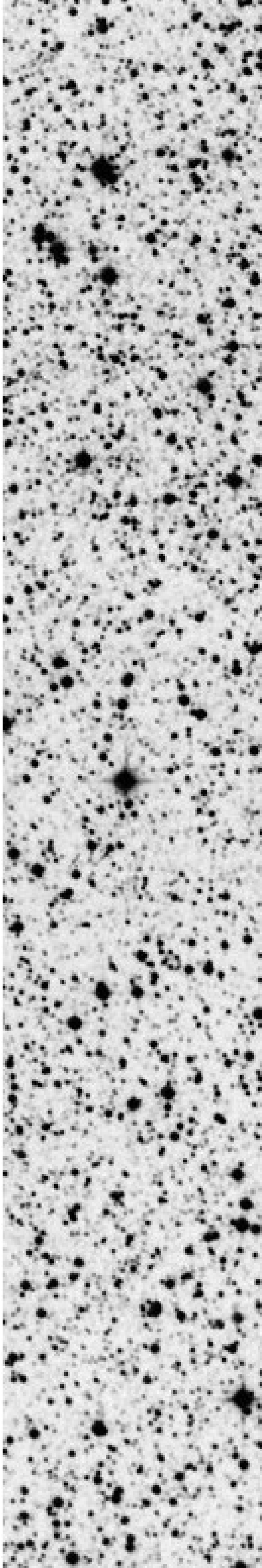
EDWARD BROWN

NOTES ON STELLAR ASTROPHYSICS



The cover image is the starfield in the direction of the second closest star to Earth, Proxima Centauri. Credit & Copyright: David Malin, UK Schmidt Telescope, DSS, AAO

© Edward Brown 2015
git revision 3257aa9d

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 were written while teaching a graduate-level astronomy course, "Stellar Astrophysics" at Michigan State University over the course of Spring semesters in 2006, 2008, 2010, 2012, and 2014. Additional notes were incorporated from a course on Nuclear Astrophysics during Spring Semester 2011.

Some of the material was inspired by a course "Stars with Lars" taught at UC-Berkeley Spring 1995 by Professor L. Bildsten, as well as a course on Statistical Mechanics taught by Professor E. Commins in Fall 1994.

Contents

List of Figures

1
The Sun on a Blackboard

To begin our study of stellar structure, let us first consider the star that we know best, our sun. The planetary orbits and the gravitational constant G tell us its mass; our knowledge of the earth-sun distance and observations tell us its radius; measurements of the solar radiant flux and spectra tell us its luminosity and temperature; and radiometric dating of meteorites tells us the age of the solar system. In summary:

$$\begin{aligned} M_\odot &= 1.99 \times 10^{33}\,\mathrm{g} \\ R_\odot &= 6.96 \times 10^{10}\,\mathrm{cm} \\ L_\odot &= 3.86 \times 10^{33}\,\mathrm{erg\,s^{-1}} \\ T_{\mathrm{eff}} &= 5780\,\mathrm{K} \\ \tau_\odot &= 4.6\,\mathrm{Gyr}. \end{aligned}$$

Moreover, the composition of the sun is well known[1]; the most abundant elements, with abundances $A > 8.0$ (here the abundance scale is relative to hydrogen, with $A \equiv \log_{10}[N_{\mathrm{el}}/N_{\mathrm{H}}] + 12$), are H(12.00), He(10.93), N(7.78), O(8.66), and C(8.39).

Another salient feature of our sun is its stability: the power output is remarkably constant, varying by less than 0.1% over several solar cycles[2], with inferred changes over 2,000 yr on a similar scale[3]. On longer timescales, evidence for liquid water over much of Earth history suggest that the power output of the sun cannot have varied greatly over its life. The first task, then, is to investigate the mechanical and thermal stability of a self-gravitating fluid.

EXERCISE 1.1 — What is the mean density of the sun? What is the luminous flux (energy/area/time) at 1 AU? What is the orbital period of a test mass just exterior to the radius of the sun?

---

[1] E. Anders and N. Grevesse. Abundances of the elements - meteoritic and solar. *Geochim. Cosmochim. Acta*, 53:197–214, 1989; and M. Asplund, N. Grevesse, and A. J. Sauval. The Solar Chemical Composition. In T. G. Barnes, III and F. N. Bash, editors, *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, volume 336 of *Astronomical Society of the Pacific Conference Series*, page 25, September 2005

[2] R. C. Willson and H. S. Hudson. The sun's luminosity over a complete solar cycle. *Nature*, 351:42–44, May 1991

[3] C. Fröhlich and J. Lean. Solar radiative output and its variability: evidence and mechanisms. *A&A Rev.*, 12:273–320, December 2004

1.1 Fluid equation of motion

Over scales that are large compared to the collisional mean free paths between particles, we can treat the fluid as a continuous medium. That is, we suppose that we can find a scale that is infinitesimal compared to the macroscopic scales, but still much larger than the scales for microscopic interactions. Thus, we can define thermodynamic quantities at a location.

Consider such a macroscopically small volume V. Its mass is $M = \int_V \rho\, \mathrm{d}V$, where ρ is the mass density. If $\boldsymbol{u}(\boldsymbol{x}, t)$ is the velocity, then the flux of mass into the element is

$$-\int_{\partial V} \rho \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{S} = \frac{\partial}{\partial t} \int_V \rho\, \mathrm{d}V$$

where the right-hand side follows from mass conservation. Using Gauss's law to transform the left-hand side into an integral over V and combining terms, we have

$$\int_V \left\{ \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{u}) \right\} \mathrm{d}V = 0.$$

Since this equation holds for any V, the integrand must vanish, and we have our first equation,

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0. \tag{1.1}$$

Our next equation is to get the analog of $\boldsymbol{F} = m\boldsymbol{a}$. Ignoring viscous effects, the net force on our fluid element (with volume V) is due to the pressure over its surface P and the gradient of the gravitational potential Φ:

$$\int_V \rho \frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2}\, \mathrm{d}V = \int_V \boldsymbol{F}\, \mathrm{d}V = -\int_V \rho \boldsymbol{\nabla}\Phi\, \mathrm{d}V - \int_{\partial V} P\, \mathrm{d}\boldsymbol{S}.$$

Transforming the second integral on the right-hand side to a volume integral, and assuming that $\boldsymbol{\nabla}\Phi$ and $\boldsymbol{\nabla}P$ vary on macroscopic lengthscales, we arrive at an equation for the acceleration,

$$\frac{\mathrm{d}^2 \boldsymbol{r}}{\mathrm{d}t^2} = -\boldsymbol{\nabla}\Phi - \frac{1}{\rho}\boldsymbol{\nabla}P. \tag{1.2}$$

where $\boldsymbol{r}(t)$ is the position of the particle so that the left-hand side is the acceleration. Here we must be careful: $\boldsymbol{u}(\boldsymbol{x}, t)$ refers to velocity of the fluid at a given point in space and a given instance of time, *not* to the velocity of a given particle. A fluid element can still accelerate even if $\partial_t \boldsymbol{u} = \boldsymbol{0}$ by virtue of moving a different location. At time t this particle has the velocity

$$\left.\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right|_t = \boldsymbol{u}(\boldsymbol{x} = \boldsymbol{r}|_t, t) \tag{1.3}$$

where we use the fact that the particle is moving along a streamline of the fluid. At a slightly later time h, the particle has moved to a location

$\boldsymbol{r}(t+h) \approx \boldsymbol{r}(t) + h\boldsymbol{u}$, and the velocity is now

$$\left.\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}t}\right|_{t+h} = \boldsymbol{u}(\boldsymbol{x} = \boldsymbol{r}|_{t+h}, t+h) \approx \boldsymbol{u} + h(\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + \partial_t\boldsymbol{u}), \tag{1.4}$$

where we evaluate the derivatives at time t. Subtracting equation (1.3) from equation (1.4) and dividing by h gives us the acceleration; inserting this into Newton's law and dividing by volume gives us *Euler's* equation of motion,

$$\partial_t\boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} = -\boldsymbol{\nabla}\Phi - \frac{1}{\rho}\boldsymbol{\nabla}P. \tag{1.5}$$

Equations (1.1) and (1.5) form the first two equations we need to describe stellar structure.

EXERCISE 1.2 — Using equation (1.1), show that equation (1.5) can be written as

$$\partial_t(\rho u_i) + \partial_j(\rho u_i u_j) = -\rho\partial_i\Phi - \partial_i P, \tag{1.6}$$

where the subscripts i denote components and repeated subscripts are understood to be summed over. Interpret the terms on the left-hand side in terms of conservation of momentum.

1.2 Estimates of solar properties

From equations (1.1) and (1.5) we are in a position to estimate, in an order-of-magnitude sense, many of the stellar properties. First, let's consider the scale for each term in equation (1.5),

$$\underset{\mathrm{I}}{\partial_t\boldsymbol{u}} \quad + \quad \underset{\mathrm{II}}{\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}} \quad = \quad \underset{\mathrm{III}}{-\boldsymbol{\nabla}\Phi} \quad - \quad \underset{\mathrm{IV}}{\frac{1}{\rho}\boldsymbol{\nabla}P}$$

For a "characteristic" velocity U and lengthscale R, we see that terms I and II are both of order $\sim U^2/R$ (the timescale is R/U). For term III, we note that $GM/R^2 = (GM/R)/R \sim U_{\mathrm{esc}}^2/R$, where U_{esc} is the escape velocity. Finally, for term IV, $(P/\rho)/R \sim c_s^2/R$, where c_s is the speed of sound. Hence the typical scales of the terms are

$$\mathrm{I} : \mathrm{II} : \mathrm{III} : \mathrm{IV} \sim U^2 : U^2 : U_{\mathrm{esc}}^2 : c_s^2$$

It is clear that the terms on the left-hand side are quite negligible, unless we are dealing with stellar explosions; in this case we must have the two terms on the right-hand side balance, and the star is in hydrostatic balance,

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho\frac{GM(r)}{r^2}. \tag{1.7}$$

Note that this does not mean that $\boldsymbol{u}$ and $\boldsymbol{a}$ are zero; it simply means that they are not important for establishing the mechanical structure of the star.

EXERCISE 1.3 — Equation (1.7) must in general be solved numerically for a real equation of state $P = P(\rho)$, but it is useful to construct a toy model to gain insight. Suppose the sun has a density profile

$$\rho(r) = \rho_0 \left(1 - \frac{r}{R_\odot}\right)$$

where ρ_0 is the central density. Further suppose that the equation of state is that of an ideal gas with mean molecular weight μ. Find the central density, pressure, and temperature in terms of $M_\odot$, $R_\odot$, and μ. How do they compare with the values for a constant density star? Evaluate them numerically for a solar composition (hydrogen mass fraction of 0.7). Keeping M and R fixed, what happens to the central temperature if the composition is transformed to pure helium? If the nuclear reaction rate depends on temperature, what would this do the luminosity, in the absence of any other changes?

A side benefit of our argument about the scaling of the terms is that $c_s \sim U_{\text{esc}} \sim (GM_\odot/R_\odot)^{1/2}$. We can use this to get an estimate of the central temperature of the sun in terms of $M_\odot$ and $R_\odot$: $T_{\odot,\text{center}} \sim 10^7$ K, assuming that the equation of state is that of an ideal gas, $P = (n_{\text{ion}} + n_e)k_\text{B}T$ (see exercise 5).

EXERCISE 1.4 — Use this scaling to get an estimate of the central temperature of the sun in terms of $M_\odot$ and $R_\odot$, assuming the composition is an ideal ionized hydrogen plasma. What is the numerical value of the temperature?

EXERCISE 1.5 — Consider a planar atmophere, in which $-\boldsymbol{\nabla}\Phi = \boldsymbol{g} = -g\boldsymbol{e}_z$ with g constant. Thus the equation of hydrostatic equilibrium (eq. [1.7]) is

$$\frac{\mathrm{d}P}{\mathrm{d}z} = -\rho g. \tag{1.8}$$

Suppose we have an isothermal ideal gas, $P = \rho k_\text{B}T/(\mu m_\text{u})$, where T is the temperature, k_B is Boltzmann's constant, and μm_u is the mass of particles in the gas (m_u is the atomic mass unit), so that the number of particles per unit volume is $N/V = \rho/(\mu m_\text{u})$. Show that for such a gas the density decreases as

$$\rho(z) = \rho(0) \exp\left(-z/H\right)$$

and find an expression for the *scale height* H. Evaluate H for conditions at sea level on Earth. Does the value make sense? Now evaluate H under conditions appropriate for the solar photosphere; in this case what is $H/R_\odot$?

A closer look at hydrostatic equilibrium

If the center of the sun is indeed at a temperature $\sim 10^7$ K, then most of the gas should be ionized. Now electrons are much lighter than ions, so we might worry that the charges might separate. If that were the case, an electric field would be established. For a pure hydrogen plasma, then, we would have *two* equations of hydrostatic equilibrium, one for the electrons and one for the protons,

$$\begin{aligned} \nabla P_p &= n_p m_p \boldsymbol{g} + n_p e \boldsymbol{E} && (1.9) \\ \nabla P_e &= n_e m_e \boldsymbol{g} - n_e e \boldsymbol{E}. && (1.10) \end{aligned}$$

Here $\boldsymbol{g} = -g\boldsymbol{e}_r$ is the gravitational acceleration and $\boldsymbol{E}$ is the electric field. Notice that if we *presume* that the plasma is charge-neutral, then $\nabla(P_p + P_e) = \rho\boldsymbol{g}$, and we can solve for the electric field $\boldsymbol{E}$.

Of course, we must have some charge separation in order to establish the electric field in the first place, but one can show that the fractional charge separation needed is self-consistently small.

EXERCISE 1.6 — Consider a fully ionized hydrogen plasma in a gravitational field in planar geometry. You may assume that both the protons and electrons each have an ideal, non-degenerate equation of state.

1. Argue that in the absence of an electric field, the protons would sink to the bottom of the atmosphere. Show that if the atmosphere is to remain charge neutral, then an electric field

   $$\boldsymbol{E} = -\frac{1}{2}\frac{m_\mathrm{u}}{e}\boldsymbol{g},$$

   must be present. Compare this field to that between the proton and electron in an atom. Could this external field be detectable, by Stark effect for example?

2. Suppose a trace ion of charge $Z'e$ and mass $A'm_\mathrm{u}$ is introduced. What is the net force on this ion?

3. In order to have an electric field, there must be some charge separation. Quantify this: define a parameter

   $$\delta \equiv \frac{n_e - n_p}{n_e + n_p}$$

   and estimate its magnitude. *Hint:* Use Poisson's equation for both the gravitational and electrostatic potentials, and the results of part (a).

A worked example: free-fall collapse

It's worthwhile to imagine what would happen if we suddenly turned off pressure support in the sun, say by having a demon replace each particle with a non-interacting cold particle. For spherically symmetric collapse, let's follow the motion of an observer on the surface. The mass interior to the observer is $M = M_\odot$, so her equation of motion is

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -\frac{GM}{r(t)^2}. \tag{1.11}$$

Multiplying both sides by $u = \mathrm{d}r/\mathrm{d}t$ and integrating gives

$$\frac{1}{2}u^2 = GM\left(\frac{1}{r} - \frac{1}{R}\right),$$

where $R = r(t=0)$. Defining $x = r/R$ gives

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \left[2\frac{GM}{R^3}\left(\frac{1}{x} - 1\right)\right]^{1/2}. \tag{1.12}$$

Now, GM/R^3 has dimension [time^{-2}]; furthermore, $M/R^3 = 4\pi\bar{\rho}/3$, where $\bar{\rho}$ is the average density at the start of collapse. Hence, we can define the **dynamical timescale** as $t_{\mathrm{dyn}} \equiv (G\bar{\rho})^{-1/2}$. For the sun, $t_{\mathrm{dyn}} \approx 1\,\mathrm{hr}$. Defining $\tau = t/t_{\mathrm{dyn}}$ in equation (1.12) gives us a math problem,

For the sun, $\bar{\rho} = 1.4\,\mathrm{g\,cm^{-3}}$, just a bit denser than you.

$$\frac{\mathrm{d}x}{\mathrm{d}\tau} = \left(\frac{8\pi}{3}\right)^{1/2}\left(\frac{1}{x} - 1\right)^{1/2}$$

which can be integrated from $x = 1$ to $x = 0$ to give

$$t_{\mathrm{collapse}} = \left(\frac{3\pi}{32}\right)^{1/2} t_{\mathrm{dyn}} \approx 0.5\,\mathrm{hr}$$

as the time for the sun to collapse if all pressure support were removed.

This is another way of looking at the derivation of eq. (1.7): if terms III and IV are out of balance by even a small amount, the characteristic time for the star to mechanically adjust is very rapid.

1.3 *Energy considerations*

For a spherically symmetric gaseous body in hydrostatic equilibrium, the mass enclosed by radius r satisfies the differential equation $\mathrm{d}m/\mathrm{d}r = 4\pi r^2\rho$. Solving for ρ, substituting into the equation for hydrostatic balance, eq. (1.7), and rearranging terms gives

$$4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{Gm(r)}{r}\frac{\mathrm{d}m}{\mathrm{d}r}.$$

Integrating both sides from $r = 0$ to $r = R$, and changing variables on the right hand side from r to m gives

$$\int_0^R 4\pi r^3 \frac{\mathrm{d}P}{\mathrm{d}r}\,\mathrm{d}r = -3\int_V P\,\mathrm{d}V = -\int_0^M \frac{Gm}{r(m)}\,\mathrm{d}m = E_{\mathrm{grav}}, \tag{1.13}$$

where we integrated the left-hand side by parts, used the fact $P(R) \ll P(0)$, and replaced $4\pi r^2\,\mathrm{d}r$ with $\mathrm{d}V$. Now the pressure is related to the internal thermal (kinetic) energy per unit volume U. For a non-relativistic ideal gas, $P = 2U/3$; for a relativistic gas, such as photons, $P = U/3$. Defining $\gamma = (P+U)/U$, we can write the total energy of our gaseous sphere as

$$\begin{aligned} E &= E_{\mathrm{th}} + E_{\mathrm{grav}} = \int U\,\mathrm{d}V - 3\int P\,\mathrm{d}V \\ &= \frac{1 - 3(\gamma-1)}{\gamma - 1}\int P\,\mathrm{d}V = \frac{3(\gamma-1)-1}{3(\gamma-1)}E_{\mathrm{grav}}. \end{aligned} \tag{1.14}$$

This is the just an application of the virial theorem to our star.

As a first example, consider a star with the pressure provided by a non-relativistic ideal gas. Then $\gamma = 5/3$ and the total energy is[4]

$$E = \frac{1}{2}E_{\mathrm{grav}} < 0.$$

[4] This is true even if the matter is degenerate.

The star is bound. As a second example, consider a star that is so luminous that radiation pressure dominates. In this case, the pressure is that of a relativistic ideal gas. Then $\gamma = 4/3$ and $E = 0$: the star is marginally bound. We must worry about the stability of very luminous stars!

Now suppose the sun were to slowly contract, such that we can still assume hydrostatic equilibrium. How long would this take? The time needed to radiate away the thermal energy defines the **Kelvin-Helmholtz timescale**,

$$t_{\mathrm{KH}} \equiv \frac{E_{\mathrm{th}}}{L} \approx \frac{GM_\odot^2}{2R_\odot L_\odot} = 16\,\mathrm{Myr}. \tag{1.15}$$

We have written "approximately" because we made the approximation that $E_{\mathrm{grav}} = -GM_\odot^2/R_\odot$; in reality it is closer to $-(3/2)GM_\odot^2/R_\odot$. The estimated timescale is much less than the age of the earth, and fossils indicate that the sun has not changed dramatically on this timescale. Hence there is an energy source needed to maintain the interior in thermal steady-state. The total energy per particle, integrated over the lifetime of the sun, is

$$\frac{\Delta E}{N} \approx \frac{L_\odot \times 4.6\,\mathrm{Gyr}}{N} \approx 0.2\,\mathrm{MeV}.$$

This is much larger than chemical reactions[5] could provide. The sun must be powered by nuclear reactions.

[5] typical energy scale is 1 eV

1.4 *Some analytical limits*

We can use the virial theorem of the previous section to set a few limits on the interior pressure and temperature of any star. First, the mass $m(r)$ inside a volume of radius r is

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

so $\mathrm{d}m/\mathrm{d}r = 4\pi r^2\rho$. Combining this with the equation of hydrostatic equilibrium gives,

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

Integrating this equation from the center, where $P = P_c$, to some radius r gives

$$P_c - P(r) = \frac{G}{4\pi}\int_0^r \frac{m\,\mathrm{d}m}{r^4}. \tag{1.16}$$

Now, the average density enclosed in a sphere of radius r is $\bar{\rho}(r) = 3m(r)/(4\pi r^3)$; solving for r and inserting in equation (1.16) gives

$$P_c - P(r) = \left(\frac{4\pi}{3}\right)^{4/3} \frac{G}{4\pi} \int_0^r \bar{\rho}(r)^{4/3} m^{-1/3}\, \mathrm{d}m. \tag{1.17}$$

Now, the density must decrease outward if the system is to be stable (you can't have heavy fluid on top of light!) and so the average density $\bar{\rho}(r)$ must also decrease outward. Hence,

$$\rho_c \geq \bar{\rho}(r) \geq \bar{\rho}(R) = \frac{3M}{4\pi R^3}.$$

Inserting this inequality into equation (1.17) and evaluating at $r = R$ gives a constraint on the central pressure,

$$\frac{3}{8\pi}\frac{GM^2}{R^4} \leq P_c \leq \frac{1}{2}\left(\frac{4\pi}{3}\right)^{1/3} \rho_c^{4/3} GM^{2/3}. \tag{1.18}$$

The critical point here is to notice the order-of-magnitude scale: $P_c \sim GM^2/R^4$. The only assumption in setting the limits (eq. [1.18]) is that the density decreases outward.

EXERCISE 1.7 — Compute the mean kinetic (thermal) energy per hydrogen nucleus in the sun, and express it in electron volts. How does this compare to the binding energy of a hydrogen atom?

1.5 *Transport of energy*

We derived that at the current luminosity, the sun would take $\sim$ 16 Myr to radiate away its internal (thermal) energy. This raises an interesting question: what sets the luminosity? To develop this idea further, let us write the luminosity as

luminosity $\sim$ (radiation energy stored in sun)/(photon escape time).

To get the radiation energy stored in the sun, we multiply the energy density of a thermal distribution of photons, at the central temperature of the sun, by the volume of the sun:

$$E_\gamma = aT_c^4 \times \frac{4\pi}{3} R_\odot^3. \tag{1.19}$$

What about the photon escape time? As a first try, suppose the sun were transparent, so that photons could freely stream out. Then the escape time would simply be $R_\odot/c$. This gives a ridiculously large luminosity.

Of course, a transparent sun would also not produce a thermal spectrum, since there would be no way for the photons to come into thermal equilibrium with the matter.

Suppose now instead that each photon can only travel a short distance ℓ before it is scattered into some random direction. In a random walk, the total distance the photon travels to escape is $R_\odot(R_\odot/\ell)$. In this case, the flux from the sun would be

$$F = \frac{L_\odot}{4\pi R_\odot^2} \sim \frac{(4\pi/3)R_\odot^3 a T_c^4}{4\pi R_\odot^2}\frac{c\ell}{R_\odot^2} = \frac{1}{3}\ell c \frac{a T_c^4}{R_\odot}. \tag{1.20}$$

This is very crude, but we can use it to estimate that $\ell \sim 10^{-3}$ cm. The average distance a photon can travel before being absorbed or scattered is called its **mean free path**. Given this value of ℓ estimated from eq. (1.20), we can estimate the total number of scatterings a photon must suffer in escaping; it is a very large number, and the sun is quite opaque.

EXERCISE 1.8 — If we regard the sun as a large cavity filled with photons, estimate the total energy stored in the radiation field. If the sun were suddenly to become completely transparent, what would be the resulting luminosity?

1.6 Summary

In summary, we've taken the observed gross properties of the sun, the equation of motion for a fluid, and the ideal gas equation of state; from these we've deduced that the sun is in hydrostatic balance, that its interior temperature is of order 10^7 K, and that it would radiate away its thermal energy and contract within about 10 Myr if there were no nuclear reactions in its core. We have developed now a crude picture of the sun: it is a mass of plasma that is in hydrostatic equilibrium, with pressure gradients supporting the inward pull of gravity. It is very opaque, and thus it acts as a reservoir for photons with a thermal, Planckian distribution. Because it is opaque, the photons leak out very slowly. This slow leakage represents a loss of thermal energy; the thermal energy is replenished by heat liberated from nuclear reactions.

What comes next is fleshing out the detailed physics implied by these considerations: an equation of state to relate the pressure to the density and temperature; photon scattering and absorption cross sections to compute the heat transport; nuclear reaction rates to determine the thermal steady-state and the gradual change in composition of the interior.

2

The Lagrangian Equations of Stellar Structure

2.1 The conservation laws

After our rapid overview, we now gather the tools needed to tackle stellar evolution. The first is to get the macroscopic equations for stellar structure. We will start from the equations expressing conservation of mass[1], momentum, and energy. We already derived the continuity (conservation of mass) equation,

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{2.1}$$

[1] In a relativistic system, we would instead start from conservation of baryon number.

and the Euler equation,

$$\partial_t \boldsymbol{u} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} = -\boldsymbol{\nabla}\Phi - \frac{1}{\rho}\boldsymbol{\nabla}P. \tag{2.2}$$

Note that if we multiply eq. (2.2) by ρ, we can rewrite it, using eq. (2.1), as

$$\partial_t(\rho\boldsymbol{u}) + \nabla \cdot [\boldsymbol{u}(\rho\boldsymbol{u})] = -\rho\boldsymbol{\nabla}\Phi - \boldsymbol{\nabla}P. \tag{2.3}$$

The left-hand side is interpreted as expressing the conservation of momentum ($\rho\boldsymbol{u}$) in the absence of forces, analogous to eq. (2.1) for the conservation of mass (ρ).

Note the general form of a conservation equation:

$$\begin{aligned}&\partial_t(\text{conserved quantity})\\ &\qquad + \nabla \cdot (\text{flux of conserved quantity}) = (\text{sources}) - (\text{sinks}).\end{aligned}$$

Because the momentum density $\rho\boldsymbol{u}$ is a vector, its flux is a tensor: $[\boldsymbol{u}(\rho\boldsymbol{u})]_{ij} \equiv \rho u_i u_j$.

THE NEXT EQUATION IS THAT OF ENERGY CONSERVATION. Here we must consider both the internal energy per unit volume $E/V = \rho\epsilon$ and the kinetic energy per unit volume $\rho u^2/2$. In this section ϵ represents the internal energy per unit mass of the fluid. In a fixed volume of the fluid the total energy is thus

$$\int_V \left(\rho\frac{1}{2}u^2 + \rho\epsilon\right)\,\mathrm{d}V.$$

The flux of energy into this volume will clearly include

$$-\int_{\partial V}\left(\frac{1}{2}\rho u^2+\rho\epsilon\right)\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}.$$

But wait, there's more! In addition, we have a conductive heat flux,

$$-\int_{\partial V}\boldsymbol{F}\cdot\mathrm{d}\boldsymbol{S}.$$

Moreover, the pressure acting on fluid flowing into our volume does work on the gas at a rate

$$-\int_{\partial V}P\boldsymbol{u}\cdot\mathrm{d}\boldsymbol{S}.$$

As a result, the net change of energy in our volume is

$$\partial_t\int_V\left(\frac{1}{2}\rho u^2+\rho\epsilon\right)\,\mathrm{d}V = \\ -\int_{\partial V}\mathrm{d}\boldsymbol{S}\cdot\left[\boldsymbol{u}\left(\frac{1}{2}\rho u^2+\rho\epsilon+P\right)+\boldsymbol{F}\right]+\int_V\left(\rho\boldsymbol{u}\cdot\boldsymbol{g}+\rho q\right). \quad (2.4)$$

On the right-hand side we've added in the work done by gravity and the heating evolved by nuclear reactions (this could also involve sinks, such as neutrinos with a long mean free path). Expressed in differential form, equation (2.4) is

$$\partial_t\left(\frac{1}{2}\rho u^2+\rho\epsilon\right)+\nabla\cdot\left[\rho\boldsymbol{u}\left(\frac{1}{2}u^2+\epsilon+\frac{P}{\rho}\right)\right]+\nabla\cdot\boldsymbol{F}=\rho q+\rho\boldsymbol{u}\cdot\boldsymbol{g}. \quad (2.5)$$

You are possibly wondering why I didn't put gravity, which can be expressed as a potential, on the left hand side of this equation. The reason is that the gravitational stresses cannot be expressed in a *locally* conservative form; it is only when integrating over all space that the conservation law appears.

EQUATIONS (2.1), (2.3), AND (2.5) ARE SUPPLEMENTED BY AN EQUATION OF STATE, which allows one to get from the pressure P, the temperature T, and the mass fractions X_i of the species present, the remaining thermodynamical quantities, such as mass density ρ and specific energy ϵ. In addition, Poisson's equation

$$\nabla^2\Phi = 4\pi G\rho, \quad (2.6)$$

specifies the gravitational acceleration $\boldsymbol{g} = -\boldsymbol{\nabla}\Phi$. We then need one more equation to specify the heat flux F. We argued in §1.5 that the typical length over which a photon travels before scattering is very small compared to the lengthscale over which the macroscopic properties of the star vary. In this case, we expect the flux to obey a conduction equation of the form

$$\boldsymbol{F} = -K\boldsymbol{\nabla}T. \quad (2.7)$$

This assumption is clearly questionable near the stellar surface, and we have left unspecified the form of K. Such an equation does, however, close the system of equations; all of the physics is then contained in the equation of state $P(\rho, T, \{X_i\})$, the rate of heating from nuclear reactions $q(\rho, T, \{X_i\})$, and the thermal conductivity $K(\rho, T, \{X_i\})$. Here $\{X_i\}$ are the mass fractions of the isotopes composing the solar plasma. We will also need a system of equations to describe how the X_i change as a result of nuclear reactions and diffusion.

2.2 *Thermodynamics of a mixture: A digression*

Specifying the composition

In this section we'll look at how one describes the composition for a multi-component plasma. To make things concrete, let's imagine a box containing a mixture of nuclei, of many different isotopes, and electrons. (To keep things simple, we'll assume complete ionization.) Each isotope species i has N_i nuclei present, and is characterized by charge number Z_i and nucleon number A_i. Charge neutrality then specifies the number of electrons,

$$N_\mathrm{e} = \sum_i Z_i N_i. \tag{2.8}$$

The total mass of the box is

$$M = m_\mathrm{e} N_\mathrm{e} + \sum_i m_i N_i, \tag{2.9}$$

where m_e and m_i are respectively the mass of an electron and a nucleus of species i. Now what is m_i? Breaking a nucleus i into Z_i protons and $A_i - Z_i$ neutrons takes a certain amount of energy, the *binding energy* B_i. We can therefore write $m_i = Z_i m_\mathrm{p} + (A_i - Z_i) m_\mathrm{n} - B_i/c^2$, where m_p and m_n are respectively the proton and neutron rest masses.

Inserting our expression for m_i into equation (2.9), dividing by the volume of the box V, and rearranging terms gives us the mass density,

$$\rho = \frac{M}{V} = \sum_i n_i \left[(A_i - Z_i)\, m_\mathrm{n} + Z_i \left(m_\mathrm{p} + m_\mathrm{e} \right) - B_i/c^2 \right]. \tag{2.10}$$

Here n_i is the number density of isotope species i, and we have used equation (2.8) to eliminate N_e. The numbers n_i are, of course, fantastically[2] large, so we scale the numbers by *Avogadro's constant*,

[2] astronomically?

$$N_\mathrm{A} = 6.0221367 \times 10^{23}\,\mathrm{mol}^{-1}. \tag{2.11}$$

If we multiply and divide the right-hand side of equation (2.10) by N_A, we then have

Recall that a **mole** is an amount of something: in 1 mol there are N_A items.

$$\rho = \sum_i \left(\frac{n_i}{N_\mathrm{A}} \right) \mathcal{A}_i, \tag{2.12}$$

where

$$\mathcal{A}_i = \left[(A_i - Z_i)\, m_\mathrm{n} + Z_i \left(m_\mathrm{p} + m_\mathrm{e}\right) - B_i/c^2\right] \times N_\mathrm{A} \tag{2.13}$$

is the *gram-molecular weight* of species i with dimensions $[\mathcal{A}] \sim [\mathrm{g \cdot mol^{-1}}]$. Strictly speaking, the gram-molecular weight actually refers to the mass of a mole of the isotope in *atomic* form; the right-hand side of eq. (2.13) is the gram-molecular weight neglecting the electronic binding energy.

Now you may wonder where the numerical value of N_A came from. It is not pulled out of thin air, but rather is defined so that 1 mol of ^{12}C has a mass of exactly 12 g. In other words, for ^{12}C $\mathcal{A} \equiv A\,\mathrm{g\,mol^{-1}}$. In fact for all nuclei, $\mathcal{A} \approx A\,\mathrm{g\,mol^{-1}}$ to better than about 1%, as demonstrated in Table 2.1. Because in CGS $\mathcal{A} \approx A$, it is customary to write $\mathcal{A} = A \times (1\,\mathrm{g\,mol^{-1}})$, so that equation (2.12) is

$$\rho = \sum_i \left(\frac{n_i}{N_\mathrm{A}} \times 1\frac{\mathrm{g}}{\mathrm{mol}}\right) A_i. \tag{2.14}$$

This only works if our unit of mass is the gram: in SI units ^{12}C has a mass of $0.012\,\mathrm{kg\,mol^{-1}}$. Equation (2.14) would be exact if A were a real number, but the custom is to just keep it as the nucleon number, which introduces an error of order one percent. Astronomers typically then *redefine* N_A to mean $N_\mathrm{A}(\text{astronomy}) \equiv N_\mathrm{A}/(1\,\mathrm{g\,mol^{-1}}) = 6.0221367 \times 10^{23}\,\mathrm{g^{-1}}$. Alternatively, one can use the atomic mass unit (symbol u) defined as 1/12 the mass of an atom of ^{12}C, so that $1\,\mathrm{u} = (1\,\mathrm{g\,mol^{-1}})/N_\mathrm{A} = 1.66054 \times 10^{-24}\,\mathrm{g}$. This puts equation (2.14) into the more obvious form $\rho = \sum n_i \times A_i m_\mathrm{u}$, with m_u having a mass of 1 u.

nuclide	A	$\mathcal{A}$	$(\lvert\mathcal{A} - A\rvert/A) \times 100$
n	1	1.00865	0.865
^{1}H	1	1.00783	0.783
^{4}He	4	4.00260	0.065
^{16}O	16	15.99491	0.032
^{28}Si	28	27.97693	0.082
^{56}Fe	56	55.93494	0.116

Table 2.1: Selected gram-molecular weights.

With the redefinition of N_A, equation (2.14) can be rewritten as

$$1 = \sum_i \left(\frac{n_i}{N_\mathrm{A}\rho}\right) A_i \equiv \sum_i Y_i A_i \tag{2.15}$$

where $Y_i \equiv n_i/\rho/N_\mathrm{A}$ is the *molar fraction*. It is customary to call $Y_i A_i$ the *mass fraction* X_i, with $\sum X_i = 1$. We can then define the mean atomic mass number,

$$\bar{A} = \frac{\sum A_i Y_i}{\sum Y_i} = \frac{1}{\sum Y_i}, \tag{2.16}$$

and mean charge number

$$\bar{Z} = \frac{\sum Z_i Y_i}{\sum Y_i} = \bar{A} \sum Z_i Y_i. \tag{2.17}$$

The molar fraction of electrons is

$$Y_e = \sum Z_i \frac{n_i}{\rho N_\mathrm{A}} = \sum Z_i Y_i = \frac{\bar{Z}}{\bar{A}}. \tag{2.18}$$

In stellar structure work, it is common to use the *mean molecular weight*, defined so that the total number of particles, including electrons, per unit volume is

$$\sum_i n_i + n_e \equiv \frac{\rho N_\mathrm{A}}{\mu}. \tag{2.19}$$

Yes, this is still the redefined N_A: μ is dimensionless. From the definition,

$$\mu = \left(\sum_i Y_i + Y_e\right)^{-1} = \left[\sum_i (Z_i + 1)\, Y_i\right]^{-1};$$

sometimes astronomers also define the mean ion molecular weight, $\mu_I = (\sum Y_i)^{-1}$, and the mean electron weight, $\mu_e = Y_e^{-1}$.

EXERCISE 2.1 — Consider a gas of ^{1}H and ^{4}He with molar hydrogen fraction Y_H. Derive expressions for the molar fraction of ^{4}He, Y_He, $\bar{A}$, $\bar{Z}$, and μ. What are the numerical value of these quantities for $Y_\mathrm{H} = 0.7$, i. e., solar?

EXERCISE 2.2 — Assume that we can describe this plasma as an ideal gas. What is the sound speed and the average kinetic energy of a particle, for a given mass density and temperature?

Thermodynamical quantities

In most textbooks on thermodynamics and statistical mechanics, the thermodynamics are formulated in terms of some sample of fixed size. For example, in the first law,

$$\mathrm{d}E = T\mathrm{d}S - P\mathrm{d}V, \tag{2.20}$$

the energy E and entropy S are extensive quantities, and scale with the number of particles N in our sample. In a fluid, however, these quantities are all functions of position. By $S(r)$, we mean that we can define a small portion of the star about the coordinate r that is large enough particles to ensure that quantities such as pressure and temperature are well-defined, but small enough that we can treat $S(r)$ as a continuous function of position when integrating over the whole star.

Using extensive quantities in fluid mechanics is cumbersome, so we instead use quantities like the energy per unit mass $\epsilon = E/(\rho V)$ or the entropy per unit mass $s = S/(\rho V)$. Since a fixed mass of fluid M occupies a volume $V = M/\rho$, we can divide the first law, eq. (2.20), by M to obtain

$$\mathrm{d}\epsilon = T\mathrm{d}s - P\mathrm{d}\left(\frac{1}{\rho}\right) = T\mathrm{d}s + \frac{P}{\rho^2}\mathrm{d}\rho. \tag{2.21}$$

The other extensive variables can be re-defined into mass-specific forms in a similar fashion.

2.3 *The equations in Lagrangian form*

The fluid equations (2.1), (2.3), and (2.5) are in **Eulerian** form; that is, they describe everything in terms of spatial coordinates and time. This is not necessarily the most convenient form for practical calculations. For example, the star can expand and contract, making the radius a function of time. Moreover, the velocity $\boldsymbol{u}$ is *not* the velocity of a given fluid element, which is why the equation of motion (eq. [1.5]) is non-linear. It is often desirable to put the fluid equations into **Lagrangian** form, in which the coordinates are some label for a fluid element and time.

In one-dimension, the transformation to Lagrangian equations is easy. At some reference time, we label the mass enclosed by a shell of radius r

$$m(r,t) = \int_0^r \rho(r',t) 4\pi r'^2 \,\mathrm{d}r', \tag{2.22}$$

as a Lagrangian coordinate m; we then transform coordinates from (r,t) to (m,t). To do this, differentiate eq. (2.22) w.r.t. r,

$$\partial_r m = 4\pi r^2 \rho,$$

and substitute for ρ in the equation of continuity (eq. [2.1]). The first term becomes

$$\partial_t \rho = \partial_t \left(\frac{1}{4\pi r^2}\partial_r m\right) = \frac{1}{4\pi r^2}\partial_r(\partial_t m),$$

while the second term becomes

$$\frac{1}{4\pi r^2}\partial_r\left(u\partial_r m\right);$$

the equation of continuity therefore becomes

$$\frac{1}{4\pi r^2}\partial_r\left(\partial_t m + u\partial_r m\right) = 0. \tag{2.23}$$

We can integrate this over r to find that $\partial_t m + u\partial_r m = f(t)$; since $m(0,t) = 0,\ \forall t$, we must have $f(t) = 0$. Now $\partial_t m + u\partial_r m = \mathrm{D}m/\mathrm{D}t = 0$, so along a streamline, m is a constant. We can therefore transform from coordinates (r,t) to (m,t) by setting

$$\left.\frac{\partial}{\partial_t}\right|_r + u\left.\frac{\partial}{\partial r}\right|_t = \left.\frac{\partial}{\partial t}\right|_m \equiv \frac{\mathrm{D}}{\mathrm{D}t} \tag{2.24}$$

$$\left.\frac{\partial}{\partial r}\right|_t = 4\pi r^2 \rho \left.\frac{\partial}{\partial m}\right|_t. \tag{2.25}$$

Here $\mathrm{D}/\mathrm{D}t \equiv (\partial/\partial t)_m$ is the Lagrangian time derivative. In deriving this change, we used the equation of continuity, which becomes

$$\frac{\partial r}{\partial m} = \frac{1}{4\pi r^2 \rho}. \tag{2.26}$$

Our equation for momentum (eq. [2.2]) becomes

$$\frac{\partial P}{\partial m} = -\frac{Gm}{4\pi r^4} - \frac{1}{4\pi r^2}\frac{\mathrm{D}u}{\mathrm{D}t}. \tag{2.27}$$

In hydrostatic balance the second term on the right-hand side is negligible. The flux equation, (eq. [2.7]) can be transformed to

$$\frac{\partial T}{\partial m} = -\frac{1}{16\pi^2 r^4 \rho K} L_r \tag{2.28}$$

Here L_r is the luminous flux at a radius r.

The energy equation (eq. [2.5]) is more complicated. We can expand the time derivative as

$$\begin{aligned}
\partial_t\left(\frac{1}{2}\rho u^2 + \rho\epsilon\right) &= \left(\frac{1}{2}u^2 + \epsilon\right)\partial_t\rho + \rho\partial_t\left[\frac{1}{2}(\boldsymbol{u}\cdot\boldsymbol{u}) + \epsilon\right] \\
&= -\left(\frac{1}{2}u^2 + \epsilon\right)\nabla\cdot(\rho\boldsymbol{u}) + \rho\boldsymbol{u}\partial_t\boldsymbol{u} + \rho\partial_t\epsilon,
\end{aligned}$$

using equation (2.1) to substitute for $\partial_t\rho$. We then use equation (2.2) to replace $\partial_t\boldsymbol{u}$, and recognizing that $\boldsymbol{u}(\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} = \boldsymbol{u}\cdot\boldsymbol{\nabla}[(1/2)u^2]$, rewrite equation (2.5) as

$$\rho\left(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\right)\epsilon + P\nabla\cdot\boldsymbol{u} = -\nabla\cdot\boldsymbol{F} + \rho q. \tag{2.29}$$

We've canceled all common factors here. Finally, we once again use equation (2.1) to set

$$P\nabla\cdot\boldsymbol{u} = -(P/\rho)(\partial_t\rho + \boldsymbol{u}\cdot\boldsymbol{\nabla}\rho) = \rho P\left(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\right)\left(\frac{1}{\rho}\right).$$

Substituting this into the left-hand of equation (2.29) and using the first law of thermodynamics (see eq. [2.21]), we obtain

$$\rho\left(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\right)\epsilon + P\nabla\cdot\boldsymbol{u} = \rho T\left(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla}\right)s. \tag{2.30}$$

For the right-hand side of equation (2.29), we expand the divergence operator in spherical symmetry and use equation (2.25) to obtain

$$-\nabla\cdot\boldsymbol{F} = -\frac{1}{r^2}\frac{\partial(r^2 F)}{\partial r} = -\rho\frac{\partial L_r}{\partial m}.$$

Putting everything together, we finally have our heat equation in Lagrangian form,

$$\frac{\partial L_r}{\partial m} = q - T\frac{\mathrm{D}s}{\mathrm{D}t}. \tag{2.31}$$

This has a simple interpretation: the change in luminosity across a mass shell is due to sources or sinks of energy and the change in the heat content of the shell.

EXERCISE 2.3 — Show that equation (2.31) can be written as

$$\begin{aligned}
\frac{\partial L_r}{\partial m} &= q - \frac{c_\rho T}{\chi_T}\left\{\frac{\mathrm{D}\ln P}{\mathrm{D}t} - \left[\chi_\rho + \chi_T(\Gamma_3 - 1)\right]\frac{\mathrm{D}\ln\rho}{\mathrm{D}t}\right\} &\quad (2.32)\\
&= q - \frac{P}{\rho(\Gamma_3 - 1)}\frac{\mathrm{D}}{\mathrm{D}t}\ln\left(\frac{P}{\rho^{\Gamma_1}}\right), &\quad (2.33)
\end{aligned}$$

where

$$\begin{aligned}
\chi_T &\equiv \frac{T}{P}\left(\frac{\partial P}{\partial T}\right)_\rho,\\
\chi_\rho &\equiv \frac{\rho}{P}\left(\frac{\partial P}{\partial \rho}\right)_T,\\
\Gamma_1 &\equiv (\partial\ln P/\partial\ln\rho)_s\,, \text{ and}\\
\Gamma_3 - 1 &\equiv \left(\frac{\partial\ln T}{\partial\ln\rho}\right)_s
\end{aligned}$$

are defined in Appendix A.1.

