4 Stellar Evolution and Stellar Remnants

So far, we have considered only stars in static equilibrium, and found that a star of a given mass and composition has a unique, fully determined, structure. However, it is now also clear that true equilibrium cannot exist. Nuclear reactions in the central regions synthesize hydrogen into helium, and over time change the initial elemental composition. Furthermore, convection may set in at some radii and mix processed and unprocessed gas. The equations of pressure, opacity, and nuclear power density all depend sensitively on the abundances. Indeed, at some point, the hydrogen fuel in the core will be largely used up, and the star will lose the energy source that produces pressure, the gradient of which supports the star against gravitational collapse. It is therefore unavoidable that stars **evolve** with time. In this chapter, we discuss the various processes that stars of different masses undergo after the main sequence, and the properties of their compact remnants—white dwarfs, neutron stars, and black holes. We then study the phenomena that can occur when such compact objects accrete material from a companion star in a binary pair.

4.1 Stellar Evolution

Stellar evolution, as opposed to equilibrium, can be taken into account by solving a series of equilibrium stellar models (called a *stellar evolution track*), in which one updates, as a function of a star's age since formation, the gradual enrichment by elements heavier than hydrogen at different radii in the star. It turns out that the observed properties of stars on the main sequence change little during the hydrogen-burning stage, and therefore they make only small movements on the H-R diagram.

From scaling arguments, we can find the dependence of the main-sequence lifetime, $t_{\rm ms}$, on stellar mass. We previously derived the observed dependence of luminosity on mass,

$$L \sim M^{\alpha}. \tag{4.1}$$

We now also know that the energy source is nuclear reactions, whereby a fraction of a star's rest mass is converted to energy and radiated away. The total radiated energy is therefore proportional to mass,

$$Lt_{\rm ms} \sim E \sim M, \tag{4.2}$$

and

$$t_{\rm ms} \sim \frac{M}{L} \sim M^{1-\alpha}. \tag{4.3}$$

For intermediate-mass stars, which obey a mass–luminosity relation with $\alpha \sim 3$, $t_{\rm ms} \sim M^{-2}$. Thus, the more massive a star, the shorter its hydrogen-burning phase on the main sequence. Detailed stellar models confirm this result. For example, the main-sequence lifetimes of stars with initial solar abundance and various masses are

$$\begin{aligned} &0.5M_\odot \rightarrow \sim 5 \times 10^{10}\text{yr};\\ &1.0M_\odot \rightarrow \sim 10^{10}\text{yr};\\ &10M_\odot \rightarrow \sim 2 \times 10^{7}\text{yr}. \end{aligned} \tag{4.4}$$

The Sun is therefore about halfway through its main-sequence lifetime. We saw that, for the most massive stars, $\alpha \sim 1$, the result of electron-scattering opacity and radiation pressure. The lifetime $t_{\rm ms}$ therefore becomes independent of mass and reaches a limiting value,

$$>30M_\odot \rightarrow \sim 3 \times 10^{6}\text{yr}. \tag{4.5}$$

The lifetimes of massive stars, $\sim 10^6$–10^7 yr, are short compared to the age of the Sun or the age of the Universe (which, as we will see in chapters 7–9, is about 14 gigayears [Gyr], where 1 Gyr is 10^9 yr). The fact that we observe such stars means that star formation is an ongoing process, as we will see in chapter 5.

Once most of the hydrogen in the core of a star has been converted into helium, the core contracts and the inner temperatures rise. As a result, hydrogen in the less-processed regions outside the core starts to burn in a shell surrounding the core. Stellar models consistently predict that at this stage there is a huge expansion of the outer layers of the star. The increase in luminosity, due to the gravitational contraction and the hydrogen shell burning, moves the star up in the H-R diagram, while the increase in radius lowers the effective temperature, moving the star to the right on the diagram (see Fig. 4.1). This is the **red giant** phase. The huge expansion of the star's envelope is difficult to explain by means of some simple and intuitive argument, but it is well understood and predicted robustly by the equations of stellar structure. The red-giant phase is brief compared to the main sequence, lasting roughly one-tenth the time, from a billion years for solar-mass stars, to only of order a million years for $\sim 10\,M_\odot$ stars, and a few 10^5 years for the most massive stars.

As the red-giant phase progresses, the helium core contracts and heats up, while additional helium "ash" is deposited on it by the hydrogen-burning shell. At some point,

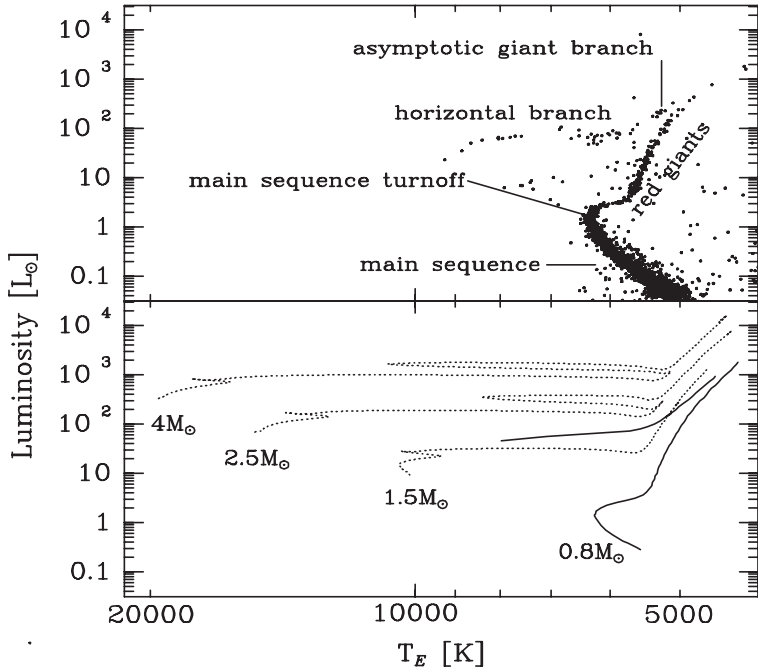

Figure 4.1 Illustration of post-main-sequence evolution on the H-R diagram. *Top:* Observed H-R diagram for stars in the *globular cluster* M3 (more on star clusters in section 5.1.5). The **main-sequence turnoff** marks the point at which stars are now leaving the main sequence and evolving on to the red-giant branch. All the stars in the cluster formed together about 13 Gyr ago, and the cluster has not experienced subsequent star formation. As a result, all stars above a certain mass, corresponding to the turnoff point, have left the main sequence, while those below that mass are still on the main sequence. The density of points in each region of the diagram reflects the amount of time spent by stars in each post-main-sequence evolution stage. *Bottom:* Theoretical stellar evolution tracks for stars of various initial main-sequence masses (with an assumed initial metal abundance of $Z = 0.0004$). Each track begins at the lower left end on the *zero-age main sequence*. After leaving the main sequence, stars evolve along, and up to the tip of, the red-giant branch. They then move quickly on the diagram to the left edge of the *horizontal branch*, where helium core burning and hydrogen shell burning take place, and evolve to the right along the horizontal branch. Once all the helium in the core has been converted to carbon and oxygen, the star rises up the "asymptotic giant branch" where double shell burning—a helium-burning shell within a hydrogen burning shell—takes place. Note the good correspondence between the theoretical track for the $0.8M_\odot$ initial-mass star (solid line) and the observed H-R diagram on top. For clarity, the theoretical horizontal and asymptotic giant branches are not shown for the other initial masses. Data credits: S.-C. Rey et al. 2001, *Astrophys. J.*, 122, 3219; and L. Girardi, et al. 2000, *Astron. Astrophys. Suppl.*, 141, 371.

the core will reach a temperature of about $T \sim 10^8$ K and a density $\rho \sim 10^4$ g cm^{-3}, where helium burning can become effective through the **triple-alpha** reaction,

$$^4\mathrm{He} + ^4\mathrm{He} + ^4\mathrm{He} \rightarrow ^{12}\mathrm{C} + \gamma(7.275\,\mathrm{MeV}). \tag{4.6}$$

Triple-alpha is the only reaction that can produce elements heavier than helium in the presence of only hydrogen and helium, because no stable elements exist with atomic mass numbers of 5 or 8. The beryllium isotope ^{8}Be, formed from the fusion of two ^{4}He nuclei,

has a lifetime of only $\sim 10^{-16}$ s. Nevertheless, a small equilibrium abundance of ^{8}Be can be established, and capture of another ^{4}He nucleus then completes the triple-alpha process. The last stage would have an extremely low probability, were it not for the existence of an excited nuclear energy level in ^{12}C, which, when added to the rest mass energy of ^{12}C, happens to have almost exactly the rest mass energies of ^{4}He $+^{8}$Be. This *resonance* greatly increases the cross section for the second stage of the reaction. In fact, from the existence of abundant carbon in the Universe (without which, of course, carbon-based life would be impossible) Hoyle predicted the existence of this excited level of ^{12}C before it was discovered experimentally.

Along with carbon production, some oxygen and neon can also be synthesized via the reactions

$$^{4}\mathrm{He} + {}^{12}\mathrm{C} \rightarrow {}^{16}\mathrm{O} + \gamma \tag{4.7}$$

and

$$^{4}\mathrm{He} + {}^{16}\mathrm{O} \rightarrow {}^{20}\mathrm{Ne} + \gamma. \tag{4.8}$$

At the same time, hydrogen continues to burn in a shell surrounding the core. When helium ignition begins, the star moves quickly on the H-R diagram to the left side of the **horizontal branch**, and then evolves more slowly to the right along this branch, as seen in Fig. 4.1. Horizontal branch evolution last only about 1% of the main-sequence lifetime. Once the helium in the core has been exhausted, the core (now composed mainly of oxygen and carbon) contracts again, until a surrounding shell of helium ignites, with a hydrogen-burning shell around it. During this brief ($\sim 10^{7}$ yr) double-shell-burning stage, the star ascends the **asymptotic giant branch** of the H-R diagram—essentially a repeat of the red-giant branch evolution, but with helium + hydrogen shell burning around an inert carbon/oxygen core, rather than hydrogen shell burning around an inert helium core.

Evolved stars undergo large mass loss, especially on the red-giant branch and on the asymptotic giant branch, as a result of the low gravity in their extended outer regions and the radiation pressure produced by their large luminosities. Mass loss is particularly severe on the asymptotic giant branch during so-called **thermal pulses**—roughly periodic flashes of enhanced helium shell burning. These mass outflows, or stellar **winds**, lead to mass-loss rates of up to $10^{-4} M_{\odot}$ yr^{-1}, which rid a star of a large fraction of its initial mass. Giants are highly convective throughout their volumes, leading to a **dredge-up** of newly synthesized elements from the core to the outer layers, where they are expelled with the winds. In these processes, and additional ones we will see below, the nuclear reactions inside post-main-sequence stars create essentially all elements in the Universe that are heavier than helium.

In stars with an initial mass of less than about $8M_{\odot}$, as the giant phase progresses, the dense matter in the core reaches equilibrium in a new state of matter called a degenerate electron gas. As we will see in the next section, regions of the core that are in this state are supported against further gravitational contraction, even in the absence of nuclear reactions. As a result, the cores of such stars do not heat up to the temperatures required for the synthesis of heavier elements, and at the end of the asymptotic giant phase they remain with a helium/carbon/oxygen core.

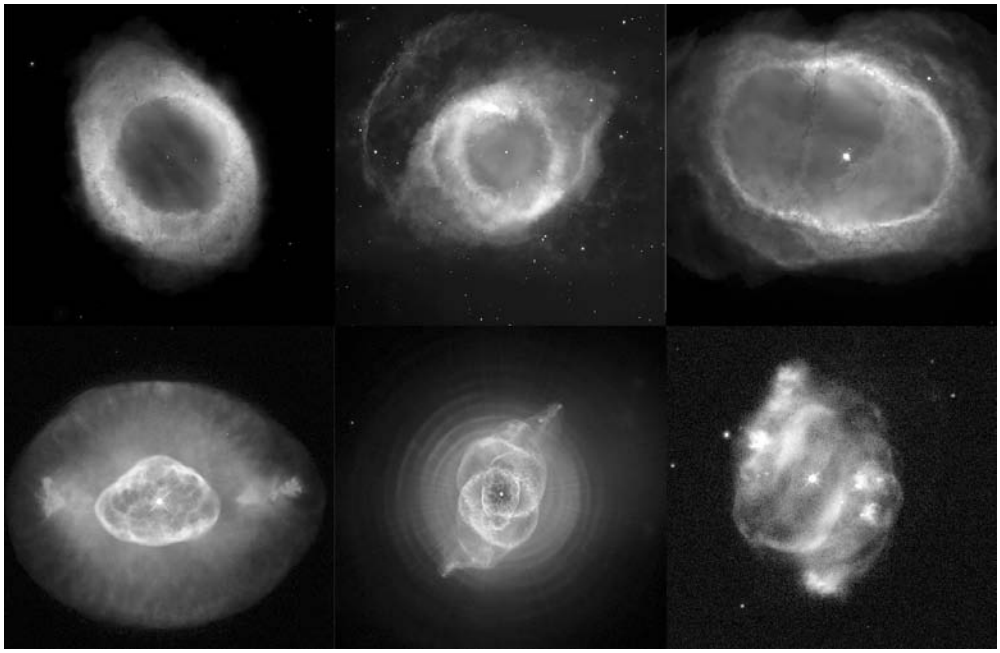

Figure 4.2 Several examples of planetary nebulae, newly formed white dwarfs that irradiate the shells of gas that were previously shed in the final stages of stellar evolution. The shells have diameters of ≈0.2–1 pc. Photo credits: M. Meixner, T. A. Rector, B. Balick et al., H. Bond, R. Ciardullo, NASA, NOAO, ESA, and the Hubble Heritage Team.

At this point, the remaining outer envelopes of the star expand to the point that they are completely blown off and dispersed. During this very brief stage ($\sim 10^4$ yr), the star is a **planetary nebula**[1] (see Fig. 4.2), in which ultraviolet photons from the hot, newly exposed, core excite the expanding shells of gas that previously constituted the outer layers of the star. Finally, the exposed remnant of the original core, called a "white dwarf," reaches the endpoint of stellar evolution for stars of this mass. In the white-dwarf region of the H-R diagram, these stars move with time to lower temperature and luminosity as they slowly radiate away their heat. White dwarfs are the subject of the next section.

Stars with an initial mass greater than about $8M_\odot$ continue the sequence of core contraction and synthesis of progressively heavier elements, which eventually (and quickly) ends in a supernova explosion. We shall return to this class of stars in section 4.3.

4.2 White Dwarfs

In the 19th and early 20th centuries, it was discovered that the nearby (2.7 pc) A-type star Sirius, the brightest star in the sky, is a visual binary, with a white dwarf companion that was named Sirius B. (In fact, Sirius B is the nearest known white dwarf, and was the first

[1] Planetary nebulae have nothing to do with planets, and the name has a purely historical origin.

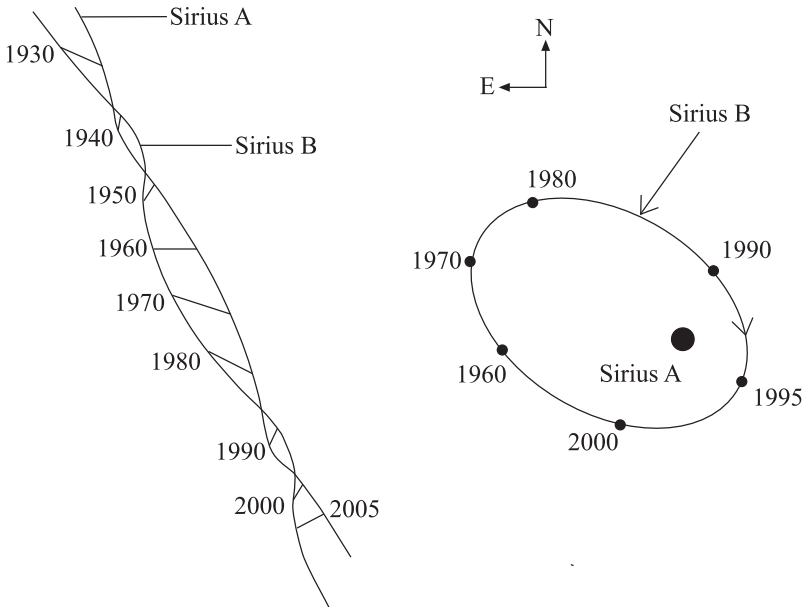

Figure 4.3 Observed motion on the sky, over the past century, of the visual binary consisting of Sirius A and its faint white dwarf companion, Sirius B. On the left are the observed positions due the orbital motions around the center of mass, combined with the proper motion of the system as a whole. On the right side, only the positions of Sirius B relative to Sirius A are shown. The maximum projected separation of the pair is 10 arcseconds. Using Kepler's law, a mass close to $1M_\odot$ is derived for the white dwarf.

one ever found.) An orbital period of about 50 years was observed (see Fig. 4.3), allowing the first measurement of the mass of a white dwarf, which turned out to be close to $1M_\odot$. Like all white dwarfs, Sirius B's low luminosity and high temperature imply a small radius of about 6000 km, i.e., less than that of the Earth. The mean density inside Sirius B is therefore of order 1 ton cm^{-3}. In this section, we work out the basic physics of white dwarfs and of matter at these extremely high densities.

