Astronomy 112: The Physics of Stars

Class 1 Notes: Observing Stars

Although this course will be much less oriented toward observations than most astronomy courses, we must always begin a study of any topic by asking what observations tell us.

With the naked eye under optimal conditions, one can distinguish $\sim 6,000$ individual stars from Earth, and in 1610 Galileo published the first telescopic observations showing that the Milky Way consists of numerous stars.

[Slide 1 – Galileo telescope image]

While these early observations are of course important, in order to study stars systematically we must be able to make quantitative measurements of their properties. Only quantitative measurements can form the nucleus of a theoretical understanding and against which model predictions can be tested.

In this first week, we will focus on how we obtain quantitative information about stars and their properties.

I. Luminosity

The most basic stellar property we can think of measuring is its luminosity – its total light output.

A. Apparent brightness and the magnitude system

The first step to measuring stars' luminosity is measuring the flux of light we receive from them. The Greek astronomer Hipparchus invented a numerical scale for describing stars' brightnesses. He described the brightest stars are being of first magnitude, the next brightest of second, etc., down to sixth magnitude for the faintest objects he could discern. In the 1800s, Pogson formalized this system, and unfortunately we are still stuck with a variant of this system today.

I say unfortunate because the magnitude system has several undesirable features. First, higher magnitudes corresponds to dimmer objects. Second, since it was calibrated off human senses, the system is, like human senses, logarithmic. Every five magnitudes corresponds to a change of a factor of 100 in brightness. In this class we will not make any further use of the magnitude system, and will instead discuss only fluxes from stars, which can be measured directly with enough accuracy for our purposes.

While fluxes are a first step, however, they don't tell us much about the stars themselves. That is because we cannot easily distinguish between stars that are bright but far and stars that are dim but close. The flux depends on the star's intrinsic luminosity and distance:

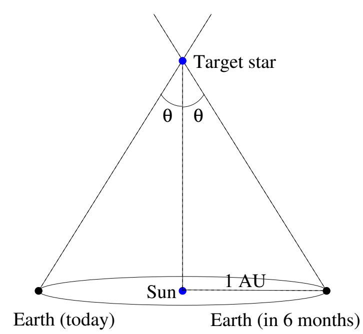
$$F = \frac{L}{4\pi r^2}.$$

From the standpoint of building a theory for how stars work, the quantity we're really interested in is luminosity. In order to get that, we need to be able to measure distances.

In terms of the magnitude system, the flux is described as an apparent magnitude. We are interested instead in the absolute magnitude, which is defined as the brightness that a star would have if we saw it from a fixed distance.

B. Parallax and distances

The oldest method, and still the only really direct one, for measuring the distance to a star is parallax. Parallax relies on the apparent motion of a distant object relative to a much more distant background as we look at it from different angles. The geometric idea is extremely simple: we measure the position of the target star today, then we measure it again in 6 months, when the Earth is on the opposite side of its orbit.



We then measure the change in the apparent position of the star, relative to some very distant background objects that don't appear to move appreciably. The change is described in terms of the parallax angle θ. For a measured change 2θ, the distance r to the target star is simply given by

$$r = \frac{1 \text{ AU}}{\tan\theta} \simeq \frac{1 \text{ AU}}{\theta}$$

where the distance between the Earth and the Sun is $1 \text{ AU} = 1.5 \times 10^{13}$ cm, and in the second step we used the small angle formula to say $\tan\theta \simeq \theta$, since in practice θ is always small.

The importance of this method of distance measurement is illustrated by the fact that in astronomy the most common unit of distance measurement is the parsec (pc), which is defined as the distance away that an object must have in order to produce a parallax shift of 1 second of arc, which is 4.85×10^{-6} radians:

$$1 \text{ pc} = \frac{1 \text{ AU}}{4.85 \times 10^{-6} \text{ rad}} = 3.09 \times 10^{18} \text{ cm} = 3.26 \text{ ly}$$

The nice thing about this definition is that the distance in parsecs is just one over the parallax shift in arcseconds.

Although this technique has been understood since antiquity, our ability to actually use it depends on being able to measure very small angular shifts. The nearest star to us is Proxima Centauri, which has a parallax of 0.88" – to put this in perspective, this corresponds to the size of a quarter at a distance of half a kilometer. As a result of this difficulty, the first successful use of parallax to measure the distance to a star outside the solar system was not until 1838, when Friedrich Bessel measured the distance to 61 Cygni.

In the 1980s and 90s, the Hipparcos satellite made parallax measurements for a large number of nearby stars – up to about 500 pc distance for the brightest stars. The Gaia satellite will push this distance out to tens of kpc, with the exact limit depending on the brightness of the target star.

Even without Gaia, the Hipparcos database provides a sample of roughly 20,000 stars for which we now know the absolute distance to better than 10% and thus, by simply measuring the fluxes from these stars, we know their absolute luminosities. These luminosities form a crucial data set against which we can test theories of stellar structure.

II. Temperature measurements

Luminosity is one of two basic direct observables quantities for stars. The other is the star's surface temperature.

