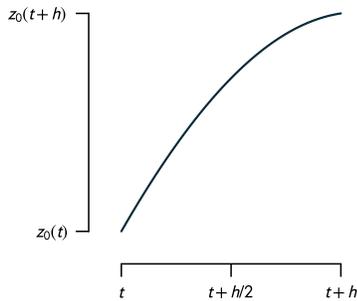


Figure A.7: Solution (A.14) for  $z_0$  in the system of equations (A.11)–(A.12) from  $t = 0.12$  to  $t + h = 0.24$ .

With the definition of a function  $f$ , we can compute the right-hand side of equation (A.7). Our problem can thus be stated as follows: given  $z(t)$ , construct an estimate for  $z(t + h)$ . If we can do this, then we can advance the solution stepwise from the initial condition  $z(t = t_0)$ . We’ll now present three algorithms for doing so, starting with the least accurate. Our example will use  $t = h = 0.12$ ; Fig. A.7 shows the solution for  $z_0(t)$  over this interval.

*Forward Euler*

The first, and simplest, method goes back to Euler. Suppose at time  $t$  we know the solution  $z(t)$  to eq. (A.7). We can expand  $z(t)$  as a Taylor series about this point to obtain

$$z(t + h) = z(t) + h \left. \frac{dz}{dt} \right|_t + \frac{h^2}{2!} \left. \frac{d^2z}{dt^2} \right|_t + \dots$$

But  $dz/dt = f(t, z)$  is a known function. So to  $\mathcal{O}(h^2)$ ,

$$\begin{aligned} z(t + h) &\approx z(t) + h \left. \frac{dz}{dt} \right|_t \\ &= z(t) + h f(t, z)|_t. \end{aligned} \tag{A.16}$$

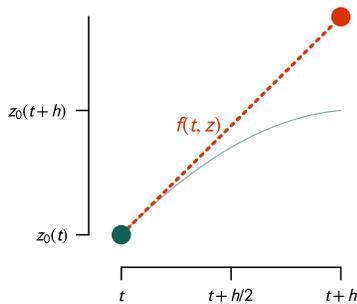


Figure A.8: Schematic of a single forward Euler step, in which we compute the slope  $f(t, z)$  and extrapolate the solution from  $t$  to  $t + h$ .

Figure A.8 displays a schematic of a forward Euler step. We first compute the slope  $f(t, z)$  and then use this to extend the solution to a point  $t + h$ . By repeating this step over and over, we can march our solution along.

How accurate is this FORWARD EULER algorithm? From its definition, the error on each step comes from truncating the Taylor series and is  $\mathcal{O}(h^2)$ . What do we mean by this? For sufficiently small  $h$ , reducing the step by a factor of 2 should reduce the error in a single step by a factor of 4. Another way to put this is that the forward Euler reproduces  $z(t)$  exactly if  $z$  is a linear function of  $t$ . Unfortunately, errors tend to compound with each step, and the smaller the stepsize, the more steps are required. To integrate over a fixed interval  $T$  takes  $T/h$  steps; if the error on a given

step is  $\mathcal{E} \sim \mathcal{O}(h^2)$ , then the total integration error will be something like  $T/h \times \mathcal{E} \sim \mathcal{O}(h)$ . We therefore call this forward Euler method a **first-order** method. Reducing the stepsize by a factor of 2 only reduces the integration error by a factor of 2.

### Second-order Runge-Kutta

The forward Euler algorithm is not accurate unless the step size  $h$  is kept small; as a consequence, a large number of steps are required, which is computationally inefficient. We can improve efficiency if the numerical solution agreed with the solution's Taylor series to a higher order in  $h$ .