4.2.1 Matter at Quantum Densities

We saw in the previous section that when the core of a star exhausts its nuclear energy supply, it contracts and heats up until reaching the ignition temperature of the next available nuclear reaction, and so on. After each contraction, the density of the core increases. At some point, the distances between atoms will be smaller than their de Broglie wavelengths. At that point, our previous assumption of a classical (rather than quantum) ideal gas, which we used to derive the equation of state, becomes invalid. To get an idea of the conditions under which this happens, recall that the de Broglie wavelength of a particle of momentum p is

$$\lambda = \frac{h}{p} = \frac{h}{(2mE)^{1/2}} \approx \frac{h}{(3mkT)^{1/2}}, \tag{4.9}$$

where we have represented the energy with the mean energy of a particle, $E \sim 3kT/2$. Since electrons and protons share the same energy, but the mass of the electron is much smaller than the mass of the proton or of other nuclei, the wavelengths of the electrons are longer, and it is the electron density that will first reach the quantum domain. At interparticle separations of order less than half a de Broglie wavelength, quantum effects should become important, corresponding to a density of

$$\rho_q \approx \frac{m_p}{(\lambda/2)^3} = \frac{8m_p(3m_e kT)^{3/2}}{h^3}. \tag{4.10}$$

For example, for the conditions at the center of the Sun, $T = 15 \times 10^6$ K, we obtain

$$\rho_q \approx \frac{8 \times 1.7 \times 10^{-24} \text{ g } (3 \times 9 \times 10^{-28} \text{ g} \times 1.4 \times 10^{-16} \text{ erg K}^{-1} \times 15 \times 10^6 \text{K})^{3/2}}{(6.6 \times 10^{-27} \text{ erg s})^3}$$

$$= 640 \text{ g cm}^{-3} \tag{4.11}$$

The central density in the Sun is $\rho \approx 150$ g cm^{-3}, and thus the gas in the Sun is still in the classical regime. Even very dense gas can remain classical, if it is hot enough. For example, for $T = 10^8$ K, i.e., $E \sim kT \sim 10$ keV,

$$\rho_q \approx 11{,}000 \text{ g cm}^{-3}. \tag{4.12}$$

Instead of the Maxwell-Boltzmann distribution, the energy distribution at quantum densities will follow Bose-Einstein statistics for bosons (particles with spin that is an integer multiple of $\hbar$) or Fermi-Dirac statistics for fermions (particles with spin that is an uneven integer multiple of $\hbar/2$). Let us develop the equation of state for such conditions.

4.2.2 Equation of State of a Degenerate Electron Gas

Heisenberg's uncertainty principle states that, due to the wave nature of matter, the position and momentum of a particle are simultaneously defined only to within an uncertainty

$$\Delta x \Delta p_x > h. \tag{4.13}$$

Similar relations can be written for each of the coordinates, x, y, and z. Multiplying the relations, we obtain

$$\Delta x \Delta y \Delta z \Delta p_x \Delta p_y \Delta p_z > h^3, \tag{4.14}$$

or

$$d^3\mathbf{p}dV > h^3. \tag{4.15}$$

The constant h^3 thus defines the six-dimensional volume of a "cell" in position–momentum phase space. The uncertainty principle implies that two identical particles

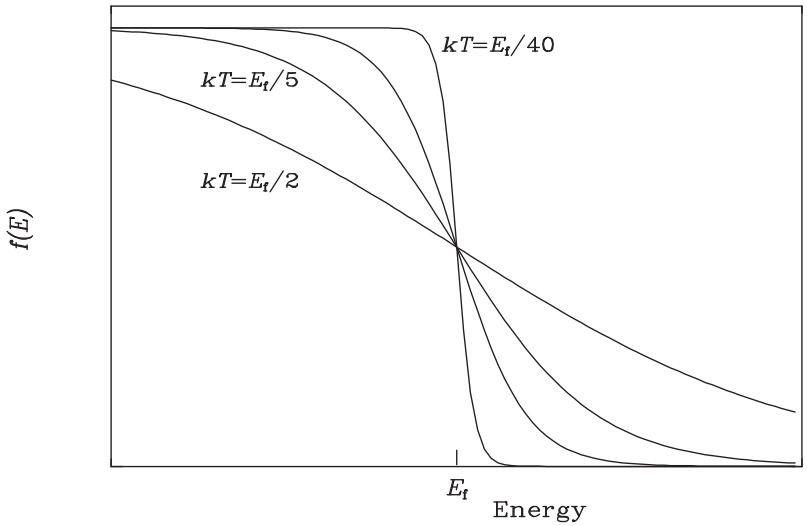

Figure 4.4 Approach to degeneracy of the Fermi-Dirac occupation fraction, $f(E)$, as $kT \to 0$, shown for $kT = E_f/2$, $E_f/5$, $E_f/10$, and $E_f/40$. At $kT \ll E_f$, all particles occupy the lowest energy state possible without violating the Pauli exclusion principle. The distribution then approaches a step function, with all energy states below E_f occupied, and all those above E_f empty.

that are in the same phase–space cell are in the same quantum state. According to Pauli's exclusion principle, two identical fermions cannot occupy the same quantum state. Thus, fermions that are closely packed and hence localized into a small volume, dV, must each have a large uncertainty in momentum, *and* have momenta $\mathbf{p}$ that are different from those of the other fermions in the volume. This necessarily pushes the fermions to large $\mathbf{p}$'s, and large momenta mean large pressure.

The **Fermi-Dirac phase–space distribution**, embodying these principles for an ideal gas of fermions, is

$$dN = \frac{2s+1}{\exp\left(\frac{E-\mu(T)}{kT}\right)+1}\frac{d^3\mathbf{p}dV}{h^3}, \tag{4.16}$$

where s is the spin of each fermion in units of $\hbar$, and $\mu(T)$ is the chemical potential[2] of the gas. When $T \to 0$, then $\mu(T)$ approaches an assymptotic value, E_f. When $kT \ll E_f$, the first term in the Fermi-Dirac distribution (the *occupation fraction*) approaches a step function (see Fig. 4.4) in which all particles occupy the lowest energy states possible without violating the Pauli principle. This means that all energy states up to an energy E_f are occupied, and all above E_f are empty. Under such conditions, the gas is said to be **degenerate**. For

[2] The chemical potential of a thermodynamic system is the change in energy due to the introduction of an additional particle, at constant entropy and volume.

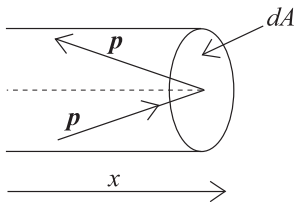

Figure 4.5 Calculation of the pressure exerted by particles of an ideal gas with momentum **p** that are reflected off the side of a container.

degenerate electrons, which are $s = 1/2$ particles, having an isotropic velocity field, the phase–space distribution will be

$$dN(p)dp = \begin{cases} 2 \times 4\pi p^2 \frac{dpdV}{h^3} & \text{if } |\mathbf{p}| \leq p_f \\ 0 & \text{if } |\mathbf{p}| > p_f \end{cases}, \tag{4.17}$$

where p_f, called the **Fermi momentum**, is the magnitude of the momentum corresponding to the **Fermi energy** E_f. Dividing by dV, we obtain the number density of electrons of a given momentum p:

$$n_e(p)dp = \begin{cases} 8\pi p^2 \frac{dp}{h^3} & \text{if } |\mathbf{p}| \leq p_f \\ 0 & \text{if } |\mathbf{p}| > p_f \end{cases}. \tag{4.18}$$

Integrating over all momenta from 0 to p_f gives a relation between the electron density and p_f:

$$n_e = \int_0^{p_f} \frac{8\pi}{h^3} p^2 dp = \frac{8\pi}{3h^3} p_f^3. \tag{4.19}$$

Next, let us derive a general expression for the pressure exerted by any ideal gas. By definition, an ideal gas consists of particles that interact only at short distances, and hence can transfer momentum only during an "impact" with another particle. Consider ideal gas particles impinging on the side of a container, with a mean interval dt between consecutive impacts (see Fig. 4.5). Set the x axis perpendicular to the surface. Particles with an x component of momentum p_x will transfer a momentum $2p_x$ to the surface with each reflection. The force per unit area due to each collision is then

$$\frac{dF_x}{dA} = \frac{2p_x}{dAdt} = \frac{2p_x v_x}{dAdx} = \frac{2p_x v_x}{dV}, \tag{4.20}$$

where $v_x = dx/dt$. The pressure is obtained by summing the forces due to all particles of all momenta:

$$P = \int_0^{\infty} dN(p) \frac{p_x v_x}{dV} dp, \tag{4.21}$$

where we have divided by 2 because, at any given time, only half of all the particles will have a v_x component toward the side of the container, rather than away from it. But

$$p_x v_x = m v_x^2 = \tfrac{1}{3} m v^2 = \tfrac{1}{3} p v, \tag{4.22}$$

where we have utilized $v_x^2 + v_y^2 + v_z^2 = v^2$ and assumed that the velocities are isotropic so that, on average, $v_x^2 = v_y^2 = v_z^2$. Since $dN/dV \equiv n$,

$$P = \frac{1}{3} \int_0^\infty n(p) p v \, dp. \tag{4.23}$$

Replacing the Maxwell-Boltzmann distribution for $n(p)$ recovers the classical equation of state,

$$P = nkT. \tag{4.24}$$

For a nonrelativistic[3] **degenerate** electron gas, however, we replace $n(p)$ with the Fermi-Dirac distribution in the degenerate limit (Eq. 4.18). Taking $v = p/m_e$, we obtain instead

$$P_e = \frac{1}{3} \int_0^{p_f} \frac{8\pi}{h^3} \frac{p^4}{m_e} dp = \frac{8\pi}{3h^3 m_e} \frac{p_f^5}{5} = \left(\frac{3}{8\pi} \right)^{2/3} \frac{h^2}{5 m_e} n_e^{5/3}, \tag{4.25}$$

where we have used Eq. 4.19 to express p_f in terms of n_e. To relate n_e to the mass density appearing in the equations of stellar structure, consider a fully ionized gas composed of a particular element, of atomic number $\mathcal{Z}$ and atomic mass number A, and a density of ions n_+. Then

$$n_e = \mathcal{Z} n_+ = \mathcal{Z} \frac{\rho}{A m_p}. \tag{4.26}$$

Substituting into Eq. 4.25, we obtain a useful form for the **equation of state of a degenerate nonrelativistic electron gas**:

$$P_e = \left(\frac{3}{\pi} \right)^{2/3} \frac{h^2}{20 m_e m_p^{5/3}} \left(\frac{\mathcal{Z}}{A} \right)^{5/3} \rho^{5/3}. \tag{4.27}$$

The important feature of this equation of state is that the electron pressure does not depend on temperature. Indeed, in our derivation of this equation, we have assumed that kT is effectively zero. (More precisely, kT is very low compared to the energy of the most energetic electrons at the Fermi energy, which are prevented from occupying lower energy states by the Pauli principle—see Problem 1.)

In a typical white dwarf, $\rho \sim 10^6$ g cm^{-3} and $T \sim 10^7$ K. White dwarfs are generally composed of material that was processed by nuclear reactions into helium, carbon, and oxygen, and therefore $\mathcal{Z}/A \approx 0.5$. Plugging these numbers into 4.27, we find

[3] Note that, although we have used nonrelativistic considerations (Eq. 4.22) to derive Eq. 4.23, it holds in the relativistic case as well. We can easily verify that, for an ultrarelativistic gas with particle energies E, by replacing p with E/c, v with c, and $n(p)dp$ with $n(E)dE$, we recover the relation $P = U/3$, which we derived in Eq. 3.74.

$$P_e \sim \frac{(6.6 \times 10^{-27}\ \text{erg s})^2}{20 \times 9 \times 10^{-28}\ \text{g}\ (1.7 \times 10^{-24}\ \text{g})^{5/3}} 0.5^{5/3} (10^6\ \text{g cm}^{-3})^{5/3}$$
$$= 3 \times 10^{22}\ \text{dyne cm}^{-2}. \tag{4.28}$$

As opposed to the electrons, the nuclei at such densities are still completely in the classical regime. The thermal pressure due to the nuclei, assuming a helium composition, is

$$P_{\text{th}} = n_+ kT = \frac{\rho}{4m_p} kT \sim \frac{10^6\ \text{g cm}^{-3} \times 1.4 \times 10^{-16}\ \text{erg K}^{-1} \times 10^7\ \text{K}}{4 \times 1.7 \times 10^{-24}\ \text{g}}$$
$$= 2 \times 10^{20} \text{dynes cm}^{-2}. \tag{4.29}$$

The degenerate electron pressure therefore completely dominates the pressure in the star.

4.2.3 Properties of White Dwarfs

Next, we can see what the degenerate electron pressure equation of state, combined with the other equations of stellar structure, implies for the properties of white dwarfs. Let us start with the relation between mass and radius.

4.2.3.1 Mass–Radius Relationship

The equations of mass continuity and hydrostatic equilibrium, expressed as scaling relations (see Eqs. 3.80 and 3.81), suggest

$$\rho \sim \frac{M}{r^3}, \tag{4.30}$$

and

$$P \sim \frac{GM\rho}{r} \sim \frac{GM^2}{r^4}. \tag{4.31}$$

The degenerate electron-gas equation of state is

$$P \sim b\rho^{5/3} \sim b\frac{M^{5/3}}{r^5}, \tag{4.32}$$

where the constant factor b is given in Eq. 4.27. Equating the pressures gives

$$r \sim \frac{b}{G} M^{-1/3}. \tag{4.33}$$

In other words, the radius of a white dwarf decreases with increasing mass. An order-of-magnitude estimate of the radius is therefore

$$r_{\text{wd}} \sim \frac{b}{G} M^{-1/3} \sim \frac{h^2}{20 m_e m_p^{5/3} G} \left(\frac{\mathcal{Z}}{A}\right)^{5/3} M^{-1/3}$$
$$\sim \frac{(6.6 \times 10^{-27}\ \text{erg s})^2 (2 \times 10^{33}\ \text{g})^{-1/3}}{20 \times 9 \times 10^{-28}\ \text{g}\ (1.7 \times 10^{-24}\ \text{g})^{5/3}\ 6.7 \times 10^{-8}\ \text{cgs}} \left(\frac{\mathcal{Z}}{A}\right)^{5/3} \left(\frac{M}{M_\odot}\right)^{-1/3}$$
$$= 1.2 \times 10^9 \text{cm} \left(\frac{\mathcal{Z}}{A}\right)^{5/3} \left(\frac{M}{M_\odot}\right)^{-1/3}, \tag{4.34}$$

i.e., about 4000 km for $\mathcal{Z}/A = 0.5$ and $M = 1M_\odot$, as deduced for observed white dwarfs from their luminosities and temperatures. A full solution of the equations of stellar structure for the degenerate gas equation of state gives

$$r_{\rm wd} = 2.3 \times 10^9 {\rm cm} \left(\frac{\mathcal{Z}}{A}\right)^{5/3} \left(\frac{M}{M_\odot}\right)^{-1/3}. \tag{4.35}$$

4.2.3.2 The Chandrasekhar Mass

The larger the white-dwarf mass that we consider, the smaller $r_{\rm wd}$ becomes, implying larger densities, and therefore larger momenta to which the electrons are pushed. When the electron velocities become comparable to the speed of light, we can no longer assume $v = p/m$ in Eq. 4.23. Instead, v, which dictates the rate at which collisions transfer momentum to the container wall, approaches c. In the *ultrarelativistic* limit, we can replace v with c. Equation 4.25 is then replaced with

$$P_e = \frac{1}{3}\int_0^{p_f} \frac{8\pi}{h^3} p^2 pc dp = \frac{8\pi c}{3h^3}\frac{p_f^4}{4}. \tag{4.36}$$

Again using Eqs. 4.19 and 4.26, we obtain the **equation of state for an ultrarelativistic degenerate spin-1/2 fermion gas**:

$$\boxed{P_e = \left(\frac{3}{8\pi}\right)^{1/3} \frac{hc}{4m_p^{4/3}} \left(\frac{\mathcal{Z}}{A}\right)^{4/3} \rho^{4/3}.} \tag{4.37}$$

Compared to the nonrelativistic case (Eq. 4.27), note the 4/3 power, but also the fact that the electron mass does not appear, i.e., this equation holds for *any* ultrarelativistic degenerate ideal gas of spin-1/2 particles. This comes about because, for ultrarelativistic particles, the rest mass is a negligible fraction of the total energy, $E = (m^2c^4 + p^2c^2)^{1/2}$, and hence $p \approx E/c$. As we go from small to large white-dwarf masses there will be a gradual transition from the nonrelativistic to the ultrarelativistic equation of state, with the power-law index of ρ gradually decreasing from 5/3 to 4/3.