A. Blackbody emission

To understand how we can measure the surface temperature of a star, we need to digress a bit into the thermodynamics of light. Most of what I am going to say here you either have seen or will see in your quantum mechanics or statistical mechanics class, so I'm going to assert results rather than deriving them from first principles.

To good approximation, we can think of a star as a blackbody, meaning an object that absorbs all light that falls on it. Blackbodies have the property that the spectrum of light they emit depends only on their temperature.

The intensity of light that a blackbody of temperature T emits at wavelength λ is given by the Planck function

$$B(\lambda, T) = \frac{2hc^2}{\lambda^5}\left(\frac{1}{e^{hc/(\lambda k_B T)} - 1}\right),$$

where $h = 6.63 \times 10^{-27}$ erg s is Planck's constant, $c = 3.0 \times 10^{10}$ cm s^{-1} is the speed of light, and $k_B = 1.38 \times 10^{-16}$ erg K^{-1} is Boltzmann's constant.

[Slide 2 – the Planck function]

If we differentiate this function, we find that it reaches a maximum at a wavelength

$$\lambda_{\max} = 0.20\frac{hc}{k_B T} = \frac{0.29 \text{ cm}}{T},$$

where in the last step the temperature is measured in K. This implies that, if we measure the wavelength at which the emission from a star peaks, we immediately learn the star's surface temperature. Even if we don't measure the full spectrum, just measuring the color of a star by measuring its flux through a set of different-colored filters provides a good estimate of its surface temperature.

The total light output by a blackbody of surface area A is

$$L = A\sigma T^4 = 4\pi R^2 \sigma T^4,$$

where the second step is for a sphere of radius R. This means that we measure L and T for a star, we immediately get an estimate of its radius. Unfortunately this is only an estimate, because star's aren't really blackbodies – they don't have well-defined solid surfaces. As a result, the spectrum doesn't look exactly like our blackbody function, and the radius isn't exactly what we infer from L and T.

B. Spectral classification

We can actually learn a tremendously larger amount by measuring the spectrum of stars. That's because a real stellar spectrum isn't just a simple continuous function like a blackbody. Instead, there are all sorts of spiky features. These were first studied by the German physicist Fraunhofer in 1814 in observations of the Sun, and for the Sun they are known as Fraunhofer lines in his honor. They are called lines because when you look at the light spread through a prism, they appear as dark lines superimposed on the bright background.

[Slide 3 – Fraunhofer lines]

Each of these lines is associated with a certain element or molecule – they are caused by absorption of the star's light by atoms or molecules in the stellar atmosphere at its surface. As you will learn / have learned in quantum mechanics, every element or molecule has certain energy levels that it can be in. The dark lines correspond to wavelengths of light where the energy of photons at that wavelength matches the difference in energy between two energy levels in some atom or molecule in the stellar atmosphere. Those photons are strongly absorbed by those atoms or molecules, leading to a drop in the light we see coming out of the star at those wavelengths.

Although you don't see it much in the Sun, in some stars there are strong emission lines as well as absorption lines. Emission lines are like absorption lines in reverse: they are upward spikes in the spectrum, where there is much more light at a given frequency than you would get from a blackbody. Emission lines appear when there is an excess of a certain species of atoms and molecules in the stellar atmosphere that are in excited quantum states. As these excited states decay, they emit extra light at certain wavelengths.

We can figure out what lines are caused by which atoms and molecules using laboratory experiments on Earth, and as a result tens of thousands of spectral lines that appear in stars have been definitively assigned to the species that produces them.

Stellar spectra show certain characteristic patterns, which lead astronomers to do what they always do: when confronted with something you don't understand, classify it! The modern spectral classification system, formally codified by Annie Jump Cannon in 1901, recognizes 7 classes for stars: O, B, A, F, G, K, M. This unfortunate nomenclature is a historical accident, but it has led to a useful mnemonic: **O**h **B**e **A** **F**ine **G**irl/guy, **K**iss **M**e. Each of these classes is subdivided into ten sub-classes from $0-9$ – a B9 star is next to an A0, an A9 is next to an F0, etc.

In the 1920s, Cecilia Payne-Gaposchkin showed that these spectra correlate with surface temperature, so the spectral classes correspond to different ranges of surface temperature. O is the hottest, and M is the coolest. Today we know that both surface temperature and spectrum are determined by stellar mass, as we'll discuss in a few weeks. Thus the spectral classes correspond to different stellar masses – O stars are the most massive, while M stars are the least massive. O stars are also the largest.

The Sun is a G star.

[Slides 4 and 5 – spectral types and colors]

In modern times observations have gotten better, and we can now see objects too dim and cool to be stars. These are called brown dwarfs, and two new spectral types have been added to cover them. These are called L and T, leading to the extended mnemonic **O**h **B**e **A** **F**ine **G**irl/guy, **K**iss **M**e **L**ike **T**hat, which proves one thing – astronomers have way too much time on their hands.

There has been a theoretical proposal that a new type of spectral class should appear for objects even dimmer than T dwarfs, although no examples of such an object have yet been observed. The proposed class is called Y, and I can only imagine the mnemonics that will generate...