It is more useful, however, to work with temperature and pressure instead of entropy. Write

$$T\frac{\mathrm{D}s}{\mathrm{D}t} = T\left(\frac{\partial s}{\partial T}\right)_P\frac{\mathrm{D}T}{\mathrm{D}t} + T\left(\frac{\partial s}{\partial P}\right)_T\frac{\mathrm{D}P}{\mathrm{D}t},$$

and use the identity (see Appendix A.1)

$$\left(\frac{\partial s}{\partial P}\right)_T = -\left(\frac{\partial s}{\partial T}\right)_P\left(\frac{\partial T}{\partial P}\right)_s$$

to obtain

$$\frac{\partial L_r}{\partial m} = q - c_P\left[\frac{\mathrm{D}T}{\mathrm{D}t} - \left(\frac{\partial T}{\partial P}\right)_s\frac{\mathrm{D}P}{\mathrm{D}t}\right]. \qquad (2.34)$$

Equations (2.26), (2.27), (2.28), and (2.34), when supplemented by an equation of state, a prescription for the thermal conductivity, and the equations for nuclear heating and neutrino cooling, form the equations for stellar structure and evolution in spherical symmetry.

3
Convection

Hot air rises, as a glider pilot or hawk can tell you. The fluid velocities in question are very subsonic, so we have hydrostatic equilibrium to excellent approximation. But the fluid motions make an enormous difference for heat transport! This state of fluid motions induced by a temperature gradient is known as *convection*. You can perform the following demonstration of the onset of convection. Brew tea, and pour the hot tea into a saucepan that is on an unlit burner. Use a straw with your thumb over the top to insert a layer of cold milk under the warm tea in the saucepan. The temperature difference between the tea and milk will inhibit their mixing. Light the burner, and watch for the development of convection—you will know it when you see it.

Figure 3.1: Onset of convection in a tea-milk mixture.

3.1 Criteria for onset of convection

To understand this process, let's consider a fluid in planar geometry and hydrostatic equilibrium,

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho g. \tag{3.1}$$

Now, imagine moving a blob of fluid upwards from r to $r + h$. We move the blob slowly enough that it is in hydrostatic equilibrium with its new surroundings, $P_b(r + h) = P(r + h)$, where the subscript b refers to "blob." We do move the blob quickly enough, however, that it does not remain in *thermal* equilibrium with its surroundings; that is, we move the blob *adiabatically*. The entropy of the blob is therefore constant, $S_b(r+h) = S_b(r) = S(r)$, and is therefore not, in general, equal to the entropy of the surrounding gas at $r+h$: $S_b(r+h) \neq S(r+h)$.

As the blob rises, it displaces some of the surrounding fluid. Archimedes tells us that if the displaced fluid is less massive than the blob, then the blob will sink. We can rephrase this in terms of the volume occupied by a unit mass of fluid V: if the volume occupied by the blob is less than the volume of an equal mass of background, then the blob will sink. Translating this into an equation: if

$$\begin{aligned} V[P(r+h), S(r+h)] - V_b[P_b(r+h), S_b(r+h)] = \\ V[P(r+h), S(r+h)] - V[P(r+h), S(r)] > 0 \end{aligned} \tag{3.2}$$

then the blob will sink. If condition (3.2) is violated, the blob will continue to rise, and the system is unstable to convection. Figure 3.2 has a cartoon of this process.

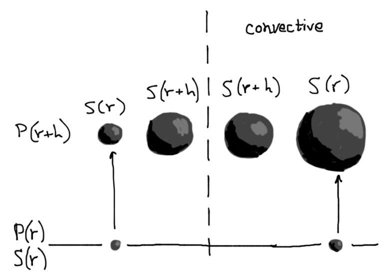


Figure 3.2: Illustration of criteria for convective instability. On the left, raising a blob a distance h adiabatically and in pressure balance with its surrounding results in a higher density $V_b < V$. This is stable: the blob will sink back. On the right, the blob is less dense and hence buoyant: it will continue to rise.

Taking h to be an infinitesimal displacement and expanding the left-hand side of equation (3.2) gives us a local condition for stability:

$$V[P(r+h), S(r)] + \left(\frac{\partial V}{\partial S}\right)_P \frac{\mathrm{d}S}{\mathrm{d}r} - V[P(r+h), S(r)] = \left(\frac{\partial V}{\partial S}\right)_P \frac{\mathrm{d}S}{\mathrm{d}r} > 0. \tag{3.3}$$

Noting that

$$\begin{aligned} \left(\frac{\partial V}{\partial T}\right)_P &= \left(\frac{\partial V}{\partial S}\right)_P \left(\frac{\partial S}{\partial T}\right)_P \\ &= \frac{C_P}{T} \left(\frac{\partial V}{\partial S}\right)_P, \end{aligned}$$

we can rewrite equation (3.3) as

$$\frac{T}{C_P} \left(\frac{\partial V}{\partial T}\right)_P \frac{\mathrm{d}S}{\mathrm{d}r} > 0.$$

Now, $(\partial V/\partial T)_P$ is positive (gas expands on being heated), so our condition for stability is simply

$$\frac{\mathrm{d}S}{\mathrm{d}r} > 0. \tag{3.4}$$

In a convectively stable star, the entropy must increase with radius. if $\mathrm{d}S/\mathrm{d}r < 0$, then convection occurs and carries high-entropy material outward, where it will eventually mix with the ambient medium. As a result, convection drives the entropy gradient toward the marginally stable configuration $\mathrm{d}S/\mathrm{d}r = 0$. If a star is fully convective and mixes efficiently, then the interior of the star lies along an adiabat.

EXERCISE 3.1 — Assuming that $\nabla \approx \nabla_{\mathrm{ad}}$ in a convective region, sketch a plot of temperature as a function of pressure for the following cases.

1. A star with a stable inner layer and a convective outer layer;
2. A star with a convective inner layer and and a stable outer layer.

Indicate on both of these plots an adiabat.

WE CAN DERIVE A CONDITION FOR CONVECTIVE STABILITY in terms of the local gradients of temperature and pressure. Writing $S = S[P(r), T(r)]$ we expand equation (3.4) to obtain

$$\frac{\mathrm{d}S}{\mathrm{d}r} = \left(\frac{\partial S}{\partial P}\right)_T \frac{\mathrm{d}P}{\mathrm{d}r} + \left(\frac{\partial S}{\partial T}\right)_P \frac{\mathrm{d}T}{\mathrm{d}r}. \tag{3.5}$$

Now, P is a monotonically decreasing function of r, which means we can use it as a spatial coordinate and write,

$$\frac{\mathrm{d}T}{\mathrm{d}r} = \left.\frac{\mathrm{d}T}{\mathrm{d}P}\right|_\star \frac{\mathrm{d}P}{\mathrm{d}r}. \tag{3.6}$$

Here $\mathrm{d}T/\mathrm{d}P|_\star$ is the slope of the $T(P)$ relation *for the stellar interior*. In particular, this is *not* a thermodynamic equality. Substituting equation (3.6) into equation (3.5), using hydrostatic equilibrium to eliminate $\mathrm{d}P/\mathrm{d}r$, and recognizing that $(\partial S/\partial T)_P = C_P/T$, we obtain

$$\frac{\mathrm{d}S}{\mathrm{d}r} = -\rho g \left[\left(\frac{\partial S}{\partial P}\right)_T + \frac{C_P}{T}\left.\frac{\mathrm{d}T}{\mathrm{d}P}\right|_\star\right]. \tag{3.7}$$

Finally, we can use the identity (see Appendix A.1)

$$\left(\frac{\partial S}{\partial P}\right)_T \left(\frac{\partial T}{\partial S}\right)_P \left(\frac{\partial P}{\partial T}\right)_S = -1 \tag{3.8}$$

to simplify equation (3.7),

$$\begin{aligned}\frac{\mathrm{d}S}{\mathrm{d}r} &= -\frac{\rho g}{P} C_P \left[\frac{P}{T}\left.\frac{\mathrm{d}T}{\mathrm{d}P}\right|_\star - \frac{P}{T}\left(\frac{\partial T}{\partial P}\right)_S\right] \\ &= -\frac{\rho g}{P} C_P \left[\nabla - \nabla_{\mathrm{ad}}\right].\end{aligned} \tag{3.9}$$

Here we have introduced the shorthand notation $\nabla \equiv \mathrm{d}\ln T/\mathrm{d}\ln P|_\star$ and $\nabla_{\mathrm{ad}} \equiv (\partial \ln T/\partial \ln P)_S$. *A mixture of uniform composition is unstable to convection if the local temperature gradient is steeper than an adiabat, i.e., if* $\nabla > \nabla_{\mathrm{ad}}$.

3.2 A second look at convective instability

Here we'll take a second-look at the convection by imagining we have a background state with $\boldsymbol{u} = 0$; we then *perturb* state by displacing fluid elements a distance $\delta\boldsymbol{r}$, and obtaining an equation of motion for $\delta\ddot{\boldsymbol{r}}$. This requires a bit of careful thought on what we are perturbing.

There are two types of perturbations. We may change a fluid quantity f at a fixed location $\boldsymbol{r}$ and time t (Fig. 3.3):

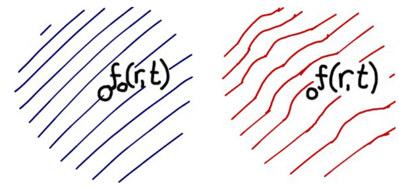


Figure 3.3: An Eulerian perturbation: we compare fluid quantities at corresponding locations.

$$\Delta f \equiv f(\boldsymbol{r}, t) - f_0(\boldsymbol{r}, t), \tag{3.10}$$

where the subscript "0" denotes the unperturbed quantity. We call Δf an *Eulerian perturbation*.

We may also change a fluid quantity f for a given fluid element; the position of this fluid element in the perturbed system is not necessarily at the same position as in the unperturbed case, however (Fig. 3.4):

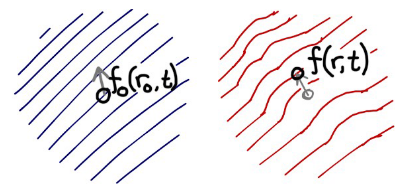


Figure 3.4: A Lagrangian perturbation: we compare fluid quantities for corresponding fluid elements.

$$\delta f \equiv f(\boldsymbol{r}, t) - f_0(\boldsymbol{r}_0, t). \tag{3.11}$$

We call δf a *Lagrangian perturbation*.

Since the fluid element is displaced $\delta\boldsymbol{r} = \boldsymbol{r} - \boldsymbol{r}_0$, we can add and subtract $f_0(\boldsymbol{r}, t)$ to eq. (3.11) and expand $f_0(\boldsymbol{r}, t)$ to first order in $\delta\boldsymbol{r}$ to obtain a relation between the two types of perturbations:

$$\delta f = \Delta f + (\delta\boldsymbol{r} \cdot \boldsymbol{\nabla}) f_0. \tag{3.12}$$

There are a few useful commutation relations that are easily proved:

$$\begin{aligned} \partial_t \Delta f &= \Delta\left(\partial_t f\right), && (3.13)\\ \boldsymbol{\nabla}\Delta f &= \Delta\boldsymbol{\nabla} f, && (3.14)\\ \frac{\mathrm{D}}{\mathrm{D}t}\delta f &= \delta\frac{\mathrm{D}f}{\mathrm{D}t}. && (3.15) \end{aligned}$$

And there are operations that do not commute:

$$\begin{aligned} \partial_t \delta f &\neq \delta\left(\partial_t f\right), && (3.16)\\ \boldsymbol{\nabla}\delta f &\neq \delta\boldsymbol{\nabla} f, && (3.17)\\ \frac{\mathrm{D}}{\mathrm{D}t}\Delta f &\neq \Delta\frac{\mathrm{D}f}{\mathrm{D}t}. && (3.18) \end{aligned}$$

One can further show that $\delta\boldsymbol{u} = (\mathrm{D}/\mathrm{D}t)\delta\boldsymbol{r}$. Finally if the fluid has unperturbed velocity $\boldsymbol{u} = 0$, then $\Delta\boldsymbol{u} = \delta\boldsymbol{u}$.

ARMED WITH THESE RELATIONS, LET US PERTURB THE MOMENTUM EQUATION by adiabatically displacing a fluid element a distance $\delta\boldsymbol{r}$. We will do this in such a way that the pressure at a fixed location does not change, i.e., $\Delta P = 0$. Of course, the pressure and density of a given fluid element will change according to the relation

$$\frac{\delta P}{P} = \Gamma_1 \frac{\delta\rho}{\rho}.$$

We'll also assume that the gravitational force does not change, $\Delta g = 0$. Our perturbed momentum equation then becomes

$$\frac{\mathrm{D}^2\delta\boldsymbol{r}}{\mathrm{D}t^2} = -\frac{1}{\rho + \Delta\rho}\boldsymbol{\nabla} P + \boldsymbol{g}.$$

Since in the unperturbed fluid $\boldsymbol{\nabla} P = \rho \boldsymbol{g}$, this equation simplifies to

$$\frac{\mathrm{D}^2 \delta \boldsymbol{r}}{\mathrm{D}t^2} = \frac{\Delta\rho}{\rho}\boldsymbol{g}. \tag{3.19}$$

Expanding,

$$\begin{aligned}\frac{\Delta\rho}{\rho} &= \frac{\delta\rho}{\rho} - \frac{1}{\rho}(\delta\boldsymbol{r}\cdot\boldsymbol{\nabla})\rho = \frac{1}{\Gamma_1}\frac{\delta P}{P} - \frac{1}{\rho}(\delta\boldsymbol{r}\cdot\boldsymbol{\nabla})\rho \\ &= \frac{1}{\Gamma_1}\frac{\Delta P}{P} + (\delta\boldsymbol{r}\cdot\boldsymbol{\nabla})\left[\frac{1}{\Gamma_1}\ln P - \ln\rho\right].\end{aligned}$$

Since by assumption $\Delta P = 0$, the radial component of equation (3.19) becomes

$$\delta\ddot{\boldsymbol{r}} = g\left[\frac{\mathrm{d}\ln\rho}{\mathrm{d}r} - \frac{1}{\Gamma_1}\frac{\mathrm{d}\ln P}{\mathrm{d}r}\right]\delta r \equiv g\mathcal{A}\delta r. \tag{3.20}$$

The quantity $\mathcal{A}$ is called the *Schwarzschild discriminant*: if $\mathcal{A} < 0$, then the motion is oscillatory with frequency $N = (-g\mathcal{A})^{1/2}$; N is called the *Brunt-Väisälä* frequency. The condition $\mathcal{A} > 0$ implies that the fluid is convectively unstable; and indeed, one can show that $\mathcal{A} > 0$ is equivalent to $\mathrm{d}S/\mathrm{d}r < 0$.

3.3 *Efficiency of Heat Transport*

A superadiabatic temperature gradient, $\nabla > \nabla_{\mathrm{ad}}$, induces convective motions. A rising blob will be hotter than its surroundings and heat will therefore be conducted from the blob to its surroundings as it rises. The efficiency by which the heat is transported determines by how much convection is able to drive the temperature gradient towards an adiabat. Clearly the gradient must be super-adiabatic to drive the convection in the first place. We shall see, however, that in stars the difference between the gradient and the adiabat are typically exceedingly small. In other words, convection is extraordinarily efficient at transporting heat.

To understand this, let's go back to equation 3.19. Write $\Delta\rho$ as stemming from differences in temperature between rising and falling blobs (recall that $\Delta P = 0$). With this substitution, we have

$$(\partial_t \boldsymbol{u} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}) = \frac{\Delta\rho}{\rho}\boldsymbol{g} = \left(\frac{\partial\ln\rho}{\partial\ln T}\right)_P \frac{\Delta T}{T}\boldsymbol{g}. \tag{3.21}$$

Our goal is to estimate the velocity of convective motions u, the departure of the temperature gradient from an adiabat ΔT, and the fraction of the total heat flux carried by convective motions from these equations.

First, the velocity. The left-hand side of equation (3.21) has a characteristic scale $\sim U^2/L$, whereas the right-hand side has a scale $g\Delta T/T$. (Recall that in an ideal gas, $(\partial\ln\rho/\partial\ln T)_P = -1$.) If we take $L \sim c_s^2/g$, a pressure scale height, than we get an estimate of the convective velocity,

$$\frac{U}{c_s} \sim \left(\frac{\Delta T}{T}\right)^{1/2}. \tag{3.22}$$

What is the heat flux carried by convection? Hot fluid rises and carries an excess of heat, per gram, of $c_P \Delta T$, giving a heat flux $\approx \rho \boldsymbol{u} c_P \Delta T$. Thus to carry a given flux F, we have

$$c_s \rho c_P T \left(\frac{\Delta T}{T}\right)^{3/2} \sim F. \tag{3.23}$$

Note that in order of magnitude, $c_P T \sim c_s^2$, so

$$\frac{U}{c_s} \sim \left(\frac{\Delta T}{T}\right)^{1/2} \sim \left(\frac{F}{\rho c_s^3}\right)^{1/3}.$$

For conditions in the solar interior, $F \ll \rho c_s^3$, and therefore the convective velocities are very subsonic. Indeed,

$$\begin{aligned}
\frac{F}{\rho c_s^3} &\sim \frac{L_\odot}{4\pi R_\odot^2} \frac{4\pi R_\odot^3}{3M_\odot} \left(\frac{R_\odot}{GM_\odot}\right)^{3/2} \\
&\sim \frac{L_\odot}{GM_\odot^2/R_\odot} \left(\frac{R_\odot^3}{GM_\odot}\right)^{1/2} \\
&\sim \frac{t_{\mathrm{dyn}}}{t_{\mathrm{KH}}} \ll 1.
\end{aligned}$$

That is, the ratio of the solar flux to what could be carried for near-sonic convective motions is of the order of the dynamical timescale to the Kelvin-Helmholtz timescale. We therefore expect that in a convective region, slow circulation will produce a temperature gradient that is very nearly adiabatic. This argument breaks down near the surface, where the cooling time of a fluid layer (the “local” Kelvin-Helmholtz timescale) can be small.

3.4 Turbulence

From the discussion of the previous section, it might seem possible, given the boundary conditions, of solving for the flow planform, that is, the velocity profile $\boldsymbol{u}(\boldsymbol{x}, t)$. This is decidedly not the case, however: the flow is turbulent, with intermittent velocity fluctuations seen over a large dynamical range of spatial and temporal scales. Modeling of such flows is a vexing problem in fluid dynamics.

To explore this topic a bit further, we need to introduce the concept of dynamical similarity. Suppose you want to optimize a wing shape for an aircraft, and you wish to test its performance in a wind tunnel. Why should you expect that the behavior of a model wing will have any relation to the full-scale one?

To see how this works, start with the Navier-Stokes equation (for simplicity, we’ll keep it in one dimension):

$$(\partial_t + u \cdot \partial_x) u = -\frac{1}{\rho} \partial_x P + \nu \partial_x^2 u. \tag{3.24}$$

Here ν is the coefficient of kinematic viscosity, with dimensions $[\nu] \sim [\text{length}]^2 \cdot [\text{time}]^{-1}$. Let's recast equation (3.24) into dimensionless form by scaling our variable: let L and U represent the characteristic length and velocity scales, and define the dimensionless variables $\tilde{x} = x/L$ and $\tilde{u} = u/U$. This choice then implicitly defines the time variable, $\tilde{t} = t \cdot U/L$. Upon changing to the variables $\tilde{x}$, $\tilde{u}$, and $\tilde{t}$, and writing the equation of state as $P = c_s^2 \rho$ (appropriate for adiabatic flow—we are ignoring heat conduction), we obtain the equation

$$(\partial_{\tilde{t}} + \tilde{u} \cdot \partial_{\tilde{x}})\tilde{u} = -\left\{\frac{c_s^2}{U^2}\right\} \partial_{\tilde{x}} \ln \tilde{\rho} + \left\{\frac{\nu}{UL}\right\} \partial_{\tilde{x}}^2 \tilde{u}. \tag{3.25}$$

Each term in this equation is dimensionless. The physical characteristics of the fluid and the scales involved are described by just two dimensionless parameters:

$\mathrm{Ma} \equiv \dfrac{U}{c_s}$ Mach number (measure of compressibility)

$\mathrm{Re} \equiv \dfrac{UL}{\nu}$ Reynolds number (measure of viscous forces)

So, if we build a model wing at a certain scale and place it in a wind tunnel with a certain velocity, then by adjusting the density and temperature (and hence the sound speed and viscosity) to the desired Ma and Re, the flow pattern in our model will faithfully replicate the flow in the actual system.

For stellar convection, $\mathrm{Ma} \ll 1$. What about Re? In typical astrophysical plasmas, the large lengthscales make Re ludicrously large. Terrestrial experiments and simulations cannot approach this regime. Experimentally, when $\mathrm{Re} \gtrsim 10^3$, then flow becomes *turbulent*: the velocity has strong intermittent fluctuations across a wide range of lengthscales and timescales. How to characterize the flow in such a case? It is useful to describe the flow in terms of correlated velocities—an "eddy"—which have some lengthscale.

Suppose we pass water through a pipe that has an embedded mesh screen (Fig. 3.5). For sufficiently large $\mathrm{Re} = UL/\nu$, where L is the mesh spacing, the downstream flow becomes turbulent. The turbulent eddies are damped. Now an eddy of size λ has an effective Reynolds number $\mathrm{Re}(\lambda) = U(\lambda)\lambda/\nu$; if this is very large, then molecular viscosity cannot be the reason for damping fluid motions on that scale. Instead what happens is that an eddy with lengthscale λ and velocity scale $U(\lambda)$ drives eddies on a smaller scale $\lambda' < \lambda$. These in terms drive still smaller eddies, which in turn drive still smaller eddies, and so on, until eventually very tiny eddies are excited, with size $\lambda_\nu \sim \nu/U(\lambda_\nu)$; and these eddies *are* damped by viscosity!

Kolmogorov argued that in steady-state, intermediate-sized eddies (i.e., those with lengthscales $\nu/U(\lambda) \ll \lambda \ll L$) are neither losing or

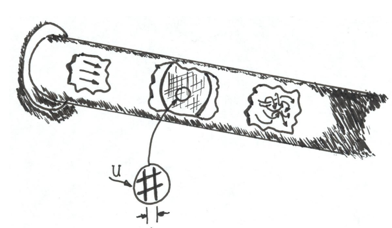


Figure 3.5: A simple mechanism for generating turbulence. A flow of water in a pipe (upstream velocity U) flows through a mesh (spacing L). If $\mathrm{Re} = UL/\nu$ is sufficiently large, the downstream flow becomes turbulent.

gaining energy and hence were transferring energy to smaller scales at the same rate as they were being driven; further, this rate at which energy is being transferred to smaller scales is just the net rate of dissipation in the fluid (which is done by the smallest eddies). The huge dynamic range in lengthscales implies that the velocity of the eddy should not depend on either L or ν, and hence $U(\lambda)$ can only be a function of λ (length) and the rate of energy dissipation per unit mass ϵ (energy/mass/time $\sim$ length^2/time^3). There is only one way to combine these quantities to form something with a dimension of length/time, and so

$$U(\lambda) \sim \epsilon^{1/3}\lambda^{1/3}. \tag{3.26}$$

This is seen experimentally: in flows with a large dynamic range of scales, the velocity spectrum follows a power-law with this slope, over an intermediate range of scales, the *inertial range*. A good example is the flow in a tidal channel[1].

[1] H. L. Grant, R. W. Stewart, and A. Moilliet. Turbulence spectra from a tidal channel. *Journal of Fluid Mechanics*, 12: 241–268, 1962

4
Polytropes and the Lane-Emden Equation

In the previous chapters, we derived the basic equations of stellar structure. To construct detailed models is clearly an involved task because there are many linked physical variables. For example, the equation of hydrostatic balance, eq. (1.7), contains both ρ and P; in order to connect these two variables we have to know the temperature, so in general we need in include an equation for heat transport and heat flux. Before diving into all that, however, we'll take a little digression to solve some simplified stellar models known as **polytropes**; these are useful not only for historical reasons, but they also allow for quick analytical calculations.

4.1 Historical Background

To understand where the term polytropic comes from, let's first consider an ideal gas in hydrostatic equilibrium, and furthermore suppose that the temperature and density lie along an adiabat. In that case we have the following relations:

$$T\rho^{1-\gamma} = \text{const}; \quad P\rho^{-\gamma} = \text{const}; \quad TP^{(1-\gamma)/\gamma} = \text{const}, \tag{4.1}$$

which you will recall from elementary thermodynamics. Here $\gamma = C_P/C_\rho$ is the ratio of specific heats. The equation of state is

$$P = \left(\frac{N_A k_B}{\mu}\right)\rho T, \tag{4.2}$$

where μ is the mean molecular weight and the quantity in parenthesis is $C_P - C_\rho = N_A k_B/\mu$. We might imagine, for example, a rising plume of hot air in Earth's atmosphere. There is one snag with this analysis for the troposphere, however: the condensation of water vapor means that one cannot hold $\mathrm{d}s = 0$ in a rising plume of hot, moist air. Attempting to model this moist convection in the Earth's troposphere motivated work in the early 1900's by Kelvin, Lane, Emden, and others to consider a more general problem, in which $T\mathrm{d}s = C\mathrm{d}T$, where C is a constant. A configuration for which this is true is called **polytropic**. An adiabat is a special case of a polytrope with $C = 0$. Writing the first law of thermodynamics as $T ds = C_\rho dT - (P/\rho^2) d\rho$, substituting for $T ds$, and using equation (4.2), we obtain

$$(C_\rho - C) \frac{dT}{T} = (C_P - C_\rho) \frac{d\rho}{\rho}.$$

This equation has a solution $T \propto \rho^{(C_P - C_\rho)/(C_\rho - C)}$. Comparing this solution with equation (4.1), we can define a *polytropic exponent*, $\gamma' = (C_P - C)/(C_\rho - C)$. Then equation (4.1) holds with γ replaced by γ'. The advantage of this approximation is that it relates density to pressure so that one can solve the equation of hydrostatic equilibrium without simultaneously having to solve for $T(r)$.

4.2 The Lane-Emden Equation and Solution

To use the polytropic equation of state, write the pressure P as

$$P(r) = K\rho^{1+1/n}(r) \tag{4.3}$$

where n and K are constants. Further define the dimensionless variable θ via

$$\rho(r) = \rho_c \theta^n(r), \tag{4.4}$$

where the subscript c denotes the central value at $r = 0$. Note that since

$$P(r) \propto \rho \times \rho^{1/n} \propto \rho\theta,$$

the quantity θ plays the role of a dimensionless temperature for an ideal non-degenerate gas.

Substitute equations (4.3) and (4.4) into Poisson's equation,

$$\nabla^2 \Phi = 4\pi G \rho, \tag{4.5}$$

and the equation for hydrostatic equilibrium,

$$\boldsymbol{\nabla} P = -\rho \boldsymbol{\nabla} \Phi, \tag{4.6}$$

to obtain the *Lane-Emden* equation for index n,

$$\xi^{-2} \frac{d}{d\xi} \left(\xi^2 \frac{d\theta}{d\xi} \right) = -\theta^n. \tag{4.7}$$

Here $\xi = r/r_n$ is the dimensionless coordinate, and

$$r_n = \left[\frac{(n+1)P_c}{4\pi G \rho_c^2} \right]^{1/2} \tag{4.8}$$

is the radial length scale.

For a stellar model described by a single polytropic relation, the appropriate boundary conditions are

$$\theta(\xi)\big|_{\xi=0} = 1, \tag{4.9}$$

$$\theta'(\xi)\big|_{\xi=0} = 0. \tag{4.10}$$

From the form of equation (4.7), it follows that $\theta(-\xi) = \theta(\xi)$, that is, the solution is *even* in ξ. A power-series solution of θ out to order ξ^6 is

$$\theta(\xi) = 1 - \frac{1}{6}\xi^2 + \frac{n}{120}\xi^4 - \frac{n(8n-5)}{15120}\xi^6 + \mathcal{O}(\xi^8) \tag{4.11}$$

There are analytical solutions for $n = 0$, 1, and 5:

$$\theta_0(\xi) = 1 - \frac{\xi^2}{6} \tag{4.12}$$

$$\theta_1(\xi) = \frac{\sin\xi}{\xi} \tag{4.13}$$

$$\theta_5(\xi) = \left(\frac{3}{3+\xi^2}\right)^{1/2}. \tag{4.14}$$

Finally, the radius of the stellar model is determined by the location of the first zero, ξ_1, where $\theta(\xi_1) = 0$. For example, if $n = 0$ (eq. [4.12]), $\xi_1 = \sqrt{6}$. Note that if $n = 5$ there is no root; $\theta_5(\xi) > 0, \forall \xi > 0$. A sample of Lane-Emden solutions for various indices is shown in Figure 4.1.

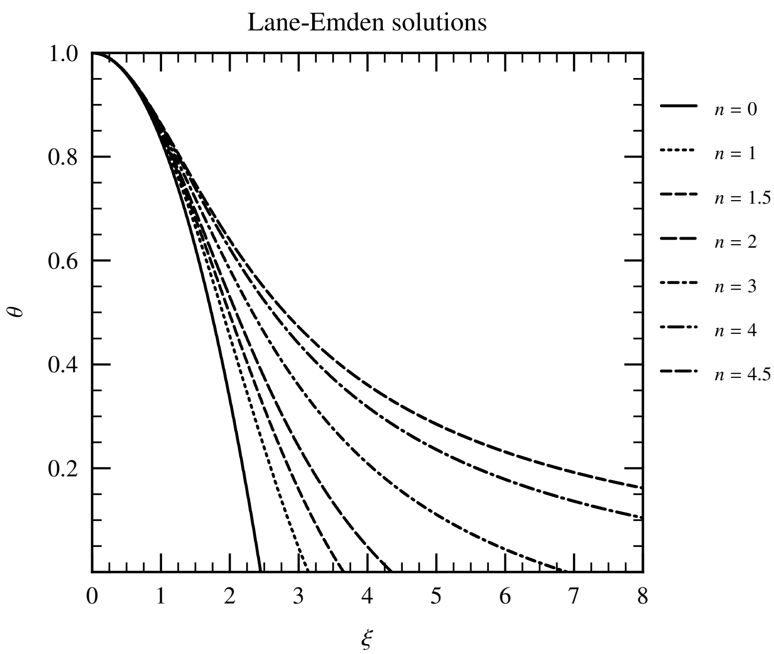


Figure 4.1: Solutions of the Lane-Emden equation for selected values of the index n.

4.3 *Some Useful Relations*

First, let's get the mass of our polytropic sphere. To do this, we write the integral

$$M = \int_0^R 4\pi r^2 \rho \,\mathrm{d}r$$

and make the substitutions $r = r_n\xi$, $R = r_n\xi_1$, and $\rho = \rho_c\theta^n(\xi)$ to obtain

$$M = 4\pi r_n^3 \rho_c \int_0^{\xi_1} \xi^2 \theta^n(\xi)\,\mathrm{d}\xi.$$

Using equation (4.7) , the integrand can be written as a perfect differential, so we get

$$M = 4\pi r_n^3 \rho_c \left(-\xi_1^2 \theta_1'\right). \tag{4.15}$$

Here I define the shorthand $\theta_1' \equiv [\mathrm{d}\theta(\xi)/\mathrm{d}\xi]_{\xi=\xi_1}$. Substituting $r_n = R/\xi_1$ and dividing by 3 allows us to get a formula relating the central density to the mean density,

$$\rho_c = \frac{3M}{4\pi R^3}\left(-\frac{\xi_1}{3\theta_1'}\right). \tag{4.16}$$

For the solutions shown in Figure 4.1, we have the following values of $\rho_c/\bar{\rho}$, as shown in Table 4.1. As the index n increases, the configuration becomes more and more concentrated toward the center.

n	0	1.0	1.5	2.0	3.0	4.0
ξ_1	2.449	3.142	3.654	4.353	6.897	14.972
$-\theta_1'$	0.8165	0.3183	0.2033	0.1272	0.04243	0.008018
$\rho_c/\bar{\rho}$	1.00	3.29	5.99	11.41	54.18	622.4

Table 4.1: Properties of the Lane-Emden solutions.

Starting from equation (4.15), we can substitute for r_n using equation (4.8) and ρ_c using equation (4.16) to get an equation for the central pressure,

$$P_c = \frac{GM^2}{R^4}\frac{1}{4\pi(n+1)(-\theta_1')^2}. \tag{4.17}$$

For an ideal gas, $P_c = (N_\mathrm{A} k_\mathrm{B}/\mu)\rho_c T_c$ with μ being the mean molecular weight, we can solve for the central temperature,

$$T_c = \left(\frac{\mu}{N_\mathrm{A} k_\mathrm{B}}\right)\left(\frac{GM}{R}\right)\frac{1}{(n+1)\xi_1(-\theta_1')}. \tag{4.18}$$

Finally, starting from equation (4.17), substituting for P_c using equation (4.3), and eliminating ρ_c using equation (4.16), we obtain a relation between mass and radius in terms of K and n,

$$M^{1-1/n} = \left[\frac{K(n+1)}{G(4\pi)^{1/n}}\xi_1^{1+1/n}\left(-\theta_1'\right)^{1-1/n}\right] R^{1-3/n}. \tag{4.19}$$

Alternatively, one could use this equation to fit K to a star of known M and R.

EXERCISE 4.1 — Derive equations (4.15)–(4.19). Explain what the mass-radius relation, eq. (4.19), means for the cases $n = 1$ and $n = 3$.

Finally, we can derive a formula for the gravitational energy of our polytropic sphere. First, we can integrate the equation for the energy,

$$E_\mathrm{grav} = -G\int_0^M \frac{m}{r}\,\mathrm{d}m,$$

by parts to obtain

$$E_{\text{grav}} = -\frac{GM^2}{2R} - \frac{1}{2}\int_0^R \frac{Gm^2}{r^2}\,\mathrm{d}r = -\frac{GM^2}{2R} - \frac{1}{2}\int_0^R \frac{\mathrm{d}\Phi}{\mathrm{d}r} m\,\mathrm{d}r. \tag{4.20}$$

If we define our zero of energy to be such that $\Phi(R) = 0$, then we can integrate by parts again to obtain

$$E_{\text{grav}} = -\frac{GM^2}{2R} + \frac{1}{2}\int_0^R \Phi\,\mathrm{d}m. \tag{4.21}$$

We can rewrite the equation of hydrostatic equilibrium as

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{\mathrm{d}\Phi}{\mathrm{d}r}\rho$$

and use equation (4.3) to eliminate P to obtain

$$\frac{\mathrm{d}\Phi}{\mathrm{d}r} = (1+n)K\rho^{1/n}.$$

Integrating from a point in the star r to R, and again using choosing $\Phi(R) = 0$, we obtain

$$\Phi(r) = -(1+n)K\rho(r)^{1/n} = -(1+n)\frac{P(r)}{\rho(r)}. \tag{4.22}$$

Inserting equation (4.22) into equation (4.21), we have

$$E_{\text{grav}} = -\frac{GM^2}{2R} - \frac{1+n}{2}\int_0^M \frac{P}{\rho}\,\mathrm{d}m = -\frac{GM}{2R} + \frac{1+n}{6}E_{\text{grav}}, \tag{4.23}$$

where we used equation (1.13) to relate the integral of P/ρ to E_{grav}. Solving equation (4.23) for E_{grav} gives us the desired result,

$$E_{\text{grav}} = -\frac{3}{5-n}\frac{GM^2}{R}. \tag{4.24}$$

Note that solutions with $n > 5$ have a positive gravitational energy.

EXERCISE 4.2 — For a fully convective star with an ideal gas equation of state, how is the polytropic constant K related to the entropy? Derive a formula for R in terms of M and s in this case. What happens to the star if the entropy increases, i.e., heat is added to it?
Hint: Recall from thermodynamics that the entropy per unit mass of an ideal gas is

$$s = \frac{k_B N_A}{\mu}\left\{\frac{5}{2} + \ln\left[\frac{\mu}{\rho N_A}\left(\frac{\mu k_B T}{2\pi N_A \hbar^2}\right)^{3/2}\right]\right\}. \tag{4.25}$$

Use the Lane-Emden solution to compute the specific entropy, per unit mass, in terms of the central temperature T and the stellar mass M: $s = s(T_c, M)$. From this expression, compute the "gravothermal" specific heat

$$c_\star = T_c \frac{\partial s(T_c, M)}{\partial T_c}, \tag{4.26}$$

and comment on its physical significance.

5
The Equation of State

In statistical equilibrium, we can describe a system of particles by a distribution function $f(\boldsymbol{p}, \boldsymbol{x})\, \mathrm{d}^3p\, \mathrm{d}^3x$, such that the number of particles is

$$N = \int \mathrm{d}^3p\, \mathrm{d}^3x\, f(\boldsymbol{p}, \boldsymbol{x}), \tag{5.1}$$

where the integration is over the phase spaces of momentum and position coordinates $(\boldsymbol{p}, \boldsymbol{x})$. In an ideal gas, the particles do not interact. In such a case, the distribution function $f = f(\boldsymbol{p})$ does not depend on position. The integration over d^3x just gives a factor of the volume, so the number density is $n = \int \mathrm{d}^3p f(\boldsymbol{p})$. From equation (5.1), we can get our other thermodynamic quantities: for example.

$$\frac{E}{V} \equiv u = \int \epsilon(\boldsymbol{p}) f(\boldsymbol{p})\, d^3p \qquad \text{energy per unit volume;} \tag{5.2}$$

$$P = \int (\boldsymbol{p} \cdot \boldsymbol{e}_z)(\boldsymbol{v} \cdot \boldsymbol{e}_z) f(\boldsymbol{p})\, d^3p \qquad \text{pressure.} \tag{5.3}$$

Here ϵ is the particle energy and v the velocity.

EXERCISE 5.1 — For an isotropic momentum distribution, show that

$$P = \frac{1}{3} \int |p||v| f(\boldsymbol{p})\, \mathrm{d}^3\boldsymbol{p}.$$

Then show that for a non-relativistic gas, $P = (2/3)E/V$, and that for a relativistic gas, $P = (1/3)E/V$.

5.1 Connection to thermodynamics

Once we have the distribution functions, we can get all of the other thermodynamic properties from the thermodynamic relations: in what follows let $N = nV$ be the total number of particles, with V being the volume of the system. The total energy is then $E = uV$, and the entropy is

S, and we have

$$F = E - TS, \qquad \text{Helmholtz free energy,} \tag{5.4}$$

$$H = E + PV, \qquad \text{Enthalpy,} \tag{5.5}$$

$$\mu N = G = F + PV, \qquad \text{Gibbs free energy.} \tag{5.6}$$

For example, we have in the non-degenerate limit that

$$\mu = k_{\mathrm{B}} T \ln K = k_{\mathrm{B}} T \ln \left[\frac{n}{g} \left(\frac{2\pi\hbar^2}{m k_{\mathrm{B}} T} \right)^{3/2} \right], \tag{5.7}$$

and so we could write the entropy per unit mass as

$$\begin{aligned} s \equiv \frac{S}{Nm} &= \frac{1}{Nm} \frac{E + PV - \mu N}{T} \\ &= \frac{k_{\mathrm{B}}}{m} \left\{ \frac{5}{2} + \ln \left[\frac{g}{n} \left(\frac{m k_{\mathrm{B}} T}{2\pi\hbar^2} \right)^{3/2} \right] \right\}. \end{aligned} \tag{5.8}$$

In this equation I have used the ideal non-degenerate values $E = (3/2) N k_{\mathrm{B}} T$, $PV = N k_{\mathrm{B}} T$ and have denoted the mass per particle as m and the degeneracy of the spin-states as g.

EXERCISE 5.2 — In an external field (i.e., gravitational) the chemical potential, which is the change in energy when the number of particles is increased, must include the potential. Consider an ideal gas in a planar atmosphere of constant gravitational acceleration g. Write the chemical potential as $\mu(z) = \mu_{\mathrm{id}} + \Phi$, where μ_{id} is the chemical potential for an ideal gas in the absence of gravity, and Φ is the gravitational potential. For an atmosphere in complete equilibrium ($\mu = \text{const}$, $T = \text{const}$), calculate the pressure as a function of position, $P = P(z)$, and show that it agrees with considerations from hydrostatic balance, equation (1.8).

5.2 *An ideal Fermi gas*

For *fermions*, particles with half-integer spin, it can be shown that

$$f(p) = \frac{g}{(2\pi\hbar)^3} \left[\exp\left(\frac{\epsilon - \mu}{k_{\mathrm{B}} T} \right) + 1 \right]^{-1}. \tag{5.9}$$

In this equation $\epsilon(p)$ is the energy of a particle, μ is the *chemical potential*, T is the temperature, and g denotes the number of particles that can occupy the same energy level (for spin-1/2 particles, $g = 2$). The connection to thermodynamics is via the relations

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N,V}, \qquad -\frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right)_{E,V},$$

and

$$TS = PV - \mu N + E;$$

these are derived in standard texts. Let's explore what happens in various limits.

Non-degenerate, non-relativistic limit

First, let's take $K \equiv \exp(\mu/k_B T) \ll 1$. (In the literature, K is called the *fugacity*.) Then in equation (5.9), we see that the exponential term dominates. If our system is isotropic, then $d^3p = 4\pi p^2 dp$, and we'll use this substitution from now on. We then have for the number density

$$n(\mu, T) = \frac{gK}{2\pi^2\hbar^3} \int_0^\infty \exp\left(-\frac{\epsilon}{k_B T}\right) p^2 dp. \tag{5.10}$$

To do the integral, notice that since $2m\epsilon = p^2$, we have $p^2\, dp = m(2m\epsilon)^{1/2}\, d\epsilon$; making the substitution $x = \epsilon/(k_B T)$, we get

$$n(\mu, T) = \frac{gK}{2\pi^2\hbar^3} \sqrt{2}(mk_B T)^{3/2} \int_0^\infty x^{1/2} e^{-x}\, dx. \tag{5.11}$$

You have all struggled with this integral in your past, but to avoid unpleasant flashbacks, I will just tell you that it is $\sqrt{\pi}/2$. So, we have our first result (but we still don't know what it means),

$$n(\mu, T) = K\left[g\left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}\right]. \tag{5.12}$$

Let's forge on a little further, though, and try to get the energy per unit voume u. Once we have u, we know we can get the pressure from the relation for a non-relativistic gas, $P = 2/3\ u$. Using equations (5.2) and (5.9),

$$u(\mu, T) = \frac{gK}{2\pi^2\hbar^3} \int_0^\infty \exp\left(-\frac{\epsilon}{k_B T}\right) \epsilon p^2 dp. \tag{5.13}$$

Let's repeat our trick of changing variables from p to $x = \epsilon(p)/k_B T$; we then have

$$u(\mu, T) = \frac{gK}{2\pi^2\hbar^3} \sqrt{2} k_B T (mk_B T)^{3/2} \int_0^\infty x^{3/2} e^{-x}\, dx. \tag{5.14}$$

Did you notice that if we integrate by parts,

$$\int_0^\infty x^{3/2} e^{-x}\, dx = \frac{3}{2}\int_0^\infty x^{1/2} e^{-x}\, dx = \frac{3}{4}\sqrt{\pi},$$

we get the integral we already solved in equation (5.11)? Putting everything together and using the expression for n (eq. [5.12]), we have

$$u(\mu, T) = \frac{3}{2}\left\{K\left[g\left(\frac{mk_B T}{2\pi\hbar^2}\right)^{3/2}\right]\right\} k_B T = \frac{3}{2} n k_B T. \tag{5.15}$$

which gives us the pressure,

$$P = \frac{2}{3}u = nk_B T. \tag{5.16}$$

Whoo-hoo! We've rediscovered the ideal gas.

Now we have to understand this chemical potential μ. We can solve equation (5.12) for μ,

$$\exp\left(\frac{\mu}{k_{\mathrm{B}}T}\right) = K = n\left[g\left(\frac{mk_{\mathrm{B}}T}{2\pi\hbar^2}\right)^{3/2}\right]^{-1}. \qquad (5.17)$$

Now K is dimensionless, a number, so the thing in [ ] must have dimensions of number density. Let's call it n_Q. Our chemical potential is then $\mu = k_{\mathrm{B}}T\ln(n/n_Q)$. To understand the significance of n_Q, let's calculate the uncertainty in position of a particle having energy $k_{\mathrm{B}}T$; from Heisenberg, we have

$$\Delta x \approx \frac{\hbar}{\Delta p} \sim \frac{\hbar}{\sqrt{mk_{\mathrm{B}}T}}$$

where I am dropping numerical factors and I've made the substitution $\Delta p \sim p \approx \sqrt{mk_{\mathrm{B}}T}$. Now what happens if I pack the particles so that on average there are g particles per box of volume $(\Delta x)^3$? In that case the density would be $n = g(\sqrt{mk_{\mathrm{B}}T}/\hbar)^3 \approx n_Q$. So, what appears in the chemical potential is the ratio of the density to that density at which the particles are packed so closely that the uncertainty in their positions is the same size as the typical inter-particle spacing. In the ideal-gas limit $K \ll 1$, which makes sense: $n \ll n_Q$, so the particles are very far apart compared to their thermal de Broglie wavelengths, and quantum effects ought to be unimportant.