This necessarily means that, as we go to higher masses, and the density increases due to the shrinking radius, the pressure support will rise more and more slowly, so that the radius shrinks even more sharply with increasing mass.[4] To see what happens as a result, let us rederive the scaling relations between mass and radius, but with an index $(4+\epsilon)/3$, and then let ϵ approach 0. Thus,

$$P \sim \rho^{(4+\epsilon)/3}, \tag{4.38}$$

so

$$\frac{M^{(4+\epsilon)/3}}{r^{4+\epsilon}} \sim \frac{M^2}{r^4}, \tag{4.39}$$

[4] Sirius B, with a mass of $1M\odot$, is among the more massive white dwarfs known, and its equation of state is already in the mildly relativistic regime. Its radius, 5880 km, is therefore smaller than would be expected based on Eq 4.35, but is fully consistent with the results of a relativistic calculation.

or

$$r^{\epsilon} \sim M^{(\epsilon-2)/3}, \tag{4.40}$$

and

$$r \sim M^{(\epsilon-2)/3\epsilon}. \tag{4.41}$$

When $\epsilon \to 0$,

$$r \to M^{-\infty} = 0. \tag{4.42}$$

In other words, at a mass high enough so that the electrons become ultrarelativistic, the electron pressure becomes incapable of supporting the star against gravity, the radius shrinks to zero (and the density rises to infinity), unless some other source of pressure sets in. We will see that, at high enough density, the degeneracy pressure due to protons and neutrons begins to operate, and it can sometimes stop the full gravitational collapse, producing objects called neutron stars.

The above argument implies that there is a maximum stellar mass that can be supported by degenerate electron pressure. It is called the **Chandrasekhar mass**. To estimate it, recall from the virial theorem that

$$\bar{P}V = -\tfrac{1}{3}E_{\rm gr}. \tag{4.43}$$

Substituting the ultrarelativistic electron degeneracy pressure for $\bar{P}$, and the usual expression for the self-energy $E_{\rm gr}$, we can write

$$\left(\frac{3}{8\pi}\right)^{1/3} \frac{hc}{4m_p^{4/3}} \left(\frac{\mathcal{Z}}{A}\right)^{4/3} \rho^{4/3} V \sim \frac{1}{3}\frac{GM^2}{r}. \tag{4.44}$$

With

$$\rho \sim \frac{M}{V} \tag{4.45}$$

and

$$V = \frac{4\pi}{3}r^3, \tag{4.46}$$

r cancels out of the equation and we obtain

$$M \sim 0.11 \left(\frac{\mathcal{Z}}{A}\right)^2 \left(\frac{hc}{Gm_p^2}\right)^{3/2} m_p. \tag{4.47}$$

A full solution of the equations of stellar structure for this equation of state gives a somewhat larger numerical coefficient, so that the Chandrasekhar mass is

$$M_{\rm ch} = 0.21 \left(\frac{\mathcal{Z}}{A}\right)^2 \left(\frac{hc}{Gm_p^2}\right)^{3/2} m_p. \tag{4.48}$$

The expression $Gm_p^2/(hc)$ that appears in the Chandrasekhar mass is a dimensionless constant that can be formed by taking a proton's gravitational self-energy, with the proton radius expressed by its de Broglie wavelength, and forming the ratio with the proton's rest energy:

$$\alpha_G \equiv \frac{Gm_p^2}{h/(m_p c)m_p c^2} = \frac{Gm_p^2}{hc}$$
$$= \frac{6.7 \times 10^{-8}\ \text{cgs}\ (1.7 \times 10^{-24}\ \text{g})^2}{6.6 \times 10^{-27}\ \text{erg s} \times 3 \times 10^{10}\ \text{cm s}^{-1}} \approx 10^{-39}. \tag{4.49}$$

The constant α_G expresses the strength of the gravitational interaction, and is the gravitational analog of the fine-structure constant,

$$\alpha_{\text{em}} = \frac{e^2}{\hbar c} \approx \frac{1}{137}, \tag{4.50}$$

which expresses the strength of the electromagnetic interaction. Equation 4.48 says that the maximum mass of a star supported by electron degeneracy pressure is, to an order of magnitude, the mass of $\alpha_G^{-3/2}$ protons (i.e., $\sim 10^{57}$ protons). Since $\mathcal{Z}/A \approx 0.5$,

$$M_{\text{ch}} = 0.21 \times 0.5^2 \times 10^{39\frac{3}{2}} \times 1.7 \times 10^{-24}\ \text{g} = 1.4 M_\odot. \tag{4.51}$$

In fact, no white dwarfs with masses higher than M_{ch} have been found.

There is also a lower bound to the masses of isolated[5] white dwarfs that have been measured, of about $0.25 M_\odot$. This, however, is a result of the finite age of the Universe, 1.4×10^{10} yr. Stars that will form white dwarfs having masses smaller than this (namely, stars that have an initial mass on the main sequence smaller than about $0.8 M_\odot$) have not yet had time to go through their main-sequence lifetimes, even if they were formed early in the history of the Universe.

4.2.3.3 White Dwarf Cooling

Due to the good thermal conduction of the degenerate electrons in a white dwarf (similar to the conduction in metals, which arises in the same way), the temperature inside a white dwarf is approximately constant with radius. The temperature can be estimated by recalling that a white dwarf forms from the contraction of a thermally unsupported stellar core, of mass M, down to the radius at which degeneracy pressure stops the contraction. Just before reaching that final point of equilibrium, from the virial theorem, the thermal energy will equal half the gravitational energy:

$$E_{\text{th}} \sim \frac{1}{2}\frac{GM^2}{r}. \tag{4.52}$$

[5] In interacting binaries, ablation by beams of matter and radiation from a companion can sometimes lower the mass of a white dwarf, or even destroy the white dwarf completely. See section 4.6.3.

For a pure helium composition, the number of nuclei in the core is $M/4m_H$, and the number of electrons is $M/2m_H$. The total thermal energy (which, once degeneracy sets it, will no longer play a role in supporting the star against gravity) is therefore

$$E_{\rm th} = \frac{3}{2}NkT = \frac{3}{2}\frac{M}{m_p}\left(\frac{1}{2}+\frac{1}{4}\right)kT = \frac{9}{8}\frac{M}{m_p}kT, \tag{4.53}$$

and so

$$kT \sim \frac{4}{9}\frac{GMm_p}{r}. \tag{4.54}$$

Substituting the equilibrium $r_{\rm wd}$ of white dwarfs from Eq. 4.34 yields

$$\begin{aligned} kT &\sim \frac{80G^2m_em_p^{8/3}}{9h^2}\left(\frac{\mathcal{Z}}{A}\right)^{-5/3}M^{4/3} \\ &= \frac{80(6.7\times10^{-8}\ {\rm cgs})^2\ 9\times10^{-28}\ {\rm g}\ (1.7\times10^{-24}\ {\rm g})^{8/3}}{9(6.6\times10^{-27}\ {\rm erg\ s^{-1}})^2}\ 0.5^{-5/3}(1\times10^{33}\ {\rm g})^{4/3} \\ &= 1.1\times10^{-8}{\rm erg}, \end{aligned} \tag{4.55}$$

for a $0.5M_\odot$ white dwarf. The temperature is thus $kT \sim 70$ keV, or $T \sim 8\times10^8$ K, and a just-formed degenerate core is a very hot object, with thermal emission that peaks in the X-ray part of the spectrum. As such, once the core becomes an exposed white dwarf, its radiation ionizes the layers of gas that were blown off in the various stages on the giant phase. As already noted, this produces the objects called planetary nebulae.

A white dwarf is an endpoint in stellar evolution, devoid of nuclear reactions. It therefore cools by radiating from its surface the thermal energy stored in the still-classical gas of nuclei within the star's volume. (The degeneracy of the electron gas limits almost completely the ability of the electrons to lose their kinetic energies.) The radiated luminosity will be

$$L = 4\pi r_{\rm wd}^2\sigma T_E^4, \tag{4.56}$$

where T_E is the effective temperature of the white-dwarf photosphere. Although electron heat conduction leads to a constant temperature over most of the volume, there is a thin nondegenerate surface layer (of order 1% of the white-dwarf radius) that insulates the star. This layer lowers T_E relative to the interior temperature and slows down the rate of energy loss.

However, to obtain a crude upper limit on the rate at which a white dwarf cools by means of its radiative energy loss, let us assume a constant temperature all the way out to the surface of the star, so that $T_E \sim T$. The radiative energy loss rate is then

$$4\pi r_{\rm wd}^2\sigma T^4 \sim \frac{dE_{\rm th}}{dt} = \frac{3Mk}{8m_p}\frac{dT}{dt} \tag{4.57}$$

(where we have included in the right-hand term only the contribution of the nuclei to the thermal energy from Eq. 4.53). Separating the variables T and t, and integrating, the cooling time to a temperature T is

$$\begin{aligned}\tau_{\rm cool} &\sim \frac{3Mk}{8m_p 4\pi r_{\rm wd}^2 \sigma 3T^3} \\ &= \frac{3 \times 1 \times 10^{33}\ {\rm g} \times 1.4 \times 10^{-16}\ {\rm erg\ K^{-1}}}{8 \times 1.7 \times 10^{-24}\ {\rm g}\ 4\pi (4 \times 10^8\ {\rm cm})^2 \times 5.7 \times 10^{-5}\ {\rm cgs} \times 3T^3} \\ &= 3 \times 10^9 {\rm yr} \left(\frac{T}{10^3\ {\rm K}}\right)^{-3}, \end{aligned} \tag{4.58}$$

where we have taken $M = 0.5M_\odot$ and $r_{\rm wd} = 4000$ km. (We have abbreviated the units of the Stefan-Boltzmann law's σ as cgs.) Alternatively, we can write the temperature as a function of time as

$$\frac{T}{10^3\ {\rm K}} \sim \left(\frac{t}{3 \times 10^9\ {\rm yr}}\right)^{-1/3} \tag{4.59}$$

Thus, even with the unrealistically efficient cooling we have assumed, it would take a $0.5M_\odot$ white dwarf several gigayears to cool to 10^3 K. In reality, the insulation of the nondegenerate surface layer results in an effective temperature that is significantly lower than the interior temperature, and hence an even lower cooling rate. Furthermore, at some point in the cooling evolution, crystallization of the nucleons inside the white dwarf takes place, and the latent heat that is released and added to the thermal balance further slows down the decline in temperature. Detailed models have been calculated that take these and other processes into account for various masses and chemical compositions of white dwarfs (a carbon/oxygen core is usually assumed, surrounded by helium and hydrogen envelopes). The models show that over 10^{10} yr, comparable to the age of the Universe, white dwarfs cannot cool below ∼3000–4000K. This explains why most white dwarfs are observed to have high temperatures, and hence their blue to white colors. The coolest white dwarfs known have effective temperatures of ∼3500 K.

4.2.3.4 Brown Dwarfs

Let us digress for a moment from the subject of stellar remnants, and use the equations we have developed to see that electron degeneracy and its consequences also dictate a minimal initial mass that a star must have to shine. Consider a newly forming star (or "protostar") composed of a collapsing cloud of hydrogen. Nuclear ignition of hydrogen requires a minimal temperature of about $T_{\rm ign} \approx 10^7$K. Recall the relation between temperature and mass of a white dwarf, obtained by arguing that the contraction of the core will halt at the radius when degeneracy pressure sets in (Eq. 4.55). However, for fully ionized hydrogen (as opposed to helium), $N = 2M/m_p$, (rather than $N = 3M/4m_p$; Eq. 4.53). There are 8/3 times more particles, and the temperature is correspondingly lower, so Eq. 4.55 becomes

$$kT \sim \frac{10G^2 m_e m_p^{8/3}}{3h^2} \left(\frac{\mathcal{Z}}{A}\right)^{-5/3} M^{4/3}. \tag{4.60}$$

If the mass of the protostar is small enough such that contraction halts before $T_{\rm ign}$ is attained, the object will never achieve true stardom on the main sequence. The limiting mass is

$$M_{\rm min} \sim (kT_{\rm ign})^{3/4} \left(\frac{10G^2 m_e m_p^{8/3}}{3h^2} \right)^{-3/4} \left(\frac{\mathcal{Z}}{A} \right)^{5/3} = 0.09 M_\odot, \tag{4.61}$$

where we have assumed $\mathcal{Z}/A = 1$, appropriate for hydrogen. A full solution of the stellar structure equations gives

$$M_{\rm min} \approx 0.07 M_\odot. \tag{4.62}$$

Such "failed stars," with masses lower than this limit, are called **brown dwarfs**. As noted in section 2.2.2, stars of this type have indeed been found, and are labeled with spectral types L and T.

4.3 Supernovae and Neutron Stars

4.3.1 Core Collapse in Massive Stars

We now return to stars with initial masses (i.e., their masses when they begin their lives on the main sequence) of about $8M_\odot$ or more. This corresponds to spectral types O and "early" B. After exhausting most of the hydrogen in their cores, such stars move to the giant branch. They then begin a sequence of steps, each consisting of the contraction and heating of the inner regions, resulting in the ignition of new nuclear reactions. As time advances, shells at various inner radii attain the temperatures and the densities required for the reactions that produce progressively heavier elements. Apart from the reactions already discussed for lower-mass stars,

$$^4\mathrm{He} + {}^{12}\mathrm{C} \rightarrow {}^{16}\mathrm{O} + \gamma \tag{4.63}$$

and

$$^4\mathrm{He} + {}^{16}\mathrm{O} \rightarrow {}^{20}\mathrm{Ne} + \gamma, \tag{4.64}$$

these massive stars can also burn carbon via the reactions

$$^{12}\mathrm{C} + {}^{12}\mathrm{C} \rightarrow {}^{20}\mathrm{Ne} + {}^4\mathrm{He} + \gamma, \tag{4.65}$$

$$^{12}\mathrm{C} + {}^{12}\mathrm{C} \rightarrow {}^{23}\mathrm{Na} + p, \tag{4.66}$$

and

$$^{12}\mathrm{C} + {}^{12}\mathrm{C} \rightarrow {}^{23}\mathrm{Mg} + n. \tag{4.67}$$

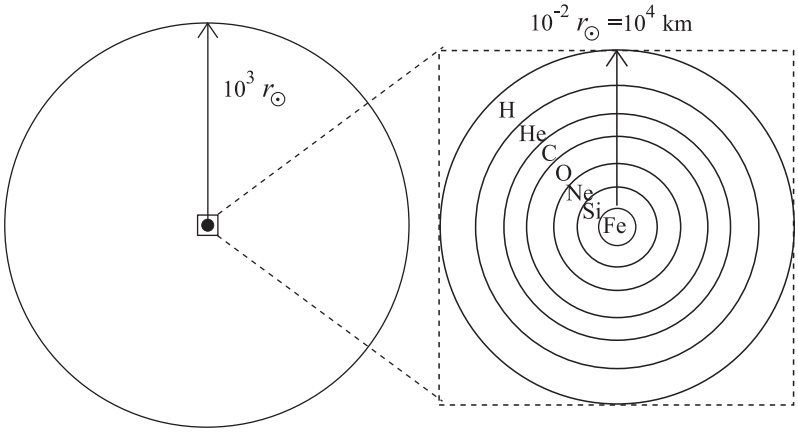

Figure 4.6 Simplified schematic view of the layered structure of a massive star and the distribution of the main elements that compose it, at the onset of core collapse and the ensuing supernova explosion.

Carbon burning is followed by neon, oxygen, and silicon burning. Each of these stages takes less and less time. For example, for a $25M_\odot$ star, the duration of each burning stage is

H $\sim 5 \times 10^6$ yr
He $\sim 5 \times 10^5$ yr
C ~ 500 yr
Ne ~ 1 yr
Si ~ 1 day.

Massive stars undergo all the stages of nuclear burning up to the production of elements in the "iron group" with atomic mass number around $A = 56$, consisting of isotopes of Cr, Mn, Fe, Co, and Ni. At this stage, the star's outer envelope has expanded to about $1000r_\odot$, and it has a dense core of radius $\sim 10^4$ km with an onion-like layered structure (see Fig. 4.6). The outer layers of this core are still burning hydrogen. Looking inward, the core consists of concentric shells composed primarily of helium, carbon, oxygen, neon, silicon, and iron, respectively.

Figure 4.7 shows, for all the chemical elements, the binding energy per nucleon (i.e., the binding energy of a nucleus divided by its mass number A). Energy can be gained by fusing or fissioning elements with low binding energy per nucleon into elements with high binding energy per nucleon. The iron group elements are the most tightly bound nuclei, and are therefore a "dead end" in nuclear energy production. Synthesis of iron-group elements into heavier elements consumes, rather than releases, thermal energy. This fact is at the root of the "iron catastrophe" that ensues.

When the central iron core continues to grow and approaches $M_{\rm ch}$, two processes begin: nuclear photodisintegration and neutronization.

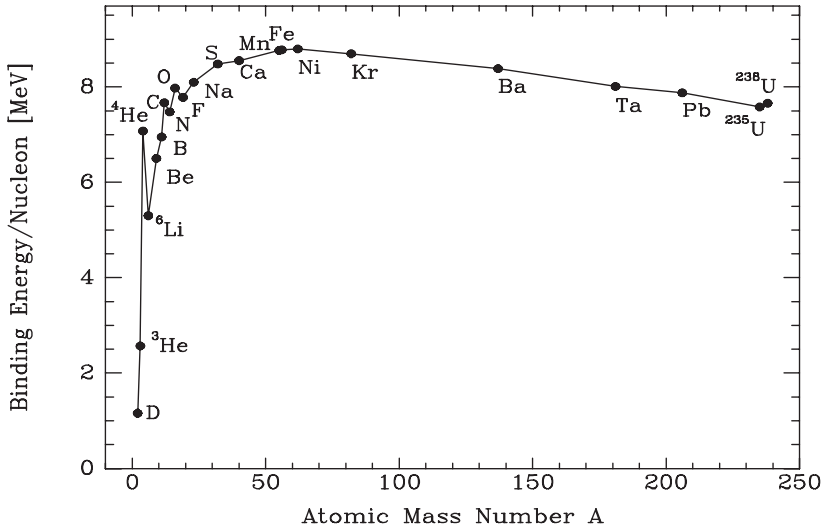

Figure 4.7 Binding energy per nucleon as a function of atomic mass number. Several elements are marked. The iron-group elements with $A \approx 56$ have the highest binding energy per nucleon, 8.8 MeV, and therefore nuclear fusion of these elements into heavier elements does not release thermal energy, but rather consumes it.