III. Chemical abundance measurements

One of the most important things we can learn from stellar spectra is what stars, or at least their atmospheres, are made of. To see how this works, we need to spend a little time discussing the physical properties of stellar atmospheres that are responsible for producing spectral lines. We'll do this using two basic tools: the Boltzmann distribution and the Saha equation. I should also mention here that what we're going to do is a very simple sketch of how this process actually works. I'm leaving out a lot of details. The study of stellar atmospheres is an entire class unto itself!

A. A quick review of atomic physics

Before we dive into how this works, let's start by refreshing our memory of quantum mechanics and the structure of atoms. Quantum mechanics tells us that the electrons in atoms can only be in certain discrete energy states. As an example, we can think of hydrogen atoms. The energy of the ground state is -13.6 eV, where I've taken the zero of energy to be the unbound state. The energy of the first excited state is $-13.6/2^2 = -3.40$ eV. The second excited state has an energy of $-13.6/3^2 = -1.51$ eV, and so forth. The energy of state n is

$$E_n = \frac{-13.6 \text{ eV}}{n^2}. \tag{1}$$

Atoms produce spectral lines because a free atom can only interact with photons whose energies match the difference between the atom's current energy level and some other energy level. Thus for example a hydrogen atom in the $n = 2$ state (the first excited state) can only absorb photons whose energies are

$$\begin{aligned}
\Delta E_{3,2} = E_3 - E_2 &= \frac{-13.6 \text{ eV}}{3^2} - \frac{-13.6 \text{ eV}}{2^2} = 1.89 \text{ eV} \\
\Delta E_{4,2} = E_4 - E_2 &= \frac{-13.6 \text{ eV}}{4^2} - \frac{-13.6 \text{ eV}}{2^2} = 2.55 \text{ eV} \\
\Delta E_{5,2} = E_5 - E_2 &= \frac{-13.6 \text{ eV}}{5^2} - \frac{-13.6 \text{ eV}}{2^2} = 2.86 \text{ eV}
\end{aligned}$$

and so forth. This is why we see discrete spectral lines. In terms of wavelength, the hydrogen lines are at

$$\lambda = \frac{hc}{\Delta E} = 656, 486, 434, \ldots \text{ nm}$$

This particular set of spectral lines corresponding to absorptions out of the $n = 2$ level of hydrogen are called the Balmer series, and the first few of them fall in the visible part of the spectrum. (In fact, the 656 nm line is in the red, and when you see bright red colors in pictures of astronomical nebulae, they are often coming from emission in the first Balmer line.)

Thus if we see an absorption feature at 656 nm, we know that is produced by hydrogen atoms in the $n = 2$ level transitioning to the $n = 3$ level. Even better, suppose we see another line, at a wavelength of 1870 nm, which corresponds to the $n = 3 \rightarrow 4$ transition. From the relative strength of those two lines, i.e. how much light is absorbed at each energy, we get a measurement of the relative abundances of atoms in the $n = 2$ and $n = 3$ states.

The trick is that there is no reason we can't do this for different atoms. Thus if we see one line that comes from hydrogen atoms, and another that comes from (for example) calcium atoms, we can use the ratio of those two lines to infer the ratio of hydrogen atoms to calcium atoms in the star. To do that, however, we need to do a little statistical mechanics, where is where the Boltzmann distribution and the Saha equation come into play. Like the Planck function, these come from quantum mechanics and statistical mechanics, and I'm simply going to assert the results rather than derive them from first principles, since you will see the derivations in those classes.

B. The Boltzmann Distribution

We'll start with the Boltzmann distribution. The reason we need this is the following problem: when we see a particular spectral line, we're seeing absorptions due to one particular quantum state of an atom – for example the strength of a Balmer line tells us about the number of hydrogen atoms in the $n = 2$ state. However, we're usually more interested in the total number of atoms of a given type than in the number that are in a given quantum state.

One way of figuring this out would be to try to measure lines telling us about many quantum states, but on a practical level this can be very difficult. For example the transitions associated with the $n = 1$ state of hydrogen are in the ultraviolet, where the atmosphere is opaque, so these can only be measured by telescopes in space. Even from space, gas between the stars tends to strongly absorb these photons, so even if we could see these lines from in the Sun, we couldn't measure them for any other star. Thus, we instead turn to theory to let us figure out the total element abundance based on measurements of one (or preferably a few) states.

Consider a collection of atoms at a temperature T. The electrons in each atom can be in many different energy levels; let E_i be the energy of the ith level. To be definite, we can imagine that we're talking about hydrogen, but what we say will apply to any atom. The Boltzmann distribution tells us the ratio of the number of atoms in state i to the number in state j:

$$\frac{N_i}{N_j} = \exp\left(-\frac{E_i - E_j}{k_B T}\right),$$

where k_B is Boltzmann's constant. Note that the ratio depends only on the difference in energy between the two levels, not on the absolute energy, which is good, since we can always change the zero point of our energy scale.