EXERCISE 5.3 — Get the first order corrections to the Maxwell-Boltzmann gas. Take the fugacity $K \ll 1$, and expand the Fermi-Dirac distribution (eq. [5.9]) to lowest order in $K\exp[-\epsilon/(k_{\mathrm{B}}T)]$. Show that

$$\begin{aligned} n(K,T) &= n_0\left(1 - 2^{-3/2}K\right) \\ u(K,T) &= \frac{3}{2}n_0k_{\mathrm{B}}T\left(1 - 2^{-5/2}K\right), \end{aligned}$$

where n_0 is the density in the limit $K \to 0$. Then derive the equation of state $P = P(n,T)$ to lowest order in K. For a given (n,T), is the pressure larger or smaller than that of the ideal Maxwell-Boltzmann limit?

Degenerate, non-relativistic limit

When $n \gtrsim n_Q$, we can no longer use the approximation $K \ll 1$, so let's go to the opposite limit, for which $\mu \gg k_{\mathrm{B}}T$. In this case, notice from equation (5.9) that

$$f(p) \approx \frac{g}{(2\pi\hbar)^3}\begin{cases} 1 & \epsilon < \mu \\ 0 & \epsilon > \mu \end{cases}. \qquad (5.18)$$

We can think of this as inserting g particles in each energy level, starting with the lowest energy level and continuing until all of the particles are used. The last particle is inserted with energy $\epsilon \approx \mu$. The only levels that will be partially filled will be those lying in a thin band $\epsilon \approx \mu \pm k_B T$. If that is the case, we can make the following approximation. Let's take the limit $T \to 0$, and define the *Fermi energy* by $\epsilon_F = \mu|_{T\to 0}$ and the *Fermi momentum* by $p_F = \sqrt{2m\epsilon_F}$. We can then write equation (5.1) as

$$n(\mu) = \frac{1}{\pi^2 \hbar^3} \int_0^{p_F} p^2 dp, \tag{5.19}$$

since the integrand is zero for $p > p_F$. The main application is for electrons, which are spin one-half, so we substitute $g = 2$. Now this is an easy integral,

$$n(\mu) = \frac{p_F^3}{3\pi^2\hbar^3} = \frac{(2m\epsilon_F)^{3/2}}{3\pi^2\hbar^3}, \tag{5.20}$$

or $\mu \approx \epsilon_F = (3\pi^2 n)^{2/3}\hbar^2/(2m)$. Let's get the energy per unit volume and the pressure,

$$u(\mu) = \frac{1}{\pi^2\hbar^3}\int_0^{p_F} \frac{p^2}{2m} p^2\, dp = \frac{p_F^5}{5\pi^2\hbar^3}. \tag{5.21}$$

Comparing this with equation (5.20), we have

$$u = \frac{3}{5} n\epsilon_F, \tag{5.22}$$

$$P = \frac{2}{5} n\epsilon_F. \tag{5.23}$$

To lowest order, neither u nor P depend on T. Substituting for ϵ_F in equation (5.23) gives us the equation of state,

$$P = \frac{2}{5}\left(3\pi^2\right)^{2/3} \frac{\hbar^2}{2m} n^{5/3}. \tag{5.24}$$

Notice that in equation (5.17), $n_Q \propto m^{3/2}$. This means that at any given temperature, n_Q for electrons is $1836^{3/2} = 80,000$ times smaller than it is for protons, not to mention helium or heavier nuclei. As a result, the electrons will become degenerate ($n \gtrsim n_Q$) at a much lower mass density than the ions. A common circumstance, then, is to have a mixture of degenerate electrons and ideal ions (we will deal with non-ideal corrections due to electric forces later).

Now, let's estimate the boundary between the non-degenerate and degenerate regimes. At a given temperature, we know in the low-density limit that the electrons obey the ideal gas law (eq. [5.16]) and in the high-density limit the electrons are degenerate (eq. [5.24]). So, let's extrapolate our two limiting expressions for the pressure and see where they meet,

$$n_e k_B T = P_{e,\mathrm{ideal}} \sim P_{e,\mathrm{deg.}} = \frac{2}{5} n_e \epsilon_F, \tag{5.25}$$

or $\epsilon_F \approx k_B T$. No surprise here. Notice that the ratio

$$\frac{\epsilon_F}{k_B T} \sim \frac{(3\pi^2)^{2/3}\hbar^2}{2m_e k_B T} n_e^{2/3} \sim \left(\frac{n_e}{n_Q}\right)^{2/3}, \tag{5.26}$$

so marking the onset of degeneracy with $\epsilon_{\mathrm{F}} \sim k_{\mathrm{B}} T$ also makes sense from that aspect as well. Our boundary in the density-temperature plane between the non-degenerate and degenerate regimes is then determined by setting $k_{\mathrm{B}} T = \epsilon_{\mathrm{F}}$,

$$T = \frac{(3\pi^2)^{2/3}\hbar^2}{2m_e k_{\mathrm{B}}} \left(\frac{Y_e \rho}{m_{\mathrm{u}}}\right)^{2/3} = 3.0 \times 10^5\,\mathrm{K}\,(Y_e\rho)^{2/3}. \tag{5.27}$$

If the temperature falls below this value, the electrons will be degenerate. Here Y_e is the electron molar fraction, or electron abundance and m_{u} is the atomic mass unit; consult §2.2 for details.

EXERCISE 5.4 — Repeat the derivation of equation (5.22) and (5.23) for a relativistic Fermi gas. What is the expression for the temperature at which the gas becomes degenerate (cf. eq. [5.27]) in this case?

5.3 *Fermi-Dirac integrals*

This condition for the onset of degeneracy, eq. (5.27), is only a rule-of-thumb; in any serious calculation we would want to calculate the electron thermal properties from the exact integrals

$$n(\mu, T) = \frac{\sqrt{2}(m k_{\mathrm{B}} T)^{3/2}}{\pi^2\hbar^3} \int_0^\infty \frac{x^{1/2}\,dx}{\exp(x-\psi)+1} \tag{5.28}$$

$$P(\mu, T) = \frac{(2 m k_{\mathrm{B}} T)^{3/2}(k_{\mathrm{B}} T)}{3\pi^2\hbar^3} \int_0^\infty \frac{x^{3/2}\,dx}{\exp(x-\psi)+1}, \tag{5.29}$$

where $\psi = \mu/(k_{\mathrm{B}} T)$. These integrals cannot be done analytically, but they occur so frequently that there are many published tables and numerical approximation schemes[1]. Specifically, the *non-relativistic Fermi-Dirac integral of order* ν is defined as

$$F_\nu(\psi) = \int_0^\infty \frac{x^\nu\,dx}{\exp(x-\psi)+1}. \tag{5.30}$$

[1] F. X. Timmes and F. D. Swesty. The Accuracy, Consistency, and Speed of an Electron-Positron Equation of State Based on Table Interpolation of the Helmholtz Free Energy. *ApJS*, 126:501, 2000

One can (numerically) invert equation (5.28) to solve for the chemical potential $\psi k_{\mathrm{B}} T$.

The general case for a relativistic Fermi gas is left as an exercise.

5.4 *Relativistic photon gas*

Photons are *bosons*—they have spin 1. For bosons, the distribution function is similar to that in equation (5.9), but with the $+1$ replaced by -1 in the denominator. In addition, photon number is not conserved: one can freely create and destroy photons. This implies that their chemical potential is zero. Also, $g = 2$ for photons: there are two independent polarization modes. Putting all of these together, we can write energy per unit volume as

$$u = \frac{1}{\pi^2 \hbar^3} \int_0^\infty \epsilon p^2 \left[\exp\left(\frac{\epsilon}{k_B T} \right) - 1 \right]^{-1} dp. \tag{5.31}$$

Now, use the fact that $p = \epsilon/c$ and change variables to $x = \epsilon/(k_B T)$ to get

$$u = \frac{k_B^4 T^4}{\pi^2 c^3 \hbar^3} \int_0^\infty \frac{x^3\, dx}{e^x - 1}.$$

This integral is a classic and is equal to $\pi^4/15$. Hence the energy per unit volume and the pressure are

$$u = \left(\frac{k_B^4 \pi^2}{15 c^3 \hbar^3} \right) T^4 = a T^4 \tag{5.32}$$

$$P = \frac{1}{3} a T^4. \tag{5.33}$$

In CGS units, $a = 7.566 \times 10^{-15}\,\mathrm{erg\,cm^{-3}\,K^{-4}}$. With this energy and pressure, we can compute the other thermodynamical quantities.

EXERCISE 5.5 — Show that the entropy of a blackbody radiation is

$$S_{rad} = \frac{4}{3} a T^3 V.$$

Now consider a mixture of blackbody radiation and an ideal gas. For the ideal gas, the entropy is just

$$S_{gas} = N \mu m_u s,$$

where s is given by Eq. (4.25). Along an adiabat, $dS = d(S_{rad} + S_{gas}) = 0$; use this to find an expression for ∇_{ad} in terms of

$$\beta \equiv \frac{P_{gas}}{P},$$

where $P = P_{gas} + P_{rad}$.
Show that the expression for ∇_{ad} has the correct limiting values for $\beta \to 0$ and $\beta \to 1$.

5.5 *Chemical Equilibrium: The Saha Equation*

Consider a reaction, $A + B + \ldots \to C + D + \ldots$. When this reaction comes into equilibrium, we are at a maximum in entropy, and the condition for equilibrium is that the energy cost, at constant entropy to run the reaction in the forward direction is the same as to run the reaction in reverse. This can be expressed in terms of chemical potentials as

$$\mu_A + \mu_B + \ldots \to \mu_C + \mu_D + \ldots \tag{5.34}$$

Note in this formalism that a reaction $2A \to B$ would be expressed as $2\mu_A = \mu_B$.

As a worked example, we consider the ionization equilibrium of hydrogen,

$$H^+ + e \to H.$$

To use equation (5.34), we need to have both sides on the same energy scale. The reaction in the exothermic direction; that is, heat is evolved if the reaction proceeds as written. This means that the right-hand side is more bound, and its minimum energy is less than that of the right-hand side. To get both sides on the same energy scale, we must subtract the binding energy, $Q = 13.6\,\mathrm{eV}$, from the right-hand side:

$$\mu_+ + \mu_- = \mu_0 - Q. \tag{5.35}$$

Another way to see why Q appears is to add the rest mass for each species to its chemical potential; collecting all terms on the right, we would then have a term $(m_0 - m_+ - m_-)c^2 = -Q$. Some ionization potentials, along with the half-ionization temperature for a given electron number density n_e, are given in Table 5.1.

element	H^-	Na	H	He
$Q/(\mathrm{eV})$	0.75	5.14	13.6	24.6
$T_{1/2}(n_e = 10^{13}\,\mathrm{cm}^{-3})/\mathrm{K}$	600	3300	8 000	13 900
$T_{1/2}(n_e = 10^{16}\,\mathrm{cm}^{-3})/\mathrm{K}$	900	5 000	11 800	20 200

Table 5.1: Selected ionization potentials and half-ionization temperatures

For a non-degenerate plasma, we can insert eq. (5.7) into eq. (5.35), divide through by $k_\mathrm{B}T$, and take the exponential to obtain

$$\frac{n_+ n_-}{n_0} = \frac{g_+ g_-}{g_0} \left(\frac{m_- k_\mathrm{B} T}{2\pi\hbar^2}\right)^{3/2} \exp\left(-\frac{Q}{k_\mathrm{B}T}\right). \tag{5.36}$$

The number density of all hydrogen in the gas is $n_0 + n_+ = n_\mathrm{H}$. Denote the ionized fraction by $x = n_+/n_\mathrm{H} = n_i/n_\mathrm{H}$, so that the left-hand side of equation (5.36) is $n_\mathrm{H}x^2/(1-x)$. In the hydrogen atom ground state, the electron spin and proton spin are either aligned or anti-aligned. These states are very nearly degenerate, so that $g_0 = 2$. Both the proton and electron have spin $1/2$; there are really only two available states, however, because of the freedom in choosing our coordinate system. As a result, $g_+ g_- = 2$ as well.

Inserting these factors into equation (5.36), and using $k_\mathrm{B} = 8.6173 \times 10^{-5}\,\mathrm{eV/K}$, we obtain

$$\frac{x^2}{1-x} = \frac{2.41 \times 10^{21}\,\mathrm{cm}^{-3}}{n_\mathrm{H}} \left(\frac{T}{10^4\,\mathrm{K}}\right)^{3/2} \exp\left(-\frac{15.78 \times 10^4\,\mathrm{K}}{T}\right). \tag{5.37}$$

This equation defines a set of points in the $\rho - T$ plane for which $x = 1/2$. We may take this set of points to mark the boundary between neutral and ionized hydrogen. At fixed density, the transition from neutral to fully ionized is very rapid.

EXERCISE 5.6 —

1. Solve equation (5.37) for a density of $10^{16}\,\mathrm{cm}^{-3}$ (a fiducial value for the solar photosphere) and find the half-ionization temperature, i.e., the temperature at which $x = 0.5$. Explain the reason for the discrepancy between the half-ionization temperature and $k_{\mathrm{B}}T = 13.6\,\mathrm{eV}$.
2. At this half-ionization temperature, what is the occupancy of the excited levels of the hydrogen atom? Do we need to worry about corrections to the ionization from these excited states?

5.6 *Coulomb interactions*

Under the conditions in a stellar interior, most of the atoms are ionized, and the stellar matter consists of positively charged nuclei and ions and negatively charged electrons. A *plasma* is defined as a gas of charged particles in which the kinetic energy of a typical particle is much greater than the potential energy due to its nearest neighbors. To make this quantitative, consider a gas with only one species present, with charge q. Let the mean spacing between particles be a; clearly the number density of such particles is $n = (4\pi a^3/3)^{-1}$. We may then take the quantity

$$\Gamma \equiv \frac{q^2}{a k_{\mathrm{B}} T} \tag{5.38}$$

as indicating the relative importance of potential to kinetic energy. In a classical plasma, $\Gamma \ll 1$. Note, however, that systems with $\Gamma > 1$ are often (confusingly) called *strongly coupled plasmas*. The meaning is usually clear from context.

EXERCISE 5.7 — Show that for a non-relativistic plasma, the magnetic interaction between two charged particles is much less than the electrostatic interaction.

Debye shielding

Imagine a typical charged particle in a plasma. Very close to the particle, we expect the electrostatic potential to be that of an isolated charge $\Phi = q/r$. Far from the particle, there will be many other particles surrounding it, and the potential is screened. For example, a positive ion will tend to attract electrons to be somewhat, on average, closer to it than other ions: we say that the ion *polarizes* the plasma. As a result of this polarization, the potential of any particular ion should go to zero much faster than $1/r$ due to the "screening" from the enhanced density of opposite charges around it.

Let's consider a plasma having many ion species, each with charge Z_i, and elections. About any selected ion j, particles will arrange themselves according to Boltzmann's law,

$$n_i(r) = n_{i0} \exp\left[-\frac{Z_i e\Phi(r)}{k_\mathrm{B}T}\right]. \tag{5.39}$$

Here n_{i0} is the density of particle i far from the charge j, and r is the distance between particles i and j. (A similar equation holds for the electrons, with Z replaced by -1.) To solve for the potential, we can use Poisson's equation,

$$\nabla^2\Phi = -4\pi \sum_i Z_i e n_i(r) + 4\pi e n_e(r). \tag{5.40}$$

Our assumption is that the term in the exponential of equation (5.39) is small, so we may expand it to first order in Φ and substitute that expansion into equation (5.40) to obtain in spherical geometry

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\Phi) = -4\pi e\left[\sum_i n_{i0}Z_i\left(1-\frac{Z_i e\Phi}{k_\mathrm{B}T}\right) - n_{e0}\left(1+\frac{e\Phi}{k_\mathrm{B}T}\right)\right].$$

The overall charge neutrality of the plasma implies that $n_{e0} = \sum_i Z_i n_{i0}$; using this to simplify the above equation gives

$$\frac{1}{r}\frac{\partial^2}{\partial r^2}(r\Phi) = \left[\frac{4\pi e^2}{k_\mathrm{B}T}\sum_i n_{i0}\left(Z_i^2 + Z_i\right)\right]\Phi \equiv \lambda_\mathrm{D}^{-2}\Phi. \tag{5.41}$$

The quantity in [ ] has dimensions of reciprocal length squared and we define it as $(1/\lambda_\mathrm{D})^2$ with λ_D being called the *Debye length*.

Multiplying equation (5.41) by r, integrating twice, and determining the constant of integration from the condition that as $r \to 0$, $\Phi \to Z_j e/r$ gives the self-consistent potential

$$\Phi = \frac{Z_j e}{r}\exp\left(-\frac{r}{\lambda_\mathrm{D}}\right). \tag{5.42}$$

The Debye length λ_D determines the size of the screening cloud around the ion.

In order for the above derivation to be valid, we require that $\lambda_\mathrm{D} \gg a$, where a is the mean ion spacing; otherwise, there won't be any charges in our cloud to screen the potential! Equivalently, we require the number of particles in a sphere of radius λ_D to be large,

$$\frac{4\pi}{3}\lambda_\mathrm{D}^3\sum_i n_i \gg 1. \tag{5.43}$$

This condition must hold if we are to treat the gas as a plasma.

EXERCISE 5.8—

1. Show that equation (5.43) is equivalent to $\Gamma \ll 1$ for a single species plasma.
2. Show that the net charge in the shielding cloud about an ion of charge Ze is $-Ze$; the shielding cloud cancels out the ion's charge.

Corrections to the ideal gas EOS

In a plasma the particles are not independent: shake one particle and other nearby particles will shake as well. In statistical mechanics, this requires introducing *correlation functions* to derive the equation of state. We'll adopt the more intuitive approach of Debye and Hückel to get the lowest-order correction to the ideal gas EOS. First, the total electrostatic energy in a volume V is

$$E_{\mathrm{Coul}} = \frac{1}{2} V \sum_j Z_j e n_j \Phi_j. \tag{5.44}$$

Here Φ_j is the potential at a particle j *due to all the other particles in the plasma*. Now, we computed the total potential around a particle (eq. [5.42]); expanding and subtracting off the self-potential of particle j gives $\Phi_j = -Z_j e/\lambda_{\mathrm{D}}$. Inserting this into equation (5.44) and expanding gives

$$E_{\mathrm{Coul}} \approx -V \left(\frac{\pi}{k_{\mathrm{B}} T}\right)^{1/2} e^3 \left[\sum_i n_{i0} \left(Z_i^2 + Z_i\right)\right]^{3/2}. \tag{5.45}$$

This energy is to be added to the kinetic energy of the gas. The effect of the electrostatic interactions is to *decrease* the energy in the gas, that is, to make it more bound.

We can't directly get the pressure from equation (5.45) because the equation isn't in terms of S and V (recall that $P = -(\partial E/\partial V)_S$) but rather in terms of T and V. In order to get the pressure, we first must find the Helmholtz free energy F. To do this, we integrate the thermodynamical identity

$$E = -T^2 \left(\frac{\partial}{\partial T}\right)_V \left(\frac{F}{T}\right)$$

and then take $P = -(\partial F/\partial V)_{T,N}$ to obtain

$$P_{\mathrm{Coul}} \approx -\frac{e^3}{3} \left(\frac{\pi}{k_{\mathrm{B}} T}\right)^{1/2} \left[\frac{\left(\langle Z^2\rangle + \langle Z\rangle\right)\rho}{\langle A\rangle m_{\mathrm{u}}}\right]^{3/2}. \tag{5.46}$$

The effect of Coulomb interactions is to decrease the pressure below the ideal gas value.

Coulomb corrections when the electrons are degenerate

The above discussion holds only when both the electrons and ions are non-degenerate. What happens when the electrons are degenerate? In that case the kinetic energy is of order the Fermi energy, not the temperature. We might think to replace $k_B T$ with ϵ_F in equation (5.38). Recalling the formula for ϵ_F from § 5.2, we have the condition for the Coulomb interactions to be weak,

$$\frac{e^2}{a\epsilon_F} = \left(\frac{4\pi n}{3}\right)^{1/3} \left(\frac{m_e e^2}{\hbar^2}\right) \frac{2}{(3\pi^2 n)^{2/3}} < 1. \tag{5.47}$$

Here m_e and n denote, respectively, the electron mass and number density. Do you recognize the quantity $\hbar^2/(m_e e^2)$? It is the Bohr radius, a_B. What is a_B doing in this equation? Well, we are looking for a quantum mechanical system in which the Coulomb interaction is comparable to the non-relativistic kinetic energy. Does that sound like any system you've seen before?

Cleaning up equation (5.47), our condition for the electrons to be weakly interacting when degenerate is

$$\left(\frac{2^{5/3}}{3\pi}\right) \left(n a_B^3\right)^{-1/3} < 1, \tag{5.48}$$

or, in terms of mass density ρ and electron fraction Y_e, $(Y_e \rho) > 0.4\,\mathrm{g\,cm^{-3}}$. As the density increases, the electron gas becomes more ideal, that is, the electrostatic interaction matters less and less.

Just to complete the discussion on electrons, what if the electrons are relativistic? In this case, $\epsilon_F = p_F c = (3\pi^2 n)^{1/3}\hbar c$, and

$$\frac{e^2}{a\epsilon_F} = \left(\frac{4}{9\pi}\right)^{1/3} \left(\frac{e^2}{\hbar c}\right) = 3.8 \times 10^{-3}. \tag{5.49}$$

In this case $\epsilon_F \propto n^{1/3}$ so the density dependence cancels. You will note the appearance of the fine structure constant $\alpha_F = e^2/(\hbar c)$, as you might have expected when dealing with relativistic electrons and electrostatics.

Under astrophysical conditions, we can almost always regard degenerate electrons as being ideal. What about the ions? They are not usually degenerate under conditions of interest. We can get a simple expression if we go to the opposite limit, in which the electrons are very degenerate. In that case, the electrons are an ideal gas and hence have uniform density. If the temperature is low enough, the ions will have $Z^2 e^2/(a k_B T) \gg 1$; in this case we might expect the ions will arrange themselves into a lattice that maximizes the inter-ionic spacing.

To get an estimate of the electrostatic energy, let's compute the energy of a charge-neutral sphere centered on a particular ion of charge Z_i. Because the electrons have a uniform density, $Y_e \rho/m_u$, we can find the radius of the sphere a by requiring it to have Z_i electrons,

$$\frac{4\pi}{3}a^3\left(Y_e\frac{\rho}{m_\mathrm{u}}\right) = Z_i, \tag{5.50}$$

or $a = [3Z_i m_\mathrm{u}/(4\pi Y_e\rho)]^{1/3}$. The potential energy of this sphere has two components. The first is due to electron-electron interactions,

$$E_{ee} = \int_0^a \frac{q(r)\,\mathrm{d}q}{r} = \frac{3}{5}\frac{Z_i^2 e^2}{a}, \tag{5.51}$$

where $q(a) = Z_i e(r/a)^3$ is the charge in a sphere of radius $r < a$. The second component of the potential energy is due to the ion-electron interaction,

$$E_{ei} = -Z_i e\int_0^a \frac{\mathrm{d}q}{r} = -\frac{3}{2}\frac{Z_i^2 e^2}{a}. \tag{5.52}$$

Combining equations (5.51), (5.52), and (5.50) gives the total electrostatic energy for a single ion-sphere,

$$E = -\frac{9}{10}\frac{Z_i^2 e^2}{a} = -\frac{9}{10}Z_i^{5/3}e^2\left(\frac{4\pi}{3}\frac{Y_e\rho}{m_\mathrm{u}}\right)^{1/3}. \tag{5.53}$$

Multiplying this by $n_i = Y_i\rho/m_\mathrm{u}$, summing over all ion species i, and defining $\langle Z^{5/3}\rangle = n^{-1}\sum Z_i^{5/3} n_i$ where $n = \sum n_i$, gives the total Coulomb energy per volume,

$$E_\mathrm{Coul} = -\frac{9}{10}nk_\mathrm{B}T\left[\frac{\langle Z^{5/3}\rangle e^2}{k_\mathrm{B}T}\left(\frac{4\pi}{3}\frac{Y_e\rho}{m_\mathrm{u}}\right)^{1/3}\right]. \tag{5.54}$$

Notice that the quantity in [ ] reduces to Γ for a single-species plasma. If we therefore define $\Gamma \equiv [\,]$ for a multi-component plasma, we have $E_\mathrm{Coul} \approx -0.9nk_\mathrm{B}T\Gamma$; the pressure is then $P_\mathrm{Coul} = E_\mathrm{Coul}/3 = -0.3nk_\mathrm{B}T\Gamma$. This holds in the limit $\Gamma \gg 1$.

EXERCISE 5.9—

1. In the zero-temperature limit (electrons are fully degenerate) use the charge-neutral sphere approximation (p. 53) to calculate the density at which completely ionized ^{56}Fe has zero pressure.
2. Estimate the pressure that would be required to compress ^{56}Fe at the density found in part 1.
3. Compute the mass-radius relation for a cold object ("a rock"). Write the total pressure as the sum of electron (zero-temperature limit) and Coulomb pressure (use the charge-neutral sphere approximation, eq. [5.54] and following text). To make this easier, use the virial scalings for density and pressure to obtain a relation between mass and radius. Find the mass having the largest radius, and express this mass in terms of fundamental physical constants. How does it compare with the mass of Jupiter? Scale the mass and radii to that of Jupiter, and plot $R(M)$ for objects composed of pure ^{1}H, pure ^{4}He, and pure ^{12}C. Also indicate on this plot the masses and radii of the Jovian planets for comparison.

6
Radiation Transport

The sun is very opaque. Were photons able to stream freely, they would exit in $\sim R_\odot/c = 2.0\,\mathrm{s}$. Given the luminosity of the sun, however, we derived that the time for the sun to radiate away its stored thermal energy is instead millions of years (see eq. [1.15]). As a result, we can regard the sun as a cavity filled with photons with a very slight leakage. This is the description commonly invoked to describe blackbody radiation, and we expect that in the interior of the sun, the radiation can be described by a photon gas in thermal equilibrium at the ambient temperature.

6.1 Description of the Radiation Field

Consider a cavity containing a gas of photons. In general we can describe the mean number of photons in this cavity as

$$N = \int f(\boldsymbol{p}, \boldsymbol{x})\, \mathrm{d}^3x\, \mathrm{d}^3p. \tag{6.1}$$

Here f is a distribution function, as described in Chapter 5; if we are in thermal equilibrium, f is of course the Bose-Einstein distribution (cf. eq. [5.31]), but our discussion here will be more general. Consider a small opening on our cavity with area $\mathrm{d}A$ and unit normal $\hat{\boldsymbol{n}}$. The energy incident on this area in a time $\mathrm{d}t$ having propagation vector along $\hat{\boldsymbol{n}}$ and propagating into solid angle $\mathrm{d}\Omega$ (see Fig. 6.1) is found by integrating equation (6.1) over a volume $\mathrm{d}^3x = c\mathrm{d}t\,\mathrm{d}A$,

$$\mathrm{d}E = \mathrm{d}A\, c\mathrm{d}t\, \left(p^2 \mathrm{d}p\, \mathrm{d}\Omega\right) h\nu\, f.$$

Since the photon momentum is $p = h\nu/c$, we have

$$I_\nu \equiv \frac{\mathrm{d}E}{\mathrm{d}t\, \mathrm{d}A\, \mathrm{d}\Omega\, \mathrm{d}\nu} = \frac{h^4\nu^3}{c^2} f. \tag{6.2}$$

This defines the *specific intensity* I_ν. It is easy to show that in the absence of interactions with matter, I_ν is conserved along a ray (see, e.g., Rybicki & Lightman).

If the photons are in thermal equilibrium, then f is the Bose-Einstein distribution, $f = (2/h^3)(\exp[h\nu/k_B T] - 1)^{-1}$, and the specific intensity becomes

$$B_\nu \equiv \frac{2h\nu^3}{c^2}\left[\exp\left(\frac{h\nu}{k_B T}\right) - 1\right]^{-1}. \tag{6.3}$$

Here B_ν is called the *Planck function*.

The energy density per frequency u_ν can be defined as $\mathrm{d}E/(c\mathrm{d}t\,\mathrm{d}A\,\mathrm{d}\nu)$, that is, the energy per unit frequency that is in a cylinder of length $c\,\mathrm{d}t$ and cross-sectional area $\mathrm{d}A$; comparing u_ν with the definition of I_ν, we see that

$$u_\nu = \frac{1}{c}\int I_\nu\,\mathrm{d}\Omega.$$

For a blackbody, $I_\nu = B_\nu$ doesn't depend on angle, and we can integrate over $\mathrm{d}\Omega$,

$$u_\nu = \frac{8\pi h\nu^3}{c^3}\left[\exp\left(\frac{h\nu}{k_B T}\right) - 1\right]^{-1}. \tag{6.4}$$

The total energy density can then be found by integrating over all frequencies, giving

$$u = \left[\frac{8\pi^5 k_B^4}{15h^3c^3}\right] T^4 \equiv aT^4$$

in agreement with what we derived from statistical mechanics, §5.4.

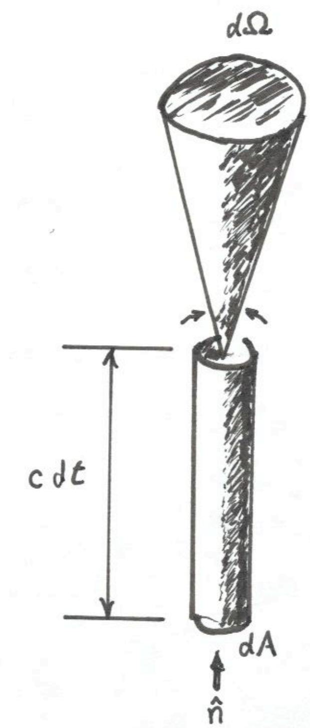


Figure 6.1: Schematic of a pencil of radiation propagating into an angle dΩ.

The next quantity to define is the *flux* of energy, along direction $\hat{\boldsymbol{k}}$, per unit time $\mathrm{d}t$, per unit area $\mathrm{d}A$, and per frequency interval $\mathrm{d}\nu$. We multiply I_ν by a direction vector $\hat{\boldsymbol{k}}$ and integrate over $\mathrm{d}\Omega$:

$$\boldsymbol{F}_\nu = \int \hat{\boldsymbol{k}} I_\nu\,\mathrm{d}\Omega. \tag{6.5}$$

Note that $\boldsymbol{F}_\nu$ is a vector; the net flux along a direction $\hat{\boldsymbol{n}}$ is

$$F_\nu = \int I_\nu(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{k}})\mathrm{d}\Omega.$$

If we take our polar angle with respect to $\hat{\boldsymbol{k}}$, then $(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{k}})\,\mathrm{d}\Omega = \cos\theta\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi$; defining the direction cosine $\mu = \cos\theta$, this becomes

$$F_\nu = \int I_\nu(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{k}})\,\mathrm{d}\Omega = \int_0^{2\pi}\int_{-1}^{1} I_\nu\mu\,\mathrm{d}\mu\,\mathrm{d}\varphi.$$

Note that if the radiation field is isotropic then $F_\nu = 0$: there must be some anisotropy in the radiation field to generate a net flux.

For blackbody radiation, if we only integrate over outgoing directions, $0 \le \mu \le 1$, as would be the case for thermal radiation emerging from a *hohlraum*,

$$F_\nu = \pi B_\nu.$$

Integrating this F_ν over all frequencies, we recover the Stefan-Boltzmann formula,

$$F = \left(\frac{ac}{4}\right)T^4 \equiv \sigma_{SB}T^4,$$

where $\sigma_{\rm SB}$ is the Stefan-Boltzmann constant.

Finally, let's look at the momentum flux along direction $\hat{\boldsymbol{\jmath}}$ being transported along direction $\hat{\boldsymbol{k}}$, per unit time dt, per unit area dA, and per unit frequency interval dν. Since for a photon, $E = pc$, we divide the energy flux by c. This is a *tensor*,

$$\mathbf{P}_\nu = \frac{1}{c}\int \hat{\boldsymbol{\jmath}}\hat{\boldsymbol{k}} I_\nu \,\mathrm{d}\Omega. \tag{6.6}$$

The net (isotropic) momentum flux along a direction $\hat{\boldsymbol{n}}$ is then

$$P_\nu = \frac{1}{c}\int (\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{k}})(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{k}}) I_\nu \,\mathrm{d}\Omega = \frac{2\pi}{c}\int_{-1}^{1} I_\nu \mu^2 \,\mathrm{d}\mu.$$

For blackbody radiation,

$$P_\nu = \frac{4\pi}{3c}B_\nu = \frac{1}{3}u_\nu$$

and we may integrate this over frequency to obtain $P = u/3$, the standard result from thermodynamics.

6.2 *Some simple estimates*

We argued in the previous section that the solar interior is quite opaque. Naively, we might imagine some radiative transition, e.g. bremsstrahlung, emitting a photon. The photon speeds away at c, but it doesn't get very far before being absorbed or scattered by another particle. A new photon, either due to emission or scattering, will be emitted at some random direction, and the whole process repeats. This is just a description of a *random walk*.

For some simple estimate, let's assume that the hop is the same for all photons, regardless of frequency or ambient temperature. If the hop length is ℓ, then we know that the total path length to get from the center to the surface is $R_\odot(R_\odot/\ell)$ and the time for this to occur is $R_\odot^2/\ell/c$. What is a good estimate for ℓ? Consider a planar electromagnetic wave incident on a collection of scatterers. If these scatterers are uncorrelated, then the probability of scattering is just the number of scatterers times the probability for scattering from a single scatterer. Define the probability of scattering as

$$\mathcal{P} = N \times \left(\frac{\text{energy scattered per unit time by one scatterer}}{\text{energy incident per unit time per unit area}}\right)\frac{1}{\mathcal{A}} \tag{6.7}$$

where $\mathcal{A}$ is the area normal to the propagation direction $\hat{\boldsymbol{k}}$ of the volume containing the N scatterers. The quantity in parenthesis is just the definition of the cross-section σ. Furthermore, if we set $\mathcal{P} = 1$, then the total number of scatters is just $N = n \times \mathcal{A} \times \ell$, where n is the number density of scatterers. Thus we define the *mean free path*,

$$\ell = \frac{1}{n\sigma}. \tag{6.8}$$

In stellar work, it is more convenient to use mass density rather than number density. Writing $n = Y\rho/m_{\mathrm{u}}$, where Y is the abundance of scatterers, we have

$$\ell = \rho^{-1}\left(\frac{m_{\mathrm{u}}}{Y\sigma}\right) \equiv (\rho\kappa)^{-1}$$

where κ is the *opacity* and has dimensions $[\kappa] \sim [\mathrm{cm}^2\,\mathrm{g}^{-1}]$.

The opacity in the stellar interior is set by a large number of processes (see §6.5): Thomson scattering, free-free absorption, atomic absorption, and photoionization. In general, the cross-section depends on the ambient temperature and density and the frequency of the photon. Over the length of a hop ℓ the temperature and density will only vary slightly. As a result, the conditions are nearly isotropic, so we indeed expect the radiation to come into thermal equilibrium with the ambient material. But the conditions are not perfectly isotropic—otherwise there would be zero net heat flux! It is the small anisotropy that gives rise to the transport of energy. Let's imagine a small cube of material, with the size of this cube being ℓ. Because we are so very nearly isotropic and in thermal equilibrium, the flux through any one face of this cube must be $(c/6)u$. Now suppose we have two adjacent cubes, with the common face of the cubes being at $x = 0$. The flux across the face has contributions from photons emitted at $x - \ell$ and $x + \ell$, so the net flux is

$$\begin{aligned} F &\approx \frac{c}{6}u(x-\ell) - \frac{c}{6}u(x+\ell) \\ &\approx -\frac{1}{3}c\ell\frac{\mathrm{d}u}{\mathrm{d}x}. \end{aligned} \tag{6.9}$$

This is a diffusion equation with coefficient $c/(3\rho\kappa)$. Our derivation is very crude, as it neglects the variation in cross section with the properties of the ambient medium and with the photon frequency. Nonetheless, this is basically the correct scenario; heat diffuses with a coefficient given by some suitably defined average over all sources of opacity.

6.3 *Equation of Transfer*

We're now ready to formalize the crude work in the previous section. In the absence of interactions with matter, the specific intensity I_ν is conserved along a ray propagating in direction $\hat{\boldsymbol{k}}$: $\mathrm{d}I_\nu/\mathrm{d}s = c^{-1}\partial_t I_\nu + \hat{\boldsymbol{k}} \cdot \boldsymbol{\nabla} I_\nu = 0$. If matter is present, it can do three things to change I_ν.

emit Matter may spontaneously emit photons and add to the beam: $\mathrm{d}I_\nu/\mathrm{d}s = \rho\epsilon_\nu/(4\pi)$. Here ϵ_ν is the energy spontaneously emitted per unit frequency per unit time per unit mass. The factor of 4π is to make this term per steradian.

absorb Photons have a chance of being absorbed or scattered out of the beam: $\mathrm{d}I_\nu/\mathrm{d}s = -\rho\kappa_\nu I_\nu$. Here the right-hand side is the energy removed from the beam along a path ds with $\kappa_\nu = \kappa_\nu^{\mathrm{abs}} + \kappa_\nu^{\mathrm{sca}}$ being the total opacity (absorption plus scattering). The dimensions of opacity are clearly $[\kappa_\nu] \sim [\mathrm{cm}^2/\mathrm{g}]$. (If we had stimulated emission, this would be a *negative* κ_ν.)

scatter Photons may be scattered into the beam from other directions: $\mathrm{d}I_\nu/\mathrm{d}s = \rho\kappa_\nu^{\mathrm{sca}}\varphi_\nu$. If the scattering is isotropic, then

$$\varphi_\nu = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} I_\nu\,\mathrm{d}\varphi\,\sin\theta\,\mathrm{d}\theta \equiv J_\nu, \tag{6.10}$$

where J_ν is the mean intensity: the scattering redistributes the energy over all angles.

Putting all these terms together gives us the *equation of transfer*,

$$\frac{1}{c}\partial_t I_\nu + \hat{\boldsymbol{k}}\cdot\boldsymbol{\nabla} I_\nu = \rho\frac{\epsilon_\nu}{4\pi} - \rho\kappa_\nu I_\nu + \rho\kappa_\nu^{\mathrm{sca}}\varphi_\nu \tag{6.11}$$

for the specific intensity I_ν.

Radiative equilibrium

The emissivity ϵ_ν and the opacity κ_ν describe how the radiation interacts with matter. A condition of steady-state is that the gas not gain or lose energy to the radiation. This requires balancing

$$(\text{energy emitted per unit volume}) = \rho\int\frac{\epsilon_\nu}{4\pi}\,\mathrm{d}\nu\,\mathrm{d}\Omega$$

with

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

or

$$\int_0^\infty\left(\frac{\epsilon_\nu}{4\pi} - \kappa_\nu^{\mathrm{abs}} J_\nu\right)\,\mathrm{d}\nu = 0. \tag{6.12}$$

Here we assume ϵ_ν does not depend on angle. We don't include scattering because it doesn't transfer energy between the radiation and the gas.

Now suppose that the level populations of the matter are in thermal equilibrium and can be described by a temperature T. In that case, detailed balance must hold, so that

$$\frac{\epsilon_\nu}{4\pi\kappa_\nu^{\mathrm{abs}}} = B_\nu(T), \tag{6.13}$$

where $B_\nu(T)$ is the Planck function. This defines *local thermodynamic equilibrium (LTE)*. If the radiation field is, in addition, described by a Planck function *at the same temperature* then we would have complete thermodynamic equilibrium.

Optical depth

Consider a ray directed into a medium in steady-state ($\partial_t \to 0$). In the absence of emission ($\epsilon_\nu = 0$) or scattering ($\kappa_\nu^{\rm sca} = 0$) equation (6.11) takes a particularly simple form:

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = -I_\nu. \tag{6.14}$$

Here we have set $\hat{\boldsymbol{k}} \cdot \boldsymbol{\nabla} = (\mathrm{d}/\mathrm{d}s)$, where $\mathrm{d}s$ is a infinetesimal along the path of the ray, and further have defined the *optical depth* as

$$\tau_\nu = \int \rho\kappa_\nu \,\mathrm{d}s. \tag{6.15}$$

Note that τ_ν is dimensionless. Taking ρ and κ_ν as given, equation (6.14) has a simple solution,

$$I_\nu(\tau_\nu) = I_\nu(0)\exp\left(-\tau_\nu\right).$$

Note that $\rho\kappa = \ell^{-1}$, so equation (6.15) is just $\tau = \int \mathrm{d}s/\ell$, i.e., it expresses distance by counting the number of mean free pathlengths traversed.

Source function

Having defined the optical depth, we can now add the emissivity ϵ_ν and scattering term (henceforth we will assume isotropic scattering) to equation (6.14) to obtain

$$\frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = S_\nu - I_\nu, \tag{6.16}$$

where

$$S_\nu \equiv \frac{1}{\kappa_\nu}\left(\frac{\epsilon_\nu}{4\pi} + \kappa_\nu^{\rm sca} J_\nu\right) \tag{6.17}$$

is the *source function*. In the absence of scattering, so that $S_\nu = \epsilon_\nu/(4\pi\kappa_\nu)$ is a known function of τ_ν, we can formally solve equation (6.16):

$$I_\nu(\tau_\nu) = I_\nu(0)\exp(-\tau_\nu) + \int_0^{\tau_\nu} S_\nu(\tau_\nu)\exp(t-\tau_\nu)\,\mathrm{d}t.$$

In the presence of scattering, S_ν depends on $J_\nu = (1/4\pi)\int I_\nu \,\mathrm{d}\Omega$, so that equation (6.16) is an *integro-differential* equation.

6.4 *Diffusion Approximation and the Rosseland Mean Opacity*

At large optical depth, such as deep in a stellar interior, the radiation field is in thermal equilibrium, so that $I_\nu = S_\nu = B_\nu$. To see this, consider the relative scales of terms in the transfer equation.

$$\underset{\rm I}{\frac{1}{c}\frac{\partial I_\nu}{\partial t}} \quad + \quad \underset{\rm II}{\hat{\boldsymbol{k}}\cdot\boldsymbol{\nabla} I_\nu} \quad = \quad \underset{\rm III}{\rho\frac{\epsilon_\nu}{4\pi}} \quad - \quad \underset{\rm IV}{\rho\kappa_\nu I_\nu} \quad + \quad \underset{\rm V}{\rho\kappa_\nu^{\rm sca} J_\nu}$$

On the left-hand side, term I scales as $I_\nu/(ct_\odot)$, where $t_\odot$ is the evolutionary timescale of the sun (Gyr), and term II scales as $I_\nu/R_\odot$. On the right-hand side, term IV scales at I_ν/ℓ. Hence the ratio of terms of term II to term IV is $\ell/R_\odot \ll 1$ and that of term I to term IV is $\ell/\mathrm{Gpc} \ll 1$. In addition, stellar properties change negligibly on scales of a mean-free path, so conditions are nearly isotopic over much of the interior and $I_\nu = J_\nu$. Hence $I_\nu = J_\nu = S_\nu$, and inserting the relation between ϵ_ν and $\kappa_\nu^{\mathrm{abs}}$ from detailed balance, eq. (6.13), into equation (6.17) implies that $S_\nu = B_\nu$.

If the radiation field is perfectly isotropic there is no flux, however, so we must have some small anisotropy. Let's write I_ν as a thermal term plus a correction,

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

Substituting this into the steady-state equation of transfer,

$$\frac{1}{\rho\kappa_\nu}\hat{\boldsymbol{k}}\cdot\boldsymbol{\nabla} I_\nu = S_\nu - I_\nu$$

and setting the term $S_\nu - B_\nu = 0$ on the right-hand side, we obtain

$$I_\nu^{(1)} = -\frac{1}{\rho\kappa_\nu}\hat{\boldsymbol{k}}\cdot\boldsymbol{\nabla} B_\nu = -\frac{1}{\rho\kappa_\nu}\frac{\partial B_\nu}{\partial T}\hat{\boldsymbol{k}}\cdot\boldsymbol{\nabla} T. \tag{6.18}$$

This is anisotropic: the energy transport is largest in the direction "down" the temperature gradient. Let's get the net flux: multiply equation (6.18) by $\hat{\boldsymbol{k}}$ to get the flux; and then take the component along a direction $\hat{\boldsymbol{n}}$ parallel to $\boldsymbol{\nabla} T$; finally replace the two dot products by the angle cosine μ, and integrate over $\mathrm{d}\Omega = 2\pi\mathrm{d}\mu$ to obtain

$$\boldsymbol{F}_\nu = -\frac{4\pi}{3}\frac{1}{\rho}\left[\frac{1}{\kappa_\nu}\frac{\partial B_\nu}{\partial T}\right]\boldsymbol{\nabla} T. \tag{6.19}$$

The quantity in [ ] deserves a closer look. First, suppose κ_ν is independent of frequency. Then equation (6.19) means that the energy transport is greatest at the frequency where $\partial B_\nu/\partial T$ is maximum, and *not* at the peak of the Planck spectrum.

EXERCISE 6.1 — Explain on physical grounds why the flux for a grey opacity would be greatest at the frequency for which $\partial B_\nu/\partial T$, rather than B_ν, is maximized.

Let us define the *Rosseland mean opacity* as

$$\kappa_\mathrm{R} \equiv \left[\frac{\int \mathrm{d}\nu\, \kappa_\nu^{-1}(\partial B_\nu/\partial T)}{\int \mathrm{d}\nu\, (\partial B_\nu/\partial T)}\right]^{-1}.$$

We can use this to integrate equation (6.19) to obtain the total radiative flux,

$$\boldsymbol{F} = -\frac{4\pi}{3}\frac{1}{\rho\kappa_\mathrm{R}}\boldsymbol{\nabla}\left[\int \mathrm{d}\nu\, B_\nu\right] = -\frac{1}{3}\frac{c}{\rho\kappa_\mathrm{R}}\boldsymbol{\nabla} aT^4. \tag{6.20}$$

This is just our formula for radiation diffusion (eq. [6.9]) that we obtained from physical arguments, but now we have an expression for the effective opacity κ_{R}.