Nuclear Photodisintegration: The temperature is high enough for energetic photons to be abundant, and they get absorbed in the endothermic (i.e., energy-consuming) nuclear reaction

$$\gamma + {}^{56}\mathrm{Fe} \rightarrow 13\,{}^{4}\mathrm{He} + 4n, \tag{4.68}$$

with an energy consumption of 124 MeV. The helium nuclei are further unbound in the process

$$\gamma + {}^{4}\mathrm{He} \rightarrow 2p + 2n, \tag{4.69}$$

consuming 28.3 MeV (the binding energy of a ^{4}He nucleus). The total energy of the star is thus reduced by $(124 + 13 \times 28.3)/56 \approx 8.8$ MeV$= 1.4 \times 10^{-5}$ erg per nucleon. With about 10^{57} protons in a Chandrasekhar mass, this corresponds to a total energy loss of 1.4×10^{52} erg, $\sim$10 times the energy radiated by the Sun over 10^{10} yr.

Neutronization: The large densities in the core lead to a large increase in the rates of processes such as

$$e^- + p \rightarrow n + \nu_e, \tag{4.70}$$

$$e^- + {}^{56}\mathrm{Fe} \rightarrow {}^{56}\mathrm{Mn} + \nu_e, \tag{4.71}$$

$$e^- + {}^{56}\mathrm{Mn} \rightarrow {}^{56}\mathrm{Cr} + \nu_e. \tag{4.72}$$

This neutronization depletes the core of electrons, and their supporting degeneracy pressure, as well as of energy, which is carried off by the neutrinos.

The two processes lead, in principle, to an almost total loss of thermal pressure support and to an unrestrained collapse of the core of a star on a free-fall timescale. For the typical core densities prior to collapse, $\rho \sim 10^9$ g cm^{-3} (calculated from stellar evolution models), this timescale is (Eq. 3.15)

$$\tau_{\rm ff} = \left(\frac{3\pi}{32G\bar{\rho}}\right)^{1/2} \sim 0.1 \text{ s}. \tag{4.73}$$

In practice, at these high densities, the mean free path for neutrino scattering becomes of order the core radius. This slows down the energy loss, and hence the collapse time, to a few seconds.

As the collapse proceeds and the density and the temperature increase, the reaction

$$e^- + p \rightarrow n + \nu_e \tag{4.74}$$

becomes common, and is infrequently offset by the inverse process of neutron decay

$$n \rightarrow p + e^- + \bar{\nu}_e, \tag{4.75}$$

leading to an equilibrium ratio of densities of

$$n_e = n_p \approx \tfrac{1}{200} n_n. \tag{4.76}$$

Thus, most of the nucleons become neutrons, and a **neutron star** forms, in which degenerate neutrons, rather than electrons, provide the pressure support against gravity.

4.3.2 Properties of Neutron Stars

The properties of neutron stars can be estimated easily by replacing m_e with m_n in Eqs. 4.34 – 4.35, describing white dwarfs. Thus,

$$r_{\rm ns} \approx 2.3 \times 10^9 \text{ cm } \frac{m_e}{m_n}\left(\frac{\mathcal{Z}}{A}\right)^{5/3}\left(\frac{M}{M_\odot}\right)^{-1/3} \approx 14 \text{ km}\left(\frac{M}{1.4M_\odot}\right)^{-1/3}. \tag{4.77}$$

Here we have set $\mathcal{Z}/A = 1$, since the number of particles contributing to the degeneracy pressure (i.e., the neutrons) is almost equal to the total number of nucleons. Since the radius of a neutron star is about 500 times smaller than that of a white dwarf, the mean density is about 10^8 times greater, i.e., $\rho \sim 10^{14}$ g cm^{-3}. This is similar to the density of nuclear matter. In fact, one can consider a neutron star to be one huge nucleus of atomic mass number $A \sim 10^{57}$.

Our estimate of the radius is only approximate, since we have neglected two effects which are important. First, at these interparticle separations, the nuclear interactions play an important role in the equation of state, apart from the neutron degeneracy pressure. The equation of state of nuclear matter is poorly known, due to our poor understanding of the details of the strong interaction. In fact, it is hoped that actual measurements of the sizes of neutron stars will provide experimental constraints on the nuclear equation

of state, which would be important input to nuclear physics. Second, the gravitational potential energy of a test particle of mass m at the surface of a $\sim 1.4 M_\odot$ neutron star, of radius $r \sim 10$ km, is a significant fraction of the particle's rest-mass energy:

$$\frac{E_{\rm gr}}{mc^2} = \frac{GM}{rc^2} \approx \frac{6.7 \times 10^{-8}\ {\rm cgs} \times 1.4 \times 2 \times 10^{33}\ {\rm g}}{10 \times 10^5\ {\rm cm}\ (3 \times 10^{10}\ {\rm cm\ s^{-1}})^2} \approx 20\%. \tag{4.78}$$

This means that matter falling onto a neutron star loses 20% of its rest mass, and the mass of the star (as measured, e.g., via Kepler's law) is 20% smaller than the total mass that composed it. Thus, a proper treatment of the physics of neutron stars needs to be calculated within the strong-field regime of general relativity. More detailed calculations, including these two effects, give a radius of about 10 km for a $1.4 M_\odot$ neutron star.

The Chandrasekhar mass,

$$M_{\rm ch} = 0.2 \left(\frac{\mathcal{Z}}{A}\right)^2 \left(\frac{hc}{G m_p^2}\right)^{3/2} m_p, \tag{4.79}$$

can be used to estimate a maximal mass for a neutron star, beyond which the density is so high that even the degenerate neutron gas becomes ultra-relativistic and unable to support the star against gravity. Again replacing the $\mathcal{Z}/A = 0.5$, appropriate for white dwarfs, with $\mathcal{Z}/A = 1$, describing neutron stars, gives a factor of 4, or

$$M_{\rm ns,max} = 1.4 M_\odot \times 4 = 5.6 M_\odot. \tag{4.80}$$

Taking into account general relativistic effects lowers this estimate to about $5 M_\odot$. This reduction come about because, in the regime of strong gravity, the pressure itself contributes significantly to the gravitational field, and thus pressure gradually loses its effectiveness in counteracting gravitation. Detailed calculations that attempt also to take into account the strong interaction of nuclear matter further lower the maximal mass to $2–3 M_\odot$, but this is still highly uncertain.

4.3.3 Supernova Explosions

The fall of the layers of matter that surrounded the core onto the surface of a newly formed neutron star sets off a shock wave that propagates outward and blows off the outer shells of the star, in what is observed as a **supernova** explosion (see Fig. 4.8). The details of how exactly this occurs are not understood yet. In fact, sophisticated numerical simulations of the collapse are presently still unable to reproduce the "explosion," i.e., the ejection of the star's outer regions. A kinetic energy of about 3×10^{51} erg is imparted to the material flying out (as determined from measurements of the mass and velocity of ejecta in supernova remnants). About 3×10^{49} erg can be observed over a period of order one month as luminous energy, driven primarily by the decay of radioactive elements synthesized during the last few moments before collapse, during the collapse, and during the explosion. Although the luminous energy is only 1% of the kinetic energy, it nevertheless makes a supernova an impressive event; the mean luminosity is of order

Figure 4.8 Optical-light image of the supernova SN1994D, below and to the left of its host galaxy, NGC 4526. For several weeks around its peak brightness, the luminosity of a supernova is comparable to that of an entire galaxy, with $L \sim 10^{10} L_\odot$. (The spikes emerging from the supernova are due to diffraction.) Photo credit: NASA, ESA, the High-Z Supernova Search Team and the Hubble Key Project Team.

$$L_{\rm SN} \sim \frac{3 \times 10^{49}\ {\rm erg}}{30\ {\rm d} \times 24\ {\rm hr} \times 60\ {\rm m} \times 60\ {\rm s}} \sim 10^{43}\ {\rm erg\ s^{-1}} \sim 3 \times 10^9 L_\odot, \tag{4.81}$$

comparable to the luminosity of an entire galaxy of stars (see chapter 6).

However, the total gravitational binding energy released in the collapse of the core to a neutron star is

$$E_{\rm gr} \sim \frac{GM^2}{r_{\rm ns}} = 5 \times 10^{53} \left(\frac{M}{1.4 M_\odot}\right)^2 \left(\frac{r_{\rm ns}}{10\ {\rm km}}\right)^{-1}\ {\rm erg}. \tag{4.82}$$

The kinetic and radiative energies are just small fractions, $\sim 10^{-2}$ and $\sim 10^{-4}$, respectively, of this energy. The bulk of the energy released in the collapse is carried away by neutrino–antineutrino pairs. The density is so high that photons cannot emerge from the star, and they undergo frequent photon–photon collisions. These produce electron–positron pairs, which form neutrino pairs:

$$\gamma + \gamma \rightarrow e^+ + e^- \rightarrow \nu_e + \bar{\nu}_e, \nu_\mu + \bar{\nu}_\mu, \nu_\tau + \bar{\nu}_\tau \tag{4.83}$$

(the μ and τ neutrinos are neutrinos related to the muon and the tauon, which are heavy relatives of the electron.) The neutrinos can pass through the star with few scatterings (see Problem 3), and can therefore drain almost all of the thermal energy.

A striking confirmation of this picture was obtained in 1987, with the explosion of Supernova 1987A in the Large Magellanic Cloud, a satellite galaxy of our Galaxy (the Milky Way; see chapter 6), at a distance of 50 kpc from Earth. This was the nearest supernova observed since the year 1604. A total of 20 antineutrinos (several of them with directional information pointing toward the supernova) were detected simultaneously in the span of a few seconds by two different underground experiments. Each experiment consisted of a detector composed of a large tank filled with water and surrounded by photomultiplier tubes. These experiments were initially designed to search for proton decay. The experiments discovered the antineutrinos, and measured their approximate energies and directions via the reaction

$$\bar{\nu}_e + p \rightarrow n + e^+, \tag{4.84}$$

by detecting the Cerenkov radiation emitted by the positrons moving faster than the speed of light in water. The typical energies of the $\bar{\nu}_e$'s were 20 MeV. The detection of 20 particles, divided by the efficiency of the experiments to antineutrino detection (which was a function of antineutrino energy), implied that a *fluence* (i.e., a time-integrated flux) of 2×10^9 cm^{-2} electron antineutrinos had reached Earth. The electron antineutrinos, $\bar{\nu}_e$'s, are just one out of six types of particles (ν_e, $\bar{\nu}_e$, ν_μ, $\bar{\nu}_\mu$, ν_τ, $\bar{\nu}_\tau$) that are produced in similar numbers and carry off the collapse energy. Thus, the total energy released in neutrinos was

$$\begin{aligned} E_{\rm neutrino} \sim\ & 2 \times 10^9\ {\rm cm}^{-2} \times 6 \times (20\ {\rm MeV} \times 1.6 \times 10^{-6} {\rm erg\ MeV}^{-1}) \\ & \times 4\pi (5 \times 10^4\ {\rm pc} \times 3.1 \times 10^{18}\ {\rm cm\ pc}^{-1})^2 \sim 10^{53}\ {\rm erg}, \end{aligned} \tag{4.85}$$

close to the total energy expected from the collapse of a stellar core.

We note here that there is an altogether different avenue for stars to pass the Chandrasekhar limit and explode, in events that are called **type Ia supernovae**. White dwarfs that are in close binaries, where mass transfer takes place from a companion onto the white dwarf, can reach $M_{\rm ch}$. At that stage, or possibly even before actually reaching $M_{\rm ch}$, thermonuclear ignition of the carbon core occurs. However, this happens under degenerate conditions, without the thermostatic effect of a classical equation of state. With a classical equation of state, a rise in temperature produces a rise in pressure that leads to an expansion of the star, a lowering of the temperature, and a decrease in the nuclear reaction rates. Instead, under degenerate conditions, the white dwarf structure is insensitive to the rise in temperature, which raises the nuclear reaction rates more and more, ending in a thermonuclear runaway that blows up the entire star. As opposed to core-collapse supernovae, which leave a neutron star remnant (or a black hole, see below), it is believed that type Ia supernovae leave no stellar remnant (see Problem 4). It is presently unknown what kind of star is the companion of the white dwarf in the systems that are the progenitors of of type Ia supernovae. It is also possible that a type-Ia explosion is actually the result

of the merger of two white dwarfs—see Problem 6. The supernova shown in Fig. 4.8 was a type Ia event. We will return later to the physics of accretion in close binaries (section 4.6) and to the use of type Ia supernovae in cosmology (chapters 7 and 9).

Finally, there exists a class of objects even more luminous than SNe, though very transient, called **gamma-ray bursts** (GRBs). These explosions release of the order of 10^{51} erg over a period of just a few seconds. As their name implies, much of this energy is released at gamma-ray frequencies, but the rapidly fading "afterglows" of the explosions can sometimes be detected at longer wavelengths on longer timescales—minutes in X-rays, days in the optical, and weeks at radio wavelengths. A GRB occurs about once a day in the observable Universe. It is now known that at least half of these explosions occur in star-forming galaxies, i.e., galaxies that have massive young stars. This argues that some GRBs result from the core collapse of massive stars of a particular type or in a particular configuration, i.e., that they are a certain kind of core-collapse supernova. In recent years, evidence has been accumulating that actually links some GRBs to supernovae observed at the same location. The nature and mechanisms of GRBs are still widely debated. The large energy outputs, as well as indirect evidence of the existence of highly relativistic bulk motions of material, suggest that GRBs involve the formation of black holes.

The material expelled by the mass outflows from giants and by both types of supernovae—core-collapse and Ia—is essentially the only source of all "heavy" elements. Except for helium and trace amounts of the next few lightest elements, which were synthesized early in the history of the Universe (as we will see in chapter 9), all *nucleosynthesis* takes place inside stars during their various evolution stages, or during their explosions as supernovae.

4.4 Pulsars and Supernova Remnants

Many neutron stars have been identified as such in their manifestation as **pulsars**. Pulsars were first discovered with radio telescopes in the 1960s as point sources of periodic pulses of radio emission, with periods of the order of $\tau \sim 10^{-3}$ to 1 s. Today, over 1000 pulsars are known. The periods of most pulsars are observed to grow slowly with time in a very regular manner. The predictability of the pulse arrival times is comparable to that of the most accurate man-made clocks. Figure 4.9 shows a typical pulsar time series. One of the best studied pulsars, which we shall use as an example, is the Crab pulsar, at the center of the Crab nebula (see Fig. 4.10). The Crab nebula, an example of a **supernova remnant**, is an expanding cloud of gaseous fragments at the same location in the sky where a bright supernova explosion was observed and documented in the year 1054 by Chinese, Japanese, and Korean astronomers. The Crab pulsar, from which pulsations are detected at radio, optical, and X-ray wavelengths, has a pulsation period of $\tau = 33$ ms, i.e., an angular frequency

$$\omega = \frac{2\pi}{\tau} = 190 \text{ s}^{-1}. \tag{4.86}$$

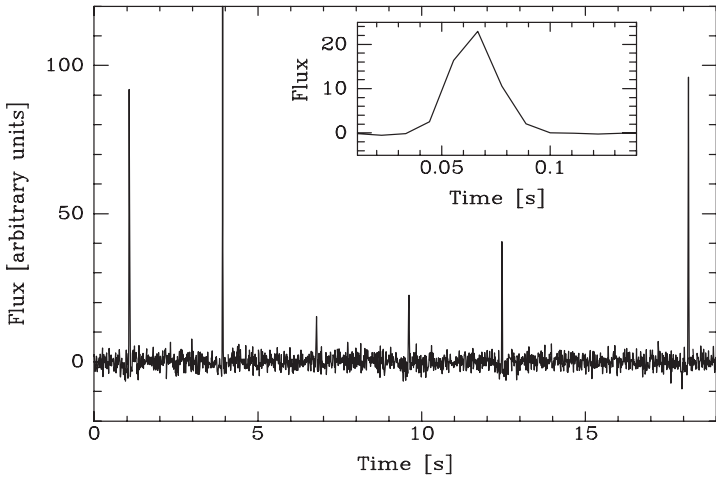

Figure 4.9 Flux at 430 MHz vs. time from PSR J0546+2441, a typical radio pulsar, over several periods. The pulse period is 2.84385038524 s (i.e., measured to 12 significant digits). Note the variable strength, and occasional disappearance of the pulses. The inset shows a zoom-in on the pulse profile, averaged over many periods. Data credit: D. Champion, see *Mon. Not. Royal. Astron. Soc.* (2005), 363, 929.

The period derivative is

$$\frac{d\tau}{dt} = \frac{1\ \text{ms}}{75\ \text{yr}} = 4.2 \times 10^{-13}, \tag{4.87}$$

or

$$\frac{d\omega}{dt} = -\frac{2\pi}{\tau^2}\frac{d\tau}{dt} = -2.4 \times 10^{-9}\ \text{s}^{-2}. \tag{4.88}$$

The total luminosity of the Crab nebula, integrated over all wavelengths, is

$$L_{\text{tot}} \approx 5 \times 10^{38}\text{erg s}^{-1}, \tag{4.89}$$

and is mostly in the form of **synchrotron radiation**, i.e., radiation emitted by relativistic electrons as they spiral along magnetic field lines.