Strictly speaking, this expression is only true if states i and j really are single quantum states. In reality, however, it is often the case that several quantum states will have the same energy. For example, in a non-magnetic atom, the energy doesn't depend on the spin of an electron, but the spin can be up or down, and those are two distinct quantum states. An atom is equally likely to be in each of them, and the fact that there are two states at that energy doubles the probability that the atom will have that energy. In general, if there are g_i states with energy E_i, then the probability of being in that state is increased by a factor of g_i, which is called the degeneracy of the state. The generalization of the Boltzmann distribution to degenerate states is

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

This describes the ratio of the numbers of atoms in any two states. Importantly, it depends only on the gas temperature and on quantum mechanical constants that we can measure in a laboratory on Earth or compute from quantum mechanical theory (although the latter is only an option for the very simplest of atoms). This means that, if we measure the ratio of the number of atoms in two different states, we can get a measure of the temperature:

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

It is also easy to use the Boltzmann distribution to compute the fraction of atoms in any given state. The fraction has to add up to 1 when we sum over all the possible states, and you should be able to convince yourself pretty quickly that this implies that

$$\frac{N_i}{N} = \frac{g_i e^{-(E_i - E_1)/(k_B T)}}{\sum_{j=1}^{N_{\rm state}} g_j e^{-(E_j - E_1)/(k_B T)}},$$

where $N_{\rm state}$ is the total number of possible states. Since the sum in the denominator comes up all the time, we give it a special name: the partition function. Thus the fraction of the atoms in a state i is given by

$$\frac{N_i}{N} = \frac{g_i e^{-(E_i - E_1)/(k_B T)}}{Z(T)},$$

where

$$Z(T) = \sum_{j=1}^{N_{\rm state}} g_j e^{-(E_j - E_1)/(k_B T)}$$

is the partition function, which depends only on the gas temperature and the quantum-mechanical structure of the atom in question.

This is very useful, because now we can turn around and use this equation to turn a measurement of the number of atoms in some particular quantum state into a measurement of the total number of atoms:

$$N = N_i \frac{Z(T)}{g_i e^{-(E_i - E_1)/(k_B T)}}.$$

Of the terms on the right, N_i we can measure from an absorption line, T we can measure based on the ratio of two lines, and everything else is a known constant.

C. The Saha Equation

The Boltzmann equation tells us what fraction of the atoms are in a given quantum state, but that's only part of what we need to know, because in the atmosphere of a star some of the atoms will also be ionized, and each ionization state produces a different set of lines. Thus for example, it turns out that most of the lines we see in the Sun that come from calcium arise not from neutral calcium atoms, but from singly-ionized calcium: Ca^+. Brief note on notation: astronomers usually refer to ionization states with roman numerals, following the convention that the neutral atom is roman numeral I, the singly-ionized state is II, the twice-ionized state is III, etc. Thus, Ca^+ is often written Ca II.

If we want to know the total number of calcium atoms in the Sun, we face a problem very similar to the one we just solved: we want to measure one line that comes from one quantum state of one particular ionization state, and use that to extrapolate to the total number of atoms in all quantum and ionization states. This is where the Saha equation comes in, named after its discoverer, Meghnad Saha. I will not derive it in class, although the derivation is not complex, and is a straightforward application of statistical mechanics. The Saha equation looks much like the Boltzmann distribution, in that it tells us the ratio of the number of atoms in one ionization state to the number in another. If we let N_i be the number of atoms in ionization state i and N_{i+1} be the number in state $i+1$, then the Saha equation tells us that

$$\frac{N_{i+1}}{N_i} = \frac{2Z_{i+1}}{n_e Z_i} \left(\frac{2\pi m_e k_B T}{h^2} \right)^{3/2} e^{-\chi/(k_B T)},$$

where Z_i and Z_{i+1} are the partition functions for ionization state i and $i+1$, n_e is the number density of free-electrons, and χ is the energy required to ionize an atom. Often the pressure is easier to measure than the electron abundance, so we use the ideal gas law to rewrite things: $P_e = n_e k_B T$. Thus

$$\frac{N_{i+1}}{N_i} = \frac{2k_B T Z_{i+1}}{P_e Z_i} \left(\frac{2\pi m_e k_B T}{h^2} \right)^{3/2} e^{-\chi/(k_B T)},$$

The calculation from here proceeds exactly as for the Boltzmann distribution: we measure the number of atoms in one particular ionization statement, then use the Saha equation to convert it into a total number in all ionization states.

D. A Worked Example: the Solar Calcium Abundance

Let's go through a real example: we will determine the ratio of calcium atoms to hydrogen atoms in the Sun. Our input to this calculation is the following: (1) the Sun's surface temperature is 5777 K, (2) the Sun's surface pressure is about 15 dyne cm^{-2}, (3) comparing a line produced by absorptions by singly-ionized calcium in the ground state (called the Ca II K line) to one produced by absorptions by the first excited state of neutral hydrogen (called the Hα line) shows that there are about 400 times as many atoms in the ground state of Ca^+ as in the first excited state of neutral H.

Let's start with the hydrogen: we're seeing the first excited state of neutral hydrogen, so let's figure out what fraction of all hydrogen atoms this represents. This has two parts: first, we need to use the Saha equation to figure out what fraction of the hydrogen is neutral, then we need to use the Boltzmann distribution to figure out what fraction of the neutral hydrogen is in the first excited state.