EXERCISE A.3 — Suppose we have managed to construct a sequence of numerical solutions  $\phi_n$  to the ODE, eq. (A.7), such that  $\phi_n = z(t_n = n \times h)$ . To find the solution  $\phi_{n+1}$  at  $t_{n+1} = (n+1)h$ , we write  $\phi_{n+1}$  in terms of the solutions at  $t = t_n$  and  $t = (n-1)h$ :

$$\phi_{n+1} = \phi_n + h [af(t_n, \phi_n) + bf(t_{n-1}, \phi_{n-1})]. \quad (\text{A.17})$$

Find the parameters  $a$  and  $b$  such that  $\phi_{n+1}$  agrees with  $z(t_{n+1})$  to second order in  $h$ , that is, so the truncation error is  $\mathcal{O}(h^3)$ . Equation (A.17) is called the **SECOND-ORDER ADAMS-BASHFORTH METHOD**.

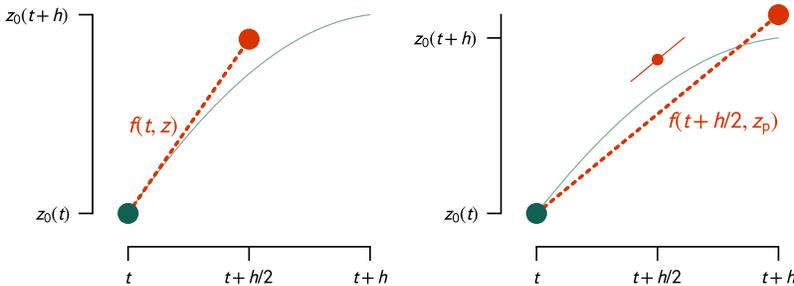


Figure A.9: In the first stage of a second-order Runge-Kutta step (left), we compute the slope  $f(t, z)$  and extrapolate the solution to the midpoint  $t + h/2$ ; in the second stage (right), we compute the slope of our approximate midpoint solution  $f(t + h/2, z_{\text{mp}})$  (left) and use this slope to extend the solution across the interval  $[t, t + h]$ .

A higher-order method starts with using forward Euler to take a step to the midpoint of the interval (Fig. A.9, left),

$$z_{\text{mp}} \left( t + \frac{h}{2} \right) = z(t) + \frac{h}{2} f[t, z(t)]. \quad (\text{A.18})$$

Here  $z_{\text{mp}}$  is our estimate of the solution at  $t + h/2$ . We then compute the derivative  $f$  at  $t + h/2, z_{\text{mp}}$ , and use this corrected value of  $f$  to take a step across the entire interval (Fig. A.9, right):

$$z(t+h) = z(t) + hf \left( t + \frac{h}{2}, z_{\text{mp}} \right). \quad (\text{A.19})$$

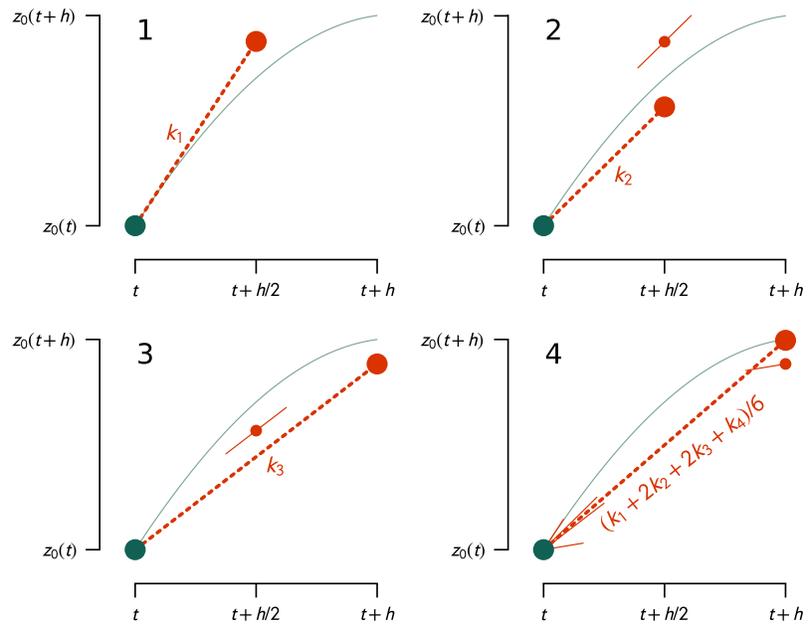
One can show that equations (A.18) and (A.19), which are known as the **SECOND-ORDER RUNGE-KUTTA** method, yield a numerical estimate  $z(t+h)$