6.5 *Sources of Opacity*

There are several processes that contribute to radiative opacity in stellar interiors. These are well described in standard texts, so we'll just briefly list them here.

Thomson scattering

Thomson scattering is scattering from non-relativistic electrons when the photon energy is sufficiently low that we can neglect the recoil of the electron. The cross-section for Thomson scattering derived in Jackson and is

$$\sigma_{\mathrm{Th}} = \frac{8\pi}{3}\left(\frac{e^2}{m_e c^2}\right)^2 = 0.665 \times 10^{-24}\,\mathrm{cm}^2. \tag{6.21}$$

The opacity for Thomson scattering is then

$$\kappa_{\mathrm{Th}} = \frac{n_e \sigma_{\mathrm{Th}}}{\rho}(0.4\,\mathrm{cm}^2\,\mathrm{g}^{-1})Y_e.$$

The factor of Y_e is because the electrons, which are much lighter than nuclei and therefore easier for an incident wave to shake, do the scattering.

EXERCISE 6.2 — Using Thomson scattering for the dominant opacity, estimate the photon diffusion time for the sun.

Free-free absorption

Another important one is free-free absorption. This is the inverse of *bremsstrahlung*, which is radiation emitted when an electron is scattered from an ion (see Figure 7.1). The procedure for calculating the opacity is to first compute the emissivity and then use detailed balance (eq. [6.13]) to obtain

$$\kappa_\nu^{\mathrm{ff}} = \frac{\epsilon_\nu}{4\pi B_\nu(T)}.$$

To calculate the emissivity, we start with the derivation of the momentum gained by an electron (eq. [7.1]). The acceleration leads to an emission of radiation; according to Larmor's formula, the power emitted is

$$P(b) = \frac{2}{3}\frac{e^2}{c^3}|\dot{\boldsymbol{v}}|^2 = \frac{2}{3}\frac{Z^2 e^6}{m_e^2 c^3 b^4}. \tag{6.22}$$

The radiation is distributed over a broad range of frequencies up to a cutoff $\nu_{\max} \sim v/b$. Integrating over a range of impact parameters and then over the distribution of electron velocities gives the emissivity, which is (restoring all of the numerical factors)

$$\rho\epsilon_\nu = 4\pi\left(\frac{2\pi}{3}\right)^{1/2} Z^2 n_I n_e \hbar c^2 \alpha\sigma_{\mathrm{Th}} \left(\frac{m_e}{k_\mathrm{B}T}\right)^{1/2} \exp\left(-\frac{h\nu}{k_\mathrm{B}T}\right)\bar{g}_{\mathrm{ff}}. \quad (6.23)$$

The velocity-averaged Gaunt factor $\bar{g}_{\mathrm{ff}}$ contains most of the details about the integration. The factor of $T^{-1/2}$ is because there is a factor of v^{-1} that appears in the integration (the collision time is $\sim b/v$).

Applying detailed balance, equation (6.13), gives the opacity as a function of frequency,

$$\kappa_\nu^{\mathrm{ff}} = \pi\left(\frac{2\pi}{3}\right)^{1/2} Z^2 \frac{n_I n_e}{\rho} c^3 \alpha\sigma_{\mathrm{Th}} \left(\frac{m_e c^2}{k_\mathrm{B}T}\right)^{1/2} \nu^{-3}\left[1-\exp\left(-\frac{h\nu}{k_\mathrm{B}T}\right)\right]\bar{g}_{\mathrm{ff}} \quad (6.24)$$

Notice that $n_I n_e/\rho = \rho/(m_\mathrm{u}^2\mu_I\mu_e)$, so that the opacity scales with density. Further, note that when taking the Rosseland mean over all frequencies, the factor of ν^{-3} introduces a factor of T^{-3}, so that $\langle\kappa_\nu^{\mathrm{ff}}\rangle \propto \rho T^{-7/2}$.

EXERCISE 6.3 — In terms of central density and temperature, under what conditions is free-free opacity more important than Thomson scattering? For what mass range of stars is free-free opacity dominant in the core? What about for Thomson scattering?

Bound-free and bound-bound absorption; Kramer's opacity

Bound-free and bound-bound transitions have a cross-section with a frequency dependence (at photon energies above threshold) of ν^{-3} as well; therefore their Rosseland averages also scale as $\rho T^{-7/2}$. An opacities with this form is known as Kramer's opacity, $\kappa = \kappa_0 \rho T^{-3.5}$. For conditions in the solar center, a good approximation is

$$\begin{aligned}\kappa \approx 0.012\,\mathrm{cm^2\,g^{-1}} &\left(\frac{\rho}{1\,\mathrm{g\,cm^{-3}}}\right)\left(\frac{T}{10^7\,\mathrm{K}}\right)^{-7/2} \\ &\times\left[(1+X_\mathrm{H})\left(X_\mathrm{H}+X_\mathrm{He}+\sum_{Z>2}\frac{X_i Z_i^2}{A_i}\right)\right]. \end{aligned} \quad (6.25)$$

The expression in [ ] is just an approximation for $Y_e\langle Z^2\rangle$ at solar composition.

EXERCISE 6.4 — Suppose both free-free absorption (κ_ν^{ff}) and Thomson scattering (κ_ν^{Th}) contribute to the opacity. Denote by $\langle\,\rangle$ the Rosseland averaging of an opacity. Does $\langle\kappa_\nu^{\mathrm{ff}}+\kappa_\nu^{\mathrm{Th}}\rangle = \langle\kappa_\nu^{\mathrm{ff}}\rangle + \langle\kappa_\nu^{\mathrm{Th}}\rangle$?
Now suppose that κ_ν^{ff} is due to free-free absorption on two different ion species (denoted below by subscripts "1" and "2") with different charge number Z. In this case does $\langle\kappa_{\nu,1}^{\mathrm{ff}}+\kappa_{\nu,2}^{\mathrm{ff}}\rangle = \langle\kappa_{\nu,1}^{\mathrm{ff}}\rangle + \langle\kappa_{\nu,2}^{\mathrm{ff}}\rangle$?

6.6 *Eddington Standard Model*

Polytropes with index $n = 3/2$ correspond to fully convective stars ($P \propto \rho^{5/3}$, the relation for an adiabat) or for white dwarfs (non-relativistic, degenerate equation of state). Another interesting case, for historical reasons, is the *Eddington Standard Model*, which is a fair approximation to main-sequence stars with $M \gtrsim M_\odot$. Suppose we write the equation of state as the sum of ideal gas and radiation pressure,

$$P = \frac{\rho k_\mathrm{B} T}{\mu m_\mathrm{u}} + \frac{1}{3} a T^4. \tag{6.26}$$

Now make the *ansatz* that

$$\frac{P_\mathrm{rad}}{P} = \frac{aT^4}{3P} = 1 - \beta = \text{const.}, \tag{6.27}$$

that is, the radiation pressure is a fixed fraction of the total pressure everywhere. Solving for T in terms of P and β,

$$T = \left[\frac{3(1-\beta)P}{a}\right]^{1/4},$$

and inserting this into equation (6.26) gives us a simple EOS,

$$P = \left[\left(\frac{k_\mathrm{B}}{\mu m_\mathrm{u}}\right)^4 \frac{3}{a}\right]^{1/3} \left[\frac{1-\beta}{\beta^4}\right]^{1/3} \rho^{4/3}. \tag{6.28}$$

This is the equation for a polytrope of index 3.

EXERCISE 6.5 — Derive an expression for β in terms of the mass of the star for the Eddington Standard Model.

Why is it at all reasonable to take β as being constant? To explore this, go back to the equation for radiative diffusion

$$F(r) = -\frac{1}{3}\frac{c}{\rho\kappa}\frac{\mathrm{d}aT^4}{\mathrm{d}r}.$$

Write the flux as $F(r) = L(r)/(4\pi r^2)$, and since pressure decreases monotonically with radius, write

$$\frac{\mathrm{d}aT^4}{\mathrm{d}r} = \frac{\mathrm{d}aT^4}{\mathrm{d}P}\frac{\mathrm{d}P}{\mathrm{d}r} = -\rho\frac{Gm(r)}{r^2}\frac{\mathrm{d}aT^4}{\mathrm{d}P}.$$

The equation of radiation transport then becomes

$$L(r) = \frac{4\pi G m(r) c}{\kappa(r)}\frac{\mathrm{d}P_\mathrm{rad}}{\mathrm{d}P}.$$

Dividing both sides by $L \cdot M/\kappa_\mathrm{Th}$ and rearranging terms,

$$\frac{\mathrm{d}P_\mathrm{rad}}{\mathrm{d}P} = \left[\frac{L\kappa_\mathrm{Th}}{4\pi G M c}\right]\left(\frac{\kappa(r)}{\kappa_\mathrm{Th}}\frac{L(r)}{L}\frac{M}{m(r)}\right). \tag{6.29}$$

Here L is the total luminosity of the star and M is the total mass. The term in [ ] is a constant (the Thomson opacity κ_{Th} doesn't depend on density or temperature) and we define the *Eddington luminosity* as $L_{\mathrm{Edd}} = 4\pi GMc/\kappa_{\mathrm{Th}}$. For the sun, $L_{\mathrm{Edd}} = 1.5\times10^{38}\,\mathrm{erg\,s^{-1}} = 3.8\times10^{4}\,L_{\odot}$. For the term ( ) on the right-hand side, note the $L(r)/m(r)$ is basically the average energy generation rate interior to a radius r. Since nuclear reactions are temperature sensitive, the heating is concentrated toward the stellar center and $L(r)/m(r)$ decreases with radius. For stars like the sun, free-free opacity is dominant, and since the free-free Rosseland opacity goes as $T^{-3.5}$, $\kappa(r)$ increases with radius. Thus, if the energy generation rate is not too temperature dependent (the reaction $p + p \to {}^{2}\mathrm{H}$ goes roughly as $T^{4.5}$ at $T = 10^{7}$ K), then the term in ( ) does not vary strongly with radius, and $\mathrm{d}P_{\mathrm{rad}}/\mathrm{d}P$ is indeed roughly constant.

EXERCISE 6.6 — You are now in a position to understand why the luminosity depends strongly on the mass. Cast the flux equation (eq. [6.20]) into dimensionless form. Assume the opacity has the functional form $\kappa = \kappa_0 \rho^{a} T^{-b}$, and scale ρ and T in terms of M and R.

1. You should be able to find a characteristic scale for the luminosity which depends on the stellar mass M and radius R, as well as on the exponents a and b. Regard κ_0 as a fitting constant, and adjust it so that you get an expression in the form
$$\frac{L}{L_{\odot}} = \left(\frac{M}{M_{\odot}}\right)^{\alpha} \left(\frac{R}{R_{\odot}}\right)^{\beta}.$$
2. If the opacity is dominated by Thomson scattering, what are α and β? What about if the opacity is Kramer's (eq. [6.25])?

7
Transport in a Plasma

7.1 Collisions

Without collisions, a plasma cannot reach thermodynamic equilibrium, and the rate of collisions mediates both the approach to equilibrium and the transport of quantities, such as heat, in a forced system. In this section, we'll make an estimate for the rate of electron-electron ion-ion, and electron-ion collisions.

To begin, let's imagine a light particle (electron) colliding with a much heavier, fixed particle (an ion), as illustrated in Figure 7.1. (This picture also applies to a pseudo particle of reduced mass scattering in a fixed potential.) Let the impact parameter be b, and the mass of the incident particle is μ. For Coulomb interactions, the force on the particle is $(q_1 q_2/r^2)\hat{\boldsymbol{r}}$. The incident momentum is p_0. Now by assumption, in our plasma most of the interactions are weak (potential energy is much less than kinetic), so let's treat the deflection of the particle as a perturbation. That is, we shall assume that $p_0 = \text{const}$ and that the effect of the interaction is to produce a perpendicular (to p_0) component of the momentum $p_\perp$. The total change in $p_\perp$ is then

$$p_\perp = \int_{-\infty}^{\infty} \mathrm{d}t\, \frac{q_1 q_2}{r^2} \sin\theta,$$

where $\sin\theta = b/r$ is the angle that the radial vector makes with the horizontal. Substituting $r = b/\sin\theta$ and $\mathrm{d}t = -\mu b \mathrm{d}\theta/p_0/\sin^2\theta$, we have

$$p_\perp = -\int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, \frac{\mu}{p_0}\frac{q_1 q_2}{b},$$

leading to the intuitive result

$$\frac{p_0 p_\perp}{2\mu} = \frac{q_1 q_2}{b}. \tag{7.1}$$

Clearly a large angle scattering occurs if $p_\perp \geq p_0$, or

$$b \leq b_0 \equiv \frac{2\mu q_1 q_2}{p_0^2}; \tag{7.2}$$

our approach is only valid for $b \gg b_0$. Note that $p_\perp/p_0 = b_0/b$. What is the rate of large angle scatterings? The cross section for a large angle scattering is $\sigma_{\rm LA} = \pi b_0^2$. Imagine a particle incident on a cylinder of length $(p_0/\mu)\mathrm{d}t$ and cross-sectional area $\mathcal{A}$. Within this cylinder there are $n \times (p_0/\mu)\mathrm{d}t\mathcal{A}$ scatterers of cross-section $\sigma_{\rm LA}$. so the probability of the particle interacting per time $\mathrm{d}t$ is

$$\frac{(\sigma_{\rm LA} \times n \times p_0/\mu)\mathcal{A}}{\mathcal{A}} = n\sigma_{\rm LA}p_0/\mu.$$

This defines the large-angle collision rate,

$$\nu_{\rm LA} = n\sigma_{\rm LA}v_0 = \frac{4\pi\mu(q_1q_2)^2}{p_0^3}. \tag{7.3}$$

Note that it goes as p_0^{-3}; fast-moving particles are hard to scatter.

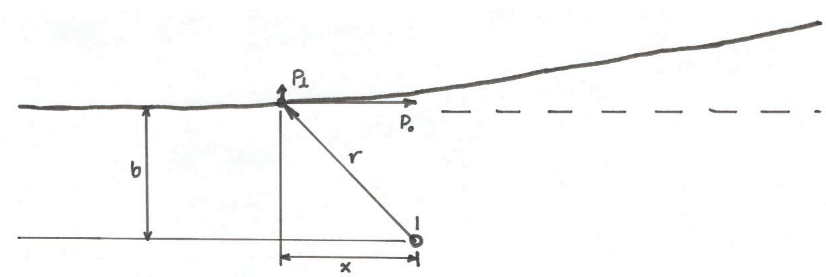


Figure 7.1: Geometry for scattering problem.

As mentioned earlier, we are in a weakly coupled plasma, so we expect large angle scatterings to be a rare occurrence. What happens instead is that the particle suffers a number of small deflections. Let's consider an impact parameter in the range $(b, b + \mathrm{d}b)$. Each deflection will have $\boldsymbol{\Delta p}_\perp$ in some random direction, so

$$\langle \boldsymbol{p}_\perp \rangle = \sum_{i=1}^{N} \boldsymbol{\Delta p}_i \approx \boldsymbol{0}.$$

What is happening is that the component of momentum perpendicular to p_0 is executing a random walk. Indeed, after N collisions,

$$\begin{aligned}\langle \boldsymbol{p}_\perp \cdot \boldsymbol{p}_\perp \rangle &= \left(\sum_{i=1}^{N} \boldsymbol{\Delta p}_i\right) \cdot \left(\sum_{i=1}^{N} \boldsymbol{\Delta p}_i\right) \\ &= \sum_{i=1}^{N} (\Delta p_i)^2 + 2\sum_{i\neq j} \boldsymbol{\Delta p}_i \cdot \boldsymbol{\Delta p}_j = N(\Delta p_\perp)^2,\end{aligned}$$

where we have assumed that all $\boldsymbol{\Delta p}_\perp$ have the same magnitude and are uncorrelated. Now N is just $\Delta t \times n \times (2\pi b\,\mathrm{d}b) \times (p_0/\mu)$: the number of particles with impact parameters between b and $b + \mathrm{d}b$ along the length of the particles path over a time Δt. Dividing by Δt and integrating over b, we have the rate of change of the perpendicular component of the momentum

$$\frac{\mathrm{d}\langle p_\perp^2\rangle}{\mathrm{d}t} = \frac{8\pi n\mu(q_1 q_2)^2}{p_0}\int_{b_\mathrm{min}}^{b_\mathrm{max}} \frac{\mathrm{d}b}{b}. \tag{7.4}$$

What are b_max and b_min, the maximum and minimum impact parameters? Clearly, if $b > \lambda_\mathrm{D}$ then the potential will be screened. Since our approximation is only good for $b > b_0$, we may take $b_\mathrm{min} = b_0$. (Our rate of scattering only depends logarithmically on $b_\mathrm{max}/b_\mathrm{min}$, so these estimates are good enough for our purposes). With this substitution,

$$\begin{aligned}\frac{\mathrm{d}\langle p_\perp^2\rangle}{\mathrm{d}t} &= \frac{8\pi n\mu(q_1 q_2)^2}{p_0}\ln\left(\frac{\lambda_\mathrm{D}}{b_0}\right)\\ &\equiv \frac{8\pi n\mu(q_1 q_2)^2}{p_0}\ln\Lambda. \end{aligned}\tag{7.5}$$

In the literature, the quantity $\ln\Lambda$ is called the Coulomb logarithm; for a plasma such as we are considering it is $\sim \ln\left(\Lambda_\mathrm{D}^3 n\right)$, the logarithm of the number of particles in a Debye sphere (see eq. [5.43]). For conditions typical of the solar center (hydrogen plasma, $\rho \gtrsim 1\,\mathrm{g\,cm^{-3}}$, $T \approx 10^7\,\mathrm{K}$), $\ln\Lambda \approx (5\text{–}10)$.

For small-angle scattering, the concept of a collision rate is fuzzy: the particle is constantly being bombarded by many tiny collisions. Setting $\mathrm{d}\langle p_\perp^2\rangle/\mathrm{d}t = p_0^2\nu$ allows us to define a deflection rate,

$$\nu \approx \frac{8\pi n\mu(q_1 q_2)^2}{p_0^3}\ln\Lambda \tag{7.6}$$

$$= \frac{8\pi n(q_1 q_2)^2}{(3k_\mathrm{B}T)^{3/2}\mu^{1/2}}\ln\Lambda. \tag{7.7}$$

Comparing equations (7.6) and (7.3), we see that many small angle scatterings are more important than single large angle scattering. Note that the ion-ion collision rate will be about $\sqrt{m_p/m_e} \approx 43$ times less than the electron-electron collision rate for a given temperature.

One can define a mean free path ℓ from equation (7.6). Consider a particle incident on a cylinder of cross-sectional area $\mathcal{A}$ and length ℓ, as illustrated in Figure 7.2. We chose ℓ so that the time for the particle to traverse it is ν^{-1}, the timescale for deflection. Thus $\ell = v_0/\nu = p_0/(\mu\nu)$. Note that for a large angle collision, we can write the probability for scattering as the total cross-section of scatterers per unit area,

$$\mathcal{P} = \frac{N\sigma}{\mathcal{A}} = \frac{n\times(\ell\mathcal{A})\sigma}{\mathcal{A}}, \tag{7.8}$$

so the particle will suffer on average a collision after traversing a distance $\ell = (n\sigma)^{-1}$. Comparing equation (7.8) with our expression for ℓ in terms of ν allows us to define an effective cross-section for small angle scattering.

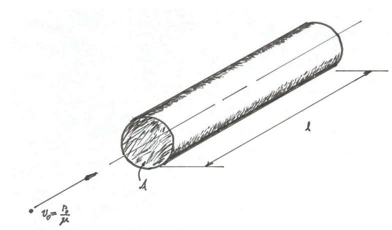


Figure 7.2: Schematic of a particle incident on a cylinder containing $n\times\ell\times\mathcal{A}$ particles.

7.2 *Transport coefficients*

We now have enough machinery to make estimates of *transport coefficients,* such as the viscosity and the thermal conductivity. Let's begin with the viscosity. Suppose we have a fluid with a gradient in the velocity, a shear, as depicted in Figure 7.3. Let the mean thermal velocity of a particle be v_0. In a time Δt, a number of particles will enter the box from the top, $(1/6)nv_0\Delta t$, and a similar number will leave the box via the top face. On average, these particles are endowed with the fluid properties of their last scattering, so the *net* momentum carried into the box across the top face is

$$\frac{1}{6}nmv_0\Delta t\Delta A\left[v_y(z_t+\ell)-v_y(z_t-\ell)\right]\approx\frac{1}{3}nmv_0\Delta t\Delta A\left.\frac{\partial v_y}{\partial z}\right|_{z_t}\ell. \tag{7.9}$$

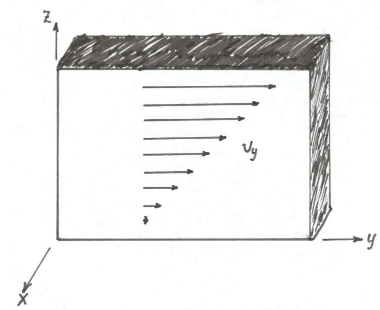


Figure 7.3: An element of fluid with a shear $\partial v_y/\partial z$.

Here ΔA is the cross-section area of our box in the xy plane and z_t is the coordinate of the top face.

A similar process occurs across the bottom face, located at coordinate $z = z_b$: the momentum flux across the bottom face is

$$\approx -\frac{1}{3}nmv_0\Delta t\Delta A\left.\frac{\partial v_y}{\partial z}\right|_{z_b}\ell. \tag{7.10}$$

Note the difference in sign: the momentum flux is positive if the y-velocity is larger below the box. Putting equations (7.9) and (7.10) together, the net change of momentum per time per unit volume $\Delta A\Delta z$ is

$$\begin{aligned}\frac{m}{\Delta A\Delta z}\frac{\Delta v_y}{\Delta t} &\approx \frac{1}{\Delta z}\frac{1}{3}\left[\left(nmv_0\ell\frac{\partial v_y}{\partial z}\right)_{z_t}-\left(nmv_0\ell\frac{\partial v_y}{\partial z}\right)_{z_b}\right]\\ &\approx \frac{\partial}{\partial z}\left(\mu\frac{\partial v_y}{\partial z}\right).\end{aligned} \tag{7.11}$$

Here we have defined $\mu = nmv_0\ell/3$ as the coefficient of dynamic viscosity.

On the left-hand side, the quantity $m/(\Delta A\Delta z)$ is just the mass density ρ, and equation (7.11), so the left-hand side is the y-component of our old friend $\rho(\partial_t\boldsymbol{u}+\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}$. On the right-hand side, we can repeat the above derivation for the transport of momentum in the x- and y-directions; if we also add back in forces from gravity and pressure gradients, we transform Euler's equation, eq. (1.5), to the *Navier-Stokes equation*,

$$\partial_t\boldsymbol{u}+\boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u}=-\boldsymbol{\nabla}\Phi-\frac{1}{\rho}\boldsymbol{\nabla}P+\frac{1}{\rho}\nabla\cdot(\mu\boldsymbol{\nabla}\boldsymbol{u}). \tag{7.12}$$

In an isothermal, incompressible fluid, one can pull μ outside the divergence operator and the last term becomes

$$\frac{\mu}{\rho}\nabla^2\boldsymbol{u}\equiv\nu\nabla^2\boldsymbol{u},$$

where ν is defined as the *coefficient of kinematic viscosity* (sorry for the overload of notation with the scattering frequency earlier!). Note that in order-of-magnitude

$$\nu \sim \frac{1}{3} v_0 \ell,$$

that is, it is roughly the thermal velocity times the mean free path.

EXERCISE 7.1 — Estimate the ratio of the pressure to viscous accelerations,

$$\frac{|\rho^{-1} \boldsymbol{\nabla} P|}{|\nu \nabla^2 \boldsymbol{u}|},$$

in equation (7.12). Express your answer in terms of a characteristic lengthscale, Mach number, and mean free path. Under what conditions are viscous effects important?

An identical procedure, but replacing the average momentum of a particle with its average thermal energy, yields an expression for the *thermal conductivity* K, such that the heat flux is

$$\boldsymbol{F} = -K \boldsymbol{\nabla} T. \tag{7.13}$$

If one writes the change in energy of a fluid element as being

$$\rho C \partial_t T = \nabla \cdot (K \boldsymbol{\nabla} T),$$

it can be seen that in order of magnitude the thermal diffusivity $\chi \equiv K/(\rho C)$, where C is the specific heat per unit mass, is $\chi \sim (1/3) v_0 \ell$. (From the form of the equation and dimensional analysis, it has to be like this.) But we need to be careful here: in a plasma with ions and electrons, the ions are responsible for momentum transport, whereas electrons, being more nimble, are more effective at heat transport. Thus the thermal diffusivity is larger than the kinematic viscosity by a factor $\sim \sqrt{m_p/m_e} \approx 43$.

EXERCISE 7.2 — Estimate the plasma thermal conductivity under conditions appropriate to the solar center. How does heat conduction by the electrons compare to that by photons?
Now suppose we wish to write a single equation for heat transport,

$$\boldsymbol{F} = -\frac{4}{3} \frac{acT^3}{\rho \kappa_{\text{total}}} \boldsymbol{\nabla} T.$$

Derive an expression for κ_{total} in terms of the electron conductivity K and the free-free opacity κ_{ff}.

8
Stellar Atmospheres

In the atmosphere of the star, the optical depth approaches unity, and we can no longer treat the radiation field as being isotropic. Let's consider the time-independent problem ($\partial_t \to 0$) of a plane-parallel atmosphere. The *optical depth* for an outward-directed ray is

$$\tau_\nu = \int_z^\infty \rho\kappa_\nu \,\mathrm{d}z'. \tag{8.1}$$

Now the optical depth just the distance divided by the mean free path. Clearly, when $\tau_\nu < 1$, a photon has a good chance of reaching a distant observer without any further interactions with the stellar matter. As a result, the intensity takes its final form around $\tau_\nu \approx 1$, and this defines the stellar *photosphere*. To get some of the basic properties of the photosphere, rewrite eq. (8.1) in differential form,

$$\frac{\mathrm{d}\tau}{\mathrm{d}z} = -\rho\kappa. \tag{8.2}$$

This is for a crude estimate, so we neglect the frequency dependence for now. We can use equation (8.2) along with hydrostatic balance to get an estimate of the photospheric pressure,

$$\frac{\mathrm{d}P}{\mathrm{d}\tau} = -\left(\frac{\mathrm{d}\tau}{\mathrm{d}z}\right)^{-1}\rho g = \frac{g}{\kappa}. \tag{8.3}$$

Thus, at $\tau \approx 1$, the pressure is $P_{\mathrm{ph}} \approx g/\kappa$. Since the flux at the photosphere is $\sigma_{\mathrm{SB}}T_{\mathrm{eff}}^4$, we would expect that the local temperature is $T \approx T_{\mathrm{eff}}$.

8.1 The Eddington Approximation

To get an analytical approximation for the atmosphere, we'll first redefine our transfer equation in terms of optical depth (eq. [6.16]). Here however, we will take the optical depth to be along the z-direction, so we define $\mu = \hat{\boldsymbol{k}}\cdot\hat{\boldsymbol{n}}$, where $\hat{\boldsymbol{n}}$ is the direction along the ray. The equation of transfer then becomes

$$\mu\frac{\partial I_\nu}{\partial \tau_\nu} = I_\nu - S_\nu, \tag{8.4}$$

where
$$S_\nu \equiv \frac{1}{\kappa_\nu}\left(\frac{\epsilon_\nu}{4\pi} + \kappa_\nu^{\rm sca} J_\nu\right) \tag{8.5}$$
is the *source function*. In local thermodynamical equilibrium (LTE), we can write $S_\nu = (1-A_\nu)B_\nu + A_\nu J_\nu$, where $A_\nu \equiv \kappa_\nu^{\rm sca}/\kappa_\nu$ is the *albedo*. Recall that $J_\nu = (4\pi)^{-1}\int d\Omega\, I_\nu$ is the angle-average of I_ν.

We noted that in thermal equilibrium, $P_\nu = c^{-1}\int_{-1}^{1} d\mu\, \mu^2 I_\nu = u_\nu/3$. This relation holds even when the radiation is not thermal, so long as it is isotropic to terms linear in μ. To make this concrete, suppose we write

$$I_\nu(\mu) = I_\nu^{(0)} + \mu I_\nu^{(1)} + \mu^2 I_\nu^{(2)} + \ldots.$$

Here we are assuming that terms marked (0) are much larger than terms marked (1), etc. To lowest order, the energy density, flux, and momentum flux are then

$$\begin{aligned}
u_\nu &= \frac{2\pi}{c}\int_{-1}^{1} d\mu\, I_\nu(\mu) = \frac{4\pi}{c} I_\nu^{(0)},\\
F_\nu &= 2\pi \int_{-1}^{1} d\mu\, \mu\, I_\nu(\mu) = \frac{4\pi}{3} I_\nu^{(1)},\\
P_\nu &= \frac{2\pi}{c}\int_{-1}^{1} d\mu\, \mu^2\, I_\nu(\mu) = \frac{4\pi}{3c} I_\nu^{(0)} = \frac{u_\nu}{3}.
\end{aligned}$$

The *Eddington approximation* then consists of treating the radiation field as if its anisotropy is linear in μ *everywhere*, so that the above relations hold; in particular, it means assuming that $P_\nu = u_\nu/3$ everywhere.

8.2 A Grey Atmosphere

Finally, to get an analytical approximation to the structure of the solar atmosphere, let's consider a grey atmosphere in LTE, i. e., one for which $\kappa_\nu^{\rm abs} = \kappa^{\rm abs}$ and $\kappa_\nu^{\rm sca} = \kappa^{\rm sca}$ are independent of frequency. Equation (8.4) can then be integrated over all frequencies to become

$$\mu \frac{\partial I}{\partial \tau} = I - S. \tag{8.6}$$

Integrating over all angles (note that we can pull the derivative wrt τ out of the integral) gives
$$\frac{1}{4\pi}\frac{\partial F}{\partial \tau} = J - S = 0. \tag{8.7}$$
Why does the right-hand side vanish? Note that $S - J = (1-A)(B-J)$. Clearly $S = J$ if $A = 1$ (a pure scattering atmosphere). If $A \neq 1$, so that there is some absorption, then the condition of detailed balance, equation (6.13), implies that $\epsilon_\nu = 4\pi\kappa^{\rm abs} B_\nu(T)$; inserting this into equation (6.12), factoring out the constant $\kappa^{\rm abs}$, and integrating over ν implies that $B - J = 0$, and hence $S - J = 0$. Note that $J = B$ does *not* necessarily imply that $I_\nu = B_\nu$!

Now multiply equation (8.6) by μ and integrate over $2\pi\,\mathrm{d}\mu$ to obtain

$$c\frac{\partial P}{\partial \tau} = F, \tag{8.8}$$

the integral over μS vanishing because it is odd in μ. Equation (8.7) implies that F is constant; hence we can integrate equation (8.8) at once to obtain

$$cP = F(\tau + \tau_0), \tag{8.9}$$

where τ_0 is a constant of integration. Of course, this does help us yet; all we have done is introduce a new variable P, the radiation pressure. This is where the Eddington approximation comes in. We set $P = u/3 = 4\pi J/(3c)$ in equation (8.9) to obtain $4\pi J = 3F(\tau + \tau_0)$. Since $J = S$, we can then write equation (8.6) as

$$\mu\frac{\partial I}{\partial \tau} = I - \frac{3}{4\pi}F(\tau + \tau_0). \tag{8.10}$$

Since F is constant, this first-order differential equation is now solvable,

$$\begin{aligned} I(\mu, \tau = 0) &= \frac{1}{\mu}\int_0^\infty \frac{3}{4\pi}F(\tau + \tau_0)e^{-\tau/\mu}\,\mathrm{d}\tau, \\ &= \frac{3}{4\pi}F(\mu + \tau_0). \end{aligned} \tag{8.11}$$

Now at $\tau = 0$, all of the flux must be outward-directed ($\mu > 0$), so $I(\mu < 0, \tau = 0) = 0$ if the star is not irradiated by another source. Note that the Eddington approximation is clearly violated here. Still, we will see later that this approximation is not too terrible.

To determine τ_0, multiply $I(\mu, \tau = 0)$ by μ and integrate equation (8.11) over all angles to find

$$F = 2\pi\int_0^1 \mu I(\mu, 0)\,\mathrm{d}\mu = \frac{1}{2}\int_0^1 3F(\mu + \tau_0)\,\mu\,\mathrm{d}\mu = F\left(\frac{1}{2} + \frac{3}{4}\tau_0\right). \tag{8.12}$$

We therefore find $\tau_0 = 2/3$. Now, since we are in LTE, $P = aT^4/3$. Further, let us define an effective temperature by the relation $F = \sigma_{\mathrm{SB}}T_{\mathrm{eff}}^4$. Substituting these definitions and the value of τ_0 into equation (8.9) gives us the atmospheric temperature structure,

$$T^4(\tau) = \frac{3}{4}T_{\mathrm{eff}}^4\left(\tau + \frac{2}{3}\right). \tag{8.13}$$

Thus $T(\tau = 0) = 2^{-1/4}T_{\mathrm{eff}}$ and $T(\tau = 2/3) = T_{\mathrm{eff}}$.

To get the spectral distribution, go back to equation (8.4) and (assuming the atmosphere has some absorption so that the matter and radiation can come into equilibrium) insert $S_\nu = B_\nu(T)$; solving for I_ν at $\tau = 0$ then gives

$$I_\nu(\mu, \tau = 0) = \frac{1}{\mu}\int_0^\infty B_\nu\left[T(\tau)\right]\,e^{-\tau/\mu}\,\mathrm{d}\tau. \tag{8.14}$$

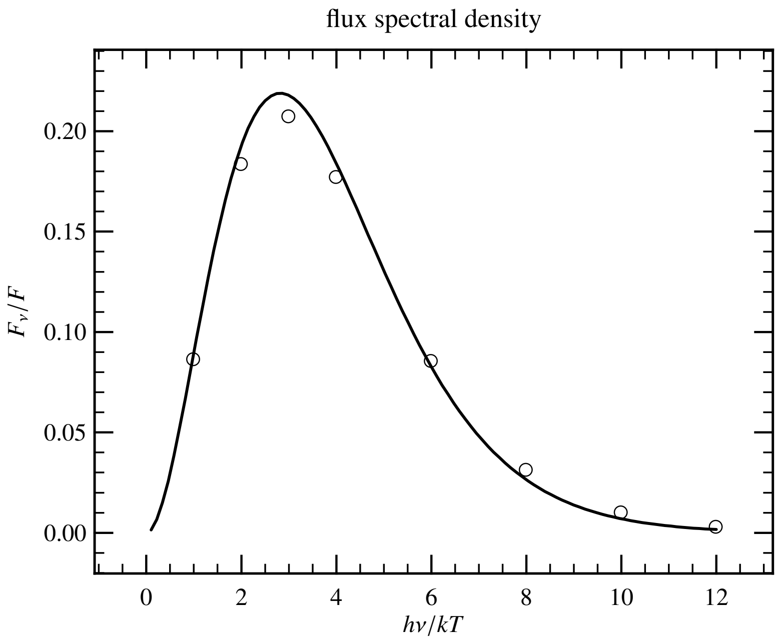


Figure 8.1: Spectral distribution from a grey atmosphere. The open circles are from Chandrasekhar, *Radiative Transfer*; the solid line is the Planck distribution.

A plot of the spectral distribution for the emergent flux is shown (*open circles*) in Fig. 8.1. For comparison, a plot of the Planck distribution (*solid line*) is also shown. Both fluxes are normalized to the total flux. Note that $I_\nu(\mu, \tau = 0)$ depends on angle; rays propagating at a slant will have a lower intensity. As a result, when we observe the sun, the edge of the visible disk appears darker than the center, a phenomenon known as *limb darkening*.

EXERCISE 8.1 — Compute the reduction in intensity I as a function of viewing angle.

8.3 Some examples

Requirement for convection in the atmosphere

Let's construct a simple atmosphere model using our temperature structure, eq. (8.13). Although we assume a grey opacity, we will let if vary with temperature and density, $\kappa(\rho, T) = \kappa_0 \rho^r T^s$. For an ideal gas, we can rewrite this in terms of pressure and temperature, $\kappa(P, T) = \kappa_0 (\mu m_\mathrm{u}/k)^r P^r T^{s-r}$. Substituting this into equation (8.3), and using equation (8.13), we obtain,

$$\frac{\mathrm{d}P}{\mathrm{d}\tau} = \frac{g}{\kappa_0} \left(\frac{k}{\mu m_\mathrm{u}}\right)^r P^{-r} \left[\frac{3}{4} T_\mathrm{eff}^4 \left(\tau + \frac{2}{3}\right)\right]^{(r-s)/4}.$$

This is easily integrated: we'll take $P(\tau = 0) = 0$ and obtain

$$P(\tau) = (\mathrm{const})\tau^{(r-s+4)/4/(1+r)}$$

so that

$$\frac{\mathrm{d}\ln P}{\mathrm{d}\tau} = \frac{r-s+4}{4(1+r)\tau}. \tag{8.15}$$

For the temperature,

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\tau} = \frac{1}{4}\frac{\mathrm{d}\ln T^4}{\mathrm{d}\tau} = \frac{1}{4(\tau+2/3)}. \tag{8.16}$$

We now combine eqn. (8.15) and (8.16) to obtain

$$\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P} = \left(\frac{1+r}{r-s+4}\right)\left(\frac{\tau}{\tau+2/3}\right). \tag{8.17}$$

For convection to happen, $\mathrm{d}\ln T/\mathrm{d}\ln P > (\partial \ln T/\partial \ln P)_s = 1/(1+n)$, where $n = 3/2$ for an ideal gas. That is,

$$n > \frac{r-s+4}{1+r} - 1 = \frac{3-s}{1+r}, \tag{8.18}$$

is required for convection to happen somewhere. Table 8.1 illustrates the behavior of $\mathrm{d}\ln T/\mathrm{d}\ln P$ for various opacity sources. The fact that the H^- opacity increases with temperature forces the temperature gradient to steepen with increasing pressure and ensures that low-mass stars have outer convective zones.

source	r	s	$\frac{3-s}{1+r}$
Thomson	0	0	3
free-free	1	$-7/2$	$13/4$
H^-	$1/2$	9	-4

Table 8.1: Right-hand side of eq. (8.18) for various opacities

An irradiated atmosphere

Many extra-solar planets are in rather tight orbits and as a result are strongly irradiated. The following example is a simplified treatment[1]. At a distance D from the star, the luminous flux is $\sigma_{\mathrm{SB}} T_\star^4 (R_\star/D)^2$. The incident intensity is then $(\sigma_{\mathrm{SB}}/\pi) W T_\star^4$, where $W = (R_\star/D)^2$, since this will give the flux when integrated over all forward directions.

An classic approximation in stellar atmospheres is to write the intensity as a sum of two streams,

$$I_\nu(\mu) = I_\nu^+ \delta\left(\mu - \frac{1}{\sqrt{3}}\right) + I_\nu^- \delta\left(\mu + \frac{1}{\sqrt{3}}\right). \tag{8.19}$$

[1] D. G. Hummer. The effect of reflected and external radiation on stellar flux distributions. *ApJ*, 257:724–732, June 1982; and I. Hubeny, A. Burrows, and D. Sudarsky. A Possible Bifurcation in Atmospheres of Strongly Irradiated Stars and Planets. *ApJ*, 594:1011–1018, September 2003

The reason for the choice of μ becomes apparent when we compute the mean intensity, the flux, and the pressure:

$$\begin{aligned} J_\nu &= \frac{1}{4\pi}\int \mathrm{d}\varphi\,\mathrm{d}\mu\; I_\nu = \frac{1}{2}\left(I_\nu^+ + I_\nu^-\right) \\ F_\nu &= \int \mathrm{d}\varphi\,\mathrm{d}\mu\; \mu I_\nu = \frac{2\pi}{\sqrt{3}}\left(I_\nu^+ - I_\nu^-\right) \\ P_\nu &= \frac{1}{c}\int \mathrm{d}\varphi\,\mathrm{d}\mu\; \mu^2 I_\nu = \frac{2\pi}{3c}\left(I_\nu^+ + I_\nu^-\right). \end{aligned}$$

You will recognize by comparing P_ν with J_ν that this formalism automatically satisfies the Eddington approximation, since $J_\nu = (c/4\pi)u_\nu$ (cf. § 6.1 and eq. [A.17]).

Following the standard method, we take successive moments of our equation of transfer (for a grey atmosphere),

$$\mu\frac{\mathrm{d}I}{\mathrm{d}\tau} = I - S,$$

to obtain

$$\begin{aligned} \frac{\mathrm{d}F}{\mathrm{d}\tau} &= 4\pi(J - S) && (8.20) \\ c\frac{\mathrm{d}P}{\mathrm{d}\tau} &= F. && (8.21) \end{aligned}$$

In LTE, $J - S = 0$ and therefore $F =$ const. We therefore integrate eq. (8.21) and use the Eddington approximation, $cP = (4\pi/3)J$, to obtain

$$J(\tau) = \frac{3}{4\pi}F\tau + J_0. \tag{8.22}$$

To determine J_0, we use our two stream approximation to write $J_0 = (\sqrt{3}/4\pi)F + I^-$. Since F is constant, we set it to its value at great depth in the star. Let us characterize F by temperature T_{int} via $F \equiv \sigma_{\mathrm{SB}}T_{\mathrm{int}}^4$. Finally, we set I^- to the incident intensity, $I^- = (\sigma_{\mathrm{SB}}/\pi)WT_\star^4$ and note that in radiative equilibrium, $J = B = (\sigma_{\mathrm{SB}}/\pi)T^4(\tau)$. Collecting terms, we have the equation for the temperature structure,

$$T^4(\tau) = \frac{3}{4}T_{\mathrm{int}}^4\left(\tau + \frac{1}{\sqrt{3}}\right) + WT_\star^4. \tag{8.23}$$

If $WT_\star \gg T_{\mathrm{int}}$, then the temperature is nearly isothermal to a depth $\tau_h \approx W(T_\star/T_{\mathrm{int}})^4$. The assumption of a grey atmosphere is, however, quite poor: the incident photons are peaked in the optical, whereas the local temperatures are in the infrared. Even taking a mean opacity is not sufficient.

8.4 Line formation and the curve of growth

Spectral lines are the diagnostics of a stellar atmosphere's temperature, pressure, and composition. We'll briefly treat here how the ambient conditions set the line shape.

The classical oscillator

Suppose we have a classical charged harmonic oscillator. The instantaneous power emitted by the oscillator is

$$P(t) = \frac{2}{3}\frac{e^2}{c^3}|\dot{\boldsymbol{u}}|^2, \tag{8.24}$$

and when averaged over a cycle is

$$\langle P(t) \rangle = \frac{e^2}{3c^3} x_0^2 \omega^4, \tag{8.25}$$

since $\dot{\boldsymbol{u}} = -\omega^2 \boldsymbol{x}_0 \cos \omega t$. Since the oscillator is radiating, it is losing energy and is damped. Let us write the damping as $\boldsymbol{F}_{\mathrm{rad}} \cdot \boldsymbol{u}$ and integrate over a cycle,

$$-\int_{t_1}^{t_2} \mathrm{d}t\, \frac{2}{3}\frac{e^2}{c^3} \dot{\boldsymbol{u}} \cdot \dot{\boldsymbol{u}} = -\frac{2}{3}\frac{e^2}{c^3} \dot{\boldsymbol{u}} \cdot \boldsymbol{u}\Big|_{t_1}^{t_2} + \frac{2}{3}\frac{e^2}{c^3} \int_{t_1}^{t_2} \mathrm{d}t\, \ddot{\boldsymbol{u}} \cdot \boldsymbol{u}.$$

The first term vanishes and we can therefore identify

$$\boldsymbol{F}_{\mathrm{rad}} = \frac{2}{3}\frac{e^2}{c^3} \ddot{\boldsymbol{u}} = -m \left(\frac{2e^2 \omega^2}{3c^3 m} \right) \boldsymbol{u}$$

as the radiation damping term with the term in parenthesis being the damping constant γ. If there is an driving electric field on our oscillator, then its equation of motion becomes

$$m\ddot{\boldsymbol{x}} = -m\omega_0^2 \boldsymbol{x} + e\boldsymbol{E} e^{i\omega t} - m\gamma \dot{\boldsymbol{x}}. \tag{8.26}$$

Using a trial function $\boldsymbol{x} \propto e^{i\omega t}$ gives

$$\boldsymbol{x} = \frac{e}{m} \frac{E e^{i\omega t}}{(\omega_0^2 - \omega^2) + i\omega\gamma}.$$

Taking the second derivative w.r.t. time of $\boldsymbol{x}$, substituting into eq. (8.24), and averaging over a cycle gives the power radiated by the oscillator,

$$\langle P(t) \rangle = \frac{e^4 \omega^4 E^2}{3c^2 m^2} \frac{1}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}.$$

Dividing $\langle P(t) \rangle$ by the incident power per unit area, $cE^2/(8\pi)$, gives the cross-section,

$$\sigma = \frac{8\pi}{3} \frac{e^4}{m^2 c^3} \frac{\omega^4}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \tag{8.27}$$

Now, for $\omega \approx \omega_0$, we can expand $(\omega_0^2 - \omega^2)^2 \approx 4\omega_0^2 (\omega_0 - \omega)^2$; furthermore, we identify $2e^2 \omega_0^2 / (3c^3 m) = \gamma$ and equation (8.27) becomes

$$\sigma = \pi \left(\frac{e^2}{mc} \right) \frac{\gamma}{(\omega_0 - \omega)^2 + (\gamma/2)^2}. \tag{8.28}$$

The line profile is Lorentzian, with a width γ. In terms of wavelength, the width is

$$\Delta\lambda = \left|\frac{d\lambda}{d\omega}\right|\gamma = \frac{2\pi c}{\omega^2}\gamma = 1.2\times 10^{-4}\,\text{Å}.$$

This width is independent of the transition frequency (it is just the classical electron radius), and it is very, very small. In a stellar atmosphere, the width is set by interactions and doppler broadening.