For the Saha equation, we need the partition functions for neutral and ionized hydrogen. For ionized hydrogen, it's trivial: ionized hydrogen is just a proton, so it has exactly one state, and $Z_{\rm II} = 1$. For neutral hydrogen, the degeneracy of state j is $2j^2$ and the energy of state j is $E_j = -13.6/j^2$ eV, so the partition function is

$$Z_{\rm I} = \sum_{j=1}^{\infty} g_j e^{-(E_i - E_1)/(k_B T)} = \sum_{j=1}^{\infty} 2j^2 \exp\left[\frac{-13.6 \text{ eV}}{k_B T}\left(1 - \frac{1}{j^2}\right)\right] \simeq 2$$

Note that only the first term contributes appreciably to the sum, because the exponential factor is tiny for $j \neq 1$: $k_B T = 0.5$ eV, so these terms are all something like e^{-20}.

Plugging into the Saha equation, we now get

$$\frac{N_{\rm II}}{N_{\rm I}} = \frac{2k_B T Z_{\rm II}}{P_e Z_{\rm I}} \left(\frac{2\pi m_e k_B T}{h^2}\right)^{3/2} e^{-13.6\text{eV}/(k_B T)} = 7.9 \times 10^{-5}$$

Thus we find that the fraction of H atoms in the ionized state is tiny, and we can treat all the atoms as neutral.

The next step is to compute the fraction that are in the first excited state, and therefore capable of contributing to the Hα line. For this we use the Boltzmann distribution:

$$\frac{N_2}{N} = \frac{2(2^2)e^{-(E_2 - E_1)/(k_B T)}}{Z_{\rm I}} = 5.1 \times 10^{-9}.$$

Thus only one in 200 million H atoms is in the first excited state and can contribute to the Hα line.

The next step is to repeat this for calcium, using first the Saha and then the Boltzmann equations. For calcium, we need some data that are available from laboratory experiments. At a temperature of 5000 – 6000 K, the partition functions for the neutral and once-ionized states are $Z_{\rm I} = 1.32$ and $Z_{\rm II} = 2.30$. The ionization potential is 6.11 eV. Thus from the Saha equation we have

$$\frac{N_{\rm II}}{N_{\rm I}} = \frac{2k_B T Z_{\rm II}}{P_e Z_{\rm I}} \left(\frac{2\pi m_e k_B T}{h^2}\right)^{3/2} e^{-6.11\text{eV}/(k_B T)} = 920$$

Thus there is much more Ca in the once-ionized state than in the neutral state – this is essentially all because of the difference between an ionization potential of 6.11 eV and 13.6 eV. It may not seem like much, but the it's in the exponential, so it makes a big difference. In fact, the second ionization potential of calcium is 11.9 eV, and as a result there is almost no twice-ionzed calcium. Thus, to good approximation, we can simply say that all the calcium is once-ionized.

For this once-ionized calcium, the next step is to compute what fraction is in the ground state that is responsible for producing the Ca II K line. The degeneracy of the ground state is $g_1 = 2$, so from the Boltzmann distribution

$$\frac{N_1}{N} = \frac{2}{Z_{\rm II}} = 0.87.$$

Note that the exponential factor disappeared here because we're asking about the ground state, so it's e^0. Thus we find that 87% of the Ca II ions are in the ground state capable of producing the Ca II K line.

Now we're at the last step. We said that there are 400 times as many atoms producing the Ca II K line as there are producing the Hα line. Now we know that the atoms producing the Ca II K line represent 87% of the all the calcium, while those responsible for producing the Hα represent 5.1×10^{-9} of the hydrogen. Thus the ratio of the total number of Ca atoms to the total number of H atoms is

$$\frac{N_{\mathrm{Ca}}}{N_{\mathrm{H}}} = 400 \frac{5.1 \times 10^{-9}}{0.87} = 2.3 \times 10^{-6}.$$

Thus hydrogen atoms outnumber calcium atoms by a factor of about 500,000.

Repeating this exercise for other elements yields their abundances. This calculation was first carried out for 18 elements by Cecilia Payne-Gaposchkin in her PhD thesis in 1925, and since then has been done for vastly more. We therefore have a very good idea of the chemical composition of the stars. By number and mass, hydrogen is by far the most abundant element, followed by helium, and with everything else a distant third.

[Slide 6 – solar system element abundances]

This technique has gotten so good that it is now generally possible to measure relative element abundances in stars to accuracies of a few percent. That's pretty impressive: just by measuring some absorption lines and knowing some atomic physics, we are able to deduce what distant stars are made of.

Class 2 Notes: Binary Stars, Stellar Populations, and the HR Diagram

In the first class we focused on what we can learn by measuring light from individual stars. However, if all we ever measured were single stars, it would be very difficult to come up with a good theory for stars work. Fortunately, there are a lot of stars in the sky, and that wealth of stars provides a wealth of data we can use to build models. This leads to the topic of our second class: what we can learn from groups of stars.