that agrees with the actual solution to  $\mathcal{O}(h^3)$ . When integrating over an interval  $T$  and taking  $T/h$  steps, the solution then has a global error  $\sim \mathcal{O}(h^2)$ . Reducing the stepsize by a factor of 2 reduces the integration error by a factor of 4.

#### Fourth-order Runge-Kutta

An even better method is the classic FOURTH-ORDER RUNGE-KUTTA algorithm. In this method, the integration of  $z$  from  $t$  to  $t + h$  is done in four steps, as illustrated in Fig. A.10.

Figure A.10: Stages of the fourth-order Runge-Kutta method.



1. A forward Euler step with slope  $k_1$  is taken to the midpoint  $t + h/2$  and the solution estimated there (Fig. A.10), just as in the second-order method.
2. This solution at the midpoint is used to make a second estimate of the slope  $k_2 = f(t + h/2, z + k_1h/2)$ . Using  $k_2$ , we take a new step from  $t$  just to the midpoint  $t + h/2$  again (Fig. A.10).
3. A new value of the slope  $f$  at the midpoint is then computed:  $k_3 = f(t + h/2, z + k_2h/2)$ . Using  $k_3$ , we then step across the full interval, from  $t$  to  $t + h$  (Fig. A.10).
4. The slope  $k_4 = f(t + h, z + hk_3)$  at the endpoint  $t + h$  is then computed. The full solution  $z(t + h)$  is then constructed from a weighted sum of

the  $k_i$ ,

$$z(t+h) \approx z(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4), \quad (\text{A.20})$$

with

$$\begin{aligned} k_1 &= f(t, z(t)) \\ k_2 &= f\left(t + \frac{h}{2}, z(t) + \frac{h}{2}k_1\right) \\ k_3 &= f\left(t + \frac{h}{2}, z(t) + \frac{h}{2}k_2\right) \\ k_4 &= f(t+h, z(t) + hk_3). \end{aligned}$$

Notice that the  $k_i$  are not independent of one another: each one depends on the intermediate value of  $z$  computed in the previous step. The fourth-order Runge-Kutta scheme “sniffs” the behavior of  $f(t, z)$  over the interval  $(t, t+h)$  and constructs a weighted approximation for  $dz/dt$ . One can show that this method produces solutions with a global truncation error  $\sim \mathcal{O}(h^4)$ . That is, reducing the stepsize by a factor of 2 should reduce the error by a factor  $2^4 = 16$ .

The fourth-order Runge-Kutta scheme is a good general-purpose basic algorithm for integrating systems of ordinary differential equations. It has several limitations, however, three of which we’ll discuss here. First, although the truncation error is  $\sim \mathcal{O}(h^4)$ , we have no way of knowing *a priori* the size of the error. Integrators with `ADAPTIVE STEPSIZES` use Runge-Kutta steps with different orders to estimate the size of the error and adjust  $h$  to maintain the accuracy to some desired tolerance. Second, the method doesn’t know anything about underlying symmetries in the systems of ODEs. `SYMPLECTIC INTEGRATORS` are constructed so that errors in position and momentum will tend to offset when computing the total energy, so that conservation of energy is maintained to high accuracy. Finally, the methods we in this section are `EXPLICIT`: the solution is advanced using an explicit formula in terms of the current values of  $t, z$ . For systems with a large dynamical range (e.g., the difference between the dynamical time and Kelvin-Helmholtz time in a star), the steps must be kept quite small, perhaps prohibitively so, to avoid the numerical solution diverging exponentially.