Suppose we model the oscillator as being started and stopped by impacts; in between impacts it just goes as $e^{i\omega_0 t}$. To get the spectrum, we take the Fourier transform,

$$F(\omega,t) = \int_0^t dt'\ \exp[i(\omega_0-\omega)t'],$$

where t is some time between impacts. Now if the impacts are distributed randomly and are uncorrelated, then the distribution of wait times follows a Poisson distribution,

$$W(t)\,dt = e^{-t/\tau}\,dt/\tau,$$

where τ is the average time between collisions. Using this to compute the energy spectrum, we obtain

$$E(\omega) = \frac{1}{2\pi\tau}\int_0^\infty dt\ F(\omega,t)F^*(\omega,t)W(t) = \frac{1}{\pi\tau}\frac{1}{(\omega_0-\omega)^2+(1/\tau)^2};$$

the line profile is again Lorentzian, with a FWHM $2/\tau$.

We might be inclined to treat the atoms as hard spheres, but this gives a large τ, or equivalently a narrow line width. We are therefore led to consider longer-range interactions for setting the intrinsic line width. Table 8.2 lists such interactions. The picture is similar to our considerations of collisions in §7.1. For a given impact parameter, the interaction perturbs the energy levels; by integrating over a distribution of impact parameters one gets the intrinsic damping. Of course, we should really use a quantum mechanical calculation. We can scale our cross-section to the classical result (eq. [8.28]), however, by writing

$$\sigma_\nu = \left(\frac{\pi e^2}{m_e c}\right) f\varphi_\nu, \tag{8.29}$$

where φ_ν is the line profile (dimension $\sim$ Hz^{-1}) and f is a dimensionless cross-section called the **oscillator strength**.

perturbation	form	source	affects
linear Stark	$C_2 r^{-2}$	e^-, p, ions	H (Hα, Hβ, ...)
quadratic Stark	$C_4 r^{-4}$	e^-	non-hydrogenic ions
van der Waals	$C_6 r^{-6}$	atoms, H	most atomic lines, esp. in cool stars

Table 8.2: Interactions in stellar atmospheres

The Curve of Growth

A classical technique in the analysis of stellar spectra is to construct the *curve of growth*, which relates the equivalent width of a line W_ν to the opacity in the line. This discussion follows Mihalas, *Stellar Atmospheres*.

Let's first get the opacity in the line. Write the cross-section for the transition $i \to j$ as

$$\sigma_\nu = \left(\frac{\pi e^2}{m_e c}\right) f_{ij}\varphi_\nu,$$

where the first term is the classical oscillator cross-section, f_{ij} is the oscillator strength and contains the quantum mechanical details of the interaction, and φ_ν is the line profile. Now recall that the opacity is given by $\kappa_\nu = n_i\sigma_\nu/\rho$, where n_i denotes the number density of available atoms in state i available to absorb a photon. Furthermore, we need to allow for *stimulated emission* from state j to state i. With this added, the opacity is (I'm writing it as χ_ν to distinguish it from the *continuum opacity*)

$$\rho\chi_\nu = \left(\frac{\pi e^2}{m_e c}\right) f_{ij}\varphi_\nu n_i \left[1 - \frac{g_i}{g_j}\frac{n_j}{n_i}\right]. \tag{8.30}$$

If we are in LTE, then the relative population of n_i and n_j follow a Boltzmann distribution,

$$1 - \frac{g_i}{g_j}\frac{n_j}{n_i} = 1 - \exp\left(-\frac{h\nu}{kT}\right).$$

This ensures we have a positive opacity. If our population were inverted, i. e., more atoms in the upper state j, then the opacity would be negative and we would have a *laser*.

Now for the line profile. In addition to damping, there is also Doppler broadening from thermal (or convective) motion. Let the line profile (here we'll switch to ν, rather than ω) be Lorentzian,

$$\varphi = \frac{\Gamma/(4\pi)}{(\nu-\nu_0)^2 + (\Gamma/[4\pi])^2}.$$

In a Maxwellian distribution, the probability of having a line-of-sight velocity in $(u, u + \mathrm{d}u)$ is

$$\mathcal{P}(u)\,\mathrm{d}u = \frac{1}{\sqrt{\pi}u_0}\exp\left(-\frac{u^2}{u_0^2}\right),$$

where $u_0 = (2kT/m)^{1/2} = 12.85\,\mathrm{km\,s^{-1}}\left(T/10^4\,\mathrm{K}\right)$ (for H) is the mean thermal velocity. The atom absorbs at a it shifted frequency $\nu(1 - u/c)$, so the mean cross section is

$$\sigma_\nu = \int_{-\infty}^{\infty} \sigma\left[\nu\left(1 - \frac{u}{c}\right)\right]\mathcal{P}(u)\,\mathrm{d}u. \tag{8.31}$$

After some algebraic manipulations, we have the cross-section

$$\begin{aligned}\sigma_\nu &= \left(\frac{\sqrt{\pi}e^2}{m_e c}\right) f_{ij} \frac{1}{\Delta\nu_D} \left\{ \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{\exp(-y^2)\,\mathrm{d}y}{(v-y)^2 + a^2} \right\} \\ &\equiv \frac{1}{\Delta\nu_D} H(a, v) \end{aligned} \tag{8.32}$$

where $\Delta\nu_D \equiv \nu u_0/c$ is the doppler width, $a = \Gamma/(4\pi\Delta\nu_D)$ is the ratio of the damping width Γ to the doppler width, and $v = \Delta\nu/\Delta\nu_D$ is the difference in frequency from the line center in units of the doppler width. The function $H(a, v)$ is called the *Voigt* function.

Let's combine the line opacity with the continuum opacity and solve the equation of transfer. For simplicity, we are going to assume pure absorption in both the continuum and the line. Under these conditions, the source function is (see the notes on the Eddington atmosphere) $S_\nu = B_\nu$, the Planck function. For a plane-parallel atmosphere, the equation of transfer is then

$$\mu \frac{\mathrm{d}I_\nu}{\mathrm{d}\tau_\nu} = I_\nu - B_\nu \tag{8.33}$$

where μ is the cosine of the angle of the ray with vertical. Solving equation (8.33) for the emergent intensity at $\tau_\nu = 0$ gives

$$I_\nu(\mu) = \frac{1}{\mu} \int_0^\infty B_\nu[T(\tau_\nu)] \exp(-\tau_\nu/\mu)\,\mathrm{d}\tau_\nu. \tag{8.34}$$

The opacity is given by

$$\kappa_\nu = \kappa_\nu^C + \chi_\nu, \tag{8.35}$$

where κ_ν^C is the continuum opacity and $\chi_\nu = \chi_0\varphi_\nu$ is the line opacity, with

$$\chi_0 = \frac{1}{\rho} \left(\frac{\pi e^2}{m_e c}\right) f_{ij} n_i \left(1 - e^{h\nu_\ell/kT}\right)$$

being the line opacity at the line center ν_ℓ.

As a further simplification, we can usually ignore the variation with ν in κ_ν^C over the width of the line. As a more suspect approximation (although it is not so bad in practice), let's assume that $\beta_\nu \equiv \chi_\nu/\kappa_C$ is independent of τ_ν. With this assumption we can write $\mathrm{d}\tau_\nu = (1+\beta_\nu)\mathrm{d}\tau$, where $\tau = -\rho\kappa^C\,\mathrm{d}z$. Finally, let's assume that in the line forming region, the temperature does not vary too much, so that we can expand B_ν to first order in τ,

$$B_\nu[T(\tau)] \approx B_0 + B_1\tau,$$

where B_0 and B_1 are constants. Inserting these approximations into equation (8.34), multiplying by the direction cosine μ and integrating over outward bound rays gives us the flux,

$$\begin{aligned} F_\nu &= 2\pi \int_0^1 \int_0^\infty [B_0 + B_1\tau] \exp\left[-\frac{\tau}{\mu}(1+\beta_\nu)\right] (1+\beta_\nu)\,\mathrm{d}\tau\,\mathrm{d}\mu \\ &= \pi\left[B_0 + \frac{2}{3}\frac{B_1}{1+\beta_\nu}\right]. \end{aligned} \tag{8.36}$$

Far from the line-center, $\beta_\nu \to 0$, implying that the continuum flux is

$$F_\nu^C = \pi \left[B_0 + \frac{2B_1}{3} \right].$$

Hence the depth of the line is

$$A_\nu \equiv 1 - \frac{F_\nu}{F_\nu^C} = A_0 \frac{\beta_\nu}{1+\beta_\nu}, \tag{8.37}$$

where

$$A_0 \equiv \frac{2B_1/3}{B_0 + 2B_1/3} \tag{8.38}$$

is the depth of an infinitely opaque ($\beta_\nu \to \infty$) line.

EXERCISE 8.2 — Explain why an infinitely opaque line ($A_0 = 0$ in eq. [8.38]) is not completely black.

Now that we have the depth of the line A_ν we can compute the *equivalent width*,

$$W_\nu \equiv \int_0^\infty A_\nu \, \mathrm{d}\nu = A_0 \int_0^\infty \frac{\beta_\nu}{1+\beta_\nu} \, \mathrm{d}\nu. \tag{8.39}$$

Let's change variables from ν to $v = \Delta\nu/\Delta\nu_D = (\nu - \nu_\ell)/\Delta\nu_D$. Since $H(a,v)$ is symmetrical about the line center, we will just integrate over $\Delta\nu > 0$, giving

$$W_\nu = 2A_0 \Delta\nu_D \int_0^\infty \frac{\beta_0 H(a,v)}{1+\beta_0 H(a,v)} \, \mathrm{d}v, \tag{8.40}$$

with $\beta_0 = \chi_0/(\kappa^C \Delta\nu_D)$.

It's useful to understand the behavior of W_ν in various limits. First, at small line optical depth ($\beta_0 \ll 1$) only the core of the line will be visible. In the core of the line, $H(a,v) \approx \exp(-v^2)$ so we insert this into equation (8.40) and expand the denominator to give

$$\begin{aligned} W_\nu^\star \equiv \frac{W\nu}{2A_0\Delta\nu_D} &= \int_0^\infty \sum_{k=1}^{\infty} (-1)^{k-1} \beta_0^k e^{-kv^2} \, \mathrm{d}v \\ &= \frac{1}{2}\sqrt{\pi}\beta_0 \left[1 - \frac{\beta_0}{\sqrt{2}} + \frac{\beta_0^2}{\sqrt{3}} - \dots \right]. \end{aligned} \tag{8.41}$$

Here $W_\nu^\star$ is the *reduced equivalent width*. Notice that since $\beta_0 \propto 1/\Delta\nu_D$ (cf. eq. [8.32]), the equivalent width W_ν is independent of $\Delta\nu_D$ in this *linear regime*. Physically, in the limit of small optical depth, each atom in state i is able to absorb photons, and the flux removed is just proportional to the number of atoms n_i.

As we increase β_0 eventually the core of the line saturates—no more absorption in the core is possible. As a result, the equivalent width should be nearly constant until there are so many absorbers that the damping wings contribute to the removal of flux. In the *saturation regime*, the Voigt function is still given by e^{-v^2}, but we can no longer assume $\beta_0 \ll 1$, so our expansion in equation (8.41) won't work. Let's go back to our integral, eq. (8.40), change variables to $z = v^2$, and define $\alpha = \ln \beta_0$ to find

$$W_\nu^\star = \frac{1}{2}\int_0^\infty \frac{z^{-1/2}}{e^{z-\alpha}+1}\,\mathrm{d}z.$$

This may not look like an improvement, but you might notice that it bears a resemblance to a Fermi-Dirac integral (see the notes on the equation of state). That means that very smart people figured out tricks to handle these integrals and all we have to do is look up what they did. In this case we have Sommerfeld to thank. In this saturation regime,

$$W_\nu^\star \approx \sqrt{\ln \beta_0}\left[1 - \frac{\pi^2}{24(\ln\beta_0)^2} - \frac{7\pi^4}{384(\ln\beta_0)^4} - \ldots\right]. \tag{8.42}$$

Note that the amount of flux removed is basically $2A_0\Delta\nu_D$: the line is maximally dark across the gaussian core.

Finally, if we continue to increase the line opacity, there will finally be so many absorbers that there will be significant flux removed from the wings. Now the form of the Voigt profile is $H(a,v) \approx (a/\sqrt{\pi})v^{-2}$, so our integral (eq. [8.40]) in this *damping regime* becomes

$$\begin{aligned} W_\nu^\star &= \int_0^\infty \left(1 + \frac{\sqrt{\pi}v^2}{\beta_0 a}\right)^{-1} \mathrm{d}v \\ &= \frac{1}{2}\left(\pi a \beta_0\right)^{1/2}. \end{aligned} \tag{8.43}$$

Note that since $a\beta_0 \propto \Delta\nu_D^{-2}$, W_ν is again independent of the doppler width in this regime.

Now that we have this "curve of growth", $W_\nu^\star(\beta_0)$, why is it useful? Since it only involves the equivalent width, it is possible to construct the curve of growth empirically without a high-resolution spectrum. Next, let's put some of the factors back into the quantities in the curve of growth. First, for a set of lines, the population of the excited state depends on the Boltzmann factor $\exp(-E/kT)$. Second, we can expand out the Doppler width in both $W_\lambda^\star$ and β_0,

$$\log\left(\frac{W_\lambda}{\Delta\lambda_D}\right) = \log\left(\frac{W_\lambda}{\lambda}\right) - \log\left(\frac{u_0}{c}\right) \tag{8.44}$$

$$\log\beta_0 = \log(g_i f_{ij}\lambda) - \frac{E}{kT} + \log(N/\kappa^C) + \log C \tag{8.45}$$

where C contains all of the constants and the continuum opacity. The temperature T is picked as a free parameter, and is picked to minimize scatter about a single curve that is assumed to fit all of the lines. What is measured then is $\log(W_\lambda/\lambda)$ and $\log(g_i f_{ij}\lambda)$; by comparing them to theoretical curves one gets an estimate of $\log(u_0/c)$, the mean velocity of atoms (may be thermal or turbulent). Since the continuum opacity κ^C usually depends on the density of H, one gets from equation (8.45) an estimate of the abundance of the line-producing element to H.

EXERCISE 8.3 — There is a subtlety involved when an atmospheric opacity is scattering-dominated, because scattering does not change the photon energy. Suppose we have an atmosphere where the Thomson scattering dominates the opacity, and the absorption of a photon is inverse bremsstrahlung (free-free), for which you can get the cross section expression from equation (6.24). Note that we do not want the Rosseland mean here, we want to know what happens to a photon of a specific frequency. Finally, we are after scalings here, so don't get hung up on the precise value of numerical prefactors.

1. Is the opacity scattering-dominated at all frequencies?
2. Trace a photon of frequency ν back into the atmosphere. How deep (in terms of the scattering optical depth) does it go before being absorbed? Is there a single well-defined photosphere for all frequencies? (*Hint: the photon is taking a random walk into the star.*)
3. Now, suppose the emergent intensity is still Planckian, but with a temperature that is the local temperature at the depth where the photon was last absorbed. Obtain an expression for T and ρ as a function of scattering optical depth, and use this to derive an approximate expression for the spectrum at high frequencies. How does it compare to a blackbody at temperature T_{eff}? *Hint*: you may find the article by Illarionov and Sunyaev (Astrophys. & Space Science **19:** 61 [1972]) helpful.

9

Contraction to the Main Sequence

9.1 The Jeans' criterion

We'll mention briefly a classic piece of stability analysis, and that is of the collapse of a homogeneous, isotropic fluid. Our equations are conservation of mass and momentum, plus Poisson's equation for the gravitational potential:

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0 \tag{9.1}$$

$$\partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u}\boldsymbol{u}) = -\rho \boldsymbol{\nabla} \Phi - \boldsymbol{\nabla} P \tag{9.2}$$

$$\nabla^2 \Phi = 4\pi G \rho. \tag{9.3}$$

Right away we run into a snag: if our system is homogenous and isotropic, then $\boldsymbol{\nabla}\Phi = 0$, since there is no preferred direction for the vector to point. In this case the left-hand side of Poisson's equation vanishes, which is inconsistent with our having a background density. This, of course, is the central point to cosmology, and if we want to do this calculation correctly we need general relativity and an expanding background universe.

Instead, we shall take an alternate route: following Jeans's lead, we simply assert that there is a background state of uniform density ρ_0. This not-quite-consistent approach still provides insight. Forging ahead, we write the density, velocity, and potential as a background piece (subscript "0") plus a perturbation (subscript "1"):

$$\begin{aligned}
\rho(\boldsymbol{x}, t) &= \rho_0 + \rho_1 \exp(i\boldsymbol{k} \cdot \boldsymbol{x} - i\omega t) \\
\boldsymbol{u}(\boldsymbol{x}, t) &= \boldsymbol{u}_1 \exp(i\boldsymbol{k} \cdot \boldsymbol{x} - i\omega t) \\
\Phi(\boldsymbol{x}, t) &= \Phi_1 \exp(i\boldsymbol{k} \cdot \boldsymbol{x} - i\omega t)
\end{aligned}$$

We take the perturbations to be adiabatic: $\mathrm{d}P = c_s^2 \mathrm{d}\rho$. Substituting these into eqns. (9.1)–(9.3), dropping all terms that are higher than first order, dotting $\boldsymbol{k}$ into the perturbed form of eq. (9.2) and using the other two equations to eliminate terms results in a dispersion relation,

$$\omega^2 = c_s^2 k^2 - 4\pi G \rho_0. \tag{9.4}$$

For sufficiently small k (long wavelengths) $\omega^2 < 0$ and the perturbations will grow exponentially. Setting $\omega = 0$ defines the **Jeans' length**,

$$\lambda_J = c_s \sqrt{\frac{\pi}{G\rho_0}}. \tag{9.5}$$

We recognize the right-hand side as just being $\sim c_s \tau_{\text{dyn}}$: equation (9.5) simply says that regions where the sound-crossing time are longer than the dynamical timescale are subject to collapse. The mass contained in a box of size λ_J is $M_J = c_s^3 (\pi/G)^{3/2}/\rho^{1/2}$. For conditions appropriate to the dense cores of molecular clouds—temperatures $\sim 10\,\text{K}$, H_2 densities $\sim 10^3\,\text{cm}^{-3}$—the Jeans mass is $M_J \sim 100\,M_\odot$.

9.2 The Hayashi Track

For a fully convective star, we can use the fact that the entropy per unit mass is the same at the photosphere as at the center to relate the surface temperature to the mass and radius of the star. Namely,

$$T_{\text{eff}} = T_c \left(\frac{P_{\text{ph}}}{P_c}\right)^{2/5}, \tag{9.6}$$

with

$$T_c = 0.5 \frac{GM\mu m_u}{kR} \tag{9.7}$$

$$P_c = 0.8 \frac{GM^2}{R^4} \tag{9.8}$$

being the central density and pressure of a polytrope of index $n = 3/2$, and P_{ph} being the root of the equation

$$P_{\text{ph}} \approx \frac{GM}{R^2} \frac{1}{\kappa_0 (\mu m_u/k)^r P_{\text{ph}}^r T_{\text{eff}}^{s-r}}. \tag{9.9}$$

(We could have used the solution for $P(\tau)$ from before, but this approximation is accurate enough to demonstrate our point.)

Now for some crazy fractions: insert equations (9.7), (9.8), and (9.9) into equation (9.6) and solve for T_{eff} to find

$$T_{\text{eff}}^{5+3r+2s} = 0.55^{5(1+r)} \kappa_0^{-2} \left(\frac{G\mu m_u}{k}\right)^{5+3r} M^{3+r} R^{3r-1}. \tag{9.10}$$

What does this say for Thomson scattering ($\kappa_0 \approx 0.4$, $r = s = 0$)? Inserting these values into eq. (9.10) gives

$$T_{\text{eff}} \approx 250\,\text{K} \left(\frac{M}{M_\odot}\right)^{3/5} \left(\frac{R}{R_\odot}\right)^{-1/5}, \tag{9.11}$$

a ridiculous value. Let's try it with H^- opacity ($\kappa_0 \approx 2.5 \times 10^{-31}$, $r = 1/2$, $s = 9$). In this case,

$$T_{\text{eff}} \approx 2200\,\text{K} \left(\frac{M}{M_\odot}\right)^{1/7} \left(\frac{R}{R_\odot}\right)^{1/49}. \tag{9.12}$$

Note the extremely weak dependence on R. This is a consequence that R is very sensitive to the entropy at the photosphere. Writing $L = 4\pi R^2 \sigma_{\mathrm{SB}} T_{\mathrm{eff}}^4$, we can solve for R and insert it into equation (9.12) to get the effective temperature in terms of mass and luminosity,

$$T_{\mathrm{eff}} \approx 2300\,\mathrm{K} \left(\frac{M}{M_\odot}\right)^{7/51} \left(\frac{L}{L_\odot}\right)^{1/102}. \tag{9.13}$$

This is a crude estimate, so don't take these numbers too seriously. What this exercise illustrates, however, is that the effective temperature for fully convective, low-mass stars is essentially independent of luminosity. On an HR diagram, these stars follow a vertical track, known as the *Hayashi track*, as they contract to the main sequence.

9.3 *Formation of a radiative core*

For low-mass stars, the opacity in the core has a Kramers'-like form, $\kappa \propto \rho T^{-7/2}$. From our scalings, we see that $\kappa \propto M^{-5/2} R^{1/2}$. As a result, as the star contracts, the central opacity decreases. If it decreases enough, then photons can carry the flux and a radiative core will develop.

To find out when a radiative core forms, let's start with our equation for the flux,

$$F(r) = -\frac{1}{3}\frac{c}{\rho\kappa}\frac{\mathrm{d}aT^4}{\mathrm{d}r},$$

and use hydrostatic balance to rewrite this as

$$L(r) = 4\pi r^2 F(r) = \frac{16\pi}{3}\frac{acT^4}{\kappa}\frac{Gm(r)}{P}\frac{\mathrm{d}\ln T}{\mathrm{d}\ln P}.$$

Now imagine we consider a tiny amount of matter δm about the center of the star. The luminosity coming out of this sphere is $\delta L(r)$, and since in a convectively stable atmosphere $\mathrm{d}\ln T/\mathrm{d}\ln P < (\Gamma_2 - 1)/\Gamma_2$, the maximum amount of energy that can be generated in the sphere δm and transported away in the absence of convection is

$$\frac{\delta L(r)}{\delta m} < \frac{16\pi}{3}\frac{Gac}{\kappa}\frac{T_c^4}{P_c}\frac{\Gamma_2 - 1}{\Gamma_2}. \tag{9.14}$$

Flashback! Do you remember doing problem 3 of chapter 2? This gave us an expression, eq. (2.33), for $\partial L/\partial m$, which we can equate with the LHS of equation (9.14):

$$q - \frac{P_c}{\rho_c(\Gamma_3 - 1)}\frac{\mathrm{D}}{\mathrm{D}t}\ln\left(\frac{P_c}{\rho_c^{\Gamma_1}}\right) < \frac{16\pi}{3}\frac{Gac}{\kappa}\frac{T_c^4}{P_c}\frac{\Gamma_2 - 1}{\Gamma_2}.$$

Let's do the time warp again: in problem 2 of chapter 4, you showed that $P/\rho^{\Gamma_1} \propto R$ for an ideal gas (i.e., the polytropic constant K); making this substitution in the derivative gives the condition for the formation of a radiative core,

$$q - \frac{N_A k_B T_c}{\mu(\Gamma_3 - 1)} \frac{\mathrm{D} \ln R}{\mathrm{D}t} < \frac{16\pi}{3} \frac{Gac}{\kappa} \frac{T_c^4}{P_c} \frac{\Gamma_2 - 1}{\Gamma_2}. \tag{9.15}$$

Stars with $M \gtrsim 0.3 M_\odot$ form a radiative core while contracting to the main-sequence; in contrast, lower-mass stars remain fully convective throughout their main-sequence lives.

EXERCISE 9.1 — We showed that low-mass pre-main-sequence stars (including brown dwarfs) are fully convective and have nearly constant effective temperatures. Use these facts to model their pre-main sequence contraction. Assume $T_\mathrm{eff} = \text{const.}$ so that the luminosity is $L = 4\pi R^2 \sigma_\mathrm{SB} T_\mathrm{eff}^4$.

1. Use the appropriate polytropic relation for the energy of the protostar and assume that the luminosity is entirely powered by contraction, i.e., the star is not yet approaching the main-sequence. Derive an equation for $R(t)$. What is the characteristic timescale for a low-mass star to contract? Scale your answer to $T_\mathrm{eff} = 3000\,\mathrm{K}$ and $M = 0.1\,M_\odot$ (i.e., get an analytical solution in terms of the variables $\tilde{T} = [T_\mathrm{eff}/3000\,\mathrm{K}]$ and $\tilde{M} = [M/0.1\,M_\odot]$).
2. Compare your findings with more elaborate calculations. You will find a review in "Theory of Low-Mass Stars and Substellar Objects," G. Chabrier and I. Baraffe, Ann. Rev. Astron. Astrophys. **38:** 337 (2000).

10
Nuclear Physics

10.1 The Nuclear Landscape

From nucleon-nucleon scattering, we find that the nuclear force operates over a range of $\lesssim 2 \times 10^{-13}$ cm $\equiv 2$ fm, where fm denotes the unit of length known as a "fermi." At low energies the nuclear force can be described as the exchange of spin-zero pions; recall from quantum field theory that the exchange of a spin-zero particle produces an attractive potential of the form $e^{-r/\lambda_\pi}/r$, where $\lambda_\pi = \hbar/(m_\pi c)$ is the Compton wavelength of the force carrier. The pion rest mass is $\approx 140\,\mathrm{MeV}/c^2$, so $\lambda_\pi \approx 1.4$ fm. The nuclear force is indeed attractive over this range, but it becomes repulsive at distances < 1 fm where higher order terms in the interaction become important.

This interaction—short range and attractive, but with a repulsive core—resembles the interaction between molecules in a fluid. Indeed, for heavy nuclei, the nucleons form a "nuclear fluid" with a characteristic density of $0.16\,\mathrm{fm}^{-3}$. The **binding energy** is defined as

$$B(N,Z) = \left[Zm_p + Nm_n - M(N,Z)\right] c^2, \tag{10.1}$$

where M is the total mass of a nucleus with N neutrons and Z protons. The total number of nucleons is $A = N + Z$, the proton rest mass is $m_p = 938.272\,\mathrm{MeV}/c^2$, and the neutron rest mass is $m_n = 939.565\,\mathrm{MeV}/c^2$. From the form of this definition $B > 0$ for bound nuclei and is ≈ 8 MeV for most nuclei. This binding of the nucleons into a "nuclear fluid" implies that one can make a crude model of the nucleus as a liquid drop, and use this model to fit the binding energy. Such a model, due to Weizäcker, is

$$-B(N,Z) = a_V A + a_S A^{2/3} + a_A \frac{(N-Z)^2}{A} + a_C \frac{Z^2}{A^{1/3}} + a_p \delta A^{-1/2}, \tag{10.2}$$

and has five terms: a bulk energy term $a_V A$ that scales with the total number of nucleons; a surface term $a_S A^{2/3}$ that corrects for the weaker binding near the surface; an asymmetry term $a_A (N-Z)^2/A$ that accounts for the energy cost to have an imbalance in the number of neutrons and protons; and a Coulomb term $a_C Z^2/A^{1/3}$ that accounts for the repulsion between other protons. The final term accounts for the pairing between like nucleons, with

$$\delta = \begin{cases} +1, & N, Z \text{ even;} \\ -1, & N, Z \text{ odd;} \\ 0, & N \text{ even, } Z \text{ odd, or } N \text{ odd, } Z \text{ even} \end{cases} .$$

Using the online fitting routine from the Institute for Nuclear Theory, I obtained the coefficients listed in Table 10.1.

coefficient	a_V	a_S	a_A	a_C	a_p
value (MeV)	-15.67	17.04	23.09	0.71	-14.55

Table 10.1: Coefficients for the Weizäcker mass formula.

EXERCISE 10.1 — In the r-process, a heavy seed nucleus captures a large number of neutrons and then decays back to stability. Suppose we start with ^{56}Fe in a bath of free neutrons, so that the iron nucleus captures 152 neutrons (with β-decays occurring as necessary to keep the nucleus bound) until it reaches the stable nucleus ^{208}Pb. Is this process exothermic or endothermic? Explain your answer.

Equation (10.2) gives a good, if somewhat crude, description of the nuclear landscape. As a first example, let's look at how the binding energy per nucleon, $-B/A$ trends with A. For simplicity, we'll ignore the pairing term. Dividing equation (10.2) by A, and denoting the neutron asymmetry by $\eta \equiv (N-Z)/(N+Z)$, we obtain

$$-\frac{B}{A} = a_V + a_S A^{-1/3} + a_A \eta^2 + \frac{a_C}{4}(1-\eta)^2 A^{2/3}. \qquad (10.3)$$

Minimizing this expression for small A, we see that the most bound nuclei (smallest $-B/A$) have $\eta \to 0$, i.e., equal numbers of neutrons and protons. As A increases, however, the Coulomb term becomes important. Expanding the last two terms and combining gives the sum of the asymmetry and Coulomb terms,

$$a_A \eta^2 + \frac{a_C}{4}(1-\eta)^2 A^{2/3} = \frac{a_C}{4}A^{2/3} - \frac{a_C}{2}A^{2/3}\eta + \left[\frac{a_C}{4}A^{2/3} + a_A\right]\eta^2.$$

For large A, this expression is minimized (although it cannot be made to vanish) for $\eta > 0$: that is, $N > Z$ and the most bound massive nuclei are neutron-rich. The nuclei for which $-B/A$ is a minimum for a fixed A define the *valley of stability*. How does the binding energy change with A along this valley? At small A, the asymmetry term dominates and $\eta^2 \ll 1$. As A increases, the surface term $a_S A^{-1/3}$ becomes smaller and B/A increases. At large A, however, the sum of the Coulomb and asymmetry terms decreases B/A. As a result, there is a peak in B/A, which is around $A = 56$.

Next, let's find the boundaries of our nuclear landscape: the most neutron-rich and proton-rich nuclei that are still bound. Define the **neutron separation energy** S_n as the energy needed to remove a neutron from a nucleus,

$$\begin{aligned} S_n(N,Z) &\equiv c^2\left\{[M(N-1,Z)+m_n]-M(N,Z)\right\} \\ &= B(N,Z)-B(N-1,Z). \end{aligned} \tag{10.4}$$

Likewise, define the **proton separation energy** as

$$\begin{aligned} S_p(N,Z) &\equiv c^2\left\{\left[M(N,Z-1)+m_p\right]-M(N,Z)\right\} \\ &= B(N,Z)-B(N,Z-1). \end{aligned} \tag{10.5}$$

If we take a nucleus (N,Z) in the valley of stability and add protons keeping N fixed, we will eventually reach a nucleus for which $S_p = 0$, that is, it costs no energy to add or to remove a proton. This defines the **proton-drip line**: nuclei more proton-rich are unstable to proton emission. Likewise, on the neutron-rich side there is the **neutron-drip line**, for which $S_n = 0$. Note that because of the pairing term there is an odd-even staggering in the S_n and S_p; it is therefore useful, sometimes, to define the two-neutron and two-proton separations energies S_{2n} and S_{2p}.

EXERCISE 10.2—

1. (a) For a fixed A, find $Z_\star(A)$ such that the binding energy per nucleon, $f = B(N = A - Z_\star, Z_\star)/A$ is maximized?
   (b) Plot $Z_\star$ vs N.
   (c) Using this $Z_\star$, plot $Y_e = Z_\star/A$ for $4 < A < 200$ and explain qualitatively any trends.
   (d) Now substitute the value of $Z_\star$ into the expression for $B(N,Z)$ and plot $B(N,Z_\star)/A$ as a function of $A = N + Z_\star$. Explain qualitatively any trends.
2. (a) For each $2 \le Z \le 82$, find the maximum value of N such that $S_n(N,Z) > 0$. Plot the values (N,Z) you find.
   (b) For each $2 \le N \le 120$, find the maximum value of Z such that $S_p(N,Z) > 0$. Plot the values (N,Z) you find.
3. Compare the plots of problems 1b, 2a, and 2b to a chart of the nuclides.

10.2 Non-resonant nuclear reactions

The situation of interest is the reaction between two nuclei, (A_1, Z_1) and (A_2, Z_2). The nuclear radius is $r_\mathrm{N} \approx A^{1/3}$ fm, and the Coulomb energy at this distance is

$$\frac{Z_1 Z_2 e^2}{r_\mathrm{N}} = \frac{Z_1 Z_2 \alpha \hbar c}{r_\mathrm{N}} \approx 1.4 Z_1 Z_2 A^{-1/3}\,\mathrm{MeV} \gg kT. \tag{10.6}$$

For nuclear reactions, typical energy scales are $\sim$ MeV and typical length scales are $\sim$ fm. In these units, $\hbar c = 197\,\mathrm{MeV\,fm}$. In the first equality in eq. (10.6), we also introduce the fine-structure constant $\alpha = e^2/(\hbar c) = 1/137$. In "nuclear units," $e^2 = \alpha\hbar c = (197\,\mathrm{MeV\,fm})/137 = 1.44\,\mathrm{MeV\,fm}$. Remember these numbers! If we scatter two nuclei together, the closest approach (cf. §7.1) is $\sim e^2/(k_\mathrm{B}T) \sim 1440\,\mathrm{fm}$ at typical stellar energies $k_\mathrm{B}T \sim 1\,\mathrm{keV}$. Clearly the cross-section for a reaction between our pair of particles is controlled by the probability of tunneling through the Coulomb potential.

For a two-body system, it is convenient to transform into a center-of-mass frame. Our problem then reduces to a one-body problem with reduced mass $m = Am_\mathrm{u}$, with $A = A_1 A_2/(A_1 + A_2)$ and incident energy $E = mv^2/2$, where v is the relative velocity of the two particles. For now, we'll neglect angular momentum ($\ell = 0$) so our scattering is s-wave. At low energies, we can form a "geometrical" cross-section from the particle wavenumber $k = p/\hbar$, with

$$\pi k^{-2} = \pi \frac{\hbar^2}{(2mE)} = 660\,\mathrm{b}\frac{1}{A}\left(\frac{\mathrm{keV}}{E}\right) \tag{10.7}$$

Here the cross-section is in units of *barns*, with $1\,\mathrm{b} = 10^{-24}\,\mathrm{cm}^2$. This is the first part of our nuclear cross-section $\sigma(E)$.

The second portion of the nuclear cross-section is the probability of tunneling through the Coulomb barrier. First, let's get the classical turning point r_E from

$$\begin{aligned} \frac{Z_1 Z_2 e^2}{r_\mathrm{E}} &= E, \\ r_\mathrm{E} &= 1440\,\mathrm{fm}\,Z_1 Z_2 \left(\frac{\mathrm{keV}}{E}\right). \end{aligned} \tag{10.8}$$

Now the wavelength is $k^{-1} = \hbar(2Am_\mathrm{u}E)^{-1/2} = \hbar c(2Am_\mathrm{u}c^2 E)^{-1/2}$ and since $m_\mathrm{u}c^2 = 932\,\mathrm{MeV}$ the wavelength $k^{-1} = 145\,\mathrm{fm}\,(\mathrm{keV}/E)^{1/2}$. The important point is that since $k^{-1} \ll r_\mathrm{E}$, we can solve the Schrödinger equation using the WKB approximation.

The WKB approximation is standard, so let me just remind you that the probability of tunneling through the barrier depends on the *action*,

$$\mathcal{P} \propto \exp\left\{\frac{2}{\hbar}\int_{r_\mathrm{E}}^{r_\mathrm{N}} \left[2m\left(\frac{Z_1 Z_2 e^2}{r} - E\right)\right]^{1/2} \mathrm{d}r\right\}. \tag{10.9}$$

To do this integral, note that $r_\mathrm{E} \gg r_\mathrm{N}$, so we can make the approximation $r_\mathrm{N} \to 0$ in the integral's lower limit; with the substitution

$$\sin\varphi = \left[2m\left(\frac{Z_1 Z_2 e^2}{r} - E\right)\right]^{1/2} \left(\frac{r}{2mZ_1 Z_2 e^2}\right)^{1/2}$$

we tame the integral and obtain

$$\mathcal{P} \propto \exp\left\{-\frac{8mZ_1Z_2e^2}{\hbar(2mE)^{1/2}}\int_0^{\pi/2}\sin^2\varphi\,\mathrm{d}\varphi\right\} = \exp\left[-\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right], \quad (10.10)$$

where

$$E_\mathrm{G} \equiv 2\pi^2 A m_\mathrm{u} c^2 \alpha^2 (Z_1Z_2)^2 = 979\,\mathrm{keV}\,A(Z_1Z_2)^2 \quad (10.11)$$

is the *Gamow energy*. Note the strong dependence on Z_1Z_2: E_G determines which reactions can occur at a given temperature. If you stare at the factor multiplying the integral in equation (10.10), you will see that $\mathcal{P} \propto \exp(-r_\mathrm{E}/\lambda)$, the exponential of ratio of the width of the forbidden region to the wavelength of the incident particle. This makes intuitive sense.

Now we have the second part of our cross-section, the probability of getting through the Coulomb barrier. This third part depends on the nuclear interactions. For non-resonant reactions, this third part does not depend strongly on energy, so it is common to define the *astrophysical S-factor* by writing the cross section as the product (geometrical) $\times$ (tunneling) $\times$ (nuclear),

$$\sigma(E) = \frac{1}{E}\exp\left[-\left(\frac{E_\mathrm{G}}{E}\right)^{1/2}\right]S(E). \quad (10.12)$$

It is easier to extrapolate the slowly varying $S(E)$ from lab energies of $> 100\,\mathrm{keV}$ down to center-of-mass energies of $\sim$ keV than it would be to fit the rapidly varying cross-section.

Now each nucleus has a Maxwellian velocity distribution,

$$n_1(\boldsymbol{v}_1)\,\mathrm{d}^3v = n_1\left(\frac{m_1}{2\pi kT}\right)^{3/2}\exp\left(-\frac{mv^2}{2kT}\right)\mathrm{d}^3v, \quad (10.13)$$

and similarly for particle 2. Let's call a particular nucleus 1 (having velocity $\boldsymbol{v}_1$) the target. By definition the cross section is

$$\frac{\text{number of reactions/target/time}}{\text{number of incident particles/area/time}},$$

so to get the number of reactions per target per time we need to multiply $\sigma(E)$ by the number of incident particles per unit area per unit time. The incident flux is just $n_2(\boldsymbol{v}_2)|\boldsymbol{v}|\,\mathrm{d}^3v_2$ where $\boldsymbol{v} = \boldsymbol{v}_2 - \boldsymbol{v}_1$. Hence the reaction rate per unit volume per unit time between a pair of particles having velocities in volumes d^3v_1 and d^3v_2 about $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is just

$$\frac{1}{1+\delta_{12}}n_1(\boldsymbol{v}_1)n_2(\boldsymbol{v}_2)\sigma(E)|\boldsymbol{v}|\,\mathrm{d}^3v_1\mathrm{d}^3v_2.$$

The factor $(1+\delta_{12})^{-1}$ is equal to $1/2$ if particles 1 and 2 are identical, and is there to avoid double-counting in that case. To get the total reaction rate per unit time, we need to integrate over the joint velocity distribution $d^3v_1\, d^3v_2$,

$$r_{12} = \frac{n_1 n_2}{1+\delta_{12}} \left[\frac{m_1 m_2}{(2\pi kT)^2}\right]^{3/2} \times \int \sigma(E) v \exp\left(-\frac{m_1 v_1^2}{2kT} - \frac{m_2 v_2^2}{2kT}\right) d^3v_1\, d^3v_2. \qquad (10.14)$$

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

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

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 (10.14) by $m_1 + m_2$ allows us to write

$$r_{12} = \frac{n_1 n_2}{1+\delta_{12}} \left(\frac{m_1+m_2}{2kT}\right)^{3/2} \left(\frac{m}{2kT}\right)^{3/2} \times \int d^3V \int d^3v\, \sigma(E) v \exp\left[-\frac{mv^2}{2kT}\right] \exp\left[-\frac{(m_1+m_2)V^2}{2kT}\right].$$

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

$$\begin{aligned} r_{12} &= \frac{1}{1+\delta_{12}} n_1 n_2 \left\{ \left(\frac{m}{2\pi kT}\right)^{3/2} \int_0^\infty \sigma(E) v \exp\left(-\frac{mv^2}{2kT}\right) 4\pi v^2\, dv \right\}. \\ &\equiv \frac{1}{1+\delta_{12}} n_1 n_2 \langle \sigma v \rangle. \end{aligned} \qquad (10.15)$$

The term in $\{\}$ is the averaging over the joint distribution of the cross-section times the velocity, and is usually denoted as $\langle \sigma v \rangle$.

Changing variables to $E = mv^2/2$ in equation (10.15) and inserting the formula for the cross-section, equation (10.12), gives

$$\langle \sigma v \rangle = \left(\frac{8}{\pi m}\right)^{1/2} \left(\frac{1}{kT}\right)^{3/2} \int_0^\infty S(E) \exp\left[-\left(\frac{E_G}{E}\right)^{1/2} - \frac{E}{kT}\right] dE. \qquad (10.16)$$

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 $\exp(-E/kT)$. As a result, the exponential will have a strong peak, and we can expand the integrand in a Taylor series about the maximum. Let

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

Then we can write

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

Here E_pk is found by solving $(\mathrm{d}f/\mathrm{d}E)|_{E=E_\mathrm{pk}} = 0$. This trick allows us to turn the integral into a Gaussian! (Before the internet, all there was to do for fun were integrals.)

Solving for E_pk, we get

$$E_\mathrm{pk} = \frac{E_\mathrm{G}^{1/3}(kT)^{2/3}}{2^{2/3}},$$

and

$$\exp\left[f(E_\mathrm{pk})\right] = \exp\left[-3\left(\frac{E_\mathrm{G}}{4kT}\right)^{1/3}\right].$$

Further,

$$\frac{1}{2}\left.\frac{\mathrm{d}^2 f}{\mathrm{d}E^2}\right|_{E=E_\mathrm{pk}} = -\frac{3}{2(2E_\mathrm{G})^{1/3}(kT)^{5/3}} = -\frac{3}{4E_\mathrm{pk}kT}.$$

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

$$\langle\sigma v\rangle = \left(\frac{8}{\pi m}\right)^{1/2}\left(\frac{1}{kT}\right)^{3/2}$$
$$\times S(E_\mathrm{pk})\exp\left[-3\left(\frac{E_\mathrm{G}}{4kT}\right)^{1/3}\right]\int_0^\infty \exp\left[-\frac{(E-E_\mathrm{pk})^2}{(\Delta/2)^2}\right]\,\mathrm{d}E \qquad (10.17)$$

How well does this approximation do? Figure 10.1 shows the integrand (*solid line*) and the approximation by a Gaussian (*dashed line*). Although the integrand is skewed to the right, the area is approximately the same. We could correct for this by taking more terms in our expansion. Consult Clayton for details.