A quick notation note to start the class: anything with a subscript $\odot$ refers to the Sun. Thus $M_\odot = 1.99 \times 10^{33}$ g is the mass of the Sun, $L_\odot = 3.84 \times 10^{33}$ erg s^{-1} is the luminosity of the Sun, and $R_\odot = 6.96 \times 10^{10}$ cm is the radius of the Sun. These are convenient units of measure for stars, and we'll use them throughout the class.

I. Mass measurements using binaries

Thus far we have figured out how to measure stars' luminosities, temperatures, and chemical abundances. However, we have not yet discussed how to measure perhaps the most basic quantity of all: stars' masses.

This turns out to be surprisingly difficult – how do you measure the mass of an object sitting by itself in space? The answer turns out to be that you don't, but that you can measure the mass of objects that aren't sitting by themselves. That leads to our last topic for today: binary stars and their myriad uses.

Roughly 2/3 of stars in the Milky Way appear to be single stars, but the remaining 1/3 are members of multiple star systems, meaning that two or more stars are gravitationally bound together and orbit one another. Of these, binary systems, consisting of two stars are by far the most common. Binaries are important because they provide us with a method to measure stellar masses using Newton's laws alone.

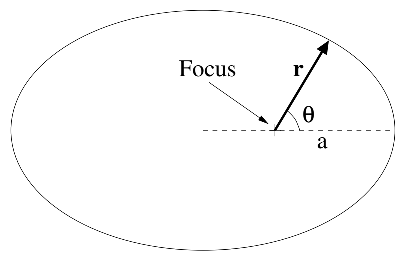
As a historical aside before diving into how we measure masses, binary stars are interesting as a topic in the history of science because they represent one of the earliest uses of statistical inference. The problem is that when we see two stars close to one another on the sky, there is no obvious way to tell if the two are physically near each other, or if it is simply a matter of two distant, unrelated stars that happen to be lie near the same line of sight. In other words, just because two stars have a small angular separation, it does not necessarily mean that they have a small physical separation.

However, in 1767 the British astronomer John Michell performed a statistical analysis of the distribution of stars on the sky, and showed that there are far more close pairs than one would expect if they were randomly distributed. Thus, while Michell could not infer that any particular pair of stars in the sky was definitely a physical binary, he showed that the majority of them must be.

A. Visual binaries

Binary star systems can be broken into two basic types, depending on how we discover them. The easier one to understand is visual binaries, which are pairs of stars which are far enough apart that we can see them as two distinct stars in a telescope.

We can measure the mass of a visual binary system using Kepler's laws. To see how this works, let's go through a brief recap of the two-body problem. Consider two stars of masses M_1 and M_2. We let $\mathbf{r}_1$ and $\mathbf{r}_2$ be the vectors describing the positions of stars 1 and 2, and $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ be the vector distance between them. If we set up our coordinate system so that the center of mass is at the origin, then we know that $M_1\mathbf{r}_1 + M_2\mathbf{r}_2 = 0$. We define the reduced mass as $\mu = M_1 M_2/(M_1 + M_2)$, so $\mathbf{r}_1 = -(\mu/M_1)\mathbf{r}$ and $\mathbf{r}_2 = (\mu/M_2)\mathbf{r}$.



The solution to the problem is that, when the two stars are at an angle θ in their orbit, the distance between them is

$$r = \frac{a(1-e^2)}{1+e\cos\theta},$$

where the semi-major axis a and eccentricity e are determined by the stars' energy and angular momentum. Clearly the minimum separation occurs when $\theta = 0$ and the denominator has its largest value, and the maximum occurs when $\theta = \pi$ and the denominator takes its minimum value. The semi-major axis is the half the sum of this minimum and maximum:

$$\frac{1}{2}[r(0) + r(\pi)] = \frac{1}{2}\left[\frac{a(1-e^2)}{1+e} + \frac{a(1-e^2)}{1-e}\right] = a$$

The orbital period is related to a by

$$P^2 = 4\pi^2 \frac{a^3}{GM},$$

where $M = M_1 + M_2$ is the total mass of the two objects.

This describes how the separation between the two stars changes, but we instead want to look at how the two stars themselves move. The distance from each of the two stars to the center of mass is given by

$$r_1 = \frac{\mu}{M_1} r = \frac{\mu}{M_1}\left[\frac{a(1-e^2)}{1+e\cos\theta}\right] \qquad r_2 = \frac{\mu}{M_2} r = \frac{\mu}{M_2}\left[\frac{a(1-e^2)}{1+e\cos\theta}\right].$$

Again, these clearly reach minimum and maximum values at $\theta = 0$ and $\theta = \pi$, and the semi-major axes of the two ellipses describing the orbits of each star are given by half the sum of the minimum and maximum:

$$a_1 = \frac{1}{2}[r_1(0) + r_1(\pi)] = \frac{\mu}{M_1}a \qquad\qquad a_2 = \frac{1}{2}[r_2(0) + r_2(\pi)] = \frac{\mu}{M_2}a.$$

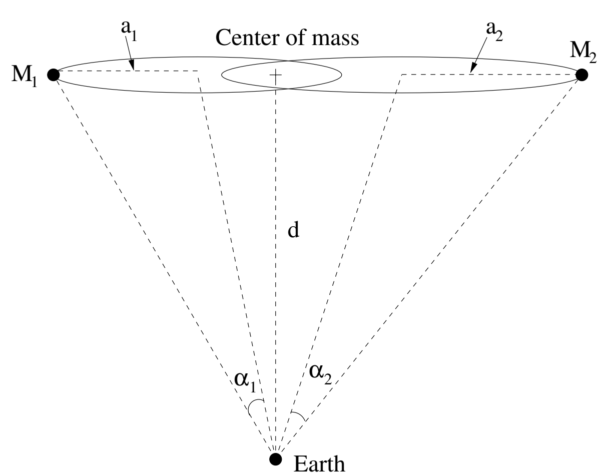
Note that it immediately follows that $a = a_1 + a_2$, since $\mu/M_1 + \mu/M_2 = 1$.