#### A.4 Cubic splines

The final commonplace numerical task we’ll discuss is interpolation of values in a table of data. In stellar physics, this is often done for opacities or equation of state: one computes, or measures, the opacity under a limited set of composition, temperature, and density and tabulates these values. From this table, we then interpolate to obtain values of the

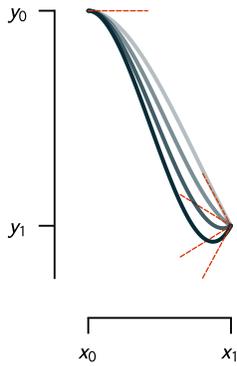


Figure A.11: Cubic polynomials between points  $(x_0, y_0)$  and  $(x_1, y_1)$ ; the slope at  $x_0$  is  $k_0 = 0$ , and the slope  $k_1$  is varied (red dotted lines).

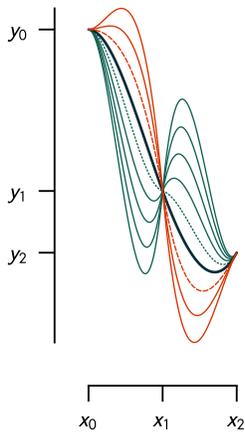


Figure A.12: Two splines spanning  $[x_0, x_1]$  and  $[x_1, x_2]$ , with the slope at the inner point,  $k_2$ , is allowed to vary.

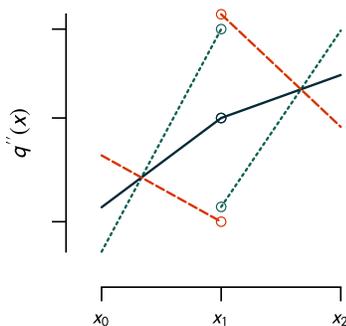


Figure A.13: Second derivatives of our spline functions.

opacity at conditions that lie between table entries. Interpolation using polynomials is discussed in Box A.1; in this section we'll illustrate a commonly used interpolation method, cubic splines.

Through any 4 points  $(x_i, y_i)_{i=0, \dots, 3}$  we can fit a unique cubic polynomial  $y = p_3(x)$  (eq. [A.6]). Alternatively, we can fit a cubic polynomial between two points  $(x_i, y_i)$  and  $(x_{i+1}, y_{i+1})$  if we also specify the first derivatives  $k_i$  and  $k_{i+1}$  at  $x_i, x_{i+1}$ : with  $m_i = (y_{i+1} - y_i)/(x_{i+1} - x_i)$  defined as the mean slope across the interval, the cubic polynomial is

$$q_i(x) = y_i \frac{x_{i+1} - x}{x_{i+1} - x_i} + y_{i+1} \frac{x - x_i}{x_{i+1} - x_i} + \frac{(x - x_i)(x_{i+1} - x)}{x_{i+1} - x_i} \times \left[ \frac{x_{i+1} - x}{x_{i+1} - x_i} (k_i - m_i) - \frac{x - x_i}{x_{i+1} - x_i} (k_{i+1} - m_i) \right]. \quad (\text{A.21})$$

An example of a cubic between  $(x_0, y_0)$  and  $(x_1, y_1)$  with  $k_0 = 0$  and various  $k_1$  is shown in Fig. A.11.

---

EXERCISE A.4 — Verify that  $q_i(x_i) = y_i$  and  $q_i(x_{i+1}) = y_{i+1}$  in eq. (A.21); also verify that  $q'_i(x_i) = k_i$  and  $q'_i(x_{i+1}) = k_{i+1}$ .