Another simplification can be made because both the Gaussian and the original integrand go to zero as $E \to 0$. As a result, we can extend the lower bound of our integral (eq. [10.17]) to $-\infty$, and obtain

$$\langle\sigma v\rangle \approx \left(\frac{8}{\pi m}\right)^{1/2}\left(\frac{1}{kT}\right)^{3/2} S(E_\mathrm{pk})\exp\left[-3\left(\frac{E_\mathrm{G}}{4kT}\right)^{1/3}\right]\frac{\Delta}{2}$$
$$= \frac{2^{13/6}}{\sqrt{3m}}\frac{E_\mathrm{G}^{1/6}}{(kT)^{2/3}}\exp\left[-3\left(\frac{E_\mathrm{G}}{4kT}\right)^{1/3}\right]S(E_\mathrm{pk}). \qquad (10.18)$$

On to some numbers. Table 10.2 lists quantities for some common reactions. A couple of notes. First, Δ/E_pk indicates how well our Gaussian

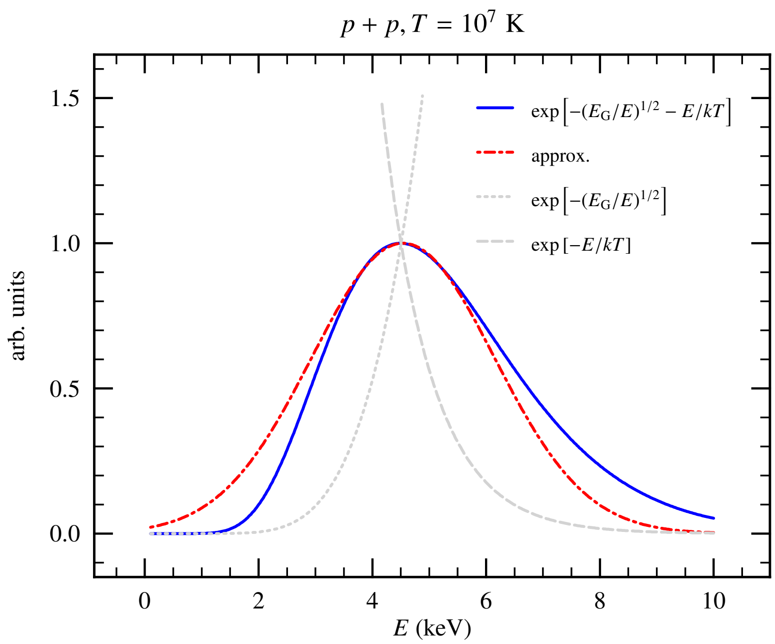


Figure 10.1: Integrand of eq. (10.16) (*solid line*) and the Gaussian (*dot-dashed line*) constructed by expanding to second order the argument of the exponential. The parameters for E_G were taken from the $p + p$ reactions ($Z_1Z_2 = 1, A = 1/2$), and the temperature is 10^7 K. Note that the grey curves, showing the two terms of the exponential, have been rescaled to fit on the same plot.

approximation works—you will see it is less than 1 in all cases. We evaluated Δ/E_{pk}, which decreases with temperature as $T^{-1/6}$, at $T = 10^7$ K. Second, the quantity $n(T)$ is the exponent if we want to approximate the reaction rate as a power-law, $r \propto T^n$. We compute this as

$$n(T) = \frac{d \ln r}{d \ln T} = -\frac{2}{3} + \left(\frac{E_G}{4kT}\right)^{1/3}, \tag{10.19}$$

as you can easily verify for yourself. In the table, the exponent is evaluated at $T = 10^7$ K; obviously n depends on temperature. Finally, note the size of $E_G/(4k)$. This makes the argument of the exponential in equation (10.18) large in absolute value, and sets the temperature scale at which a given reaction comes into play.

Reaction	$p + p$	$p + {}^3\text{He}$	${}^3\text{He} + {}^3\text{He}$	$p + {}^7\text{Li}$	$p + {}^{12}\text{C}$
A	1/2	3/4	3/2	0.88	0.92
Z_1Z_2	1	2	4	3	6
E_G (MeV)	0.489	2.94	23.5	7.70	32.5
$E_G/(4k)$ (GK)	1.4	8.5	68.0	22.0	94.0
$E_{pk}\|_{T=10^7\,\text{K}}$ (keV)	4.5	8.2	16.3	11.3	18.2
$\Delta/E_{pk}\|_{T=10^7\,\text{K}}$	1.0	0.75	0.53	0.64	0.50
$n(T = 10^7\,\text{K})$	4.6	8.8	18.3	12.4	20.5

Table 10.2: Parameters for non-resonant reactions

10.3 Resonances

This section contains my condensed notes on resonances following Blatt and Weisskopf's excellent text. Other treatments of the subject, such as that in Iliad's and in Clayton, mostly follow their approach. In this section, we shall often make use of the notation $X(a,b)Y$ to mean the reaction $X + a \to b + Y$.

Orbitals

Nuclei exhibit shell effects: one can often treat the nucleons as independent particles occupying orbitals determined by a mean force. Unlike in the atomic case, the spin-orbit term in the Hamiltonian, $-a\boldsymbol{L}\cdot\boldsymbol{S}$, is quite strong. Since the total angular momentum is $\boldsymbol{J} = \boldsymbol{L} + \boldsymbol{S}$, we have

$$\boldsymbol{L}\cdot\boldsymbol{S} = \frac{1}{2}\left(\boldsymbol{J}\cdot\boldsymbol{J} - \boldsymbol{L}\cdot\boldsymbol{L} - \boldsymbol{S}\cdot\boldsymbol{S}\right),$$

and hence states with larger $\boldsymbol{J}$ have a lower energy. The strong $\boldsymbol{L}\cdot\boldsymbol{S}$ coupling leads to the presence of "gaps" in the energy spectra; nuclei that have have filled (either neutrons or protons) shells up to this gap are unusually bound and the nucleon number (either neutron or proton) is termed a *magic number*. The magic numbers are 2, 8, 20, 28, 50, 82, and 126. For example, ^{16}O (8 protons, 8 neutrons) and ^{40}Ca (20 protons, 20 neutrons) are doubly magic and hence more strongly bound than other nuclides of similar mass.

We label the orbitals as $n\ell_j$, where n is the radial quantum number, ℓ is the orbital angular momentum ($s, p, d, f, \ldots$), and j is the total angular momentum. The first few orbitals are listed in Table 10.3. Each orbital has $2j+1$ nucleons, and a fully occupied orbital has $\boldsymbol{J} = 0$. For example, we would expect that the ground state of ^{13}C (6 protons, 7 neutrons) to have closed $1s_{1/2}$ and $1p_{3/2}$ shells for both neutrons and protons, and the remaining neutron would then occupy the $1p_{1/2}$ shell.

Orbital	Number, $2j+1$, in orbital	Total number
$1d_{3/2}$	4	20
$2s_{1/2}$	2	16
$1d_{5/2}$	6	14
$1p_{1/2}$	2	8
$1p_{3/2}$	4	6
$1s_{1/2}$	2	2

Table 10.3: Neutron orbitals

The remaining quantum number is parity ($\boldsymbol{\pi}$), which is conserved under the strong force. The parity of a nucleon orbital is $(-1)^{\ell}$, where the angular momentum number ℓ must be summed over all nucleons. Since a closed shell has an even number of nucleons, it must have positive parity. For example, the ground state of ^{17}O has 8 protons filling the $1s_{1/2}$, $1p_{3/2}$, and $1p_{1/2}$ shells 8 neutrons in closed shells, so the remaining neutron must be in the $1d_{5/2}$ shell ($\ell = 2$): the angular momentum and parity of the ground state must therefore be $j^\pi = \frac{5}{2}^+$. As a second example, ^{14}Nhas 6 protons in closed shells and 6 neutrons in closed shells, with the remaining proton and neutron both in $1p_{1/2}$ orbitals. Hence the angular momentum of the ground state could either be 0 or 1; it turns out that the symmetric state has lower energy, so $j = 1$. The parity of the ground state is $(-1)^{1+1} = 1$, so $J^\pi(^{14}\mathrm{N}) = 1^+$.

The angular momentum matters because it sets the possible range of relative angular momenta that the incoming particles can have. Writing the wave function as $\psi(r) = (u_\ell(r)/r)Y_{\ell 0}$ and substituting into Schrödinger's equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = E\psi,$$

gives the following equation for u_ℓ,

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}u_\ell + \left\{k^2 - \frac{\ell(\ell+1)}{r^2} - \frac{2m}{\hbar^2}\frac{Z_a Z_X e^2}{r}\right\}u_\ell = 0.$$

Here k is the wavenumber for the particle at $r \to \infty$, $E = \hbar^2k^2/2m$. Note the presence of the "centrifugal barrier,"

$$\frac{\hbar^2}{2mr^2}\ell(\ell+1) = 20.9\,\mathrm{MeV}\frac{m_\mathrm{u}}{m}\left(\frac{\mathrm{fm}}{r}\right)^2\ell(\ell+1):$$

there is a price to pay if the particles must have a high relative ℓ.

Formation of a compound nucleus

Even orbitals with positive energy can be long-lived; suppose we excite a proton in ^{11}C to an orbital just above the threshold for decay into $p + {}^{10}$B. Although this state has positive energy and can decay, the particle has to tunnel through the coulomb barrier, and potentially an angular momentum barrier if s-wave emission is forbidden. We saw that the probability of getting through the Coulomb barrier (for s-wave) is given by eq. (10.10). If this is very small, then we can imagine a classical particle oscillating back and forth in the well, which is does many times because the probability of escaping each time it approaches the barrier is so small. Thus, if there is a substantial energy barrier impeding escape, the classical oscillation period P is much less than the lifetime of the state τ. Now the classical oscillation period depends inversely on the spacing D between energy levels, $P \sim \hbar/D$, as you can verify for an infinite square well potential. Hence if the probability of tunneling is sufficiently small,

$$\frac{P}{\tau} \sim \frac{\hbar}{D\tau} = \frac{\Gamma}{D} \ll 1,$$

where the width of the state to particle emission is $\Gamma = \hbar/\tau$. Hence for reactions at low excitation energies involving light nuclei with widely separated levels, we can look at captures into discrete levels in the compound nucleus.

Derivation of resonant cross-section

We're now ready to do some heavy lifting and derive the cross-section for a resonant reaction. The nuclear force is short-ranged, so we can define a "channel radius" R exterior to which our potential is purely Coulomb. Our strategy is just like doing transmission resonances in quantum mechanics: we'll solve for the wave function exterior to R and match it to the wave function inside R. Since we don't completely know the form of the potential inside R, our expression will have terms that must be experimentally constrained.

The reaction is $a+X$; the coulomb potential is $Z_a Z_X e^2/r$ and the relative angular momentum of a and X is ℓ. Let m be the reduced mass of a and X. Our wave function is then $\psi(r) = (u_\ell(r)/r)Y_{\ell 0}$; substituting this into the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V(r)\psi = E\psi,$$

gives the following equation for u_ℓ,

$$\frac{\mathrm{d}^2}{\mathrm{d}r^2}u_\ell + \left\{k^2 - \frac{\ell(\ell+1)}{r^2} - \frac{2m}{\hbar^2}\frac{Z_a Z_X e^2}{r}\right\}u_\ell = 0.$$

Here k is the wavenumber for the particle at $r \to \infty$, $E = \hbar^2 k^2/2m$. There are two solutions to this differential equation; the solutions are known as Coulomb wave functions, F_ℓ and G_ℓ. The regular solution F_ℓ vanishes as $r \to 0$, and G_ℓ blows up at the origin. For $kr \gg 1$, the Coulomb wave functions go over to

$$F_\ell(r) \simeq \sin\left[kr - \frac{1}{2}\ell\pi - \gamma\ln(2kr) + \sigma_\ell\right] \qquad (10.20)$$

$$G_\ell(r) \simeq \cos\left[kr - \frac{1}{2}\ell\pi - \gamma\ln(2kr) + \sigma_\ell\right]. \qquad (10.21)$$

Here the parameter σ_ℓ is a Coulomb phase shift, and the parameter

$$\gamma = \frac{1}{2\pi}\left(\frac{E_G}{E}\right)^{1/2}$$

contains the Gamow energy. This shouldn't be too surprising: since F_ℓ and G_ℓ are the exact solutions to the motion of a particle in a Coulomb potential, they must behave like what we found using a WKB approximation in some limit.

For $r \geq R$, we can write our solution in terms of outgoing waves, u_ℓ^+, and incoming waves, u_ℓ^-; here the outgoing and incoming waves are defined as

$$\begin{aligned} u_\ell^+ &= e^{-i\sigma_\ell}\left[G_\ell + iF_\ell\right] \\ u_\ell^- &= e^{i\sigma_\ell}\left[G_\ell - iF_\ell\right]. \end{aligned}$$

At large distances these go over to plane waves, and so we can determine the coefficients,

$$u_\ell = \frac{\sqrt{\pi}}{k} i^{2\ell+1}(2\ell+1)^{1/2}\left[u_\ell^- - \eta_\ell u_\ell^+\right]. \tag{10.22}$$

The effect of the nucleus is to affect the outgoing waves via the coefficient η_ℓ. Furthermore, if we integrate the current over a large sphere we obtain the cross-section for the reaction (if the current is zero, then every particle that enters the sphere leaves it, so there is no reaction),

$$\sigma_\ell = \frac{\pi}{k^2}(2\ell+1)\left[1 - |\eta_\ell|^2\right]. \tag{10.23}$$

To determine η_ℓ, we need to find the logarithmic derivative

$$\alpha_\ell \equiv R\left.\frac{u_\ell'}{u_\ell}\right|_{r=R}$$

where $u_\ell' = \mathrm{d}u_\ell/\mathrm{d}r$. For now α_ℓ is undetermined, since it depends on the wave function inside the nucleus. It is useful to define the following quantities:

$$\frac{u_\ell^-}{u_\ell^+} = \frac{G_\ell - iF_\ell}{G_\ell + iF_\ell} e^{2i\sigma_\ell} \equiv e^{2i\xi} \tag{10.24}$$

$$R\left.\frac{u_\ell^{+\prime}}{u_\ell^+}\right|_{r=R} = R(G_\ell' G_\ell + F_\ell' F_\ell)v_\ell + ikRv_\ell \equiv \Delta_\ell + is_\ell. \tag{10.25}$$

In the last equation, we made use of the fact that $G_\ell F_\ell' - F_\ell G_\ell' = k$ for all r and defined

$$v_\ell = \left(G_\ell^2 + F_\ell^2\right)_{r=R}^{-1}.$$

To see the significance of v_ℓ, note that the ratio of an outgoing wave at $r \to \infty$ and that at $r = R$ is

$$\frac{|u_\ell(\infty)|^2}{|u_\ell(R)|^2} = \frac{1}{G_\ell^2(R) + F_\ell^2(R)},$$

where the numerator follows from the asymptotic forms, eqns. (10.20) and (10.21). This ratio, however, is just the probability of the wave transmitting through the potential barrier, and for $\ell = 0$ is approximately what we found earlier (eq. [10.10]) using the WKB approximation.

Evaluating α_ℓ for the solution defined at $r \geq R$, eq. (10.22), and using our definitions, eq. (10.24) and (10.25), we determine the phase shift,

$$\eta_\ell = \frac{\alpha_\ell - \Delta_\ell + is_\ell}{\alpha_\ell - \Delta_\ell - is_\ell} e^{2i\xi}.$$

Using this to evaluate the reaction cross-section, eq. (10.23), we have

$$\sigma_\ell = \frac{\pi}{k^2}(2\ell+1)\frac{-4s_\ell \Im\alpha_\ell}{(\Re\alpha_\ell - \Delta_\ell)^2 + (\Im\alpha_\ell - s_\ell)^2} \tag{10.26}$$

Notice that the reaction cross-section vanishes if $\alpha_\ell \in \mathbb{R}$, i.e., $\Im\alpha_\ell = 0$. So far, our efforts may just look like we are reshuffling terms for no apparent reason, but there is a method to the algebraic madness. We see explicitly, for example, that the cross-section is proportional to the penetration through the term s_ℓ in the numerator.

To make further progress, we have to evaluate α_ℓ, and this requires making some constraints on the form of the wavefunction at $r < R$. Although the precise form of the potential is unknown, we do know that it is a rather deep well. Just inside the surface R, we expect the radial wavefunction to be composed of spherical waves,

$$u_\ell(r<R) = A\left\{\exp(-iKr) + e^{2i\zeta}e^{-2q}\exp(iKr)\right\}.$$

Our reasoning for this form is as follows. In general the state will be a standing wave, which we can write as a sum of incoming and outgoing waves. The jump in the potential at $r = R$ introduces a phase shift, which we parameterize by $e^{2i\zeta}$. If the state decays by another channel than simply re-emitting the particle, i.e. if a reaction occurrs, then the amplitude of the outgoing wave will be less than that of the incoming wave; we parameterize this by the factor e^{-2q}. We expect that $q \ll 1$, because otherwise the state wouldn't be long-lasting and have a well-defined energy.

We can factor $u_\ell(r<R)$,

$$\begin{aligned} u_\ell(R) &= \left(2Ae^{i\zeta}e^{-q}\right)\left\{\frac{\exp(-iKR - i\zeta + q) + \exp(iKR + i\zeta - q)}{2}\right\} \\ &= C\cos(KR + \zeta + iq). \end{aligned}$$

Using this to evaluate the logarithmic derivative,

$$\alpha_\ell = -KR\tan(KR + \zeta + iq). \tag{10.27}$$

Now, recall that $K \gg k$; the wavenumber inside the nuclear potential well is much larger than that outside. The only way to smoothly join two waves with such discrepant wavenumbers is to have them match where $u'_\ell = 0$, i.e., where $\alpha_\ell = 0$. Let's expand α_ℓ about such a point with energy ϵ_r and $q = 0$:

$$\alpha_\ell \approx \left.\frac{\partial\alpha_\ell}{\partial\epsilon}\right|_{q=0,\epsilon=\epsilon_r}(\epsilon - \epsilon_r) + \left.\frac{\partial\alpha_\ell}{\partial q}\right|_{q=0,\epsilon=\epsilon_r} q.$$

Substituting this into equation (10.26) gives

$$\sigma = \frac{\pi}{k^2}(2\ell+1)\left\{\frac{4KRqs_\ell}{[(\partial_\epsilon\alpha_\ell)(\epsilon-\epsilon_r) - \Delta_\ell]^2 + [-KRq - s_\ell]^2}\right\}. \tag{10.28}$$

We expect that near a resonance, the cross-section will have a Lorentzian profile, with a total width

$$\Gamma = \sum_i \Gamma_i + \Gamma_\gamma$$

that is the sum of particle decay widths Γ_i and widths for radiative transitions Γ_γ. The entrance channel width will be proportional to $s_\ell = kRv_\ell$. Furthermore, a reaction will take place if the nucleus does something other than decay in the entrance channel, and this will occur with probability $\Gamma - \Gamma_a$.

Motivated by these considerations, we note that if we define the width for decay in the entrance channel as

$$\Gamma_a = -\frac{2s_\ell}{\partial_\epsilon \alpha_\ell} = -\frac{2kR\,v_\ell}{\partial_\epsilon \alpha_\ell}, \tag{10.29}$$

the reaction width as (it can be shown that in general $\partial_\epsilon \alpha_\ell < 0$)

$$\Gamma_r = \Gamma - \Gamma_\alpha = -\frac{2KR\,q}{\partial_\epsilon \alpha_\ell}, \tag{10.30}$$

and the "observed" resonance energy as

$$\epsilon_{r,\mathrm{obs}} = \epsilon_r + \frac{\Delta_\ell}{\partial_\epsilon \alpha_\ell},$$

then the cross-section becomes

$$\sigma = \frac{\pi}{k^2}(2\ell + 1)\frac{\Gamma_a \Gamma_r}{(\epsilon - \epsilon_{r,\mathrm{obs}})^2 + (\Gamma/2)^2}. \tag{10.31}$$

This is the cross-section for a compound nucleus to form in channel a and decay by any other channel. To get the cross-section for a specific exit channel b, we must multiply by the branching ratio Γ_b/Γ_r. Finally, for an unpolarized incident beam, we need to multiply the cross section by a statistical factor

$$\omega = \frac{2J + 1}{(2J_a + 1)(2J_X + 1)} \tag{10.32}$$

to account for the fraction of angular momentum states in the target X and beam a that can enter the level with the appropriate angular momentum ℓ.

You might be troubled by our identification of the particle width, eq. (10.29), and reaction width, eq. (10.30). Notice that if we were to solve the problem of a quasi-bound state leaking out of its well, we would encounter similar equations, and this would lead to the identification of the level width. We can make the argument more plausible by the following consideration. Let's consider $\partial_\epsilon \alpha_\ell$, evaluated about the point where $\alpha_\ell = 0$ and $q = 0$. The general form of α_ℓ is a tangent function (eq. [10.27]), so if we increase the energy by the spacing between levels D, the phase of the tangent must increase by π. Hence,

$\partial_\epsilon \alpha_\ell|_{KR+\zeta=n\pi,q=0} \sim -KR\pi/D$. Substituting this into eq. (10.29) and rearranging,

$$\Gamma_a \sim \hbar \left(\frac{4k}{K}\right)\left(\frac{D}{2\pi\hbar}\right) v_\ell.$$

Now, when a plane wave is incident on a step in the potential, the transmission coefficient across the step is roughly k/K, as you can verify by solving a one-dimensional Schrödinger equation. We saw earlier that classical oscillation period for a particle in a well is $2\pi\hbar/D$. Hence

$$\begin{aligned}\Gamma_a \sim \hbar &\times (\text{oscillation frequency})\\ &\times(\text{transmission across potential step at nuclear surface})\\ &\times(\text{probability of penetrating coulomb, centrifugal barrier}).\end{aligned}$$

But this is precisely what we would write down for $\hbar\times$(rate of decay in channel a)—the particle has a small probability on each oscillation to penetrate the potential jump and barrier.

A worked example

Consider the reaction ${}^{10}\mathrm{B}(p,\alpha){}^{7}\mathrm{Be}$. We can think of this reaction proceeding first via the formation of a compound nucleus in an excited state, ${}^{11}\mathrm{C}^\star$. It can be shown that the cross-section for the formation of ${}^{11}\mathrm{C}^\star$ via proton capture is proportional to the width for the decay of that state via proton emission:

$$\sigma(p + {}^{10}\mathrm{B} \to {}^{11}\mathrm{C}^\star) \propto \Gamma_p. \tag{10.33}$$

The second stage of the reaction is the decay of the ${}^{11}\mathrm{C}^\star$ into $\alpha + {}^{7}\mathrm{Be}$. Because of the short timescale $\sim 10^{-23}$ s for nuclear interactions—roughly the crossing time for a nucleon in a 20 MeV by 2 fm well—we make the assumption that the decay of an excited state does not depend on how the state was formed. If Γ_α represents the decay ${}^{11}\mathrm{C}^\star \to \alpha + {}^{7}\mathrm{Be}$, then

$$\sigma(p + {}^{10}\mathrm{B} \to \alpha + {}^{7}\mathrm{Be}) \propto \Gamma_p\Gamma_\alpha.$$

The full expression for the cross section is

$$\sigma(p + {}^{10}\mathrm{B} \to \alpha + {}^{7}\mathrm{Be}) = \frac{\pi}{k^2}(2\ell+1)\omega\frac{\Gamma_p\Gamma_\alpha}{(\epsilon - \epsilon_{r,\mathrm{obs}})^2 + (\Gamma/2)^2}. \tag{10.34}$$

Here the first term $\pi(2\ell+1)k^{-2}$ is the geometrical cross-section and Γ represents the total decay width of the excited state in the compound nucleus. The factor ω is the statistical factor

$$\omega = \frac{2J+1}{(2J_p+1)(2J_{10}+1)}$$

that accounts for the fraction of angular momentum states that can enter the level with the appropriate angular momentum ℓ. For our example, the reaction ${}^{10}\mathrm{B}(p,\alpha){}^{7}\mathrm{Be}$ can proceed via the $J^\pi = 5/2^+$ level at 8.70 MeV in ^{11}C. The angular momentum and parity of ^{10}B and the proton are $J^{\pi}(^{10}\mathrm{B}) = 3^{+}$, $J^{\pi}(p) = 1/2^{+}$. These two spins can add to either $7/2^{+}$ (multiplicity of 8) or $5/2^{+}$ (multiplicity of 6). As a result, they can enter into the resonance with $\ell = 0$, and they will do so with 6/14 probability. The cross section for this reaction is then

$$\sigma(p + {}^{10}\mathrm{B} \to {}^{7}\mathrm{Be} + \alpha) = \frac{6}{14}\frac{\pi}{k^2}\frac{\Gamma_p\Gamma_\alpha}{(\epsilon - \epsilon_{8.70})^2 + (\Gamma/2)^2}.$$

This is multiplied by the entrance channel velocity v and integrated over the thermal distribution to produce the reaction rate.

10.4 Inverse Rates

Consider a photodisintegration reaction $Y(\gamma, a)X$. We could compute the rate for this by first computing the excitation $Y \to Y^*$ by absorption of a photon followed by decay through the a-channel. There is another, easier, way to compute the thermally averaged rate, however, if we already have an expression for the forward reaction $X(a,\gamma)Y$. Suppose we allow our plasma, consisting of a, X, and Y to come into thermal balance. In such a plasma, the composition does not change with time, so our forward and inverse rates must balance:

$$n_X n_a \langle\sigma v\rangle_{Xa} = n_Y \lambda, \tag{10.35}$$

where λ is the photodissociation rate. Since we are in thermal equilibrium, however, we also have a relation between the chemical potentials,

$$\mu_X + \mu_a = \mu_Y.$$

For an ideal gas, the chemical potentials are given by eq. (5.7). Rearranging terms gives and taking the exponential gives us the equation

$$\frac{n_X n_a}{n_Y} = \frac{n_X^Q n_a^Q}{n_Y^Q}\exp\left(-\frac{Q}{k_\mathrm{B}T}\right). \tag{10.36}$$

In this expression we have substituted $Q = (m_a + m_X - m_Y)c^2$, and

$$n_i^Q = g_i\left(\frac{m_i k_\mathrm{B}T}{2\pi\hbar^2}\right)^{3/2}.$$

Substituting eq. (10.36 into eq. (10.35) gives us an expression for the photodissociation rate λ in terms of the forward rate,

$$\begin{aligned}\lambda &= \frac{n_X n_a}{n_Y}\langle\sigma v\rangle_{Xa} \\ &= \frac{g_X g_a}{g_Y}\left(\frac{m_\mathrm{u} k_\mathrm{B}T}{2\pi\hbar^2}\right)^{3/2}\left(\frac{A_X A_a}{A_Y}\right)^{3/2}\exp\left(-\frac{Q}{k_\mathrm{B}T}\right)\langle\sigma v\rangle_{Xa}.\end{aligned} \tag{10.37}$$

This expression is specific to this type of reaction, but similar formula can be generated for other types of reactions, such as $X(a,b)Y$.

10.5 Plasma corrections to the reaction rate

In stars, the nuclear reactions do not occur in isolation, but rather in the midst of a plasma. The effect of these ambient charges is to screen the long-range Coulomb interaction. This in turn perturbs the penetration factors and hence the reaction rates. We now derive the lowest order correction [1].

[1] E. E. Salpeter. Electrons Screening and Thermonuclear Reactions. *Australian Journal of Physics*, 7:373–+, September 1954

The first thing to consider are the typical scales involved. In a plasma, the mean interparticle spacing is

$$a = \left(\frac{3Am_{\mathrm{u}}}{4\pi\rho}\right)^{1/3} = 1.6\times 10^{4}\,\mathrm{fm}\times A^{1/3}\left(\frac{100\,\mathrm{g\,cm^{-3}}}{\rho}\right)^{1/3},$$

and the Debye length (cf. eq. [5.41]) is

$$\lambda_{\mathrm{D}} = 2.8\times 10^{4}\,\mathrm{fm}\times\left(\frac{T}{10^{7}\,\mathrm{K}}\right)^{1/2}\left(\frac{100\,\mathrm{g\,cm^{-3}}}{\rho}\right)^{1/2}\left(\frac{A}{\langle Z^{2}+Z\rangle}\right)^{1/2}.$$

Both of these lengths, for typical conditions in a stellar plasma, are much larger than the size of the classically forbidden region through which the particle must tunnel: at $E_{\mathrm{pk}} = 10\,\mathrm{keV}$, this length is $r_{\mathrm{E}} = 144\,Z_1Z_2\,\mathrm{fm} \ll a < \lambda_{\mathrm{D}}$. The nuclear scale is $\sim$ fm and is much smaller that all of these.

The penetration factor (eq. [10.9]) depends on the potential in the barrier; given that $r_{\mathrm{E}} \ll \lambda_{\mathrm{D}}$, we may expand the potential, eq. (5.42), and write

$$\mathcal{P} \propto \exp\left\{\frac{2}{\hbar}\int_{r_{\mathrm{E}}}^{r_{\mathrm{N}}}\left[2m\left(\frac{Z_1Z_2e^2}{r} - \frac{Z_1Z_2e^2}{\lambda_{\mathrm{D}}} - E\right)\right]^{1/2}\mathrm{d}r\right\}. \tag{10.38}$$

Since λ_{D} doesn't depend on either E or r, the effect of the screening potential is just to change the zero point of the energy scale. Assuming the S-factor doesn't depend sensitively on energy, the effect of screening on the rate is just to multiply the integrand in equation (10.15) by

$$f_{\mathrm{scr}} = \exp\left(\frac{Z_1Z_2e^2}{\lambda_{\mathrm{D}}k_{\mathrm{B}}T}\right). \tag{10.39}$$

Since this factor doesn't depend on E, it simply multiplies the rate $\langle\sigma v\rangle$. Note that since $\lambda_{\mathrm{D}} > a$ and since $e^2/(ak_{\mathrm{B}}T) \ll 1$ in a plasma, this screening factor is a small correction to the rate.

10.6 Equations for Chemical Evolution

Now that we have a reaction cross-section, we can write down the equations describing the chemical evolution of the star, and the nuclear heating. To make this concrete, let's look at the reaction $p + p \to e^{+}\nu_e + {}^{2}\mathrm{H}$. The reaction rate is (cf. eqs. [10.15] and [10.39])

$$r_{pp} = \frac{1}{2}n_{\mathrm{H}}^{2}f_{pp}^{\mathrm{scr}}\langle\sigma v\rangle_{pp}.$$

Each reaction destroys 2 protons, so we can write down our equation for the change in n_H,

$$\partial_t n_\mathrm{H} + \nabla\cdot(n_\mathrm{H}\boldsymbol{u}) = -n_\mathrm{H}^2 f_{pp}^\mathrm{scr}\langle\sigma v\rangle_{pp}. \tag{10.40}$$

Likewise, the *pp* reaction produces deuterium, which is in turn destroyed by $\mathrm{D}+p\to{}^3\mathrm{He}$:

$$\partial_t n_\mathrm{D} + \nabla\cdot(n_\mathrm{D}\boldsymbol{u}) = \frac{1}{2}n_\mathrm{H}^2 f_{pp}^\mathrm{scr}\langle\sigma v\rangle_{pp} - n_\mathrm{D}n_\mathrm{H}f_{Dp}^\mathrm{scr}\langle\sigma v\rangle_{Dp}. \tag{10.41}$$

Let's write each number density in terms of its abundance: $n_\mathrm{H} = Y_\mathrm{H}N_\mathrm{A}\rho$. We can then use the equation of mass continuity (eq. [2.1]) to simplify these equations to

$$\begin{aligned}(\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\,Y_\mathrm{H} &= -Y_\mathrm{H}^2\rho f_{pp}^\mathrm{scr}\left[N_\mathrm{A}\langle\sigma v\rangle\right]_{pp} &(10.42)\\ (\partial_t + \boldsymbol{u}\cdot\boldsymbol{\nabla})\,Y_\mathrm{D} &= \frac{1}{2}Y_\mathrm{H}^2\rho f_{pp}^\mathrm{scr}\left[N_\mathrm{A}\langle\sigma v\rangle\right]_{pp}\\ &\quad - Y_\mathrm{D}Y_\mathrm{H}\rho f_{Dp}^\mathrm{scr}\left[N_\mathrm{A}\langle\sigma v\rangle\right]_{Dp}. &(10.43)\end{aligned}$$

The left-hand side of these equations are just the Lagrangian time derivatives. On the right-hand sides, the quantities in [ ] are just functions of temperature and are compiled into rate libraries, such as REACLIB. We will need equations like (10.42) and (10.43) for each species in our star. A collection of such equations is known as a **reaction network**.

Each reaction is specified by a Q-value, which is just the energy deposited into the gas by the reactions. If there is no neutrino released, this will just be the change in nuclear binding energy, but for reaction like $p+p$ one has to account for the energy carried off by the neutrino. For the two reactions we consider here, the total heating rate, per unit mass, is

$$\begin{aligned}q &= \frac{1}{\rho}\left(\frac{Q_{pp}}{2}n_\mathrm{H}^2 f_{pp}^\mathrm{scr}\langle\sigma v\rangle_{pp} + Q_{DP}n_\mathrm{D}n_\mathrm{H}f_{Dp}^\mathrm{scr}\langle\sigma v\rangle_{Dp}\right) &(10.44)\\ &= \frac{Q_{pp}N_\mathrm{A}}{2}Y_\mathrm{H}^2\rho f_{pp}^\mathrm{scr}\left[N_\mathrm{A}\langle\sigma v\rangle\right]_{pp}\\ &\quad + Q_{Dp}N_\mathrm{A}Y_\mathrm{D}Y_\mathrm{H}\rho f_{Dp}^\mathrm{scr}\left[N_\mathrm{A}\langle\sigma v\rangle\right]_{Dp}. &(10.45)\end{aligned}$$

The equations for the change in chemical composition and the heating rate close our system of equations describing the structure and evolution of the star. We are now ready to discuss the life cycle of stars in detail.

11
Hydrogen Burning and the Main Sequence

11.1 Hydrogen burning via pp reactions: the lower main sequence

In a contracting pre-main sequence star, the reaction $^2\mathrm{H}(p,\gamma)^3\mathrm{He}$ proceeds rapidly owing to the small Coulomb barrier; in fact, this reaction can occur in objects as small as $\approx 12\,M_{\mathrm{Jupiter}}$. The small primordial abundance of deuterium, however, prevents this reaction from doing anything more than slowing contraction slightly. The reaction $p+p$ is much slower, because there is no bound nucleus $^2\mathrm{He}$; the only possible way to form a nucleus is to have a weak interaction as well, giving the reaction $p(p,e^+\nu_e)^2\mathrm{H}$.

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

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

The S-factor for the $p+p$ reaction is very small, and as a result the characteristic temperature for this reaction to occur is $\approx 1.5\times 10^7\,\mathrm{K}$; 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 $^2\mathrm{H}(p,\gamma)^3\mathrm{He}$. The nucleus $^4\mathrm{Li}$ is unbound with a lifetime of 10^{-22} s; the nucleus $^6\mathrm{Be}$ is likewise unbound ($\tau \sim 5\times 10^{-21}$ s). As a result, the next reaction that can occur is $^3\mathrm{He}(^3\mathrm{He},2p)^4\mathrm{He}$. Despite having a much greater Gamow energy than $p+p$ (see Table 10.2), this reaction still is much faster than $p+p$ owing to the small weak cross-section.

In addition to capturing another $^3\mathrm{He}$, it is also possible that

$$\begin{aligned}
^3\mathrm{He} + {}^4\mathrm{He} &\rightarrow {}^7\mathrm{Be}+\gamma \\
^7\mathrm{Be} + e^- &\rightarrow {}^7\mathrm{Li}+\nu_e \qquad (\tau = 53\,\mathrm{d}) \\
^7\mathrm{Li} + p &\rightarrow 2\,{}^4\mathrm{He}+\gamma; \qquad (11.1)
\end{aligned}$$

furthermore, at slightly higher temperatures $^7\mathrm{Be}$ can capture a proton instead of an electron, giving the third branch

$$^7\mathrm{Be} + p \rightarrow {}^8\mathrm{B}+\gamma$$

$$\begin{aligned} {}^{8}\mathrm{B} &\rightarrow {}^{8}\mathrm{Be} + e^{+} + \nu_e \qquad (\tau = 770\,\mathrm{ms}) \\ {}^{8}\mathrm{Be} &\rightarrow 2\,{}^{4}\mathrm{He} \qquad (\tau = 10^{-16}\,\mathrm{s}). \end{aligned} \tag{11.2}$$

The end result of these chains is the conversion of hydrogen to helium, although the amount of energy carried away by neutrinos differs from one chain to the next.

EXERCISE 11.1 — Compute the mass of H, in units of solar masses, that must be converted into ^{4}He in order to supply the solar luminosity over 10^{10} yr.

11.2 *Hydrogen burning via the CNO cycle: the upper main sequence*

As we saw in the previous section, the smallness of the $p+p$ cross-section means that captures onto heavier nuclei can be competitive at stellar temperatures. Let's get a rough estimate of how charged a nucleus can be before the Coulomb barrier makes the reaction slower than $p + p$. Assuming $A = 2Z$, and taking the S-factor for $p + p$ to be 10^{-22} times smaller that that for $p + {}^{\mathrm{A}}\mathrm{Z}$ gives us the rough equation

$$10^{-22} \exp\left(-\frac{33.81}{T_6^{1/3}}\right) \approx \exp\left(-\frac{41.47 Z^{2/3}}{T_6^{1/3}}\right),$$

where the factors in the exponentials come from the peak energy for the reaction (see eq. [10.19]), and $T_6 \equiv (T/10^6\,\mathrm{K})$. Solving for Z, we see that at $T_6 = 10$, proton captures onto ^{12}C have a comparable cross-section to $p + p$; at $T_6 = 20$, proton captures onto ^{16}O have a comparable cross-section.

Thus at temperatures slightly greater than that in the solar center, the following catalytic cycle becomes possible.

reaction	$\log[(\tau/\mathrm{yr}) \times (\rho X_H/100\,\mathrm{g\,cm^{-3}})]$
${}^{12}\mathrm{C}(\mathbf{p},\gamma){}^{13}\mathrm{N}$	3.82
${}^{13}\mathrm{N}(,e^{+}\nu_e){}^{13}\mathrm{C}$	$\tau = 870\,\mathrm{s}$
${}^{13}\mathrm{C}(\mathbf{p},\gamma){}^{14}\mathrm{N}$	3.21
${}^{14}\mathrm{N}(\mathbf{p},\gamma){}^{15}\mathrm{O}$	5.89
${}^{15}\mathrm{O}(,e^{+}\nu_e){}^{15}\mathrm{N}$	$\tau = 178\,\mathrm{s}$
${}^{15}\mathrm{N}(\mathbf{p},{}^{4}\mathbf{He}){}^{12}\mathrm{C}$	1.50

As indicated by the boldfaced symbols, this cycle takes in 4 protons and releases 1 helium nucleus. The reaction timescales are evaluated at a temperature of 20 MK. The reaction ${}^{14}\mathrm{N}(p,\gamma){}^{15}\mathrm{O}$ is by far the slowest step in the cycle; as a result, all of the CNO elements are quickly converted into ^{14}N in the stellar core, and this reaction controls the rate of heating. At $T = 2 \times 10^{7}$ K, $\mathrm{d}\ln\epsilon_{\mathrm{CNO}}/\mathrm{d}\ln T = 18$; in contrast the $p + p$ reaction has a temperature exponent of only 4.5.

The strong temperature sensitivity of the CNO cycle has a profound effect on the properties of the upper main sequence. Dividing equation (2.28) by equation (2.27), we have

$$\frac{\mathrm{d}T}{\mathrm{d}P} = \frac{3}{16\pi G m}\frac{\kappa}{acT^3}L_r. \tag{11.3}$$

For stars with masses $\gtrsim M_\odot$, the structure roughly follows a polytrope of index $n = 3$. We can insert the relations $T \sim M/R$ and $P \sim M^2/R^4$ into equation (11.3) and scale to the solar luminosity to obtain

$$L \approx L_\odot (M/M_\odot)^3. \tag{11.4}$$

On the other hand, we can integrate our equation for the heating rate per unit mass, $\epsilon \approx \epsilon_0 \rho T^n$, over the star; inserting the scalings for T and ρ and normalizing to solar values, we obtain $L \approx L_\odot (M/M_\odot)^{2+n}(R/R_\odot)^{-3-n}$. Equating this with L from eq. (11.4), we find that on the upper main sequence,

$$\frac{R}{R_\odot} \approx \left(\frac{M}{M_\odot}\right)^{(n-1)/(n+3)}. \tag{11.5}$$

For $n = 18$, this gives $R \sim M^{0.81}$. Since $L = L_\odot (M/M_\odot)^3 = (R/R_\odot)^2 (T_{\mathrm{eff}}/T_{\mathrm{eff},\odot})^4$, we can obtain a relation between T_{eff} and L on the upper main sequence,

$$\left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right) = \left(\frac{L}{L_\odot}\right)^{0.12}. \tag{11.6}$$

The fact that T_{eff} is so insensitive to L is a consequence that the radius increases with mass, which follows from the central temperature being roughly constant. The strong temperature sensitivity of the CNO cycle ensures that the central temperature varies only slightly over a large range of luminosity.

EXERCISE 11.2 — Suppose that there were no CNO cycle, and hydrogen could only be consumed via the PP chains. Estimate the effective temperature-luminosity relation for the upper main-sequence in this case. Would it be observationally distinguishable from the CNO-dominated upper MS?

A second effect on the stellar structure is that the luminosity is generated in a very small region concentrated about the center. Inserting $P_{\mathrm{rad}} = (a/3)T^4$ and $L_{\mathrm{Edd}} = 4\pi G M c/\kappa$ into equation (11.3) and solving for $\nabla = \mathrm{d}\ln T/\mathrm{d}\ln P$, we obtain

$$\nabla = \frac{1}{4}\frac{P}{P_{\mathrm{rad}}}\frac{L}{L_{\mathrm{Edd}}}\left(\frac{L_r}{L}\right)\left(\frac{M}{M(r)}\right). \tag{11.7}$$

For the Eddington standard model, $P/P_{\mathrm{rad}} \approx 2600(M/M_\odot)^{-2}$, and $L/L_{\mathrm{Edd}} \approx 2.7\times 10^{-5}(M/M_\odot)^2$. Inserting these factors and using the criteria for convective stability, $\nabla < (\partial \ln T/\partial \ln P)_S = 2/5$, we see that if

$$\frac{L_r}{L} > 23\frac{M(r)}{M}$$

we have convective instability. Thus, if the luminosity is produced in the innermost 4% (by mass) of the star, the core will be convective. The strong temperature sensitivity of the CNO reactions ensure that this is the case, and so the cores of upper main sequence stars have convective zones.

This convective zone changes the structure of the star, so that $L \sim M^{3.5}$ rather than the M^3 scaling used above. It also means the star can burn more of the hydrogen in its interior. The hotter $T_{\rm eff}$ means, however, that the H^- opacity is not important and the surface layers of upper main sequence stars are radiative. Table 11.1 gives a summary of the properties of main sequence stars.

characteristic	lower ($M \lesssim M_\odot$)	upper ($M \gtrsim M_\odot$)
hydrogen burning	pp	CNO
opacity	Kramers	Thomson
core	radiative, $\approx 0.1M$	convective, $\approx 0.2M$
envelope	convective	radiative

Table 11.1: Characteristics of main-sequence stars

EXERCISE 11.3 — In this exercise, we are going to examine how the core of a solar-mass star approaches a steady conversion of 4 hydrogen nuclei to 1 helium nucleus via the PPI chain

$$\begin{aligned} p + p &\rightarrow e^{+} + \nu_e + {}^{2}\mathrm{H} \\ {}^{2}\mathrm{H} + p &\rightarrow {}^{3}\mathrm{He} \\ {}^{3}\mathrm{He} + {}^{3}\mathrm{He} &\rightarrow {}^{4}\mathrm{He} + p + p. \end{aligned}$$

Denote the abundances of protons, deuterium, ^{3}He, and ^{4}He by Y_p, Y_d, Y_3, and Y_4, respectively. Furthermore, evaluate the rates at a fiducial central temperature and density (obtained with the MESA stellar evolution code) $T_{c,\odot} = 1.35 \times 10^{7}\,\mathrm{K}$, $\rho_{c,\odot} = 83.2\,\mathrm{g\,cm^{-3}}$:

$$\begin{aligned} \lambda_{pp} \equiv \rho N_{\mathrm{A}} \langle \sigma v \rangle (p + p \rightarrow {}^{2}\mathrm{H}) &= 4.40 \times 10^{-18}\,\mathrm{s^{-1}} \\ \lambda_{pd} \equiv \rho N_{\mathrm{A}} \langle \sigma v \rangle (p + {}^{2}\mathrm{H} \rightarrow {}^{3}\mathrm{He}) &= 2.58 \times 10^{-2}\,\mathrm{s^{-1}} \\ \lambda_{33} \equiv \rho N_{\mathrm{A}} \langle \sigma v \rangle ({}^{3}\mathrm{He} + {}^{3}\mathrm{He} \rightarrow {}^{4}\mathrm{He} + p + p) &= 3.40 \times 10^{-9}\,\mathrm{s^{-1}} \end{aligned}$$

Take the screening factors to be unity.

1. First, let's consider the build-up of deuterium. Start from equation (10.43):

$$\frac{\mathrm{d}}{\mathrm{d}t} Y_d = \frac{1}{2} Y_p^2 \lambda_{pp} - Y_d Y_p \lambda_{pd}. \tag{11.8}$$

   Assume that over the timescale to establish the PP chain, the abundance of hydrogen Y_p is constant, and that all the λ are constant as well. Under these assumptions, solve for $Y_d(t)$ and show that it approaches a constant value

$$\left.\frac{Y_d}{Y_p}\right|_{\mathrm{equil.}} = \frac{1}{2} \frac{\lambda_{pp}}{\lambda_{pd}}.$$

   What is this abundance? What is the timescale to reach this equilibrium?