We can measure the mass of a visual binary using Kepler's laws. To remind you, there are three laws: first, orbits are ellipses with the center of mass of the system at one focus. Second, as the bodies orbit, the line connecting them sweeps out equal areas in equal times – this is equivalent to conservation of angular momentum. Third, the period P of the orbit is related to its semi-major axis a by

$$P^2 = 4\pi^2 \frac{a^3}{GM},$$

where M is the total mass of the two objects.

With that background out of the way, let's think about what we can actually observe. We'll start with the simplest case, where the orbits of the binary lie in the plane of the sky, the system is close enough that we can use parallax to measure its distance, and the orbital period is short enough that we can watch the system go through a complete orbit. In this case we can directly measure four quantities, which in turn tell us everything we want to know: the orbital period P, the angles subtended by the semi-major axes of the two stars orbits, α_1, and α_2, and the distance of the system, d.



The first thing to notice is that we can immediately infer the two stars' mass ratio just from the sizes of their orbits. The semi-major axes of the orbits are $a_1 = \alpha_1 d$ and $a_2 = \alpha_2 d$. We know that $M_1 r_1 \propto M_2 r_2$, and since $a_1 \propto r_1$ and $a_2 \propto r_2$ by the argument we just went through, we also know that $M_1 a_1 \propto M_2 a_2$. Thus it immediately follows that

$$\frac{M_1}{M_2} = \frac{a_2}{a_1} = \frac{\alpha_2}{\alpha_1}.$$

Note that this means we can get the mass ratio even if we don't know the distance, just from the ratio of the angular sizes of the orbits.

Similarly, we can infer the total mass from the observed semi-major axes and period using Kepler's 3rd law:

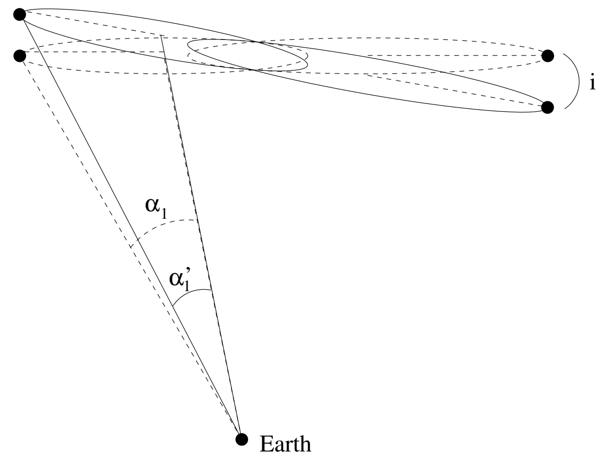
$$M = 4\pi^2 \frac{(a_1 + a_2)^3}{GP^2} = 4\pi^2 \frac{(\alpha_1 d + \alpha_2 d)^3}{GP^2},$$

where everything on the right hand side is something we can observe. Given the mass ratio and the total mass, it is of course easy to figure out the masses of the individual stars. If we substitute in and write everything in terms of observables, we end up with

$$M_1 = 4\pi^2 d^3 \frac{(\alpha_1 + \alpha_2)^2}{GP^2} \alpha_2 \qquad\qquad M_2 = 4\pi^2 d^3 \frac{(\alpha_1 + \alpha_2)^2}{GP^2} \alpha_1.$$

This is the simplest case where we see a full orbit, but in fact we don't have to wait that long – which is a good thing, because for many visual binaries the orbital period is much longer than a human lifetime! Even if we see only part of an orbit, we can make a very similar argument. All we need is to see enough of the orbit that we can draw an ellipse through it, and then we measure α_1 and α_2 for the inferred ellipse. Similarly, Kepler's second law tells us that the line connecting the two stars sweeps out equal areas in equal times, so we can infer the full orbital period just by measuring what fraction of the orbit's area has been swept out during the time we have observed the system.

The final complication to worry about is that we don't know that the orbital plane lies entirely perpendicular to our line of sight. In fact, we would have to be pretty lucky for this to be the case. In general we do not know the inclination of the orbital plane relative to our line of sight. For this reason, we do not know the angular sizes α_1' and α_2' that we measure for the orbits are different than what we would measure if the system were perfectly in the plane of the sky. A little geometry should immediately convince you that that $\alpha_1' = \alpha_1 \cos i$, where i is the inclination, and by convention $i = 0$ corresponds to an orbit that is exactly face-on and $i = 90°$ to one that is exactly edge-on. The same goes for α_2. (I have simplified a bit here and assumed that the tilt is along the minor axes of the orbits, but the same general principles work for any orientation of the tilt.)