---

Now suppose we have a sequence of  $N + 1$  points  $x_0, x_1, \dots, x_N$  with data values  $y_0, y_1, \dots, y_N$ . On each interval  $[x_i, x_{i+1}]$  we can construct a spline  $q_i$ , subject to the requirement that the splines and their first derivatives are continuous at the interior points. As an example, we take one of the curves from Fig. A.11 and cut it into two intervals. We then clamp the derivatives at the endpoints— $k_0$  and  $k_2$ —and allow the first derivative at the interior point,  $k_1$ , to vary, with the requirement that the first derivative is continuous at that point (Fig. A.12). The dark curve shows the case where the first derivative is fixed to the value from original spline shown in Fig. A.11.

As the slope  $k_1$  is varied, the splines to left and right become more tortuous. The second derivative of the spline characterizes this tight bending. In Fig. A.13 we plot the second derivative for three cases: one where  $k_1$  is fixed to the value from the original spline (solid line), and two with  $k_1$  less than (dashed line) and greater than (dotted line) this value. Note that in general the second derivative is not continuous at  $x_1$ . In contrast, the original, least contorted, spline does have a continuous second derivative.

This motivates constructing a smooth curve by requiring that both the first and second derivatives be continuous at the junction points  $x_1, \dots, x_{N-1}$ . Let's check if this gives us enough conditions. For  $N + 1$  points  $x_0, x_1, \dots, x_N$  with data values  $y_0, y_1, \dots, y_N$ , there are  $N$  splines:  $q_0(x)$  on  $x_0 \leq x \leq x_1$ ,  $q_1(x)$  on  $x_1 \leq x \leq x_2$ , and so on to  $q_{N-1}(x)$  on  $x_{N-1} \leq x \leq x_N$ . Each spline has 4 free parameters, so we have a total of  $4N$  parameters. To solve for these parameters, we have the following

conditions:

$$\begin{aligned} q_i(x_i) &= y_i, \quad i = 0, \dots, N-1 && (N \text{ conditions}) \\ q_i(x_{i+1}) &= y_{i+1}, \quad i = 0, \dots, N-1 && (N \text{ conditions}) \\ q'_i(x_i) &= q'_{i-1}(x_i), \quad i = 1, \dots, N-1 && (N-1 \text{ conditions}) \\ q''_i(x_i) &= q''_{i-1}(x_i), \quad i = 1, \dots, N-1 && (N-1 \text{ conditions}). \end{aligned}$$

Our format of the spline, eq. (A.21), automatically satisfies the first two of these. Adding in the continuity of the first and second derivatives brings us to a total of  $4N - 2$  equations and leaves us with two free parameters, namely  $k_0$  and  $k_N$ , the derivatives at the endpoints. We could specify the slopes at the ends (known as a **CLAMPED SPLINE**), but we usually don't have any way of knowing them *a priori*. A popular choice is to set the second derivative at  $x_0$  and  $x_N$  to zero (known as a **NATURAL SPLINE**).

Let's consider the case of a natural spline. Taking the second derivative of eq. (A.21) and evaluating at  $x = x_i$  and  $x = x_{i+1}$  gives

$$q''_i(x_i) = -\frac{2}{x_{i+1} - x_i} [2k_i + k_{i+1} - 3m_i]; \quad (\text{A.22})$$

$$q''_i(x_{i+1}) = \frac{2}{x_{i+1} - x_i} [k_i + 2k_{i+1} - 3m_i]. \quad (\text{A.23})$$

Defining  $\Delta_i = (x_{i+1} - x_i)^{-1}$  and equating second derivatives at the interior points  $x_i$ ,  $i = 1, \dots, N-1$  then yields the following  $N-1$  equations for the  $k_i$ ,  $i = 1, \dots, N-1$ :

$$\Delta_{i-1}k_{i-1} + 2(\Delta_{i-1} + \Delta_i)k_i + \Delta_i k_{i+1} = 3(m_{i-1}\Delta_{i-1} + m_i\Delta_i) \quad (\text{A.24})$$