2. Now consider the evolution of ^{3}He via production by $d + p$ and destruction via ${}^{3}\mathrm{He} + {}^{3}\mathrm{He}$:

$$\frac{\mathrm{d}}{\mathrm{d}t} Y_3 = Y_p Y_d \lambda_{pd} - Y_3^2 \lambda_{33}. \tag{11.9}$$

   Again, under the assumption that Y_p is constant, solve this equation for $Y_3(t)$. Use the value of the equilibrium value of Y_d, and show that Y_3 approaches a constant value

$$\left.\frac{Y_3}{Y_p}\right|_{\mathrm{equil.}} = \left(\frac{1}{2} \frac{\lambda_{pp}}{\lambda_{33}}\right)^{1/2}.$$

   What is this value? What is the timescale for the abundance of ^{3}He to reach 99% of this equilibrium value? Is the assumption that deuterium is at its equilibrium abundance a valid one?

3. Using the equilibrium value of ^{3}He, show that the rate of helium production via the ${}^{3}\mathrm{He} + {}^{3}\mathrm{He}$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} Y_4 = \frac{1}{4} Y_p^2 \lambda_{pp}.$$

12
Post-Main Sequence Evolution: Low-mass Stars

12.1 The Triple-Alpha Reaction

The consumption of ^{4}He is hindered by the lack of stable nuclei with mass numbers $A = 5$ and $A = 8$. The nucleus ^{8}Be is, however, long-lived by nuclear standards: its decay width is $\Gamma = 68\,\mathrm{eV}$, implying a decay timescale $\hbar/\Gamma = 9.7 \times 10^{-17}\,\mathrm{s}$. (For comparison, the lifetime of ^{5}Li is $\sim 10^{-22}\,\mathrm{s}$.) Thus if the reaction $^4\mathrm{He} + {}^4\mathrm{He} \to {}^8\mathrm{Be}$ can proceed quickly enough, a small amount of ^{8}Be can accumulate allowing the reaction $^8\mathrm{Be} + {}^4\mathrm{He} \to {}^{12}\mathrm{C}$ to proceed.

The reaction $2\,{}^4\mathrm{He} \to {}^8\mathrm{Be}$ is endothermic, with $Q = -92\,\mathrm{keV}$. As a result, the peak energy for Coulomb barrier transmission (see the discussion following eq. [10.16]) must reach

$$E_{\mathrm{pk}} = \frac{E_{\mathrm{G}}^{1/3}(kT)^{2/3}}{4^{1/3}} = -Q.$$

Substituting $E_{\mathrm{G}} = 979\,\mathrm{keV} A (Z_1 Z_2)^2$ and solving for T gives $T = 1.2 \times 10^8\,\mathrm{K}$ as the temperature required to build up any substantial amount of ^{8}Be. To reach such a temperature requires a helium core mass of $\gtrsim 0.45\,M_{\odot}$.

Once a sufficient temperature is reached, the reaction $2\,{}^4\mathrm{He} \to {}^8\mathrm{Be}$ comes into equilibrium with the decay, $^8\mathrm{Be} \to 2\,{}^4\mathrm{He}$. We can use a Saha-like equation (cf. eqns. [5.35] and [5.36]) to get the abundance of ^{8}Be. Writing

$$2\mu_4 = \mu_8 - Q$$

and substituting the expression for μ, eq. (5.7) gives

$$\frac{n_8}{n_4} = \left(\frac{2\pi\hbar^2}{m_{\mathrm{u}} kT}\right)^{3/2} \left(\frac{8}{4^2}\right)^{3/2} \frac{X_4 \rho}{4 m_{\mathrm{u}}} \exp\left(-\frac{92\,\mathrm{keV}}{kT}\right). \tag{12.1}$$

Here we use n_8 and n_4 to mean the number densities of, respectively, ^{8}Be and ^{4}He; also X_4 denotes the mass fraction of helium. Scaling eq. (12.1) to $\rho = \rho_5 10^5\,\mathrm{g\,cm^{-3}}$ and $T = T_8 10^8\,\mathrm{K}$, we obtain

$$\frac{n_8}{n_4} = 2.8 \times 10^{-5} T_8^{-3/2} X_4 \rho_5 \exp\left(-\frac{10.68}{T_8}\right). \tag{12.2}$$

At $X_4\rho_5 = 1$ and $T_8 = 1$, $n_8/n_4 = 6.5 \times 10^{-10}$.

If the reaction $^8\mathrm{Be} + {}^4\mathrm{He} \to {}^{12}\mathrm{C}$ were non-resonant, the reaction rate for this n_8 would be far too slow to account for the amount of $^{12}\mathrm{C}$ synthesized in stars. Hoyle proposed, therefore, that there should be an excited state of the $^{12}\mathrm{C}$ nucleus into which the reaction would proceed. Both $^4\mathrm{He}$ and $^8\mathrm{Be}$ have spin and parity $J^\pi = 0^+$; hence for s-wave capture (angular momentum $\ell = 0$), the state in $^{12}\mathrm{C}$ should also have $J^\pi = 0^+$. What energy should the level have? The Q-value for $^8\mathrm{Be} + {}^4\mathrm{He} \to {}^{12}\mathrm{C}$ is 7.367 MeV; the Gamow energy for this reaction is $E_\mathrm{G} = 1.67 \times 10^5$ keV and hence the peak energy is $E_\mathrm{pk} = 146$ keV, with a width $\Delta = 4(E_\mathrm{pk} kT/3)^{1/2} = 82$ keV. The proposed level should therefore have an energy within 2Δ of 7.513 MeV.

Such a level was indeed detected by Fowler, with $J^\pi = 0^+$ and $E = 7.654$ MeV. Radiative decay from this level is hampered: the ground state also has $J^\pi = 0^+$, so the decay is forbidden; and the decay to the $J^\pi = 2^+$ state at 4.44 MeV has a decay width of only $\Gamma_\mathrm{rad} = 3.67$meV. This level's primary decay is indeed back to $^8\mathrm{Be} + {}^4\mathrm{He}$. If the forward rate is fast enough, however, than a population of $^{12}\mathrm{C}$ in this excited state can accumulate. The total rate to the ground state would then be $n_{12*}\Gamma_\mathrm{rad}/\hbar$, where n_{12*} is the number density of $^{12}\mathrm{C}$ nuclei in the excited state.

To compute n_{12*}, we again can use the equation for chemical equilibrium: $\mu_8 + \mu_4 = \mu_{12*} - Q_*$. Here $Q_* = -287$ keV is the difference in energy between the $^4\mathrm{He}$ and $^8\mathrm{Be}$ nuclei and the energy of $^{12}\mathrm{C}$ *in the excited state*. Again using eq. (5.7) to expand μ, we obtain

$$n_{12*} = \frac{n_{Q,12*}}{n_{Q,4} n_{Q,8}} n_4 n_8 \exp\left(-\frac{287\,\mathrm{keV}}{kT}\right). \tag{12.3}$$

Substituting for n_8 using equation (12.1), this becomes

$$n_{12*} = \left(\frac{2\pi\hbar^2}{m_\mathrm{u} kT}\right)^3 \left(\frac{12}{4^3}\right)^{3/2} \left(\frac{X_4\rho}{4m_\mathrm{u}}\right)^3 \exp\left(-\frac{397\,\mathrm{keV}}{kT}\right). \tag{12.4}$$

Multiplying eq. (12.4) by $\Gamma_\mathrm{rad}/\hbar$ and scaling to ρ_5, T_8, we obtain the net rate, per unit volume, at which $3\,{}^4\mathrm{He} \to {}^{12}\mathrm{C}$,

$$\mathrm{rate} = 4.37 \times 10^{31}\,\mathrm{cm^{-3}\,s^{-1}} \frac{(X_4\rho_5)^3}{T_8^3} \exp\left(-\frac{44.0}{T_8}\right). \tag{12.5}$$

Multiplying the rate by $Q = 7.275$ MeV, the net Q-value, gives the volumetric heating rate $\rho\epsilon_{3\alpha}$, or

$$\epsilon_{3\alpha} \approx 5.1 \times 10^{21}\,\mathrm{erg\,g^{-1}\,s^{-1}} \frac{X_4^3\rho_5^2}{T_8^3} \exp\left(-\frac{44.0}{T_8}\right). \tag{12.6}$$

The temperature exponent is $\mathrm{d}\ln\epsilon_{3\alpha}/\mathrm{d}\ln T = 44.0/T_8 - 3$, that is, $\epsilon_{3\alpha} \sim T^{41}$ at $T = 10^8$ K.

The helium flash and the horizontal branch

The extreme temperature sensitivity of the triple-alpha reaction motivates a return to analyzing the stability of reactions in a stellar environment. We saw in an earlier problem that the "gravithermal" specific heat

$$C_\star \equiv T \left.\frac{\partial S}{\partial T}\right|_M < 0$$

for an ideal gas. The physical cause is that an increase in entropy leads to an increase in radius, and the resulting $P\,\mathrm{d}V$ work results in a reduced central temperature.

For conditions in low-mass stars at the time of helium ignition (nearly pure He at $T \approx 10^8\,\mathrm{K}$, $\rho \approx 10^5\,\mathrm{g\,cm^{-3}}$); at that density the temperature at $k_\mathrm{B}T = \epsilon_\mathrm{F}$ is (eq. [5.27]) $4.1 \times 10^8\,\mathrm{K}$. Thus, He ignition takes place under semi-degenerate conditions. To understand how the star responds, let's assume a homologous expansion—that is, one in which the ratios $r(m)/R(M)$ remains constant. In other words, we assume that the structure of the star retains its functional form and we are merely "rescaling" our radial length. To compute a gravithermal specific heat, we use Jacobians (see Appendix A.1) to transform $S(T,P)$ to $S(T,M)$:

$$\begin{aligned}
T\left(\frac{\partial S}{\partial T}\right)_M &= T\frac{\partial(S,M)}{\partial(T,M)}\\
&= T\frac{\partial(S,M)}{\partial(T,P)}\frac{\partial(T,P)}{\partial(T,M)}\\
&= T\left(\frac{\partial S}{\partial T}\right)_P - T\left(\frac{\partial S}{\partial P}\right)_T\left(\frac{\partial M}{\partial T}\right)_P\left(\frac{\partial P}{\partial M}\right)_T\\
&= C_P\left[1 - \left(\frac{\partial T}{\partial P}\right)_S\left(\frac{\partial P}{\partial T}\right)_M\right],
\end{aligned} \tag{12.7}$$

where in the last identity we have used

$$\left(\frac{\partial T}{\partial P}\right)_S\left(\frac{\partial S}{\partial T}\right)_P\left(\frac{\partial P}{\partial S}\right)_T = -1$$

and a similar expression relating P, T, and M. We could continue to use this technique of Jacobians to further transform $(\partial P/\partial T)_M$; but an easier method is to expand the equation of state as

$$\ln P = \chi_\rho \ln\rho + \chi_T \ln T, \tag{12.8}$$

where $\chi_\rho \equiv (\partial \ln P/\partial \ln\rho)_T$, and similarly for χ_T. In an homologous expansion or contraction, the polytropic index stays constant, so that from equations (4.16) and (4.17) we have

$$\left(\frac{\partial \ln P}{\partial \ln R}\right)_M = -4, \qquad \left(\frac{\partial \ln\rho}{\partial \ln R}\right)_M = -3. \tag{12.9}$$

Using these relations, we have

$$\begin{aligned} \left(\frac{\partial \ln T}{\partial \ln R}\right)_M &= \chi_T^{-1}\left[\left(\frac{\partial \ln P}{\partial \ln R}\right)_M - \chi_\rho \left(\frac{\partial \ln \rho}{\partial \ln R}\right)_M\right] \\ &= \frac{-4+3\chi_\rho}{\chi_T}. \end{aligned}$$

We can use this along with equations (12.8) and (12.9) to obtain

$$\begin{aligned} \left(\frac{\partial \ln P}{\partial \ln T}\right)_M &= \left(\frac{\partial \ln P}{\partial \ln R}\right)_M \left(\frac{\partial \ln T}{\partial \ln R}\right)_M^{-1} \\ &= \frac{4\chi_T}{4-3\chi_\rho}. \end{aligned} \tag{12.10}$$

Inserting eq. (12.10) into equation (12.7), we finally get the expression for the gravithermal specific heat under a homologous expansion or contraction,

$$C_\star = C_P\left[1 - \nabla_{\mathrm{ad}}\frac{4\chi_T}{4-3\chi_\rho}\right]. \tag{12.11}$$

For an ideal gas, $\chi_T = \chi_\rho = 1$ and $\nabla_{\mathrm{ad}} = 2/5$, so that $C_\star < 0$, as required for stability. As the core becomes degenerate, however, $\chi_\rho \to 5/3$ and $\chi_T \to 0$; as a result, the addition of heat to the core causes the temperature to rise, causing the rate of heating from nuclear reactions to increase even further. The ignition of helium in low-mass stars is therefore somewhat unstable and proceeds via "flashes."

Eventually, the burning of helium heats the core enough that degeneracy is lifted, and the star settles onto its "helium main-sequence." On an HR diagram, low-mass stars with core helium burning lie on the "horizontal branch."

The Asymptotic Giant Branch

After exhaustion of helium burning in the core, the star has a semi-degenerate C/O core, surrounded by a He burning shell with a superincumbent H-burning shell. Burning by these shells adds to the mass of the C/O core. As on the giant branch, the luminosity increases due to the high pressure at the edge of the core, and this high-luminosity produces a deep convective envelope, so that the star again lies along the Hayashi line in the HR diagram.

Because of the large gravitational acceleration at the edge of the core, the burning shells become thin in radial extent. This tends to make the burning unstable and leads to *thermal pulses*. To understand why the burning is unstable, let's revisit equation (12.11). This equation still holds for expansion or contraction of a thin layer, with one critical difference: the volume of the shell is $4\pi r_c^2 D$, where D is the thickness of the shell and r_c is the core radius, which is fixed. Hence if we expand or contract the shell, keeping the mass in the shell fixed, the logarithmic change in density is

$$\mathrm{d}\ln\rho = -\frac{\mathrm{d}D}{D} = -\frac{\mathrm{d}R}{D} = -\frac{R}{D}\mathrm{d}\ln R.$$

Replacing the coefficient 3 of χ_ρ with R/D in equation (12.11) gives us the specific heat during a homologous expansion or contraction of a shell of thickness D,

$$C_{\star,\mathrm{shell}} = C_P\left[1 - \nabla_{\mathrm{ad}}\frac{4\chi_T}{4-(R/D)\chi_\rho}\right]. \tag{12.12}$$

Even for an ideal gas, if $D < R/4$, say, then $C_{\star,\mathrm{shell}} > 0$ and the burning in the shell is thermally unstable.

When the burning shells are thin, the temperature, and hence rate of burning, become dependent on the local gravitational acceleration. Because the underlying core is degenerate, this means the luminosity depends almost entirely on the core mass: an empirical fit is[1]

$$\frac{L}{L_\odot} = 5.9\times 10^4\left(\frac{M_c}{M_\odot} - 0.52\right). \tag{12.13}$$

The rate at which mass is added to the core is set by the rate at which ^{1}H is processed into ^{4}He,

$$\frac{\mathrm{d}M_c}{\mathrm{d}t} = \frac{L}{q}. \tag{12.14}$$

Here q is the mass-specific energy release from the fusion of 4 ^{1}H into ^{4}He: $q = 26.72\,\mathrm{MeV}/(4m_\mathrm{u}) = 6.68\times 10^{18}\,\mathrm{erg\,g^{-1}}$. Thus as the core mass grows, the luminosity increases and the rate at which mass is added to the core increases. The high luminosity drives a strong wind from the stellar envelope. An empirical fit to the mass loss rate is[2]

$$\frac{\mathrm{d}M}{\mathrm{d}t} = -8.0\times 10^{-13}\,M_\odot\,\mathrm{yr}^{-1}\left(\frac{L}{L_\odot}\frac{g_\odot}{g}\frac{R_\odot}{R}\right). \tag{12.15}$$

Here $g = GM/R^2$ is the gravitational acceleration at the stellar surface. The coefficient has been increased beyond that in the original formula to fit the higher mass-loss rates observed from supergiants[3]. The envelope is thus consumed at the base by hydrogen and helium burning shells and expelled at the top by a radiative wind. This process ends when the envelope is consumed, leaving behind a degenerate C/O white dwarf that gradually cools.

[1] B. Paczyński. Evolution of Single Stars. I. Stellar Evolution from Main Sequence to White Dwarf or Carbon Ignition. *Acta Astronomica*, 20:47, 1970

[2] D. Reimers. On the absolute scale of mass-loss in red giants. I - Circumstellar absorption lines in the spectrum of the visual companion of Alpha-1 HER. *A&A*, 61:217–224, October 1977

[3] K.-P. Schröder and E. Sedlmayr. The galactic mass injection from cool stellar winds of the 1 to 2.5 $M_\odot$ stars in the solar neighbourhood. *A&A*, 366:913–922, February 2001

EXERCISE 12.1 — This exercise gives an illustration of the physics that sets the white dwarf initial mass function. Note that the numbers listed here are rather crude.

1. Combine equations (12.13) and (12.14) into a differential equation for the core mass as a function of time. Solve the equation. Since we don't know the initial core mass M_{c0} at the end of core He burning (other than that we are assuming it is greater than $0.52\,M_\odot$), let's leave that as a free parameter in the problem. What is the characteristic timescale for the core to increase in mass?
2. Solve equation (12.15) for the *total* stellar mass as a function of time, assuming the initial total mass at the start of the AGB phase is M_0. To make the problem concrete, assume that the surface effective temperature on the AGB is fixed at 4000 K, and use the result of problem 1 to get the luminosity as a function of initial core mass, M_{c0}.
3. For a star that starts its AGB phase with $M_0 = 1.0\,M_\odot$ and $M_{c0} = 0.55\,M_\odot$, what is the final white dwarf mass?

13

Production of Heavy Elements

To explore some features of the r-process, let's make a simple calculation of the abundances of a single isotopic chain of zirconium[1]. Suppose we assume that between any two neutron-rich isotopes of zirconium, $^{\mathrm{A+1}}\mathrm{Zr}$ and $^{\mathrm{A}}\mathrm{Zr}$, the reactions (n, γ) and (γ, n) are in equilibrium. Show that the ratio of the densities of these two isotopes, denoted by n_{A+1} and n_A, is given by

$$\ln\left(\frac{n_{A+1}}{n_A}\right) = \ln\left[\frac{g_{A+1}}{g_A}\left(\frac{A+1}{A}\right)^{3/2}\frac{n_n}{q_n}\right] + \frac{S(^{\mathrm{A+1}}\mathrm{Zr})}{k_{\mathrm{B}}T}. \qquad (13.1)$$

[1] Why zirconium? No particular reason, other than it is not close to a neutron magic number, so that a liquid-drop mass formula won't be too terrible.

Here $S(^{\mathrm{A+1}}\mathrm{Zr})$ is the neutron separation energy of $^{\mathrm{A+1}}\mathrm{Zr}$ (see eq. 10.4) and

$$\frac{1}{q_n} = \frac{1}{g_n}\left(\frac{2\pi\hbar^2}{m_n k_{\mathrm{B}}T}\right)^{3/2}.$$

Now we'll compute an approximate abundance distribution for $^{96}\mathrm{Zr}$–$^{116}\mathrm{Zr}$. Since we are only after a crude approximation, let's set $g_{A+1}/g_A = 1$ and $(A+1)/A = 1$ in eq. (13.1). Show that with these approximations

$$\begin{aligned}
\ln\left(\frac{n_A}{n_{96}}\right) &= \ln\left(\frac{n_A}{n_{A-1}}\right) + \ln\left(\frac{n_{A-1}}{n_{A-2}}\right) + \ln\left(\frac{n_{A-2}}{n_{A-3}}\right) + \ldots + \ln\left(\frac{n_{97}}{n_{96}}\right) \\
&= (A-96)\ln\left(\frac{n_n}{q_n}\right) + \frac{1}{k_{\mathrm{B}}T}\sum_{i=97}^{A} S(^{\mathrm{i}}\mathrm{Zr}) \qquad (13.2)
\end{aligned}$$

There is one free parameter, namely n_{96}, but this can be eliminated by requiring a normalized distribution,

$$\sum n_A = 1.$$

Compute the normalized abundances for this isotopic chain for a temperature and neutron density ($T = 1.5 \times 10^9\,\mathrm{K}, n_n = 10^{23}\,\mathrm{cm}^{-3}$). You may find it easiest to write a program to do this (please turn in your code). For the separation energy, use the liquid drop model, eq. (10.2) with coefficients taken from table 10.1. Make a table of the separation energy and abundance as a function of neutron number. Where does the abundance distribution peak? If you were to let this distribution β-decay until a stable nucleus were reached, what would you produce? Redo your calculation for $n_n = 10^{21}\ \mathrm{cm^{-3}}$ and $n_n = 10^{25}\ \mathrm{cm^{-3}}$ (same $k_\mathrm{B}T$). How does this shift the abundance pattern?

14
Galactic Chemical Evolution

14.1 The Initial Mass Function

As originally formulated[1], the initial mass funciton (IMF) is the number of stars, per unit volume, that have formed per logarithmic (base 10) mass interval:

$$\xi(\lg m) \equiv \frac{\mathrm{d}(N_\star/V)}{\mathrm{d}\lg m}. \tag{14.1}$$

[1] E. E. Salpeter. The Luminosity Function and Stellar Evolution. *ApJ*, 121:161–+, January 1955

Here $m \equiv M/M_\odot$. The IMF is derived from an observed *luminosity function*—the amount of starlight in a given waveband emitted per unit mass—and stellar models. As might be expected, this function is not well-constrained, but it is roughly a power law for $m > 1.0$, and at lower masses it flattens out. One such formulation[2] for the solar neighborhood is

$$\xi(\lg m) = \begin{cases} A_0 \exp\left[-\frac{(\lg m - \lg m_c)^2}{2\sigma^2}\right], & m < 1.0 \\ A_1 m^b & m > 1.0 \end{cases}. \tag{14.2}$$

[2] Gilles Chabrier. Galactic stellar and substellar initial mass function. *Publ.Astron.Soc.Pac.*, 115:763–796, 2003

Here $A_0 = 0.158$, $A_1 = 4.43 \times 10^{-2}$, $m_c = 0.079$, $\sigma = 0.69$, and $b = -1.3 \pm 0.3$. The coefficients A_0, A_1 are in $\lg M_\odot^{-1}\,\mathrm{pc}^{-3}$.

EXERCISE 14.1 — For the IMF given in eq. (14.2), what fraction of the stars formed with $m > 1.0$ will end their lives as core-collapse supernovae, if the mass threshold for forming a cc SNe is $8.0\,M_\odot$? What is the fraction if the mass threshold is $12.0\,M_\odot$?

The main-sequence lifetime is a rapidly decreasing function of mass: for $m > 1.0$, it goes roughly like $\tau_{\mathrm{MS}} \approx 10.0\,\mathrm{Gyr}\,m^{-2.5}$. For stars with lifetimes comparable to or longer than the age of the galactic disc τ_{G}, all stars that were ever formed are still on the main sequence, so that the IMF is identical to the *present day mass function* (PDMF) φ. For more massive stars, however, we only see those that were formed a time $\tau_{\mathrm{MS}}(m)$ ago.

Let's define a birthrate $B(t)$ as the number of stars per unit volume formed at a time t. If we make the *ansatz* that the IMF doesn't depend on time, then we can define a creation function,

$$C(\lg m, t) \equiv \xi(\lg m) \frac{B(t)}{\int_0^{\tau_G} B(t)\,\mathrm{d}t}. \tag{14.3}$$

Here the birthrate is normalized to the total number of stars formed over the age of the galactic disk. The present-day mass function is then

$$\varphi(\lg m) = \int_{\max(0,\tau_G-\tau_{MS})}^{\tau_G} C(\lg m, t)\mathrm{d}t.$$

For a constant birthrate over the age of the disk, the integral is trivial and

$$\varphi(\lg m) = \begin{cases} \xi(\lg m)\frac{\tau_{MS}(m)}{\tau_G} & \tau_{MS} < \tau_G \\ \xi(\lg m) & \tau_{MS} > \tau_G \end{cases}$$

As the galaxy ages, the stellar population becomes increasingly dominated by long-lived, low-mass stars. Empirically, the Milky Way birthrate has in fact been more or less constant (deviations less than a factor of 2) over the life of the galactic disk. The timescale for converting the present supply of gas into stars is $\sim (1\text{–}5)$Gyr.

For an IMF, we can define, for each generation of stars, a *lock-up fraction*, which is the amount of gas that is not eventually returned to the interstellar medium. Clearly this will include all stars with $\tau_{MS}(m) > \tau_G$, as well as the remnant mass, $m_{\rm rem}(m)$, of the remaining stars. For stars with mass $< 8\,M_\odot$, the mass of the white dwarf as a function of the progenitor's mass is fairly well known; more massive stars leave behind either a neutron star, for which observed masses (in binaries!) are 1–2 $M_\odot$; for black holes the remnant mass is more uncertain.

EXERCISE 14.2 — Suppose the IMF is simply a power-law, $\xi(\lg m) \propto m^{-1.3}$, for $0.1 < m < 120$. On average, how many stars are formed out of one solar mass of gas?

Delay time

Type Ia supernovae are observed in elliptical galaxies, which typically have an old stellar population and no ongoing star formation. There must be a substantial delay, then, between the time the progenitor was born and the supernovae. To see how this works, let's define a *delay time distribution* $\mathcal{D}(\tau)$ and a *realization probability* $A_{\rm Ia}(t)$. These are defined as follows: if $N_\star(t)$ is the total number of stars formed at time t, then define $N_\star(t)A_{\rm Ia}(t)$ as the total number of SNe Ia that will ever result from this generation of stars. The SNe Ia rate at the current time, $t = \tau_G$, is then

$$\Gamma_{\rm Ia}(t=\tau_G) = \int_{\tau_{\min}}^{\min(\tau_G,\tau_{\max})} \int_{\lg m_{\min}}^{\lg m_{\max}} C(\lg m, t-\tau)A_{\rm Ia}(t-\tau)\mathcal{D}(\tau)\,\mathrm{d}\lg m\,\mathrm{d}\tau. \tag{14.4}$$

In this definition, the delay time distribution is normalized to unity.

As an example, suppose all the stars are born in a burst of star formation at $t = 0$, so that $B(t) = \delta(t)$, then the SNe Ia rate at late times is

$$\Gamma_{\mathrm{Ia}}(t = \tau_{\mathrm{G}}) = \int_{\tau_{\min}}^{\min(\tau_{\mathrm{G}},\tau_{\max})} N_{\star}\delta(t - \tau)A_{\mathrm{Ia}}(t - \tau)\mathcal{D}(\tau)\,\mathrm{d}\tau = N_{\star}A_{\mathrm{Ia}}\mathcal{D}(t).$$

Notice that if $\mathcal{D}(t)$ is a very broad function of time, then the SNe Ia rate is proportional, for this case of a rapid burst of star formation at early times, to the total number of stars in the galaxy.

EXERCISE 14.3 — Consider the SNe Ia rate, eq. (14.4), following a burst of star formation, $B(t) = \delta(t)$, but now suppose that the delay time for each mass is just the main-sequence lifetime, and that A_{Ia} is independent of mass. That is, for stars with mass $m = M/M_{\odot} < 8$, we assume that some fraction A_{Ia} will become SNe Ia, and that the time for a particular mass star to evolve to explosion is just its main-sequence lifetime $\tau(m) = \tau_{\mathrm{MS}}(m)$. Show that, for $\xi(\lg m) \propto m^{-1.3}$ and $\tau_{\mathrm{MS}} = 10.0\,\mathrm{Gyr}\,m^{-2.5}$, the Ia rate is $\Gamma_{\mathrm{Ia}} \propto t^{-0.5}$, for $t > \tau_{\mathrm{MS}}(m = 8)$.

15
Neutron Stars

15.1 Cooling neutron stars

First, let's get some estimates of the neutron star. A solar mass has of order 10^{57} nucleons. If we pack those nucleons so that the mean spacing is $\approx$ 1 fm, then the radius of our object is $\sim 10^{19}$ fm $\sim$ 10 km. The current best observational constraints put the radius at around 12 km. Observed masses (from radio pulsars in binaries) range from around $1.1\,M_{\odot}$ up to $1.97\,M_{\odot}$.

Although the matter in the core of a neutron star is not an ideal gas, we'll start by ignoring the nuclear force and imagining that the core consists of cold neutrons, proton, and electrons in a charge-neutral plasma that is in β-equilibrium. This translates into three conditions:

1. Cold: The chemical potentials are just the Fermi energies, $\mu = E_{\mathrm{F}}$, with

$$E_{\mathrm{F}} = \frac{\hbar^2}{2m}\left(3\pi^2 n\right)^{2/3}, \qquad \text{non-relativistic} \tag{15.1}$$

$$E_{\mathrm{F}} = \hbar c\left(3\pi^2 n\right)^{1/3}, \qquad \text{relativistic.} \tag{15.2}$$

2. Charge-neutral: $n_e = n_p$.

3. Beta-equilibrium: The reaction $n \leftrightarrow p + e$ is in equilibrium, so $\mu_n = \mu_p + \mu_e - Q$, where $Q = (m_n - m_p - m_e)c^2 = 0.782$ MeV.

The proton fraction (and therefore the electron fraction) is indeed quite small at low density (near nuclear saturation density) where the protons and neutrons are non-relativistic (exercise). The proton-to-neutron ratio increases with density. The reactions to maintain β-equilibrium emit a neutrino: $n \to p + e + \bar{nu}_e$ and $p + e \to n + \nu_e$. After the first few seconds, the temperatures are well below 10 MeV, and the neutrino mean free path is larger than the stellar radius. Since the star is degenerate, the source of free energy for the neutrinos cannot come from contraction, and the neutron star must cool.

EXERCISE 15.1 — Find the equilibrium proton-to-neutron ratio, $x = n_p/n_n$. The electrons, being light, are always relativistic. Show that x is indeed vary small but increases with density in the limit of both neutrons and protons being non-relativistic. In the opposite limit, all three particles relativistic, show that $x \to 1/8$. *Hint:* neglect the electron mass and the difference between the neutron and proton masses; use the values $\hbar c = 197.327\,\mathrm{MeV\,fm}$ and $m_n c^2 = 939.565\,\mathrm{MeV}$, and scale the neutron density n_n to the nuclear saturation density $n_0 = 0.16\,\mathrm{fm}^{-3}$.

These reactions (named *Urca* reactions by Gamow) would indeed cause the neutron star to cool exceedingly rapidly, but there is one hitch: the reactions are blocked! As a result, neutrino emission proceeds via the modified Urca reactions:

$$\begin{aligned} n + n &\to n + p + e + \bar{\nu}_e \qquad (15.3)\\ n + p + e &\to n + n + \nu_e. \qquad (15.4) \end{aligned}$$

EXERCISE 15.2 — Show that the reactions

$$\begin{aligned} n &\to p + e + \bar{\nu}_e \\ p + e &\to n + \nu_e \end{aligned}$$

cannot simultaneously conserve momentum and energy if the n, p, and e are on their respective Fermi surfaces. Take the neutrons and protons to be non-relativistic, the electrons to be relativistic, and presume that the plasma is charge-neutral and in β-equilibrium. Neglect the electron rest mass and the difference between the neutron and proton rest masses. Recall that the momentum of a particle on its Fermi surface is $p_\mathrm{F} = \hbar(3\pi^2 n)^{1/3}$, so that $E_\mathrm{F} = p_\mathrm{F}^2/(2m)$ if non-relativistic and $E_\mathrm{F} = p_\mathrm{F} c$ if relativistic.

15.2 Accreting neutron stars

Basics of Mass Transfer

Suppose we have two objects orbiting a common center of mass. By convention, the more massive object is known as the *primary* and will be denoted by a subscript "1"; the less massive object is the *secondary* and is denoted by a subscript "2". From Kepler's law, the orbital separation is

$$a^3 = G(M_1 + M_2)\left(\frac{P}{2\pi}\right)^2, \qquad (15.5)$$

with a numerical value $a = 0.51\,R_\odot (m_1 + m_2)^{1/3}(P/\mathrm{hr})^{2/3}$, where we've scaled our masses to solar values, $m = M/M_\odot$.

In a co-rotating frame, there is an equipotential surface with a saddle point, the *inner Lagrange point*, between the two stars. This surface forms two lobes, the *Roche* lobes, that meet at this point. See the crude sketch in Figure 15.1. Although the Roche lobes are not spherical, we can define the radius of an equivalent spherical volume; for the secondary, this is

$$\frac{R_2}{a} \approx 0.462 \left(\frac{M_2}{M_1 + M_2}\right)^{1/3}. \tag{15.6}$$

As an aside, equations (15.6) and (15.5) imply that the average density of the secondary, if it fills its Roche lobe, is

$$\bar{\rho} = \frac{3M_2}{4\pi R_2^3} \approx 111\,\mathrm{g\,cm^{-3}} \left(\frac{\mathrm{hr}}{P}\right)^2.$$

and does not depend explicitly on the masses of the two stars.

If matter is transferred from M_2 to M_1 how does the system respond? Let's first write down the angular momentum of the system,

$$J = (M_1 a_1^2 + M_2 a_2^2)\omega = M_1 M_2 \left(\frac{Ga}{M_1 + M_2}\right)^{1/2}. \tag{15.7}$$

Here we've used equation (15.5) and the relations $a_1 = aM_2/(M_1 + M_2)$, $a_2 = aM_1/(M_1 + M_2)$. Let's make the assumption that no mass is lost from the system and that the mass transfer $\dot{M}$ is from M_2 to M_1:

$$\begin{aligned} \dot{M}_1 + \dot{M}_2 &= 0; \\ \dot{M}_2 = -\dot{M} &< 0. \end{aligned}$$

Taking the logarithm of equation (15.7) and differentiating, we then obtain

$$\frac{\dot{a}}{a} = 2\frac{\dot{J}}{J} + 2\left(\frac{\dot{M}}{M_2}\right)\left(1 - \frac{M_2}{M_1}\right). \tag{15.8}$$

We see explicitly that for $M_2 < M_1$, the response of mass transfer is to increase the orbital separation a if there is no external torque on the system.

What about the size of the lobe that the secondary inhabits? There are two countervailing tendencies: a increases, which acts to increase R_2, but M_2 decreases, which acts to decrease R_2. Taking the logarithm of eq. (15.6),

$$\frac{\dot{R}_2}{R_2} = \frac{\dot{a}}{a} + \frac{1}{3}\frac{\dot{M}_2}{M_2} = 2\frac{\dot{J}}{J} + 2\frac{\dot{M}}{M_2}\left(\frac{5}{6} - \frac{M_2}{M_1}\right). \tag{15.9}$$

Hence for $M_2 < (5/6)M_1$, the volume of the secondary's Roche lobe increases; if the secondary doesn't expand in response to mass loss and there are no external torques, then the secondary will lose contact with the inner Lagrange point and mass transfer will cease. Alternatively, if $M_2 > (5/6)M_1$, then the Roche lobe will clamp down on the secondary; this tends to drive the mass transfer at an even greater rate, and the process is unstable.

In general, there are three physical mechanisms for driving mass transfer.

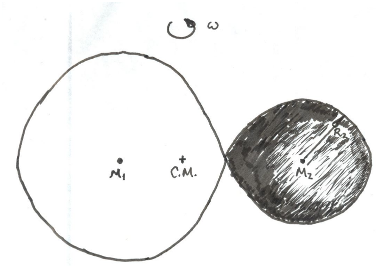


Figure 15.1: Sketch of the Roche lobes. Here the secondary, M_2, is depicted as filling its lobe.

Gravitational radiation For $P \lesssim 2\,\mathrm{hr}$, gravitational radiation from the orbit produces a negative torque on the system:

$$\frac{\dot{J}}{J} = -\frac{32}{5}\frac{G^3}{c^5}\frac{M_1 M_2 (M_1 + M_2)}{a^4}. \tag{15.10}$$

Note that for this short of an orbital period, the binary consists of two degenerate stars (e.g., WD-WD, NS-WD).

Magnetic braking At somewhat longer periods $P \lesssim 1\,\mathrm{d}$, the companion can be a main sequence star that has a tidally locked rotation period. Main-sequence stars have winds, and these winds carry angular momentum. Because of the tidal locking, this also introduces a negative torque on the system.

Evolution of the secondary For wider binary orbits, the secondary star can make contact with the Roche lobe as it becomes a giant star.

Finally, if the secondary is sufficiently evolved, it may have a strong enough wind that accretion can still occur, even if the orbit is so wide that the secondary doesn't fill its Roche lobe.

EXERCISE 15.3 — Suppose we have a $1.6\,M_\odot$ neutron star and a $0.8\,M_\odot$ companion. The companion is evolved and has an effective temperature $T_{\mathrm{eff}} = 3000\,\mathrm{K}$. We'll assume that T_{eff} is fixed. On the giant branch, the luminosity is powered by hydrogen shell burning, the ashes of which are added to the core (i.e., the core mass M_c increases due to hydrogen burning). The energy released, per gram of hydrogen consumed, is $Q = 6 \times 10^{18}\,\mathrm{erg\,g^{-1}}$. For such an evolved giant the luminosity depends mainly on the core mass M_c and may be approximated as

$$\frac{L}{L_\odot} = 10^{6.3}\left(\frac{M_c}{M_\odot}\right)^8.$$

For this system, find the orbital period P and the mass transfer rate $\dot{M}$ (in units of solar masses per year) if the secondary has a core mass $M_c = 0.2\,M_\odot$.

Matter that crosses the inner Lagrange point will find itself in orbit about the primary. The material still has enough angular momentum that it won't fall directly onto the primary, and so an *accretion disk* will form. In order for the matter to accrete, there must be enough friction in the disk so that the gravitational energy can be radiated away and angular momentum transported outward.

EXERCISE 15.4 — Suppose our binary consists of two white dwarfs with a short orbital period, less than 2 hours, so that gravitational radiation produces a torque on the system according to eq. (15.10). Because the secondary is also degenerate, its radius depends on mass as

$$R_2 = K\left(\frac{M_2}{M_\odot}\right)^{-1/3},$$

where $K \approx 2 \times 10^9$ cm.

1. Show that if the secondary just fills its Roche lobe, then $M_2 \propto P^{-1}$ and find M_2 if $P = 1$ hr.
2. Show that
$$\frac{\dot{M}}{M_2} = -\frac{\dot{J}/J}{2/3 - M_2/M_1}$$
3. Using the above relations and eq. (15.10), find $\dot{M}$ if $M_1 = 1\,M_\odot$. Scale the orbital period to units of 1 hour.

The Eddington Limit

There is a characteristic luminosity at which the pressure from radiation balances the gravitational force. This tends to act as a limit to accretion. To derive this limit, known as the *Eddington luminosity*, consider a spherically symmetric shell of matter. Radiation enters the shell and scatters isotropically (Thomson scatters) from electrons. The momentum flux (momentum per unit time per unit area) entering the shell is just

$$\boldsymbol{P} = \frac{1}{c}\frac{L}{4\pi r^2}\boldsymbol{e}_r.$$

Here L is the luminosity and r is the radius of the shell. Since the scattering is presumed isotropic, the rate at which the fluid element's momentum changes (the impulse imparted to it by the radiation) is just

$$\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}t} = \boldsymbol{P}\sigma_{\mathrm{Th}} n_e,$$

where σ_{Th} is the cross-section for Thomson scattering. Since $\mathrm{d}\boldsymbol{p}/\mathrm{d}t$ is just a force per unit volume, we can balance it with the gravitational force per unit volume,

$$\boldsymbol{f}_g = -\frac{GM\langle A\rangle m_{\mathrm{u}} n_{\mathrm{ion}}}{r^2}\boldsymbol{e}_r,$$

and solve for L to obtain

$$L_{\mathrm{Edd}} = \frac{4\pi GMc}{(\sigma_{\mathrm{Th}}/m_{\mathrm{u}})Y_e}. \tag{15.11}$$

Here I have used $n_e = \langle Z\rangle n_{\mathrm{ion}}$ and $Y_e = \langle Z\rangle/\langle A\rangle$. Note that L_{Edd} is independent of distance from the star. Numerically,

$$L_{\mathrm{Edd}} = 1.3 \times 10^{38}\,\mathrm{erg\,s^{-1}}\left(\frac{M}{M_\odot}\right)Y_e^{-1} = 3.2 \times 10^4 L_\odot\left(\frac{M}{M_\odot}\right)Y_e^{-1}.$$

We can now look at what the luminosity supplied by accretion is. First there is the just the gravitational energy release. The gravitational release, per nucleon, is

$$\frac{GMm_{\mathrm{u}}}{R} = \begin{cases} 0.069\,\mathrm{MeV} & \text{WD with } R = 2\times 10^{9}\,\mathrm{cm} \\ 138\,\mathrm{MeV} & \text{NS with } R = 10^{6}\,\mathrm{cm} \end{cases}$$

In both cases we used $M = 1\,M_{\odot}$. For hydrogen-rich accretion onto a white dwarf, steady nuclear burning (assuming it exists!) will dwarf the luminosity from accretion. The situation is reversed for a neutron star, for which the energy from nuclear burning is a perturbation to the large release of gravitational binding energy.

We can compute the accretion rate at which the luminosity supplied by either nuclear burning or release of gravitational binding energy would equal the Eddington value:

$$\dot{M}_{\mathrm{Edd}} = \begin{cases} \frac{4\pi GMc}{Q(\sigma_{\mathrm{Th}}/m_{\mathrm{u}})Y_e} & \text{nuclear, } L = \dot{M}Q \\ \frac{4\pi Rc}{(\sigma_{\mathrm{Th}}/m_{\mathrm{u}})Y_e} & \text{gravitational, } L = \frac{GM\dot{M}}{R} \end{cases}.$$

For nuclear burning, taking $M = 1\,M_{\odot}$, $Q = 6\times 10^{18}\,\mathrm{erg\,g^{-1}}$ (hydrogen burning), and $Y_e = 1$ gives $\dot{M}_{\mathrm{Edd}} = 3.3\times 10^{-7}\,M_{\odot}\,\mathrm{yr^{-1}}$. For gravitational power, taking $R = 10^{6}\,\mathrm{cm}$ and $Y_e = 1$ gives $\dot{M}_{\mathrm{Edd}} = 1.5\times 10^{-8}\,M_{\odot}\,\mathrm{yr^{-1}}$. Observed systems do in fact have inferred mass transfer rates that are typically less than these values.

EXERCISE 15.5 — Consider a star surrounded by an accretion disk of matter. In the disk, which consists of ionized gas, a fluid element gradually spirals inward, so that at each point in the disk, it is approximately in a circular orbit. Show that under these conditions, the energy radiated, per unit mass, by the fluid element over its lifetime in the disk is $GM/(2R)$.

Structure of an Accreting Neutron Star Envelope

This section is similar in spirit to lectures[1] given at a NATO Advanced Studies Institute by Prof. Bildsten, KITP.

As a worked example, we shall consider the thermonuclear stability of hydrogen/helium accretion onto a neutron star. For accretion from a disk, the luminosity (cf. problem 5) is $GM\dot{M}/(2R)$. The remainder of the gravitational binding energy is liberated at a boundary layer where the disk becomes incorporated into the star. Setting this equal to thermal emission gives an effective temperature

$$T_{\mathrm{eff}} = \left(\frac{GM\dot{M}}{2R\sigma_{\mathrm{SB}}\mathcal{A}}\right)^{1/4} = 10^{7}\,\mathrm{K}\left(\frac{M}{1.4\,M_{\odot}}\,\frac{10^{6}\,\mathrm{cm}}{R}\,\frac{10^{13}\,\mathrm{cm}^{2}}{\mathcal{A}}\,\frac{\dot{M}}{10^{-9}\,M_{\odot}\,\mathrm{yr}^{-1}}\right)^{1/4}$$

[1] L. Bildsten. Thermonuclear burning on rapidly accreting neutron stars. In A. Alpar, R. Buccheri, and J. van Paradijs, editors, *The Many Faces of Neutron Stars*, volume 515, page 419, Dordrecht, 1998. NATO ASI ser. C, Kluwer

Here we have scaled the emitting area to $\mathcal{A}$ rather than presuming it is emitted from the entire surface. This effective temperature is, in energy units, $k_B T_{\text{eff}} \approx$ keV, which is in the X-ray region of the spectrum.

The gravitational acceleration at the neutron star surface is

$$g = \frac{GM}{R^2}\left(1 - \frac{2GM}{Rc^2}\right)^{-1/2} \approx 2.4 \times 10^{14}\,\text{cm}\,\text{s}^{-2}$$

for a neutron star with $M = 1.4\,M_\odot$, $R = 10^6$ cm. The factor in parenthesis is a correction due to general relativity. The redshift, i.e., the fractional increase in wavelength of a photon emitted from the surface to a distant observer, is $\Delta\lambda/\lambda = [1 - 2GM/(Rc^2]^{-1/2} - 1 \approx 0.3$. Given the surface gravity and the temperature, we can estimate the pressure scale height,

$$H = \frac{P}{\rho g} = \frac{k_B T}{\mu m_u g} \approx 5\,\text{cm}.$$

Because $H/R \ll 1$, we can make two simplifying assumptions: first, we can treat our atmosphere in planar geometry, and second we can boost to a local frame and perform our calculations in Euclidian space-time.