4.4.1 Identification of Pulsars as Neutron Stars

To see that the Crab pulsar (and other pulsars) are most plausibly identified with spinning neutron stars, let us consider possible mechanisms for producing periodicity of the observed magnitude and regularity. Three options that come to mind, of astronomical phenomena associated with periodicity, are binaries, stellar pulsations, and stellar rotation. For binary orbits, the angular frequency, masses, and separation are related by Kepler's law,

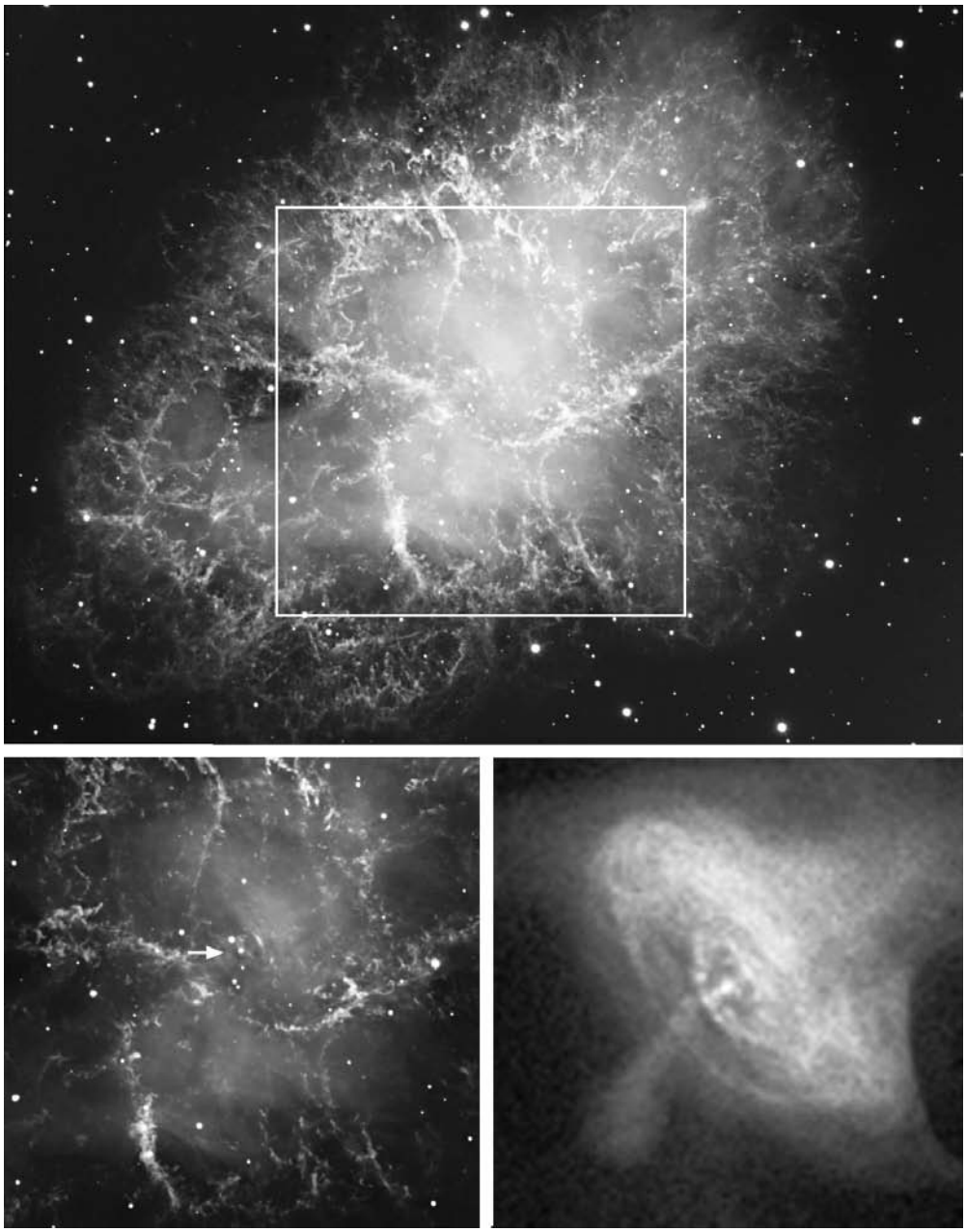

Figure 4.10 The Crab nebula, the remnant of a core-collapse supernova that exploded in the year 1054, at a distance of 2 kpc. *Top:* Image in optical light. Image scale is 4 pc on a side. *Bottom:* Zoom in on the area marked in the top photo, in optical light (left), with the pulsar at the center of the remnant indicated by the arrow; and in X-rays (right), showing the pulsar, bidirectional jets, and a toroidal structure formed by synchrotron emission from energetic particles. Note that a similar emission morphology is faintly discernible also in the optical image on the left. Photo credits: European Southern Observatory; and NASA/CXC/ASU and J. Hester et al.

$$\omega^2 = \frac{G(M_1 + M_2)}{a^3}, \tag{4.90}$$

or

$$a = \frac{[G(M_1 + M_2)]^{1/3}}{\omega^{2/3}} = \frac{[6.7 \times 10^{-8}\ \text{cgs}\ (4 \times 10^{33}\ \text{g})]^{1/3}}{(190\ \text{s}^{-1})^{2/3}} = 2 \times 10^7\ \text{cm} = 200\ \text{km}, \tag{4.91}$$

where we have assumed two solar-mass objects and inserted the Crab pulsar's frequency. The separation a is much smaller than the radii of normal stars or of white dwarfs. Only a pair of neutron stars could exist in a binary at this separation. However, general relativity predicts that two such masses orbiting at so small a separation will lose gravitational binding energy via the emission of gravitational waves (see Problems 5 and 6). This loss of energy will cause the separation between the pair to shrink, and the orbital frequency to **grow**, contrary to the observation that the pulsar frequencies decrease with time. Thus, orbital motion of stellar-mass objects cannot be the explanation for pulsars.

A second option is stellar pulsations. Stars are, in fact, observed to pulsate regularly in various modes, with the pulsation period dependent on density as[6] $\tau \propto \rho^{-1/2}$. Normal stars oscillate with periods between hours and months, and white dwarfs oscillate with periods of 100 to 1000 s. Neutron stars, which are 10^8 times denser than white dwarfs, should therefore pulsate with periods 10^4 times shorter, i.e., less than 0.1 s. However, the most common period for pulsars is about 0.8 s. There is thus no known class of stars with the density that would produce the required pulsation period.

Finally, let us assume that the rapid and very regular pulsation is produced via anisotropic emission from a rotating star. The fastest that a star can spin is at the angular frequency at which centrifugal forces do not break it apart. This limit can be found by requiring that the gravitational force on a test mass m, at the surface, be greater than the centrifugal force:

$$\frac{GMm}{r^2} > m\omega^2 r, \tag{4.92}$$

or

$$\frac{M}{r^3} > \frac{\omega^2}{G}, \tag{4.93}$$

and therefore

$$\bar{\rho} = \frac{3M}{4\pi r^3} > \frac{3\omega^2}{4\pi G} = \frac{3(190\ \text{s}^{-1})^2}{4\pi \times 6.7 \times 10^{-8}\ \text{cgs}} = 1.3 \times 10^{11}\ \text{g cm}^{-3}, \tag{4.94}$$

for the Crab pulsar. Thus, if the Crab pulsar is a spinning star and is not flying apart, its mean density must be at least five orders of magnitude larger than that of a white dwarf,

[6] It is easy to see from a dimensional argument that this must be the case for radial pulsations. Consider a star that is "squeezed" radially, and then released. The restoring force due to the pressure has dimensions of pressure times area, $F \sim PA \sim (GM\rho/r)r^2$, where we have used the equation of hydrostatic equilibrium (Eq. 3.19) to express the dimensions of the pressure. Equating this to the mass times the acceleration, $Ma \sim Mr/\tau^2$, gives the required result. Note that the pulsation period, $\tau \sim (G\rho)^{-1/2}$, is essentially the same as the free-fall timescale, Eq. 3.15.

but consistent with that of a neutron star. Note that the pulsars with the shortest periods known, of about 1 ms (rather than the Crab's 33 ms), must have mean densities 1000 times larger to avoid breaking apart, i.e., $\sim 10^{14}$ g cm^{-3}. This is just the mean density we predicted for neutron stars.

Next, let us presume that the luminosity of the Crab nebula is powered by the pulsar's loss of rotational energy as it spins down. (The observed luminosity of the pulsar itself, $\sim 10^{31}$ erg s^{-1}, is much too small to be the energy source of the extended emission.) Since

$$E_{\rm rot} = \tfrac{1}{2} I\omega^2, \tag{4.95}$$

where I is the moment of inertia,

$$L_{\rm tot} = -\frac{dE_{\rm rot}}{dt} = -I\omega\frac{d\omega}{dt}. \tag{4.96}$$

For an order-of-magnitude estimate, let us use the moment of inertia of a constant-density sphere, $I = \frac{2}{5}Mr^2$. Then

$$Mr^2 \sim -\frac{5}{2}\frac{L_{\rm tot}}{\omega d\omega/dt} = -\frac{5 \times 5 \times 10^{38}\ {\rm erg\ s^{-1}}}{2 \times 190\ {\rm s^{-1}}(-2.4 \times 10^{-9}\ {\rm s^{-2}})} = 3 \times 10^{45}\ {\rm g\ cm^2}. \tag{4.97}$$

A $1.4M_\odot$ neutron star of radius 10 km has just this value of Mr^2:

$$Mr^2 = 1.4 \times 2 \times 10^{33}\ {\rm g} \times (10^6\ {\rm cm})^2 = 2.8 \times 10^{45}\ {\rm g\ cm^2}. \tag{4.98}$$

By comparison, a normal star like the Sun has Mr^2 of order 10^9 larger than the value in Eq. 4.98. Conservation of angular momentum, $J = I\omega$, then dictates that when a stellar core of solar mass and solar radius collapses to a radius of about 10 km, it will spin up by a factor of order 10^9. The rotation period of the Sun is 25.4 days, or 2×10^6 s, which is typical of main-sequence stars. Collapse of a stellar core to neutron-star proportions is thus expected to produce an object with a spin period of order milliseconds, as observed in pulsars.

Thus, we see that if we identify pulsars as rapidly spinning stars, then their spin rate is that expected from the collapse of the cores of main-sequence stars to neutron star dimensions; their mean densities are those of neutron stars; and their loss or rotational energy accounts for the luminosity of the supernova ejecta in which they are embedded, if they have the moments of inertia of neutron stars. Finally, the location of pulsars at the sites of some historical supernovae, an explosion that is expected to accompany the formation of a neutron star (in terms of the energy released, even if the details of the explosion are not yet fully understood), leaves little doubt that pulsars are indeed neutron stars.

4.4.2 Pulsar Emission Mechanisms

The details of how pulsars produce their observed periodic emission are still a matter of active research. However, it is widely accepted that the basic phenomenon is the rotation of a neutron star having a magnetic field axis that is misaligned with respect to the star's

rotation axis by some angle θ. A spinning magnetic dipole radiates an electromagnetic luminosity

$$L = \frac{1}{6c^3} B^2 r^6 \omega^4 \sin^2\theta, \tag{4.99}$$

where B is the magnetic field on the surface of the star, at a radius r, on the magnetic pole. Solving Eq. 4.99 for B, with the observed properties of the Crab, a typical neutron-star radius, and $\sin\theta \approx 1$,

$$B \sim \frac{(6c^3 L)^{1/2}}{r^3\omega^2 \sin\theta} \sim \frac{[6(3\times10^{10}\ \mathrm{cm\ s^{-1}})^3 \times 5 \times 10^{38}\ \mathrm{erg\ s^{-1}}]^{1/2}}{(10^6\ \mathrm{cm})^3 (190\ \mathrm{s^{-1}})^2 \times 1} \sim 8\times10^{12}\ \mathrm{G}. \tag{4.100}$$

Magnetic fields of roughly such an order of magnitude are expected when the ionized (and hence highly conductive) gas in a star is compressed during the collapse of the iron core. The originally small magnetic field of the star (e.g., ~1 G in the Sun) is "frozen" into the gas. When the gas is compressed, the flux in the magnetic field lines is amplified in proportion to r^{-2}, corresponding to $\sim 10^{10}$ between the core of a main sequence star and a neutron star.

In a process that is not yet fully agreed upon, the complex interactions between magnetic and electric fields, particles, and radiation in the neighborhood of the neutron star power the nebula, and also lead to the emission of radiation in two conical beams in the direction of the magnetic axis. As the star spins and the magnetic axis precesses around the rotation axis, each beams traces an annulus of angular radius θ on the sky, as seen from the neutron star (see Fig. 4.11). Distant observers who happen to lie on the path of these "lighthouse beams" detect a pulse once every rotation, when the beam sweeps past them. This implies, of course, that we can detect only a fraction of all pulsars, namely those for which the Earth lies in the path of one of the beams.

Evidence that magnetic dipole radiation is the basic emission mechanism can be found from the age of the Crab pulsar. If such radiation is leading to the pulsar's loss of rotational energy, then, combining Eqs. 4.96 and 4.99, we find

$$\frac{dE_{\mathrm{rot}}}{dt} = I\omega\frac{d\omega}{dt} \propto \omega^4, \tag{4.101}$$

and

$$\frac{d\omega}{dt} = C\omega^3. \tag{4.102}$$

The constant C can be determined from the present values of $d\omega/dt$ and ω,

$$C = \frac{\dot{\omega}_0}{\omega_0^3}. \tag{4.103}$$

Separating variables in Eq. 4.102 and integrating, we obtain for the age of the pulsar

$$t_{\mathrm{pulsar}} = \frac{\omega_0^3}{2\dot{\omega}_0}\left(\frac{1}{\omega^2} - \frac{1}{\omega_i^2}\right), \tag{4.104}$$

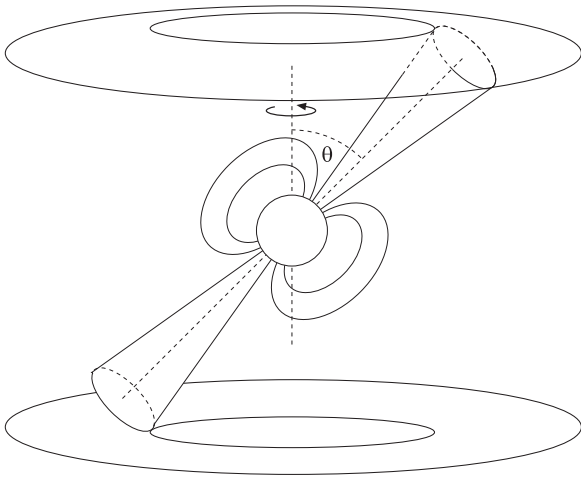

Figure 4.11 Schematic model of a pulsar. Biconical beams of radiation emerge along the magnetic axis of a neutron star. The magnetic axis is inclined by an angle θ to the star's spin axis. Observers in a direction, as viewed from the star, that is within one of the two annular regions swept out by the beams as the star rotates will detect periodic pulses.

where ω_i is the initial angular frequency of the neutron star upon formation. Thus, an upper limit on the current age of the Crab pulsar is obtained by taking $\omega = \omega_0$ and $\omega_i = \infty$,

$$t_{\text{pulsar}} < \frac{\omega_0}{2\dot{\omega}_0} = \frac{190 \text{ s}^{-1}}{2 \times 2.4 \times 10^{-9} \text{ s}^{-2}} = 4 \times 10^{10} \text{ s} = 1260 \text{ yr}. \tag{4.105}$$

This limit is consistent with the historical age, 950 yr, of the supernova of the year 1054. The pulsar age will *equal* 950 yr if we set $\tau_i = 2\pi/\omega_i = 2.5$ ms, close to the expected spin rate of newly formed neutron stars.

4.4.3 Neutron Star Cooling

As already noted, according to the above picture, we observe only a fraction of all pulsars, those for which the Earth is in the rotating pulsar beam. More significantly, pulsars slow down and lose their rotational energies, and as a result, at some point in time, will become undetectable as pulsars. However, there should exist a large population of old, spun-down, neutron stars—the remnants of all massive stars that have undergone core collapse to this state. In section 4.2.3.3, we saw that the small surface areas of white dwarfs result in very long cooling times. The surfaces of neutron stars, smaller by five orders of magnitude compared to those of white dwarfs, mean that old neutron stars will be "stuck" at temperatures of order 10^5 K, with thermal radiation peaking at photon energies of tens of electron volts (called the *extreme UV* range).

Detailed calculations of neutron star cooling are considerably more uncertain than those for white dwarfs, partly due to the poorly constrained equation of state on nuclear matter, which leads to uncertainty in the structure and composition of a neutron star. A cooling

calculation also needs to take into account many different physical processes, not all fully understood, that may play a role under the extreme conditions of gravity, temperature, density, and magnetic field inside and near the surface of a neutron star. Interstellar gas atoms falling onto a neutron-star surface also have an effect, and are likely to heat it to X-ray temperatures. To date, only several candidate isolated old neutron stars have been found in X-ray surveys. The small surface areas of neutron stars mean that their optical luminosities are very low, and hence such objects can be found only when they are near enough. X-ray surveys do reveal a large population of accreting neutron stars in binary systems, called *X-ray binaries*, which we will study in section 4.6.