Setting the second derivatives to zero at  $x_0, x_N$  gives the additional equations

$$2k_0 + k_1 = 3m_0 \quad (\text{A.25})$$

$$k_{N-1} + 2k_N = 3m_{N-1}. \quad (\text{A.26})$$

This system of equations is termed a **TRIDIAGONAL** system because  $i$  contains only  $k_{i-1}$ ,  $k_i$ , and  $k_{i+1}$ . It can be efficiently solved with two iterations over the equations (see Box A.3). After solving for  $k_i$ , if we then wish to interpolate a value  $y(x)$ , we simply need to find  $i$  such that  $x_i < x < x_{i+1}$  and then insert  $x_{i,i+1}$ ,  $y_{i,i+1}$ , and  $k_{i,i+1}$  into equation (A.21) to interpolate  $y(x)$ .

Fig. A.14 illustrates a spline for the function  $\cos(2\pi x)/2$  using just 5 points. The largest deviation is at the ends where our imposition of a vanishing second derivative bows the spline (dotted curve) upwards from the true function (solid curve).

A third choice is the **NOT-A-KNOT** condition, in which we also require the *third* derivative at  $x_1$  and  $x_{N-1}$  to be continuous.

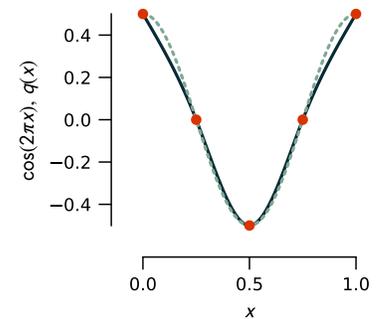


Figure A.14: Spline fit (dotted curve) to the function (solid curve)  $\cos(2\pi x)/2$  using 5 evenly spaced points.







# Bibliography

- B. P. Abbott et al. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016. DOI: 10.1103/PhysRevLett.116.061102.
- B. P. Abbott et al. GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral. *Phys. Rev. Lett.*, 119(16):161101, October 2017. DOI: 10.1103/PhysRevLett.119.161101.
- B. P. Abbott et al. Multi-messenger observations of a binary neutron star merger. *The Astrophysical Journal*, 848(2):L12, oct 2017. DOI: 10.3847/2041-8213/aa91c9.
- S. M. Adams, C. S. Kochanek, J. R. Gerke, K. Z. Stanek, and X. Dai. The search for failed supernovae with the Large Binocular Telescope: confirmation of a disappearing star. *MNRAS*, 468:4968–4981, July 2017. DOI: 10.1093/mnras/stx816.
- W. David Arnett, John N. Bahcall, Robert P. Kirshner, and Stanford E. Woosley. Supernova 1987a. *Annual Review of Astronomy and Astrophysics*, 27(1):629–700, 1989. DOI: 10.1146/annurev.aa.27.090189.003213. URL <https://doi.org/10.1146/annurev.aa.27.090189.003213>.
- R. C. Bohlin and R. L. Gilliland. Hubble Space Telescope Absolute Spectrophotometry of Vega from the Far-Ultraviolet to the Infrared. *AJ*, 127(6):3508–3515, June 2004. DOI: 10.1086/420715.
- Richard P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, 1973.
- Bradley W. Carroll and Dale A. Ostlie. *An Introduction to Modern Astrophysics*. Addison-Wesley, 2d edition, 2006.
- J. P. Cox. *Theory of Stellar Pulsation*. Princeton University Press, 1980.
- H. T. Cromartie, E. Fonseca, S. M. Ransom, et al. Relativistic Shapiro delay measurements of an extremely massive millisecond pulsar. *Nature Astronomy*, 4:72–76, January 2020. DOI: 10.1038/s41550-019-0880-2.