In a planar geometry, there is a simple relation between the time since a fluid element was deposited and its local pressure. If we assume that accreted matter spreads uniformly over the surface, than every square centimeter sees mass being added at a rate $\dot{m} = \dot{M}/(4\pi R^2)$. Thus after a time t a fluid element will find itself under a column with a mass per unit area $\dot{m} \times t = \int_z^\infty \rho\,\mathrm{d}z$. Using hydrostatic equilibrium, we can transform $\rho\,\mathrm{d}z = -\mathrm{d}P/g$, so that

$$P(t) = \dot{m}tg.$$

Here t is the time since the fluid element was accreted. At an accretion rate of $\dot{M} = 10^{-9}\,M_\odot\,\text{yr}^{-1}$ onto a 10 km neutron star, the pressure reached when the hydrogen is half-consumed is $P = 1.8 \times 10^{22}\,\text{dyn}\,\text{cm}^{-2}$. The quantity $y \equiv \int_z^\infty \rho\,\mathrm{d}z$ is known as the *column density*.

Hydrogen burning Let's consider the accretion of a (roughly solar) mixture of ^{1}H, ^{4}He, and ^{12}C, with mass fractions $X_\text{H} : X_\text{He} : X_\text{CNO} = 0.70 : 0.28 : 0.02$. At temperatures $> 10^7$ K, the ^{1}H is consumed via the hot CNO cycle,

$$^{12}\text{C}(p,\gamma)^{13}\text{N}(p,\gamma)^{14}\text{O}(\beta^+\nu)^{14}\text{N}(p,\gamma)^{15}\text{O}(\beta^+\nu)^{15}\text{N}(p,\alpha)^{12}\text{C}.$$

The two slowest reactions are the decays of ^{14}O ($\tau_{1/2} = 71$ s) and ^{15}O ($\tau_{1/2} = 122$ s). As a result, the cycle is independent of temperature and thermally stable. How long would it take to consume all of the hydrogen? For each CNO nucleus, there are

$$n_\text{H}/n_\text{CNO} = \frac{0.7}{0.02/12} = 420$$

hydrogen nuclei, and each trip around the cycle consumes 4 hydrogen nuclei. As a result, it take 105 trips around the CNO cycle to deplete the hydrogen. The time to complete one cycle is $(71\,\mathrm{s} + 122\,\mathrm{s})/\ln 2 = 278\,\mathrm{s}$, so the time to deplete all of the hydrogen is 2.92×10^4 s—about 1/3 of a day.

To get the temperature structure of the accreted envelope, we use the flux equation, eq. (8.36),

$$F = -\frac{1}{3}\frac{c}{\rho\kappa}\frac{\mathrm{d}aT^4}{\mathrm{d}r}. \tag{15.12}$$

We can again transform $\rho\,\mathrm{d}r = -\mathrm{d}P/g$; moreover, if all of the heat released by fusing hydrogen to helium comes out (and assuming there is no other source of heat in the neutron star interior), the flux will just be $QX_\mathrm{H}\dot{m}$ and will be constant. Integrating eq. (15.12) and neglecting T_eff and P(photosphere) gives us a relation for the temperature,

$$T^4 = \frac{3\kappa F y}{ac}.$$

For the parameters we are considering, $F = QX_\mathrm{H}\dot{m} = 2.1\times10^{22}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$, and the temperature at the depth where the H is half-consumed is $T = 2.9 \times 10^8$ K.

The cooling timescale for this atmosphere is

$$t_\mathrm{th} = \frac{yC_PT}{F} \approx 208\,\mathrm{s} \times \frac{y_8T_8}{F_{22}}$$

Here we are using the shorthand notation, $y_8 \equiv y/(10^8\,\mathrm{g\,cm^{-2}})$, and similarly for T_8, F_{22}, and so on. The thermal timescale can be compared to the accretion timescale,

$$t_\mathrm{ac} = \frac{y}{\dot{m}} = 10^4\,\mathrm{s} \times y_8\dot{m}_4^{-1}.$$

Because $t_\mathrm{th} \ll t_\mathrm{ac}$, the compression is slow and far from adiabatic.

As the fluid element is compressed and heated, the lifetime of a ^{4}He nucleus against the triple-alpha reaction becomes shorter than the accretion timescale and helium ignites. If protons are present, then there will be additional heating from $^{12}\mathrm{C}(p,\gamma)$. The triple-alpha reaction is strongly temperature sensitive and therefore is susceptible to an thermal instability.

Helium ignition To evaluate the thermal instability, we approximate the cooling rate as

$$\varepsilon_\mathrm{cool} \approx \frac{C_PT}{t_\mathrm{th}} = \frac{acT^4}{3\kappa y^2} \tag{15.13}$$

and compare it against the heating rate from the triple-alpha reaction, $\varepsilon_{3\alpha}$. We then apply a constant-pressure thermal perturbation about the steady-state, $\varepsilon_{\text{cool}} = \varepsilon_{3\alpha}$; upon expanding this equation we arrive at a condition for instability:

$$\left.\frac{\partial \ln \varepsilon_{3\alpha}}{\partial \ln T}\right|_y > \left.\frac{\partial \ln \varepsilon_{\text{cool}}}{\partial \ln T}\right|_y . \tag{15.14}$$

The triple-alpha heating rate is

$$\varepsilon_{3\alpha} \approx (5.3 \times 10^{21}\,\text{erg}\,\text{s}^{-1}\,\text{g}^{-1}) \times \frac{\rho_5^2 X_{\text{He}}^3}{T_8^3} \exp\left(-\frac{44}{T_8}\right).$$

If we ignore the variation in density, $\rho = 10^5\,\text{g}\,\text{cm}^{-3} \times \rho_5$, with temperature, then equation (15.14 becomes

$$\frac{44}{T_8} - 3 > 4; \tag{15.15}$$

put another way, a thermal instability is expected to occur if $T < 6.3 \times 10^8$ K at the depth where helium is consumed. This is satisfied for all $\dot{m} < \dot{m}_{\text{Edd}}$.

The remaining question is the depth at which the triple-alpha reaction ignites. For a rapidly accreting neutron star, the flux is greater, and therefore the envelope has a steeper temperature gradient[2]; roughly, for $\dot{m} \gtrsim 0.1\dot{m}_{\text{Edd}}$, the helium ignition occurs in a bath of protons, so that an rp-process ensues. The consumption of hydrogen releases more energy and requires waiting for $\beta+$-decays, this tends to produce a longer-lasting X-ray burst that is more energetic, for the amount of fuel accreted. At lower accretion rates, the hot CNO cycle consumes all of the available hydrogen before the helium ignites; as a result, the burst is more rapid, and less energetic for the amount of matter accreted. Note that the pure helium burst may still release a greater total amount of energy if it ignites at a greater depth, i.e., if the amount of matter accumulated between bursts is greater.

[2] L. Bildsten. Thermonuclear burning on rapidly accreting neutron stars. In A. Alpar, R. Buccheri, and J. van Paradijs, editors, *The Many Faces of Neutron Stars*, volume 515, page 419, Dordrecht, 1998. NATO ASI ser. C, Kluwer

16
Stellar Pulsations

In this chapter, we'll linearize the perturbed continuity and momentum equations and solve for the frequencies of the normal modes for a star.

16.1 Adiabatic, radial pulsations

Imagine that we perturb our fluid in some way. As described in section 3.2, we can describe the *Eulerian* perturbation in some fluid property *f*:

$$\Delta f \equiv f(\boldsymbol{r}, t) - f_0(\boldsymbol{r}, t), \tag{16.1}$$

where the subscript "0" denotes the unperturbed quantity. Said another way, Δf describes the change, under our perturbation, in some property of the fluid at a fixed location.

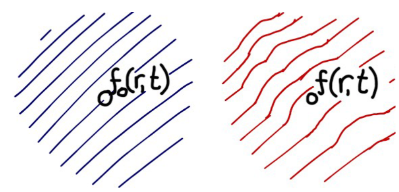


Figure 16.1: An Eulerian perturbation: we compare quantities at corresponding locations.

We may also describe our perturbation as a *Lagrangian* one, where we compare the same fluid element in both the perturbed and unperturbed systems:

$$\delta f \equiv f(\boldsymbol{r}, t) - f_0(\boldsymbol{r}_0, t). \tag{16.2}$$

Under a Lagrangian perturbation the fluid element in the perturbed system in general has a different position $\boldsymbol{r}$ than in the unperturbed system, $\boldsymbol{r}_0$.

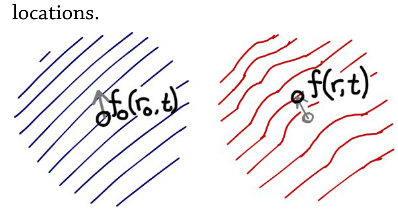


Figure 16.2: A Lagrangian perturbation: we compare quantities for corresponding fluid elements.

The two perturbations are related to one another via

$$\delta f = \Delta f + (\delta \boldsymbol{r} \cdot \boldsymbol{\nabla}) f_0. \tag{16.3}$$

There are a few useful commutation relations that are easily proved:

$$\begin{aligned} \partial_t \Delta f &= \Delta\left(\partial_t f\right), & (16.4)\\ \boldsymbol{\nabla} \Delta f &= \Delta \boldsymbol{\nabla} f, & (16.5)\\ \frac{\mathrm{D}}{\mathrm{D}t}\delta f &= \delta \frac{\mathrm{D}f}{\mathrm{D}t}. & (16.6) \end{aligned}$$

And there are operations that do not commute:

$$\begin{aligned} \partial_t \delta f &\neq \delta\left(\partial_t f\right), & (16.7)\\ \boldsymbol{\nabla} \delta f &\neq \delta \boldsymbol{\nabla} f, & (16.8)\\ \frac{\mathrm{D}}{\mathrm{D}t}\Delta f &\neq \Delta \frac{\mathrm{D}f}{\mathrm{D}t}. & (16.9) \end{aligned}$$

One can further show that $\delta \boldsymbol{u} = (\mathrm{D}/\mathrm{D}t)\delta \boldsymbol{r}$. Also, if the fluid has unperturbed velocity $\boldsymbol{u} = 0$, then $\Delta \boldsymbol{u} = \delta \boldsymbol{u}$. Finally, for purely radial motion, we can introduce the Lagrangian mass coordinate m, in which case $\partial_m \delta f = \delta(\partial_m f)$ and $\partial_m \Delta f \neq \Delta(\partial_m f)$.

WE ARE NOW READY TO USE THESE COMMUTATION RELATIONS TO DERIVE a *linear adiabatic wave equation*. By linear, we mean that we shall only keep terms to first order in δ. By adiabatic, we mean that we shall only consider the equations of continuity and momentum, and we shall relate the density and pressure perturbations via

$$\frac{\delta P}{P} = \Gamma_1 \frac{\delta \rho}{\rho}. \tag{16.10}$$

Here $\Gamma_1 \equiv (\partial \ln P/\partial \ln \rho)_s$.

For simplicity, we'll start with purely radial oscillations. First, let's perturb the equation of continuity, expressed in Lagrangian form (eq. [2.26]),

$$\frac{\partial \ln r}{\partial m} = \frac{1}{4\pi r^3 \rho}.$$

We apply a Lagrangian perturbation to both sides of this equations and expand the right-hand side to first order in δr and $\delta \rho$. Since δ and ∂/∂_m commute, we can interchange them:

$$\begin{aligned}
\frac{\partial}{\partial m}\left(\frac{\delta r}{r}\right) &= \delta\left(\frac{\partial \ln r}{\partial m}\right) \\
&= \delta\left(4\pi r^3 \rho\right)^{-1} \\
&= \left(4\pi r^3 \rho\right)^{-1}\left(-3\frac{\delta r}{r} - \frac{\delta \rho}{\rho}\right).
\end{aligned}$$

Moving $(4\pi r^3 \rho)$ to the left-hand side of the equation, and recognizing that

$$4\pi r^3 \rho \frac{\partial}{\partial m} = r\frac{\partial m}{\partial r}\frac{\partial}{\partial m} = r\frac{\partial}{\partial r},$$

we have our first equation,

$$r\frac{\partial}{\partial r}\left(\frac{\delta r}{r}\right) = -3\frac{\delta r}{r} - \frac{\delta \rho}{\rho}. \tag{16.11}$$

Next, we can perturb the force equation (eq. [2.27])

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

If the unperturbed state is taken to have $\mathrm{D}r_0/\mathrm{D}t = \mathrm{D}^2 r_0/\mathrm{D}t^2 = 0$, then a similar linearization yields

$$\rho r \frac{\mathrm{D}^2}{\mathrm{D}t^2}\left(\frac{\delta r}{r}\right) = -\frac{\partial P}{\partial r}\left(4\frac{\delta r}{r} + \frac{\delta P}{P}\right) - P\frac{\partial}{\partial r}\left(\frac{\delta P}{P}\right), \tag{16.12}$$

which is our second equation.

To proceed further, we write

$$\frac{\delta r}{r} = \zeta(r)\exp(i\sigma t),$$

so that the left-hand side of equation (16.12) becomes $-\rho r\sigma^2\zeta(r)e^{i\sigma t}$, and we can make the substitution $\partial_r(\delta r/r) \to e^{i\sigma t}(\mathrm{d}\zeta/\mathrm{d}r)$. We additionally eliminate $\delta\rho/\rho$ from equation (16.11) using the adiabatic condition, eq. (16.10) and make use of the zeroth-order momentum equation $\mathrm{d}P/\mathrm{d}r = -\rho Gm/r^2$ to obtain

$$\frac{\mathrm{d}}{\mathrm{d}r}\zeta = -\frac{1}{r}\left(3\zeta + \frac{1}{\Gamma_1}\frac{\delta P}{P}\right) \qquad (16.13)$$

$$\frac{\mathrm{d}}{\mathrm{d}r}\left(\frac{\delta P}{P}\right) = \frac{1}{\lambda_P}\left[\left(4+\sigma^2\frac{r^3}{Gm}\right)\zeta + \frac{\delta P}{P}\right]. \qquad (16.14)$$

Here we introduce the pressure scale height (in the unperturbed system) $\lambda_P \equiv -(\mathrm{d}\ln P/\mathrm{d}r)^{-1} = Pr^2/(\rho Gm)$. Multiply equation (16.13) by $\Gamma_1 Pr^4$ and then differentiate with respect to r, using equation (16.14) to eliminate the spatial derivative of $\delta P/P$ and equation (16.13) to eliminate $\delta P/P$ to obtain

$$\frac{\mathrm{d}}{\mathrm{d}r}\left[\Gamma_1 Pr^4\frac{\mathrm{d}}{\mathrm{d}r}\zeta\right] + \left\{r^3\frac{\mathrm{d}}{\mathrm{d}r}\left[(3\Gamma_1-4)P\right]\right\}\zeta + \sigma^2(r^4\rho)\zeta = 0. \qquad (16.15)$$

Notice here that we have *not* assumed that Γ_1 is a constant.

Equation (16.15) has the form

$$\mathcal{L}\zeta(r) + \sigma^2 w(r)\zeta(r)$$

where

$$\mathcal{L} \equiv \frac{\mathrm{d}}{\mathrm{d}r}\left[u(r)\frac{\mathrm{d}\zeta}{\mathrm{d}r}\right] + q(r)\zeta(r),$$

with $u(r) = \Gamma_1 Pr^4$,

$$q(r) = r^3\frac{\mathrm{d}}{\mathrm{d}r}\left[(3\Gamma_1-4)P\right],$$

and $w(r) = r^4\rho$. For the imposed boundary conditions, there will in general be solutions for only certain eigenvalues σ^2. Note that $u(r) > 0$ on the interval $0 < r < R$. Furthermore, we require that ζ and $\mathrm{d}\zeta/\mathrm{d}r$ be finite at $r = 0$ and $r = R$, which means that if ζ_i and ζ_j are solutions of eq. (16.15), then

$$u(r)\zeta_i^*\frac{\mathrm{d}\zeta_j}{\mathrm{d}r}\bigg|_{r=0} = u(r)\zeta_i^*\frac{\mathrm{d}\zeta_j}{\mathrm{d}r}\bigg|_{r=R} = 0. \qquad (16.16)$$

Using these boundary conditions and the form of the operator $\mathcal{L}$, we find that

$$\int_0^R \mathrm{d}r\;\zeta_i^*\mathcal{L}\zeta_j = u\zeta_i^*\frac{\mathrm{d}\zeta_j}{\mathrm{d}r}\bigg|_{r=0}^{r=R} - \int_0^R \mathrm{d}r\;\frac{\mathrm{d}\zeta_i^*}{\mathrm{d}r}u\frac{\mathrm{d}\zeta_j}{\mathrm{d}r} + \zeta_j q(r)\zeta_i^*$$

$$
\begin{aligned}
&= \quad -u(r)\zeta_j \frac{\mathrm{d}\zeta_i^*}{\mathrm{d}r}\bigg|_{r=0}^{r=R} + \int_0^R \mathrm{d}r\, \zeta_j \frac{\mathrm{d}}{\mathrm{d}r}\left[u \frac{\mathrm{d}}{\mathrm{d}r}\zeta_i^*\right] + \zeta_j q(r)\zeta_i^* \\
&= \quad \int_0^R \mathrm{d}r\, \zeta_j \mathcal{L}\zeta_i^*.
\end{aligned}
$$

The operator $\mathcal{L}$ is thus Hermitian. As a result, the eigenvalues σ^2 are real and denumerable. There is a minimum eigenvalue σ_0^2. The eigenfunctions corresponding to these eigenvalues are orthogonal in the following sense: if σ_i^2 and σ_j^2 are eigenvalues of equation (16.15) and ζ_1, ζ_2 their corresponding eigenfuctions, then

$$
\int_0^R \mathrm{d}r\, w(r)\zeta_i\zeta_j = \int_0^R \mathrm{d}r\, r^4\rho\zeta_i\zeta_j = 0 \quad \text{if } \sigma_i^2 \neq \sigma_j^2. \tag{16.17}
$$

Solutions with larger eigenvalues have more nodes.

16.2 Adiabatic non-radial pulsations

Now that we've warmed up with the purely radial pulsations, we'll do the more general case. Rather than going through the separation of variables in spherical coordinates, we'll keep things simple and cartesian. This amounts to looking at a small box in the star. We will also make an *ansatz* that the fluid perturbations don't change the gravitational potential[1]. In this case, the (constant) gravitational acceleration $\boldsymbol{g} = -g\boldsymbol{e}_r$ defines the local vertical, so we will separate our equations into a radial direction, labeled by "r", and a transverse direction, labeled by "t".

[1] This is known as the **Cowling approximation**.

Let us first perturb the equation of continuity,

$$
\frac{\partial\rho}{\partial t} + \nabla\cdot(\rho\boldsymbol{u}) = \frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\nabla\cdot\boldsymbol{u} = 0. \tag{16.18}
$$

We take our unpertubed state to be independent of time with $\boldsymbol{u} = \mathbf{0}$. If we take the Lagrangian perturbation of eq. (16.18), we have

$$
\frac{\mathrm{D}}{\mathrm{D}t}\left(\delta\rho + \rho\nabla\cdot\boldsymbol{\xi}\right) = 0,
$$

with $\boldsymbol{u} = \mathrm{D}\boldsymbol{\xi}/\mathrm{D}t$. Setting the constant of integration to 0 reduces this to

$$
\frac{\delta\rho}{\rho} = -\nabla\cdot\boldsymbol{\xi}.
$$

The linearization of the momentum equation,

$$
\frac{\mathrm{D}^2\boldsymbol{\xi}}{\mathrm{D}t^2} = -\frac{1}{\rho}\boldsymbol{\nabla}\Delta P + \frac{\Delta\rho}{\rho}\boldsymbol{g}, \tag{16.19}
$$

looks very similar to what we did in deriving a condition for convection, § 3.2. This time, however, we won't impose the condition that $\Delta P = 0$. When we were looking at convective instabilities, we were interested in low-frequency perturbations, in which the pressure has time to equilibrate. Keeping the terms with ΔP, the perturbed momentum equation becomes

$$\frac{\mathrm{D}^2\boldsymbol{\xi}}{\mathrm{D}t^2} = -\frac{1}{\rho}\boldsymbol{\nabla}\Delta P + \underbrace{\boldsymbol{g}\frac{1}{\Gamma_1}\frac{\Delta P}{P} + \boldsymbol{g}(\boldsymbol{\xi}\cdot\boldsymbol{\nabla})\overbrace{\left[\frac{1}{\Gamma_1}\ln P - \ln\rho\right]}^{-\mathcal{A}}}_{\text{buoyancy from density perturbation}}. \qquad (16.20)$$

If $\Delta P \to 0$, this reduces to equation (3.20) with $\mathcal{A}$ being the **Schwarzschild discriminant**.

To decompose our perturbed system into normal modes, with the perturbed quantities varying in time as $\exp(i\omega t)$, we first note that in the unperturbed system $\ln P$ and $\ln\rho$ depend only on r; and nothing depends on the transverse directions. As a result, we can impose periodic boundary conditions in the transverse direction and write

$$\begin{aligned}\Delta P(\boldsymbol{x},t) &= \Delta P(r)\exp(i\boldsymbol{k}_t\cdot\boldsymbol{x}_t)\exp(i\omega t),\\ \Delta\boldsymbol{\xi}(\boldsymbol{x},t) &= \boldsymbol{\xi}(r)\exp(i\boldsymbol{k}_t\cdot\boldsymbol{x}_t)\exp(i\omega t).\end{aligned}$$

Here $\boldsymbol{k}_t$ is a transverse wavenumber and $\boldsymbol{x}_t$ are the transverse coordinates. The transverse component of equation (16.20) is thus

$$\rho\omega^2\boldsymbol{\xi}_t = i\boldsymbol{k}_t\Delta P, \qquad (16.21)$$

and the radial equation of motion is

$$\rho\omega^2\xi_r = \partial_r\Delta P + \frac{g\rho}{\Gamma_1 P}\Delta P + \rho N^2\xi_r. \qquad (16.22)$$

We've re-introduced the Brunt-Väisälä frequency[2]

[2] We take Γ_1 to be constant.

$$N^2 = -g\frac{\mathrm{d}}{\mathrm{d}r}\mathcal{A} = g\left(\frac{1}{\Gamma_1}\frac{\mathrm{d}}{\mathrm{d}r}\ln P - \frac{\mathrm{d}}{\mathrm{d}r}\ln\rho\right),$$

just as was done in analyzing convection.

We still have more unknowns than equations. Next, we'll use the perturbed equation of continuity,

$$\frac{\delta\rho}{\rho} = -\nabla\cdot\boldsymbol{\xi} = -\partial_r\xi_r - \boldsymbol{k}_t\cdot\boldsymbol{\xi}_t. \qquad (16.23)$$

For adiabatic perturbations,

$$\begin{aligned}\frac{\delta\rho}{\rho} &= \frac{1}{\Gamma_1}\frac{\delta P}{P}\\ &= \frac{1}{\Gamma_1}\left(\frac{\Delta P}{P} + \boldsymbol{\xi}\cdot\boldsymbol{\nabla}\ln P\right)\\ &= \frac{1}{\Gamma_1}\frac{\Delta P}{P} - \xi_r\frac{1}{\Gamma_1\lambda_P}.\end{aligned}$$

Here we used $\boldsymbol{\xi} \cdot \boldsymbol{\nabla} \ln P = \xi_r (\mathrm{d}/\mathrm{d}r) \ln P = -\xi_r \rho g / P = -\xi_r / \lambda_P$, where λ_P is the pressure scale height. It makes physical sense that this scale will enter: modes with radial wavelengths much less than H shouldn't be affected by the background stratification.

Inserting the expression for $\delta\rho/\rho$ in equation (16.23) and using equation (16.21) to eliminate $\boldsymbol{\xi}_t$, we obtain

$$\partial_r \xi_r = \frac{1}{\omega^2 \Gamma_1 P} \left(\frac{k_t^2 \Gamma_1 P}{\rho} - \omega^2 \right) \Delta P + \frac{\xi_r}{\Gamma_1 \lambda_P}. \tag{16.24}$$

The quantity $S^2 = k_t^2 \Gamma_1 P/\rho = k_t^2 c_s^2$ is called the **Lamb frequency**.

Collecting equations (16.24) and (16.22), we now have two coupled first order differential equations for $\xi_r(r)$ and $\Delta P(r)$,

$$\partial_r \xi_r = \frac{1}{\Gamma_1 P} \frac{S^2 - \omega^2}{\omega^2} \Delta P + \frac{\xi_r}{\Gamma_1 \lambda_P} \tag{16.25}$$

$$\partial_r \Delta P = \rho(\omega^2 - N^2) \xi_r - \frac{\Delta P}{\Gamma_1 \lambda_P}. \tag{16.26}$$

From the form of these equations, let's try a solution in the form

$$\Delta P(r) = \Delta P e^{Kr}; \qquad \xi_r(r) = \xi_r e^{Kr}. \tag{16.27}$$

If K is imaginary, then we have oscillatory solutions. If K is real, then we must choose K so that the radial functions decay exponentially, or **evanesce**. Substituting equation (16.27) into eq. (16.26) gives

$$\Delta P = \left[\frac{\rho(\omega^2 - N^2)}{1 + 1/K\lambda_P} \right] \frac{\xi_r}{K}.$$

Substituting this expression for ΔP into eq. (16.25) and eliminating ξ_r gives us the dispersion equation,

$$K^2 = \frac{k_t^2}{S^2} \frac{(S^2 - \omega^2)(\omega^2 - N^2)}{(1 + 1/K\lambda_P)\omega^2} + \frac{1}{K\lambda_P}, \tag{16.28}$$

which relates the mode frequency ω to the transverse wavenumber k_t and the radial wavenumber K.

In the limit where the radial wavelength is much less than a pressure scale height, $|K|\lambda_P \gg 1$, equation (16.28) simplifies to

$$K^2 = \frac{k_t^2 (S^2 - \omega^2)(\omega^2 - N^2)}{S^2 \omega^2}. \tag{16.29}$$

If $N^2 < \omega^2 < S^2$, $K^2 > 0$ and the radial perturbations evanesce. To have a mode, we require either that $\omega^2 < N^2, S^2$ or $\omega^2 > N^2, S^2$.

The analysis is similar in spherical coordinates, except that instead of plane waves, $e^{\boldsymbol{k}_t \cdot \boldsymbol{x}_t}$, we'll have spherical harmonics $Y_{\ell m}$. In the above dispersion relation, make the substitution $k_t^2 \to \ell(\ell+1)/r^2$ and $S^2 \to S_\ell^2$.

EXERCISE 16.1 — Use dimensional analysis and virial estimates to sketch S (for $\ell = 1$) and N on a plot of frequency against radius for the sun. Indicate where modes with $\omega^2 < N^2, S^2$ can propagate, and likewise for $\omega^2 > N^2, S^2$.

A
Technical Notes

A.1 Thermodynamical derivatives

A common task in stellar physics is transforming between different derivatives with respect to different thermodynamical quantities. For example, you may have expressions for $(\partial\kappa/\partial T)_\rho$ and $(\partial\kappa/\partial\rho)_T$, but you need $(\partial\kappa/\partial T)_P$ and $(\partial\kappa/\partial P)_T$. There is a straightforward way to handle transforming from (ρ, T) space to (P, T) space, and that is using Jacobians. Despite the utility of this technique, it is not commonly discussed in astrophysical texts[1].

[1] L. D. Landau and E. M. Lifshitz. *Statistical Physics, part 1*. Pergamon Press, Oxford, 3 edition, 1980

The *Jacobian* is defined as the determinant of a matrix of partial derivatives,

$$\begin{aligned} \frac{\partial(a,b)}{\partial(c,d)} &\equiv \det\begin{bmatrix} \left(\frac{\partial a}{\partial c}\right)_d & \left(\frac{\partial a}{\partial d}\right)_c \\ \left(\frac{\partial b}{\partial c}\right)_d & \left(\frac{\partial b}{\partial d}\right)_c \end{bmatrix} \\ &= \left(\frac{\partial a}{\partial c}\right)_d \left(\frac{\partial b}{\partial d}\right)_c - \left(\frac{\partial a}{\partial d}\right)_c \left(\frac{\partial b}{\partial c}\right)_d . \end{aligned} \tag{A.1}$$

Because interchanging any the rows (or the columns) causes the determinant to change sign,

$$\frac{\partial(b,a)}{\partial(c,d)} = -\frac{\partial(a,b)}{\partial(c,d)} \tag{A.2}$$

and

$$\frac{\partial(a,b)}{\partial(d,c)} = -\frac{\partial(a,b)}{\partial(c,d)}. \tag{A.3}$$

Further,

$$\frac{\partial(a,s)}{\partial(c,s)} = \left(\frac{\partial a}{\partial c}\right)_s \left(\frac{\partial s}{\partial s}\right)_a - \left(\frac{\partial a}{\partial s}\right)_c \left(\frac{\partial s}{\partial c}\right)_s = \left(\frac{\partial a}{\partial c}\right)_s, \tag{A.4}$$

and

$$\frac{\partial(a,b)}{\partial(a,b)} = \left(\frac{\partial a}{\partial a}\right)_b \left(\frac{\partial b}{\partial b}\right)_a - \left(\frac{\partial a}{\partial b}\right)_a \left(\frac{\partial b}{\partial a}\right)_b = 1. \tag{A.5}$$

Hence we can write thermodynamical derivative in terms of Jacobians, for example,

$$\left(\frac{\partial T}{\partial P}\right)_S = \frac{\partial(T,S)}{\partial(P,S)}. \tag{A.6}$$

Finally, when multiplying two Jacobians, one can "cancel" identical upper and lower parts,

$$\frac{\partial(a,b)}{\partial(c,d)}\frac{\partial(c,d)}{\partial(s,t)} = \frac{\partial(a,b)}{\partial(s,t)}, \tag{A.7}$$

as can be readily checked by expanding out both the left and right hand sides.

Common thermodynamic derivatives

Certain thermodynamical derivatives occur commonly in working with fluids. The first set express how the pressure relates to the pair ρ, T:

$$\chi_T \equiv \frac{T}{P}\left(\frac{\partial P}{\partial T}\right)_\rho, \qquad \chi_\rho \equiv \frac{\rho}{P}\left(\frac{\partial P}{\partial \rho}\right)_T. \tag{A.8}$$

For a fixed composition the equation of state can always be expanded about a point $P_0(\rho_0, T_0)$ as

$$P = P_0\left(\frac{\rho}{\rho_0}\right)^{\chi_\rho}\left(\frac{T}{T_0}\right)^{\chi_T}, \tag{A.9}$$

or equivalently

$$\frac{\mathrm{d}P}{P} = \chi_\rho\frac{\mathrm{d}\rho}{\rho} + \chi_T\frac{\mathrm{d}T}{T}.$$

Here we are implicitly keeping the composition fixed.

The next set concern how quantities transform under adiabatic changes. For a general equation of state with fixed composition, define the following:

$$\Gamma_1 \equiv \frac{\rho}{P}\left(\frac{\partial P}{\partial \rho}\right)_s = \left(\frac{\partial \ln P}{\partial \ln \rho}\right)_s; \tag{A.10}$$

$$\frac{\Gamma_2 - 1}{\Gamma_2} \equiv \left(\frac{\partial \ln T}{\partial \ln P}\right)_s \equiv \nabla_{\mathrm{ad}}; \tag{A.11}$$

$$\Gamma_3 - 1 \equiv \left(\frac{\partial \ln T}{\partial \ln \rho}\right)_s. \tag{A.12}$$

The nomenclature is historical. These quantities are not independent: for example, one can show (see exercise 2) that

$$\Gamma_1 = \left[\chi_\rho + \chi_T\left(\Gamma_3 - 1\right)\right].$$

Furthermore,

$$\begin{aligned}\Gamma_3 - 1 &= \frac{P}{\rho T}\frac{\chi_T}{c_\rho},\\ \frac{\Gamma_2 - 1}{\Gamma_2} &= \frac{\Gamma_3 - 1}{\Gamma_1}.\end{aligned}$$

Furthermore, one can show that the specific heat at constant pressure is $c_P = (\Gamma_1/\chi_\rho)c_\rho$.

A worked example: derivatives of the opacity

Here's a simple worked example of how we can use these identities. Suppose we have an opacity $\kappa(\rho, T)$ that is expressed in terms of density and temperature, but we need the quantity $(\partial\kappa/\partial T)_P$. We can express $(\partial\kappa/\partial T)_P$ as

$$\begin{aligned}
\left(\frac{\partial\kappa}{\partial T}\right)_P &= \frac{\partial(\kappa, P)}{\partial(T, P)} \\
&= \frac{\partial(\kappa, P)}{\partial(T, \rho)}\frac{\partial(T, \rho)}{\partial(T, P)} \\
&= \left(\frac{\partial\rho}{\partial P}\right)_T \left[\left(\frac{\partial\kappa}{\partial T}\right)_\rho \left(\frac{\partial P}{\partial\rho}\right)_T - \left(\frac{\partial\kappa}{\partial\rho}\right)_T \left(\frac{\partial P}{\partial T}\right)_\rho\right] \\
&= \left(\frac{\partial\kappa}{\partial T}\right)_\rho - \left(\frac{\partial\kappa}{\partial\rho}\right)_T \frac{\partial(\rho, T)}{\partial(P, T)}\frac{\partial(P, \rho)}{\partial(T, \rho)} \\
&= \left(\frac{\partial\kappa}{\partial T}\right)_\rho - \left(\frac{\partial\kappa}{\partial\rho}\right)_T \frac{\chi_T}{\chi_\rho}\frac{\rho}{T}. \qquad \text{(A.13)}
\end{aligned}$$

We have to add to $(\partial\kappa/\partial T)_\rho$ a term taht allows for the density to change with temperature at constant pressure.

EXERCISE A.1—

1. Show that
$$\left(\frac{\partial T}{\partial P}\right)_S \left(\frac{\partial S}{\partial T}\right)_P \left(\frac{\partial P}{\partial S}\right)_T = -1$$

2. Show that
$$\left(\frac{\partial P}{\partial s}\right)_\rho = \frac{P}{c_\rho}\chi_T, \qquad \left(\frac{\partial P}{\partial\rho}\right)_s = \frac{P}{\rho}\left[\chi_\rho + \chi_T(\Gamma_3 - 1)\right].$$

A.2 *Moments of the radiant intensity*

In the notes on radiative transport (§ 6.1) we have tended to use quantities derived from the specific intensity with a readily interpretable physical meaning, such as the energy density, flux, and radiation pressure. Often, however, it is useful to make this more formal by defining *moments* of the specific intensity, which are just weighted angular averages. For example, integrating I_ν over all angles and dividing by 4π gives

$$J_\nu \equiv \frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\, \mathrm{d}\theta\, I_\mu = \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu\, I_\nu. \qquad \text{(A.14)}$$

Here $\mu = \cos\theta$. For the first moment, we can multiply I_ν by a unit vector $\boldsymbol{k}$, and then dot that into the unit directional vector and integrate over all directions,

$$H_\nu \equiv \frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\,\mathrm{d}\theta\, I_\mu\, \boldsymbol{k}\cdot\boldsymbol{n} = \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu\,\mu I_\nu. \tag{A.15}$$

Finally, we can multiply I_ν by a tensor $\boldsymbol{kk}$; contracting this along $\boldsymbol{n}$ gives

$$K_\nu \equiv \frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\varphi \int_0^{\pi} \sin\theta\,\mathrm{d}\theta\, I_\mu\, (\boldsymbol{k}\cdot\boldsymbol{n})^2 = \frac{1}{2}\int_{-1}^{1} \mathrm{d}\mu\,\mu^2 I_\nu. \tag{A.16}$$

If we further integrate equations (A.14)–(A.16) over all frequencies, we will obtain expressions for the energy density, flux, and radiation pressure,

$$u = \frac{4\pi}{c}J, \quad F = 4\pi H, \quad P = \frac{4\pi}{c}K. \tag{A.17}$$

A.3 Shocks

In a shock, the properties of the fluid change over a scale of a few mean free paths. Over this distance, the fluid properties—density, pressure, temperature—are not well characterized by smooth, differentiable functions. To understand how shocks arise, imagine a long tube filled with a gas and fitted with a piston at one end. The piston is pushed into the gas; this creates a compressed region of higher density and pressure, and a compression wave propagates into the cylinder. If we take the disturbance to be propagating in the x-direction, then at time t the density will look schematically like the profile in Figure A.1 (*top panel*).

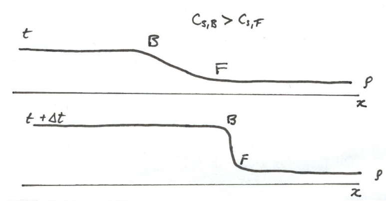


Figure A.1: Schematic of a disturbance steepening as it propagates. The plots show the density ρ at time t (*top panel*) and a time Δt later (*bottom panel*) for a disturbance propagating along the x-direction. Because the sound speed is greater in the compressed region, the "back" of the disturbance B moves faster than the front F: as a result, the disturbance steepens.

Information about the disturbance is carried by acoustic waves. In the compressed region, however, the sound speed is higher. As a result the back of the transition region (labeled "B" in Fig. A.1) travels at a faster velocity that the front "F" of the disturbance, so that over time the disturbance steepens (Fig. A.1, *bottom panel*). This steepening continues until the thickness of the transition is small enough that diffusive effects—i.e., viscosity—can balance the steepening. This gives a characteristic width to the front. On macroscopic scales, the transition is essentially a discontinuity and is termed a *shock front*.

To illustrate the properties of a shock, we'll use our thought experiment of a piston being pushed into a tube filled with gas, as illustrated in Fig. A.2. The piston moves to the right with velocity u. We assume that the piston has been pushing the fluid for a while at constant velocity, so that any transients have died down, leaving a simple flow structure at $t = 0$ (Fig. A.2, *top panel*): far from the piston, the fluid is at rest, with density ρ_0, pressure p_0, and internal energy ϵ_0; a shock propagates into this fluid with velocity D; between the shock and the piston the fluid moves with the same velocity, u, as the piston and has density ρ_1, pressure p_1, and ϵ_1.

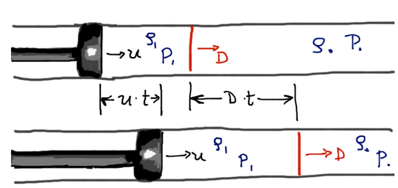


Figure A.2: Schematic of a piston driving a shock. In this schematic, the shock propagates at velocity D.

In a time t, the piston has moved (Fig. A.2, *botton panel*) a distance ut, and the shock has moved a distance Dt; the mass of the "shocked" fluid has therefore increased by $\rho_1(D-u)t$. This must equal the mass swept up by the shock, namely $\rho_0 Dt$. We therefore have our first relation,

$$\rho_1(D-u) = \rho_0 D. \tag{A.18}$$

The fluid swept up by the shock now has velocity u and thus a momentum $(\rho_0 Dt)u$. This increase in momentum must equal the net impulse imparted to the fluid, $(p_1 - p_0)t$. We therefore have our second relation,

$$p_1 - p_0 = D\rho_0 u. \tag{A.19}$$

The fluid swept up by the shock has a change in total energy, $(\rho_0 Dt)(\epsilon_1 + u^2/2 - \epsilon_0)$. This must equal the work done on the fluid by the piston, which is the force times displacement $p_1 ut$. We therefore have our third relation,

$$\rho_0 D\left[\epsilon_1 - \epsilon_0 + \frac{1}{2}u^2\right] = p_1 u. \tag{A.20}$$

To abstract the problem, we transform to a frame moving with the shock. In this frame the upstream velocity ahead of the shock has velocity $u_0 = -D$; the downstream velocity is $u_1 = u - D$. The difference in velocity is $|u_0 - u_1| = u$. In this frame, equations (A.18), (A.19), and (A.20) take on the more familiar forms,

$$\begin{aligned} \rho_1 u_1 &= \rho_0 u_0, & \text{(A.21)}\\ p_1 + \rho_1 u_1^2 &= p_0 + \rho_0 u_0^2, & \text{(A.22)}\\ (\rho_1 u_1)\left[\epsilon_1 + \frac{1}{2}u_1^2 + \frac{p_1}{\rho_1}\right] &= (\rho_0 u_0)\left[\epsilon_0 + \frac{1}{2}u_0^2 + \frac{p_0}{\rho_0}\right]. & \text{(A.23)} \end{aligned}$$

Recognize these? Compare with equations (2.1), (2.3), and (2.5), and recognize that equations (A.21), (A.22), (A.23) just express that the downstream mass flux, momentum flux, and energy flux equal their upstream counterparts.

Equations (A.21), (A.22), (A.23), when supplemented with an equation of state, are sufficient to allow solution for p_1 and ρ_1 in terms of p_0, ρ_0 and u. To explore the properties of the shock, we adopt a simple ideal adiabatic equation of state: $p = (\gamma - 1)\rho\epsilon = K\rho^\gamma$ with sound speed $c^2 = \gamma p/\rho$. The enthalpy is then

$$\epsilon + \frac{p}{\rho} = \frac{\gamma}{\gamma - 1}\frac{p}{\rho}.$$

Using this equation of state, we can reduce equations (A.21), (A.22), (A.23) into expressions in terms of the pressure ratio $P = p_1/p_0$:

$$\text{density ratio,} \qquad \frac{\rho_1}{\rho_0} = \frac{(\gamma+1)P + (\gamma-1)}{(\gamma-1)P + (\gamma+1)}; \tag{A.24}$$

$$\text{entropy increase,} \qquad s_1 - s_0 = c_v \ln\left[\frac{p_1}{p_0}\left(\frac{\rho_0}{\rho_1}\right)^\gamma\right]; \tag{A.25}$$

$$\text{temperature ratio,} \qquad \frac{T_1}{T_0} = \frac{p_1}{p_0}\frac{\rho_0}{\rho_1}; \tag{A.26}$$

$$\text{upstream Mach,} \qquad \left(\frac{u_0}{c_0}\right)^2 = \frac{(\gamma+1)P + (\gamma-1)}{2\gamma}; \tag{A.27}$$

$$\text{downstream Mach,} \qquad \left(\frac{u_1}{c_1}\right)^2 = \frac{(\gamma-1)P + (\gamma+1)}{2\gamma P}. \tag{A.28}$$

In the limiting case of a strong shock, $P \gg 1$, the density increase is finite: $\rho_1/\rho_0 \to (\gamma+1)/(\gamma-1)$; for $\gamma = 5/3$, $\rho_1/\rho_0 \to 4$. The temperature and entropy jumps across the shock, however, scale with P and are arbitrarily large. The upstream Mach number is likewise proportional to the pressure ratio, $\mathrm{Ma}_0^2 \to (\gamma+1)P/(2\gamma)$, but the downstream flow is subsonic, $\mathrm{Ma}_1^2 \to (\gamma-1)/(2\gamma)$. The shock transforms the ordered (low-entropy) supersonic upstream flow into the disordered (high-entropy) subsonic downstream flow.

Bibliography

E. Anders and N. Grevesse. Abundances of the elements - meteoritic and solar. *Geochim. Cosmochim. Acta*, 53:197–214, 1989.

M. Asplund, N. Grevesse, and A. J. Sauval. The Solar Chemical Composition. In T. G. Barnes, III and F. N. Bash, editors, *Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis*, volume 336 of *Astronomical Society of the Pacific Conference Series*, page 25, September 2005.

L. Bildsten. Thermonuclear burning on rapidly accreting neutron stars. In A. Alpar, R. Buccheri, and J. van Paradijs, editors, *The Many Faces of Neutron Stars*, volume 515, page 419, Dordrecht, 1998. NATO ASI ser. C, Kluwer.

Gilles Chabrier. Galactic stellar and substellar initial mass function. *Publ.Astron.Soc.Pac.*, 115:763–796, 2003.

C. Fröhlich and J. Lean. Solar radiative output and its variability: evidence and mechanisms. *A&A Rev.*, 12:273–320, December 2004.

H. L. Grant, R. W. Stewart, and A. Moilliet. Turbulence spectra from a tidal channel. *Journal of Fluid Mechanics*, 12:241–268, 1962.

I. Hubeny, A. Burrows, and D. Sudarsky. A Possible Bifurcation in Atmospheres of Strongly Irradiated Stars and Planets. *ApJ*, 594:1011–1018, September 2003.

D. G. Hummer. The effect of reflected and external radiation on stellar flux distributions. *ApJ*, 257:724–732, June 1982.

L. D. Landau and E. M. Lifshitz. *Statistical Physics, part 1*. Pergamon Press, Oxford, 3 edition, 1980.

B. Paczyński. Evolution of Single Stars. I. Stellar Evolution from Main Sequence to White Dwarf or Carbon Ignition. *Acta Astronomica*, 20:47, 1970.

D. Reimers. On the absolute scale of mass-loss in red giants. I - Circumstellar absorption lines in the spectrum of the visual companion of Alpha-1 HER. *A&A*, 61:217–224, October 1977.

E. E. Salpeter. Electrons Screening and Thermonuclear Reactions. *Australian Journal of Physics*, 7:373–+, September 1954.

E. E. Salpeter. The Luminosity Function and Stellar Evolution. *ApJ*, 121: 161–+, January 1955.

K.-P. Schröder and E. Sedlmayr. The galactic mass injection from cool stellar winds of the 1 to 2.5 $M_\odot$ stars in the solar neighbourhood. *A&A*, 366:913–922, February 2001.

F. X. Timmes and F. D. Swesty. The Accuracy, Consistency, and Speed of an Electron-Positron Equation of State Based on Table Interpolation of the Helmholtz Free Energy. *ApJS*, 126:501, 2000.

R. C. Willson and H. S. Hudson. The sun's luminosity over a complete solar cycle. *Nature*, 351:42–44, May 1991.