4.5 Black Holes

In the case of a stellar remnant with a mass above the maximum allowed mass of a neutron star, no mechanism is known that can prevent the complete gravitational collapse of the object. In fact, general relativity predicts that even if some new form of pressure sets in at high densities, the gravitational field due to such pressure will overcome any support the pressure gradient provides, and the collapse of the star to a singularity, or **black hole** is unavoidable.

As its name implies, matter or radiation cannot escape from a black hole. An incorrect derivation, giving the correct answer, of the degree to which a mass must be compressed to become a black hole can be obtained by requiring that the escape velocity, v_e, from a spherical mass of radius r be greater than c (and hence nothing can escape),

$$\frac{GM}{r} > \frac{1}{2}v_e^2 = \frac{1}{2}c^2, \tag{4.106}$$

and therefore the **Schwarzschild radius** is

$$r_s = \frac{2GM}{c^2} = 3\ \mathrm{km}\frac{M}{M_\odot}. \tag{4.107}$$

Photons cannot escape from an object with a mass M that is concentrated within a radius smaller than r_s. The above derivation is incorrect because the kinetic energy of a photon is not $mc^2/2$, nor is the gravitational potential accurately described by the Newtonian limit, GM/r.

A correct derivation of r_s, which we shall only outline schematically, begins with the Einstein equations of general relativity,

$$G_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu}. \tag{4.108}$$

The Einstein equations relate the geometry and curvature of spacetime to the distribution of mass–energy. $T_{\mu\nu}$ is the *energy–momentum tensor*. It is represented by a 4×4 matrix, and each of its indices runs over the four spacetime coordinates. This is the "source" term in the equations and includes mass–energy density and pressure. $G_{\mu\nu}$ is the *Einstein*

tensor consisting of combinations of first and second partial derivatives, with respect to the spacetime coordinates, of the **metric**, $g_{\mu\nu}$. (A more detailed description of $T_{\mu\nu}$ and $G_{\mu\nu}$ is given in chapter 8.) The metric describes the geometry of spacetime via the *line element*

$$(ds)^2 = \sum_{\mu,\nu} g_{\mu\nu} dx_\mu dx_\nu, \tag{4.109}$$

where ds is the interval between two close spacetime events. For example, the metric (familiar from special relativity) that describes spacetime in a flat (*Euclidean*) region of space, far from any mass concentration, is the **Minkowski metric**, with a line element

$$(ds)^2 = (cdt)^2 - (dx)^2 - (dy)^2 - (dz)^2. \tag{4.110}$$

The nonzero terms of the 4×4 matrix describing $g_{\mu\nu}$ in this case are

$$g_{00} = 1, \;\; g_{11} = -1, \;\; g_{22} = -1, \;\; g_{33} = -1. \tag{4.111}$$

In spherical coordinates, the Minkowski metric has the form,

$$(ds)^2 = (cdt)^2 - (dr)^2 - (rd\theta)^2 - (r \sin\theta d\phi)^2, \tag{4.112}$$

i.e.,

$$g_{00} = 1, \;\; g_{11} = -1, \;\; g_{22} = -r^2, \;\; g_{33} = -r^2 \sin^2\theta. \tag{4.113}$$

Since $G_{\mu\nu}$ and $T_{\mu\nu}$ are symmetric 4×4 tensors (e.g., $G_{\mu\nu} = G_{\nu\mu}$), there are only 10, rather than 16, independent Einstein equations, and a zero-divergence condition on $T_{\mu\nu}$ (implying local energy conservation) further reduces this to six equations.

A solution of the Einstein equations for the geometry of spacetime in the vacuum surrounding a static, spherically symmetric, mass distribution, as viewed by an observer at infinity (i.e., very distant from the mass) is the **Schwarzschild metric**:

$$\boxed{(ds)^2 = \left(1 - \frac{2GM}{rc^2}\right)(cdt)^2 - \left(1 - \frac{2GM}{rc^2}\right)^{-1}(dr)^2 - (rd\theta)^2 - (r\sin\theta d\phi)^2,} \tag{4.114}$$

where r, θ, and ϕ are spherical coordinates centered on the mass, and t is the time measured by the distant observer. The time shown by any clock can be found from the **proper time** τ, defined as

$$d\tau \equiv \frac{ds}{c}. \tag{4.115}$$

For a clock at rest (i.e., $dr = d\theta = d\phi = 0$),

$$d\tau = \left(1 - \frac{2GM}{rc^2}\right)^{1/2} dt = \left(1 - \frac{r_s}{r}\right)^{1/2} dt. \tag{4.116}$$

Consider now a stellar remnant that is compact enough that its radius is within r_s, and hence the Schwarzschild metric (which applies only in vacuum) describes spacetime in

the vicinity of r_s. When a clock is placed at $r \to r_s$, $d\tau$ approaches zero times dt. During a time interval of, say, $dt = 1$ s, measured by a distant observer, the clock near r_s advances by much less. In other words, clocks appear (to a distant observer) to tick more and more slowly as they approach r_s, and to stop completely at r_s. This is called **gravitational time dilation**.

The electric and magnetic fields of a light wave emitted by a source near r_s will also appear to oscillate more slowly due to the time dilation, and therefore the frequency of light will decrease, and its wavelength λ will increase, relative to the wavelength λ_0 of light emitted by the same source far from the black hole. This **gravitational redshift** is

$$\frac{\lambda}{\lambda_0} = \left(1 - \frac{2GM}{rc^2}\right)^{-1/2} = \left(1 - \frac{r_s}{r}\right)^{-1/2}. \tag{4.117}$$

When the light source is at r_s, the wavelength becomes infinite and the energy of the photons, hc/λ, approaches zero.

In general relativity, once we know the metric that describes spacetime, we can find the trajectories of free-falling particles and of radiation. In particular, massless particles and light move along **null geodesics**, defined as paths along which $ds = 0$. Setting $ds = 0$ in Eq. 4.114, the *coordinate speed* of a light beam moving in the radial direction ($d\theta = d\phi = 0$) is

$$\frac{dr}{dt} = \pm c\left(1 - \frac{2GM}{rc^2}\right) = \pm c\left(1 - \frac{r_s}{r}\right). \tag{4.118}$$

At $r \gg r_s$, the speed is $\pm c$, as expected. However, as light is emitted from closer and closer to r_s, its speed appears to decline (again, to a distant observer), going to zero at r_s. Gravity works effectively as an index of refraction, with $n = \infty$ at r_s. As a result, no information can emerge from a radius smaller than r_s, which constitutes an **event horizon** around the black hole. We have thus rederived (correctly, this time) the Schwarzschild radius and its main properties.

Because of gravitational time dilation, a star collapsing to a black hole, as viewed by a distant observer, appears to shrink in progressively slower motion, and gradually appears to "freeze" as it approaches its Schwarzschild radius. In fact, it takes an infinite time for the star to cross r_s, and therefore, formally, black holes do not exist, in terms of distant static observers such as ourselves. (They certainly can exist, even in the "present" of observers who are near enough to a black hole.) However, for all practical purposes, there are no differences in observed properties between such "frozen stars" and truly collapsed black holes. This comes about because, as a source of light falls toward r_s, the rate at which photons from the source reach the observer declines as $(1 - r_s/r)^{1/2}$. Furthermore, the energy of each photon declines due to the gravitational redshift also as $(1 - r_s/r)^{1/2}$. The equation of motion for a radially free-falling light source, $r(t)$, can be roughly estimated by noting that, as the source approaches r_s, it will move with a velocity close to c, and hence its geodesic (i.e., its path in spacetime) will be close to that of the null geodesic of

photons. Let us take the negative solution in Eq. 4.118 (the source is falling to smaller radii). Separating the variables,

$$cdt = -\frac{rdr}{r - r_s}, \tag{4.119}$$

changing variables to $x = r - r_s$, and integrating gives

$$c(t - t_0) = -\int \frac{x + r_s}{x} dx = -(x + r_s \ln x) = -[r - r_s + r_s \ln (r - r_s)]. \tag{4.120}$$

As $r - r_s \to 0$, the logarithmic term becomes dominant, and we can write the equation of motion, $r(t)$, as viewed by the distant observer, as

$$r - r_s \sim \exp\left[-\frac{c(t - t_0)}{r_s}\right]. \tag{4.121}$$

Inserting this dependence into the expression for the decline in the observed photon emission rate due to gravitational time dilation, we find that

$$\frac{dN_{\rm ph}}{dt} \sim \left(1 - \frac{r_s}{r}\right)^{1/2} = \left(\frac{r - r_s}{r}\right)^{1/2} \propto \exp\left(-\frac{ct}{2r_s}\right) \tag{4.122}$$

(where in the last step we have substituted $r \sim r_s$ in the denominator, and $r - r_s$ in the numerator using Eq. 4.121).

For example, for a $5M_\odot$ stellar core undergoing its final collapse, the characteristic time is

$$\frac{2r_s}{c} = \frac{2 \times 5 \times 3 \text{ km} \times 10^5 \text{ cm km}^{-1}}{3 \times 10^{10} \text{ cm s}^{-1}} = 10^{-4} \text{ s} = 0.1 \text{ ms}. \tag{4.123}$$

Thus, after a mere 20 ms, the photon rate will decline by a factor exp(–200) $= 10^{-87}$. The photon emission rate from a Sun-like star emitting in the optical range, at a typical photon energy of $h\nu = 1$ eV, is of order

$$\frac{dN_{\rm ph}}{dt} \sim \frac{L}{h\nu} \sim \frac{3.8 \times 10^{33} \text{ erg s}^{-1}}{1 \text{ eV} \times 1.6 \times 10^{-12} \text{ erg eV}^{-1}} \sim 10^{45} \text{ s}^{-1}. \tag{4.124}$$

A factor of 10^{87} decrease in photon flux implies that, after just 20 ms, the photon emission rate from the star will decrease to $\sim 10^{-42}$ s^{-1}. The time between the emission of consecutive photons will thus be $\sim 10^{42}$ s, many orders of magnitude larger than the age of the Universe, which is of order 10^{10} yr $\sim 10^{17}$ s. The "frozen star" is truly "black," and no photons emerge from it after a timescale of milliseconds.

Theoretically, quantum mechanics allows an exception to this rule, and small amounts of so-called *Hawking radiation* can escape a black hole, even causing it to "evaporate" completely if it is small enough. However, it is unclear if black hole evaporation has any astronomical relevance.

Observationally, there are many objects considered to be stellar-mass black hole candidates, consisting of members of binary systems in which the minimum mass of one of the members is significantly larger than $3M_\odot$, yet a main-sequence or giant star of such mass

is not seen. Presumably, black holes form from the core collapse of stars with an initial mass above some threshold (which is currently thought to be about $25M_\odot$). In some of these binary systems, accretion of matter onto the black hole is taking place. Such systems are discussed in more detail in section 4.6. Finally, there is evidence for the existence of *supermassive* black holes, with masses of $\sim 10^6$–$10^9 M_\odot$, in the centers of most large galaxies. These are discussed in chapter 6.

4.6 Interacting Binaries

Until now, stars were the only luminous objects we considered. However, there exists an assortment of objects that are powered not by nuclear reactions, but by the accretion of matter onto a gravitational potential well. Objects in this category include pre-main-sequence stars, interacting binaries, active galactic nuclei and quasars, and possibly some types of supernovae and gamma-ray bursts. While all these objects are rare relative to normal stars, they are interesting and important for many physical and observational applications. The physics of accretion is similar in many of these objects. In this section, we will focus on interacting binaries, which are the best-studied accretion powered objects.

As already noted, many stars are in binary systems.[7] Pairs with an orbital period of less than about 10 days are usually in orbits that are circular, "aligned" (i.e., the spin axes of the two stars and the orbital plane axis are all parallel), and synchronized (i.e., each star completes a single rotation about its axis once per orbit, and thus each star always sees the same side of its companion star). This comes about by the action of the strong **tidal forces** that the stars exert on each other at small separations. The force per unit mass on a mass element at the surface of a star, at distance Δr from the center, due to the mass M_1 of the star itself is

$$\frac{F_{\rm grav}}{m} = \frac{GM_1}{(\Delta r)^2}. \tag{4.125}$$

The tidal force on this mass element, due to the influence of the second star of mass M_2 at a distance r (assuming $\Delta r \ll r$) is

$$\frac{F_{\rm tide}}{m} = GM_2\left(\frac{1}{r^2} - \frac{1}{(r+\Delta r)^2}\right) \approx \frac{2GM_2\Delta r}{r^3}. \tag{4.126}$$

The ratio between the forces is

$$\frac{F_{\rm tide}}{F_{\rm grav}} = \frac{2M_2}{M_1}\left(\frac{\Delta r}{r}\right)^3. \tag{4.127}$$

Thus, the larger $\Delta r/r$, the more tidal distortion of the shapes of the stars occurs, such that they become two ovals pointing at each other. As long as the stars are not tidally locked (i.e.,

[7] Current evidence is that the binary fraction among stars depends on stellar mass, with most of the massive stars being in binaries, but most low-mass stars being single. About one-half of solar-mass stars are in binaries.

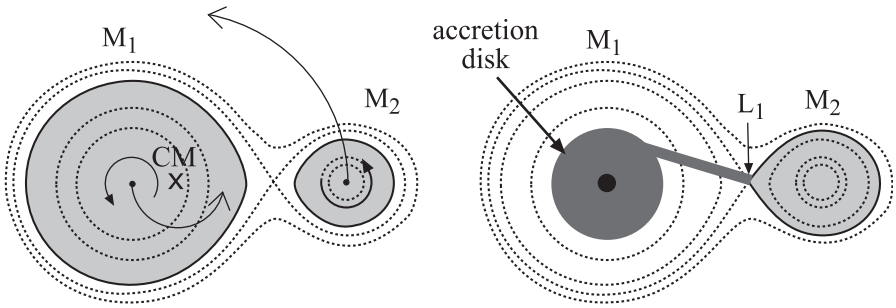

Figure 4.12 Equipotential surfaces (dotted curves) in the corotating reference frame of a binary system with mass ratio $M_1/M_2 = 5$. *Left:* In this example, both stars are inside their respective Roche lobes, but are tidally distorted. Loss of energy to tidal friction will cease only when the orbits about the center of mass become circularized, and aligned and synchronized with the rotations, so that there is no motion in the corotating frame. *Right:* Here, the secondary star, on the right, fills its Roche lobe. Matter flows through the L_1 point and falls onto the primary star, which is now a compact object. Viewed from an inertial frame, the falling material possesses angular momentum, and hence an accretion disk is formed around the compact primary star.

synchronized and circularized), energy is continuously lost to friction while the different parts of each star are deformed during the orbit. Once tidal locking is achieved, everything appears stationary in a reference frame rotating at the binary frequency, and the system achieves its minimum energy.[8]

If we draw the surfaces of constant potential energy in the rotating frame of such a binary, the isopotential surfaces close to each of the stars will be approximately spherical, but at larger radii they are more and more oval shaped, due to the gravitational pull of the companion (see Fig. 4.12). There is one particular isopotential surface for which projections onto any plane passing through the line connecting the stars traces a "figure 8", i.e., the surface is pinched into two pointed "lobes" that connect at a point between the two stars. These are called **Roche lobes** and the point where they connect is the **first Lagrange point**, L_1. At L_1, the gravitational forces due to the two stars, and the centrifugal force in the rotating frame due to rotation about the center of mass, all sum up to zero.[9]

In any star, surfaces of constant gas density and pressure will be parallel to surfaces of constant potential (which is why isolated stars are spherical). Thus, a member of a close binary that evolves and grows in radius, e.g., into a red giant, will have a shape that

[8] The same kind of tidal deformation is applied by the Sun and the Moon to the Earth, especially to the Earth's liquid water surface layer. The deformation is maximal when the three bodies are approximately aligned, during full Moon and new Moon. During one daily Earth rotation, a point on the Earth goes through two "high tide" locations and two "low tide" locations. Due to the loss of energy to tidal friction, the Earth–Moon system is by now largely circularized, but only partly synchronized. On the one hand, the Moon's orbital and rotation periods are exactly equal, and hence we always see the same ("near") side of the Moon. Although the Moon is solid, synchronization was achieved by means of the solid tidal stresses and deformations imposed on it by the Earth. The Earth's rotation, on the other hand, is not yet synchronized with either the Sun's or the Moon's orbital periods. See Problem 8 for some quantitative assessments of ocean tides.

[9] Note that L_1 is generally *not* at the center of mass. The center of mass is closer to the more massive star in the binary system, while L_1 is closer to the less massive star. Only in equal-mass binaries do the two points coincide.

is increasingly teardrop shaped. If the star inflates enough to fill its Roche lobe, stellar material at the L_1 point is no longer bound to the star, and can fall onto the companion. Three configurations are thus possible:

In a **detached binary** neither of the stars fills its Roche lobe; in a **semi-detached binary** one of the stars fills its Roche lobe; and in a **contact binary** both stars fill their Roche lobes. In the last case the binary system looks like a single, peanut-shaped object with two stellar cores and a common envelope.