- Michael C. Cushing, J. Davy Kirkpatrick, Christopher R. Gelino, et al. The Discovery of Y Dwarfs using Data from the Wide-field Infrared Survey Explorer (WISE). *ApJ*, 743:50, December 2011. DOI: 10.1088/0004-637X/743/1/50.
- Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman Lectures on Physics*, volume 3. Addison-Wesley, 1989.
- Jan-Vincent Harre and René Heller. Digital color codes of stars. *arXiv e-prints*, art. arXiv:2101.06254, January 2021.
- Ejnar Hertzsprung. Über die Sterne der Unterabteilungen c und ac nach der Spektralklassifikation von Antonia C. Maury. *Astronomische Nachrichten*, 179(24):373, January 1909. DOI: 10.1002/asna.19081792402.
- A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins. Observation of a Rapidly Pulsating Radio Source. *Nature*, 217(5130): 709–713, February 1968. DOI: 10.1038/217709a0.
- R. A. Hulse and J. H. Taylor. Discovery of a pulsar in a binary system. *ApJ*, 195:L51–L53, January 1975. DOI: 10.1086/181708.
- G. H. Jacoby, D. A. Hunter, and C. A. Christian. A library of stellar spectra. *ApJS*, 56:257–281, October 1984. DOI: 10.1086/190983.
- J. D. Kirkpatrick, I. N. Reid, J. Liebert, et al. Dwarfs Cooler than “M”: The Definition of Spectral Type “L” Using Discoveries from the 2 Micron All-Sky Survey (2MASS). *ApJ*, 519:802–833, July 1999.
- Greg Kopp and Judith L. Lean. A new, lower value of total solar irradiance: Evidence and climate significance. *Geophys. Res. Lett.*, 38(1): L01706, January 2011. DOI: 10.1029/2010GL045777.
- Francis LeBlanc. *An Introduction to Stellar Astrophysics*. John Wiley & Sons, Ltd, West Sussex, UK, 2010.
- Brian Luzum et al. The IAU 2009 system of astronomical constants. *Celestial Mechanics and Dynamical Astronomy*, 110(4):293–304, August 2011. DOI: 10.1007/s10569-011-9352-4.
- Michael Metcalf, John Reid, and Malcolm Cohen. *Modern Fortran Explained*. Numerical Mathematics and Scientific Computation. Oxford University Press, 2018.
- D. Mihalas. *Stellar Atmospheres*. W. H. Freeman, 2d edition, 1978.
- Bill Paxton, Lars Bildsten, Aaron Dotter, Falk Herwig, Pierre Lesaffre, and Frank Timmes. Modules for experiments in stellar astrophysics (MESA). *ApJS*, 192:3, January 2011.

- C. H. Payne. *Stellar Atmospheres; a Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars*. PhD thesis, RADCLIFFE COLLEGE., 1925.
- M. A. C. Perryman et al. The Hipparcos Catalogue. *A&A*, 500:501–504, July 1997.
- Henry Norris Russell. Relations Between the Spectra and Other Characteristics of the Stars. *Popular Astronomy*, 22:275–294, May 1914.
- J. H. Taylor and J. M. Weisberg. A new test of general relativity - Gravitational radiation and the binary pulsar PSR 1913+16. *ApJ*, 253:908–920, February 1982. DOI: 10.1086/159690.
- Jagdish K. Tuli. Nuclear wallet cards. National Nuclear Data Center, Brookhaven National Laboratory, 2011.
- Don A. Vandenberg and P. A. Denissenkov. Constraints on the Distance Moduli, Helium and Metal Abundances, and Ages of Globular Clusters from their RR Lyrae and Non-variable Horizontal-branch Stars. III. M55 and NGC 6362. *ApJ*, 862(1):72, July 2018. DOI: 10.3847/1538-4357/aaca9b.
- T. A. Weaver, G. B. Zimmerman, and S. E. Woosley. Presupernova evolution of massive stars. *ApJ*, 225:1021–1029, Nov 1978. DOI: 10.1086/156569.
- Steven Weinberg. *To Explain the World*. HarperCollins, New York, 2015.
- S. E. Woosley, A. Heger, and T. A. Weaver. The evolution and explosion of massive stars. *Reviews of Modern Physics*, 74:1015–1071, November 2002. DOI: 10.1103/RevModPhys.74.1015.