In the semi-detached case there is always transfer of matter from the Roche-lobe-filling star to its companion. Different observational phenomena result, depending on the nature of the receiving star. If it is a main-sequence star, an *Algol-type* binary system results. If the receiving star is a white dwarf, the resulting phenomena are called *cataclysmic variables*, *novae*, and *type Ia supernovae*. If the receiver is a neutron star or black hole, the system is called an *X-ray binary*.

Viewed from an inertial reference frame, the accreted material possesses angular momentum having the direction of the system's orbital angular momentum. (In the rotating frame, the matter experiences a Coriolis force as it falls toward the receiving star.) If the receiving star is compact, the accreted material will not reach the surface immediately, but rather go into orbit around the star. The gas particles on different coplanar, elliptical orbits will collide with each other, and eventually an **accretion disk** forms around the receiving star.

4.6.1 Accretion Disks

In an accretion disk, particles move on approximately circular orbits, and lose energy and angular momentum due to viscous interaction with particles moving along orbits at adjacent radii. The particles therefore slowly drift to progressively smaller radii, until reaching the surface of the star (or the Schwarzschild radius, if the accretor is a black hole). The frictional heat is radiated away. Although the exact process by which viscous friction operates in accretion disks is still a matter of debate, we can nonetheless derive some general properties of these objects.

Consider a mass element, dM, in the accretion disk around a star of mass M. To fall from a circular Keplerian orbit of radius $r + dr$ to an orbit at radius r, the mass element must lose some potential energy. Half of the lost potential energy is necessarily converted to additional kinetic energy at the smaller radius with its higher Keplerian velocity, and the remaining half can be converted to heat.[10] The gain in thermal energy of the mass element will thus be

$$dE_{\text{th}} = \frac{1}{2}\left(\frac{GMdM}{r} - \frac{GMdM}{r+dr}\right), \tag{4.128}$$

[10] Note that this result, while following directly from Newtonian mechanics for a particle in a circular orbit, is just another instance of the virial theorem for a classical nonrelativistic system of particles in gravitational equilibrium—in this case a system of one particle.

where we neglect the gravitational self-binding energy of the disk itself. Assuming that the hot gas radiates its thermal energy as a blackbody at the same radius where the gravitational energy is liberated, the luminosity from an annulus in the disk will be

$$dL = \frac{dE}{dt} = \frac{1}{2} GM \frac{dM}{dt} \left(\frac{1}{r} - \frac{1}{r + dr} \right) = \frac{1}{2} GM\dot{M} \frac{dr}{r^2} = 2(2\pi r) dr \sigma T^4, \tag{4.129}$$

where $\dot{M}$ is the mass accretion rate through a particular annulus of the disk, σ is the Stefan-Boltzmann constant, and the factor of 2 on the right-hand side is because the area of the annulus includes both the "top" and the "bottom." Taking the two right-hand terms and isolating T, we find for the temperature profile of an accretion disk

$$T(r) = \left(\frac{GM\dot{M}}{8\pi\sigma} \right)^{1/4} r^{-3/4}. \tag{4.130}$$

In a steady state, $\dot{M}$ must be independent of r (otherwise material would pile up in the disk, or there would be a shortage of material at small radii), and must equal the accretion rate of mass reaching the stellar surface. Thus, $T \propto r^{-3/4}$, meaning that the inner regions of the disk are the hottest ones, and it is from them that most of the luminosity emerges. The total luminosity of an accretion disk with inner and outer radii $r_{\rm in}$ and $r_{\rm out}$ is found by integrating over the luminosity from all annuli,

$$L = \int_{r_{\rm in}}^{r_{\rm out}} 2(2\pi r) \sigma T^4(r) dr = \frac{1}{2} GM\dot{M} \left(\frac{1}{r_{\rm in}} - \frac{1}{r_{\rm out}} \right). \tag{4.131}$$

This result could have, of course, been obtained directly from conservation of energy.[11] If $r_{\rm out} \gg r_{\rm in}$, the result simplifies further to

$$L = \frac{1}{2} \frac{GM\dot{M}}{r_{\rm in}}. \tag{4.132}$$

It is instructive to evaluate the **radiative efficiency** of accretion disks by dividing the luminosity above by $\dot{M}c^2$, the hypothetical power that would be obtained if all the accreted rest mass were converted to energy:

$$\eta = \frac{1}{2} \frac{GM}{c^2 r_{\rm in}}. \tag{4.133}$$

If the accreting object is, e.g., a $1.4 M_\odot$ neutron star with an accretion disk reaching down to the stellar surface at a radius of 10 km, then $r_{\rm in}$ is about 2.5 times the Schwarzschild radius, $r_s = 2GM/c^2$ (Eq. 4.107), that corresponds to such a mass (recall that $r_s \approx 3$ km for

[11] Note that, in addition to energy conservation, a full treatment of accretion disk structure must also conserve angular momentum. The angular momentum per unit mass of a disk particle at radius r, in a circular Keplerian orbit with velocity v_c, is $J/m = r v_c = \sqrt{GMr}$. Thus, a particle descending to an orbit at smaller r must get rid of angular momentum by transfering it to other particles in the disk. Those particles gain angular momentum, and hence move to larger radii. Some of the gravitational energy released by the inflow will power this outflow of matter, at the expense of the energy that can be radiated by the disk.

$1M_\odot$). The rest-mass-to-radiative energy conversion efficiency is then about 0.10. For black-hole accretors, it turns out from solution of the general relativity equations of motion that gas particles have a *last stable orbit* at which they can populate the accretion disk. At smaller radii, a particle quickly spirals in and crosses the event horizon, carrying its remaining kinetic energy with it. The last stable orbit for a nonrotating black hole[12] is at $3r_s$. Accretion disks around such black holes will therefore have an efficiency of $1/12 \approx 0.08$, somewhat lower than accretion disks around neutron stars. (A solution of the problem using the correct general relativistic, rather than Newtonian, potential, gives an efficiency of 0.057). The point to note, however, is that, in either case, the efficiency is an order of magnitude higher than the efficiency of the nuclear reactions operating in stars, $\eta = 0.007$ or less. Furthermore, only a tiny fraction of a main-sequence star's mass is involved at any given time in nuclear reactions, whereas an accretion disk can extract energy with high efficiency from all of the mass being channeled through it. Under appropriate conditions, accretion disks can therefore produce high luminosities.

Let us calculate the typical luminosities and temperatures of accretion disks in various situations. In **cataclysmic variables**, the accretor is a white dwarf, with a typical mass of $1M_\odot$ and a radius of 10^4 km. A typical accretion rate[13] is $10^{-9}M_\odot$ yr^{-1}. This produces a luminosity of

$$\begin{aligned} L &= \frac{1}{2}\frac{GM\dot{M}}{r_{\rm in}} = \frac{6.7\times10^{-8}\ \text{cgs}\times 2\times10^{33}\ \text{g}\times10^{-9}\times 2\times10^{33}\ \text{g}}{2\times 3.15\times10^{7}\ \text{s}\times10^{9}\ \text{cm}} \\ &= 4\times10^{33}\ \text{erg s}^{-1} \approx L_\odot. \end{aligned} \tag{4.134}$$

The luminosity from the accretion disk thus completely overpowers the luminosity of the white dwarf. The disk luminosity can be much greater than that of the donor star (for low-mass main-sequence donors, the most common case), comparable to the donor star (for intermediate-mass main sequence stars) or much smaller than the donor luminosity (for high-mass main sequence and red-giant donors). At the inner radius (which dominates the luminosity from the disk) the temperature is (Eq. 4.130)

$$\begin{aligned} T(r) &= \left(\frac{GM\dot{M}}{8\pi\sigma}\right)^{1/4} r^{-3/4} \\ &= \left(\frac{6.7\times10^{-8}\ \text{cgs}\times 2\times10^{33}\ \text{g}\times10^{-9}\times 2\times10^{33}\ \text{g}}{3.15\times10^{7}\ \text{s}\times 8\pi\times 5.7\times10^{-5}\ \text{cgs}}\right)^{1/4} (10^9\ \text{cm})^{-3/4} \\ &= 5\times10^{4}\ \text{K}. \end{aligned} \tag{4.135}$$

[12] A black hole is fully characterized by only three parameters—its mass, its spin angular momentum, and its electric charge (the latter probably not being of astrophysical relevance, because astronomical bodies are expected to be almost completely neutral). Spacetime around a rotating black hole is described by a metric called the Kerr metric, rather than by the Schwarzschild metric. Black-hole spin is accompanied by the general relativistic phenomenon of "frame dragging," in which spacetime outside the event horizon rotates with the black hole. In a rotating black hole, the last stable orbit and the event horizon are at smaller radii than in the nonrotating case.

[13] The accretion rate can be limited by the rate at which the donor star transfers mass through the L_1 point, by the efficiency of the viscous process that causes material in the accretion disk to fall to smaller radii, or by the radiation pressure of the luminosity resulting from the accretion process—see section 4.6.2.

The thermal spectrum from the disk therefore peaks in the far UV part of the spectrum, and is usually different from the spectrum of the main-sequence or red-giant donor star (which of course generally has a red spectrum). The integrated spectrum of the system will therefore have at least two distinct components.

When the orbits of cataclysmic variables are sufficiently inclined to our line of sight, monitoring the total light output over time, as the systems rotate, reveals changes due to mutual eclipses by the various components: the donor star, the accretion disk, and sometimes a "hot spot" where the stream of matter from the donor hits the disk. The changing projected area of the distorted donor star also affects the light output. Analysis of such data allows reconstructing the configurations and parameters of these systems. In addition to the periodic variability induced by eclipses and changes in orientation, accreting systems reveal also aperiodic variability, i.e., variations with a "noise-like" character. These variations likely arise from an unstable flow of the material overflowing the donor's Roche lobe, causing changes in $\dot{M}$, as well as from instabilities and flares in the accretion disk itself.

In a class of cataclysmic variable called **novae** there are also outbursts of luminosity during which the system brightens dramatically for about a month. The outbursts occur once every 10–10^5 yr, as a result of rapid thermonuclear burning of the hydrogen-rich (and hence potentially explosive) accreted material that has accumulated on the surface of the white dwarf. Assuming again an accretion rate of $10^{-9} M_\odot$ yr^{-1}, over a period of 1000 yr, a mass of $10^{-6} M_\odot$ will cover the surface of the white dwarf. If completely ignited, it yields an energy

$$E_{\rm nova} = 0.007mc^2 = 0.007 \times 10^{-6} \times 2 \times 10^{33}\ {\rm g} \times (3 \times 10^{10}\ {\rm cm\ s^{-1}})^2 \approx 10^{46}\ {\rm erg}. \quad (4.136)$$

When divided by a month (2.5×10^6 s), this gives a mean luminosity of 4×10^{39} erg s^{-1} = $10^6 L_\odot$, i.e., 10^6 times the normal luminosity of the accretion disk. In reality, only partial processing of the accreted hydrogen takes place, and the energy is also partly consumed in unbinding some material from the underlying white dwarf. On the other hand, for longer recurrence times between outbursts, the mass of accumulated hydrogen can be larger than assumed above. The gamma-ray spectra of novae reveal emission from the radioactive decay of elements that are synthesized in these explosions, providing direct evidence of the process at hand.

As discussed in section 4.3.3, under certain conditions (likely involving the reaching of the Chandrasekhar mass by the accreting white dwarf) an extreme, runaway version of the nova eruption, called a **type Ia supernova**, occurs. In such an event, a large fraction of the white dwarf mass (i.e., of order $1 M_\odot$ of carbon, rather than the $10^{-6} M_\odot$ of hydrogen in the nova case) is ignited and is explosively synthesized into iron-group elements. The total energy is, correspondingly, 10^6 times larger than that of a nova, i.e., 10^{51-52} erg. As in the core-collapse supernova explosions that end the life of massive stars, the ratio of kinetic to luminous energy is about 100, and thus type Ia supernovae, with a luminosity of about $10^{10} L_\odot$, can outshine their host galaxies for a period of about a month (see Problem 4). Although core-collape supernovae and type Ia supernovae have similar luminous and kinetic energy outputs, one should remember that in core-collapse supernovae

99% of the energy is carried away by neutrinos, and therefore core-collapse supernovae are intrinsically far more energetic events. Type Ia supernovae have a narrow range of observed optical luminosities, probably as a result of the fact that they generally involve the explosion of about $1.4M_\odot$ of white dwarf material. These supernovae are therefore very useful as "standard candles" for measuring distances. In chapters 7 and 9 we will see how they have been used in this application.

When the receiving star in an interacting binary is a neutron star or a black hole, the inner radius of the accretion disk is of order 10 km, rather than 10^4 km, and therefore the luminosity is much greater than in a white-dwarf accretor. For example, scaling from Eq. 4.134, if the accretor is a $1.4M_\odot$ neutron star with the same accretion rate, the accretion-disk luminosity is of order 10^{37} erg s^{-1}. The temperature at the inner radius, scaling as $M^{1/4}r^{-3/4}$ (Eq. 4.135), is $T = 10^7$ K. The emission therefore peaks in the X-rays, and hence the name **X-ray binaries.** In reality, due to the extreme matter and radiation densities, temperatures, and magnetic fields near the surface of a neutron star, the accretion disk may not actually reach the surface, and accreting material is sometimes channeled to the poles, forming a hot-spot where it hits the surface. In addition to the thermal emission from the accretion disk, other, nonthermal, radiation components are observed in such systems, e.g., synchrotron emission from relativistic electrons spiraling along magnetic field lines. Some accreting white dwarfs also possess strong magnetic fields that funnel the accretion flow directly onto hot spots on the white dwarf. Such *magnetic cataclysmic variables* also appear then as X-ray sources.

4.6.2 Accretion Rate and Eddington Luminosity

The above discussion shows that the properties of accreting systems are largely determined by three parameters, M, $\dot{M}$, and $r_{\rm in}$. M and $r_{\rm in}$ are limited to particular values by the properties of stars and stellar remnants. However, the accretion rate, $\dot{M}$, also cannot assume arbitrarily large values. To see this, consider an electron at a radius r in an ionized gas that is taking part in an accretion flow toward some compact object of mass M. The accretion flow produces a luminosity per frequency interval L_ν, and therefore the *density* of photons with energy $h\nu$ at r is

$$n_{\rm ph} = \frac{L_\nu}{4\pi r^2 ch\nu}. \tag{4.137}$$

The rate at which photons of this energy are scattered via Thomson scattering on the electron is

$$R_{\rm scat} = n_{\rm ph}\sigma_T c, \tag{4.138}$$

where σ_T is the Thomson scattering cross section. Each scattering event transfers, on average, a momentum $p = h\nu/c$ to the electron. The rate of momentum transfer to the electron, i.e., the force exerted on it by the radiation, is then

$$\frac{dp}{dt} = R_{\rm scat}\frac{h\nu}{c} = \frac{L_\nu\sigma_T}{4\pi r^2 c}. \tag{4.139}$$

The total radiative force on the electron is obtained by integrating over all frequencies ν,

$$F_{\rm rad} = \frac{L\sigma_T}{4\pi r^2 c}. \tag{4.140}$$

The electron would be repelled from the accreting source of luminosity, were it not for the gravitational attraction of the accreting object. This force will be much greater on protons than on electrons. However, the Coulomb attraction between electrons and protons prevents their separation, and therefore the gravitational attraction on a proton effectively operates on neighboring electrons as well. The attractive force on the electron is therefore

$$F_{\rm grav} = \frac{GMm_p}{r^2}. \tag{4.141}$$

The accretion flow, and its resulting luminosity, can proceed only if the radiative force does not halt the inward flow of matter, i.e., $F_{\rm rad} < F_{\rm grav}$. Equating the two forces, using Eqs. 4.140 and 4.141, we obtain the maximum luminosity possible in a system powered by accretion,

$$\boxed{L_E = \frac{4\pi cGMm_p}{\sigma_T}} \tag{4.142}$$

$$= \frac{4\pi \times 3 \times 10^{10} \times 6.7 \times 10^{-8}\ {\rm cgs} \times 2 \times 10^{33}\ {\rm g} \times 1.7 \times 10^{-24}\ {\rm g}}{6.7 \times 10^{-25}\ {\rm cm}^2} \frac{M}{M_\odot}$$

$$= 1.3 \times 10^{38}\ {\rm erg\ s^{-1}} \frac{M}{M_\odot} = 6.5 \times 10^4\ L_\odot \frac{M}{M_\odot}.$$

This limiting luminosity is called the **Eddington luminosity**.

Recalling our derivation, above, of a luminosity of order 10^{37} erg s^{-1} from an accretion disk around a $1.4M_\odot$ neutron star with an accretion rate $\dot{M} = 10^{-9}M_\odot$ yr^{-1}, we see that an accretion rate, say, 100 times larger would bring the system to a luminosity of several times L_E, and is therefore impossible. This is not strictly true, since in the derivation of L_E we have assumed spherical accretion and an isotropically radiating source. Both assumptions fail in an accretion disk, which takes in matter along an equatorial plane, and radiates preferentially in directions perpendicular to that plane. Nevertheless, detailed models of accretion disk structure show that disks become unstable when radiating at luminosities approaching L_E. The Eddington limit is therefore a useful benchmark even for nonspherical accreting systems. Finally, note that L_E applies to systems undergoing steady-state accretion. Objects of a given mass can have higher luminosities (see, e.g., the luminosities of novae and supernovae that we calculated above), but then an outflow of material is unavoidable, the object is disrupted, and the large luminosity must be transient.

4.6.3 Evolution of Interacting Binary Systems

The transfer of mass between members of interacting binaries can have drastic effects on both members. We recall that isolated neutron stars power their pulsar emission and

their surrounding supernova remnant emission at the expense of their rotational energy, and thus gradually slow down. A neutron star in a binary system, if accreting matter from its companion, under suitable conditions can gain angular momentum, which can spin the pulsar back up. Thus, many pulsars in binary systems are spinning at millisecond frequencies, i.e., close to the maximal spin possible for a neutron star, and have negative period derivatives, $\dot{P}$ (if they are still being spun up by the accretion; see Problem 9). The neutron star can also affect the donor star. The jets and beams present in pulsars may hit one side of the donor star (the binaries are tidally locked), heat it, ablate it, or completely destroy it. Several examples of such *black-widow pulsars* are known, in which an old millisecond pulsar has no companion, or in which the companion is a white dwarf of much too small a mass to have evolved in isolation from the main sequence (i.e., white dwarfs of such mass form after a time that is much greater than the age of the Universe).

The transfer of mass and angular momentum in an interacting binary can also lead to complex evolution of the parameters of the system, such as binary separation and accretion rate. Changes in those parameters can then affect the future evolution of the system. Let us see how this works. The orbital angular momentum of a circular binary composed of masses M_1 and M_2 with separation a is

$$J = I\omega = \mu a^2 \omega, \tag{4.143}$$

where I is the moment of inertia, and μ is the reduced mass,

$$\mu = \frac{M_1 M_2}{M_1 + M_2}. \tag{4.144}$$

(For simplicity, we will ignore the spin angular momentum of the stars.) Substituting ω from Kepler's law (Eq. 2.35),

$$\omega^2 = \frac{G(M_1 + M_2)}{a^3}, \tag{4.145}$$

we get

$$J = \mu\sqrt{G(M_1 + M_2)a}. \tag{4.146}$$

Assuming conservation of total mass and angular momentum, the time derivative of J equals zero,

$$\frac{dJ}{dt} = \sqrt{G(M_1 + M_2)}\left(\frac{d\mu}{dt}\sqrt{a} + \frac{\mu}{2\sqrt{a}}\frac{da}{dt}\right) = 0, \tag{4.147}$$

or

$$-\frac{2}{\mu}\frac{d\mu}{dt} = \frac{1}{a}\frac{da}{dt}. \tag{4.148}$$

Expressing $\dot{\mu}$ in terms of its constituent masses gives

$$\frac{d\mu}{dt} = \frac{1}{M_1 + M_2}\left(\frac{dM_1}{dt}M_2 + M_1\frac{dM_2}{dt}\right). \tag{4.149}$$

However, conservation of mass means that $\dot{M}_1 = -\dot{M}_2$, and hence

$$\frac{d\mu}{dt} = \frac{\dot{M}_1}{M_1 + M_2}(M_2 - M_1). \tag{4.150}$$

Replacing in Eq. 4.148, we finally get

$$2\dot{M}_1\frac{M_1 - M_2}{M_1 M_2} = \frac{1}{a}\frac{da}{dt}. \tag{4.151}$$

Equation 4.151 determines how the period and separation of the system evolve, depending on the constituent masses, the accretion rate, and its sign. For example, consider a system that starts out with two close main sequence stars, with $M_1 > M_2$. M_1 will therefore be the first to become a red giant, fill its Roche lobe, and transfer mass to M_2. Since M_1 loses mass, $\dot{M}_1$ is negative. From Eq. 4.151, $\dot{a}$ is then negative. In other words, the two stars approach each other. The decrease in separation a means that the Roche lobe around M_1 moves to a smaller radius, and the accretion rate grows further. If this trend is not interrupted (e.g., by the end of the giant stage of M_1), the system reaches a common envelope stage. Evolution resumes once M_1 becomes a white dwarf, or at a later stage, when M_2 becomes a red giant, if it fills its Roche lobe. Accretion will now be in the opposite sense, and $\dot{M}_1$ is therefore positive. If, despite the earlier accretion phase and the individual stellar evolution, M_1 is still larger than M_2, then $\dot{a}$ will now be positive. If the Roche lobe size of M_2 overtakes the star's radius, accretion will stop. Alternatively, if by this time $M_2 > M_1$, the two stars will again approach each other and there may be a second common envelope phase. Obviously, there are many other possible evolution paths, depending on the initial parameters. Moreover, in reality stars lose mass throughout their evolution by means of winds, and therefore the total mass and angular momentum of a binary system will generally not be conserved, opening further binary evolution paths.

Problems

1. In a fully degenerate gas, all the particles have energies lower than the Fermi energy. For such a gas we found (Eq. 4.19) the relation between the density n_e and the Fermi momentum p_f:

$$n_e = \frac{8\pi}{3h^3}p_f^3.$$

   a. For a nonrelativistic electron gas, use the relation $p_f = \sqrt{2m_e E_f}$ between the Fermi momentum, the electrom mass m_e, and the Fermi energy E_f, to express E_f in terms of n_e and m_e.

b. Estimate a characteristic n_e under typical conditions inside a white dwarf. Using the result of (a), and assuming a temperature $T = 10^7$ K, evaluate numerically the ratio $E_{\rm th}/E_f$, where $E_{\rm th}$ is the characteristic thermal energy of an electron in a gas of temperature T, to see that the electrons inside a white dwarf are indeed degenerate.

2. Cold, planetary-mass objects such as Jupiter are mostly devoid of internal thermal energy sources, as is the case of white dwarfs. However, planets are supported against gravity by repulsive atomic electrostatic forces rather than by free electron degeneracy pressure. Estimate the maximum mass that can be supported by atomic electrostatic forces, as follows.

a. Approximate the typical pressure inside a planet by means of the electrostatic Coulomb energy density due to each atom's repulsion of its adjacent atoms. Ignore the effect of nonadjacent atoms, whose charges are screened. Assume a pure atomic-hydrogen composition. Assume further that the atoms are distributed on a static grid of constant density with separations r, and hence there are six neighboring atoms surrounding each atom, with the centers of their electron clouds separated by $\sim r$ from the center of the electron cloud of the central atom.

b. Express the planet radius in terms of the planet mass M, the hydrogen atom mass, m_H, and the "rigid" interatomic spacing r, and then write the gravitational binding energy density of the planet in terms of these parameters.

c. Equate the electrostatic energy density you found to the gravitational binding energy density. The interatomic spacing r should cancel out from the equation (why?). Find the mass at which this equality occurs, and compare to Jupiter's mass, $M_J \approx 0.001 M_\odot$.

Answer: $8M_J$. For larger masses, the gravitational energy density will overcome the atomic electrostatic repulsion, the planet radius will stop growing with mass as fast as $M^{1/3}$, the density will increase, and quantum degeneracy pressure of the electrons will set in as the main source of pressure. From there on, the planet's radius will decrease as its mass increases, as $M^{-1/3}$ (Eq. 4.34).

3. Most of the energy released in the collapse of a massive star to a neutron star (a core-collapse supernova) is in the form of neutrinos.

a. If the just-formed neutron star has a mass $M = 1.4M_\odot$ and a radius $R = 10$ km, estimate the mean nucleon density, in cm^{-3}. Find the mean free path, in cm, of a neutrino inside the neutron star, assuming the density you found and a cross section for scattering of neutrinos on neutrons of $\sigma_{\nu n} = 10^{-42}$ cm^2.

b. How many seconds does it take a typical neutrino to emerge from the neutron star in a random walk?

Hint: Neutrinos travel at a velocity close to c. Recall that the radial distance d covered in a random walk of N steps, each of length l, is $d = \sqrt{N}\, l$.

c. Twelve electron antineutrinos from Supernova 1987A were detected by the Kamiokande neutrino detector in Japan. This experiment consisted of a tank filled with 3 kton of water, and surrounded by photomultiplier tubes. The photomultipliers

detect the Cerenkov radiation emitted by a recoiling positron that is emitted after a proton absorbs an antineutrino from the supernova. Estimate how many people on Earth could have perceived a flash of light, due to the Cerenkov radiation produced by the same process, when an antineutrino from the supernova traveled through their eyeball. Assume that eyeballs are composed primarily of water, each weighs about 10 g, and that the Earth's population was 5 billion in 1987.

4. Type Ia supernovae are probably thermonuclear explosions of accreting white dwarfs that have approached or reached the Chandrasekhar limit.

a. Use the virial theorem to obtain an expression for the mean pressure inside a white dwarf of mass M and radius R.

b. Use the result of (a) to estimate, to an order of magnitude, the speed of sound, $v_s = \sqrt{dP/d\rho} \sim \sqrt{P/\rho}$, inside a white dwarf. In an accreting white dwarf with a carbon core that has reached nuclear ignition temperature, a nuclear burning "flame" encompasses the star at the sound velocity or faster. Within how much time, in seconds, does the flame traverse the radius of the white dwarf, assuming $R = 10^4$ km, $M = 1.4M_\odot$? Note that this **sound-crossing timescale** is $\sim(G\rho)^{-1/2}$, which is also the free-fall timescale (Eq. 3.15.)

c. Calculate the total energy output, in ergs, of the explosion, assuming that the entire mass of the white dwarf is synthesized from carbon to nickel, with a mass-to-energy conversion efficiency of 0.1%. Compare this energy to the gravitational binding energy of the white dwarf, to demonstrate that the white dwarf explodes completely, without leaving any remnant.

d. Gamma rays from the radioactive decays $^{56}\mathrm{Ni} \rightarrow {}^{56}\mathrm{Co} + \gamma \rightarrow {}^{56}\mathrm{Fe} + \gamma$ drive most of the optical luminosity of the supernova. The atomic weights of ^{56}Ni and ^{56}Fe are 55.942135 and 55.934941, respectively. Calculate the total energy radiated in the optical range during the event. Given that the characteristic times for the two radioactive decay processes are 8.8 days and 111 days, respectively, show that the typical luminosity is $\sim 10^{10} L_\odot$.

5. General relativity predicts that accelerated masses radiate **gravitational waves**, thereby losing energy, in analogy to the emission of electromagnetic radiation by accelerated charges. There is indirect evidence for the existence of such waves from the orbital time evolution of some binary pulsars. If gravitational radiation were also responsible for the loss of rotational energy E_{rot} of isolated pulsars (e.g., the Crab pulsar), then a dependence

$$\frac{dE_{\mathrm{rot}}}{dt} \propto \omega^6$$

would be expected, where ω is the angular rotation velocity.

a. Under the above assumption, find an expression for $\omega(t)$.

b. For the time dependence found in (a), derive an upper limit for the age of the Crab pulsar. Given that the supernova that marked the Crab's formation occurred in the year 1054, is gravitational radiation a viable braking mechanism for the Crab pulsar?

6. A type Ia supernova is thought to be the thermonuclear explosion of an accreting white dwarf that goes over the Chandrasekhar limit (see Problem 4). An alternative scenario, however, is that supernova Ia progenitors are white dwarf binaries that lose orbital energy to gravitational waves (see Problem 5) until they merge, and thus exceed the Chandrasekhar mass and explode.

a. Show that the orbital kinetic energy of an equal-mass binary with separation a and individual masses M is

$$E_{\rm k} = \frac{GM^2}{2a},$$

and the total orbital energy (kinetic plus gravitational) is minus this amount.

b. The power lost to gravitational radiation by such a system is

$$\dot{E}_{\rm gw} = -\frac{2c^5}{5G}\left(\frac{2GM}{c^2 a}\right)^5.$$

By equating to the time derivative of the total energy found in (a), obtain a differential equation for $a(t)$, and solve it.

c. What is the maximum initial separation that a white-dwarf binary can have, if the components are to merge within 10 Gyr? Assume the white dwarfs have $1M_\odot$ each, and the merger occurs when $a = 0$.
Answer: 0.016 AU.

7. A star of mass m and radius r approaches a black hole of mass M to within a distance $d \gg r$.

a. Using Eq. 4.127, express, in terms of m, r, and M, the distance d at which the Newtonian radial tidal force exerted by the black hole on the star equals the gravitational binding force of the star, and hence the star will be torn apart.

b. Find the black-hole mass M above which the tidal disruption distance, d, is smaller than the Schwarzschild radius of the black hole, and evaluate it for a star with $m = M_\odot$ and $r = r_\odot$. Black holes with masses above this value can swallow Sun-like stars whole, without first tidally shredding them.
Answer: $10^8 M_\odot$.

c. Derive a Newtonian expression for the **tangential** tidal force exerted inward on the star, in terms of m, r, M, and d, again under the approximation $r \ll d$. The combined effects of the radial tidal force in (a) and and the tangential tidal force in (c) will lead to "spaghettification" of stars, or other objects that approach the black hole to within the disruption distance.
Hint: Remember that the star is in a *radial* gravitational field, and hence there is a tangential component to the gravitational force exerted on regions of the star that are off the axis defined by the black hole and the center of the star. The tangential component can be found by noting that the small angle between the axis and the edge of the star is $\approx r/d$.

8. Practitioners of some schools of yoga are warned not to perform yoga during the full or the new Moon, citing the tidal effect of the Moon at those times on other "watery bodies" such as the oceans. Let us investigate this idea.

a. Verify the dramatic tidal effect of the Moon on the oceans by using Eq. 4.127 to calculate the ratio of the tidal "lifting force" $F_{\rm tide}$ and the Earth's gravitational attraction $F_{\rm grav}$, for a point mass on the surface of the Earth and on the Earth–Moon line. Use this ratio to estimate the change in water height, in cm, between high tide and low tide, due to the moon alone. Repeat the calculation for the tidal effect due to the Sun alone.

Hint: The surface of the oceans traces an equipotential surface, $gR =$ constant, where R is the distance from the Earth's center to the ocean surface at every point, and g is the effective gravitational acceleration at every point on the surface. Translate the tidal-to-gravitational force ratio into a relative change in g between a point at high tide (which experiences the full tidal force) and a point at low tide (which experiences no tidal force), and thus derive the relative change in R.

Answers: 77 cm due to Moon, 33 cm due to Sun.

Note: While the Sun and Moon are the drivers of ocean tides, a reliable calculation of tides at a particular Earth location must take into account additional factors, including the varying distances between Earth, Moon, and Sun (due to their elliptical orbits), their inclined orbital planes, the latitude, coastline shape, beach profile, ocean depth, water viscosity and salinity, and prevailing ocean currents.

b. Calculate by how much (in milligrams) you are lighter when the full or new Moon is overhead or below, compared to when it is rising or setting, assuming your body weight is 50 kg.

c. Calculate in dynes and in gram-force (i.e., in dynes divided by the gravitational acceleration $g = 980$ cm s^{-2}) the tidal **stretch** exerted on your body by the Moon plus the Sun when you are standing up with the full or new moon overhead. Assume your body weight is 50 kg and your height is 180 cm.

9. A spinning neutron star of mass $M = 1.4\ M_\odot$, constant density, and radius $R = 10$ km has a period $P = 1$ s. The neutron star is accreting mass from a binary companion through an accretion disk, at a rate of $\dot{M} = 10^{-9}\ M_\odot$ yr^{-1}. Assume that the accreted matter is in a circular Keplerian orbit around the neutron star until just before it hits the surface, and once it does then all of the matter's angular momentum is transferred onto the neutron star.

a. Derive a differential equation for $\dot{P}$, the rate at which the neutron-star period decreases.

b. Solve the equation to find how long will it take to reach $P = 1$ ms, the maximal spin rate of a neutron star.

Hint: Calculate the Keplerian velocity of the accreted material a moment before it hits the neutron star surface, and use it to derive the angular momentum per unit mass of this material, J/m. The angular momentum of a rotating object with moment of

inertia I is $I\omega$. The rate of change of the star's angular momentum is just the rate at which it receives angular momentum from the accreted matter, i.e.,

$$\frac{d}{dt}(I\omega) = \dot{M}\frac{J}{m}.$$

The moment of inertia of a constant-density sphere is $I = \frac{2}{5}MR^2$. Solve for the angular acceleration $\dot{\omega}$, neglecting changes in the neutron star's mass and radius. (This will be justified by the result.) From the relation $P = 2\pi/\omega$, derive $\dot{P}$. This "spin-up" process explains the properties of old, "millisecond pulsars," some of which, indeed, have negative $\dot{P}$.

Answer: 2.6×10^8 yr. Over this time, the neutron star mass increases by 18%, and its radius decreases by 5%, justifying the approximation of constant mass and radius.

10. A compact accreting object of mass M is radiating at the Eddington luminosity (Eq. 4.142) corresponding to that mass,

$$L_E = \frac{4\pi cGMm_p}{\sigma_T} = 1.3 \times 10^{38}\ \mathrm{erg\ s^{-1}}\frac{M}{M_\odot}.$$

An astronaut wearing a white space suit is placed at rest at an arbitrary distance from the compact object. Assuming that the cross-sectional area of the astronaut's body is $A = 1.5\ \mathrm{m}^2$, find the maximum allowed mass m of the astronaut, in kg, if the radiation pressure is to support her from falling onto the compact object.

Hint: By definition, a proton at any distance from this object will float, its gravitational attraction to the object balanced by the radiation pressure on nearby electrons. Consider the astronaut as a particle with mass m and cross section equal to her geometrical cross section, $2A$ (the factor of 2 is because her suit is white, so every photon reflection transfers twice the photon momentum). Compare m to m_p and $2A$ to σ_T.

Answer: 77 kg.