This doesn't affect our estimate of the mass ratios, since $\alpha_1/\alpha_2 = \alpha_1'/\alpha_2'$, but it does affect our estimate of the total mass, because $a \propto \alpha_1$. Thus if we want to write out the total mass of a system with an inclined orbit, we have

$$M = 4\pi^2 \frac{(\alpha_1' d + \alpha_2' d)^3}{GP^2 \cos^3 i}.$$

We get stuck with a factor $\cos^3 i$ in the denominator, which means that instead of measuring the mass, we only measure a lower limit on it.

Physically, this is easy to understand: if we hold the orbital period fixed, since we can measure that regardless of the angle, there is a very simple relationship between the stars' total mass and the size of their orbit: a bigger orbit corresponds to more massive stars. However, because we might be seeing the ellipses at an angle, we might have underestimated their sizes, which corresponds to having underestimated their masses.

B. Spectroscopic binaries

A second type of binary is called a spectroscopic binary. As we mentioned earlier, by measuring the spectrum of a star, we learn a great deal about it. One thing we learn is its velocity along our line of sight – that is because motion along the line of sight produces a Doppler shift, which displaces the spectrum toward the red or the blue, depending on whether the star is moving away from us or toward us.

However, we know the absolute wavelengths that certain lines have based on laboratory experiments on Earth – for example the Hα line, which is produced by hydrogen atoms jumping between the 2nd and 3rd energy states, has a wavelength of exactly 6562.8 Å. (I could add several more significant digits to that figure.) If we see the Hα line at 6700 Å instead, we know that the star must be moving away from us.

The upshot of this is that you can use a spectrum to measure a star's velocity. In a binary system, you will see the velocities change over time as the two stars orbit one another. On your homework, you will show how it is possible to use these observations to measure the masses of the two stars.

The best case of all is when a pair of stars is observed as both a spectroscopic and a visual binary, because in that case you can figure out the masses and inclinations without needing to know the distance. In fact, it's even better than that: you can actually calculate the distance from Newton's laws!

Unfortunately, very few star systems are both spectroscopic and visual binaries. That is because the two stars have to be pretty far apart for it to be possible to see both of them with a telescope, rather than seeing them as one point of light. However, if the two stars are far apart, they will also have relatively slow orbits with relatively low velocities. These tend to produce Doppler shifts that are too small to measure. Only for a few systems where the geometry is favorable and where the system is fairly close by can we detect a binary both spectroscopically and visually. These systems are very precious, however, because then we can measure everything about them. In particular, for our purposes, we can measure their masses absolutely, with no uncertainties due to inclination or distance.

II. The HR Diagram

We've already learned a tremendous amount just by looking at pairs of stars, but we can learn even more by looking at larger populations. The most basic and important tool we have for studying stellar populations is the Hertzsprung-Russell diagram, or HR diagram for short. Not surprisingly, this diagram was first made by Hertzspring and Russell, who actually created it independently, in 1911 for Ejnar Hertzsprung and in 1913 for Henry Norris Russell.

A. The Observer's HR Diagram

The HR diagram is an extremely simple plot. We simply find a bunch of stars, measure their luminosities and surface temperatures, and make a scatter plot of one against the other. Since surface temperatures and absolute luminosities are often expensive to measure in practice, more often we plot close proxies to them. In place of total luminosity we put the luminosity (or magnitude) as seen in some particular range of wavelengths, and in place of surface temperature we plot the ratio of the brightness seen through two different filters – this is a proxy for color, and thus for surface temperature. For this reason, we sometimes also refer to these diagrams as color-magnitude diagrams, or CMDs for short.

An important point to make is that it is only possible to make an HR diagram for stars whose distances are known, since otherwise we don't have a way of measuring their luminosities. The largest collection of stars ever placed on an HR diagram comes from the Hipparcos catalog, which we discussed last time – a collection of stars near the Sun whose distances have been measured by parallax.

[Slide 1 – Hipparcos HR diagram]

This particular HR diagram plots visual magnitude against color. A note on interpreting the axes: this HR diagram was made using two different filters: B and V. Here B stands for blue, and it is a filter that allows preferentially blue light to pass. V stands for visible, and it is a filter that allows essentially all visible light to pass. On the y axis in this plot is absolute visual magnitude is basically the logarithm of the luminosity in the V band, which is close to the total luminosity because stars put out most of their power at visible wavelengths. The approximate conversion between mass and luminosity is shown on the right y axis. Note that, since the magnitude scale is backward and higher magnitude corresponds to lower luminosity, the scale on the left is reversed – magnitude decreases upward, so that brighter stars are near the top, as you would intuitively expect.

On the x axis is B magnitude minus V magnitude. Since magnitudes are a logarithmic scale, $B - V$ is a measure of the ratio of the star's luminosity in blue light to its total luminosity. Since magnitudes go in the opposite of the sensible direction (bigger numbers are dimmer), a high value of $B - V$ corresponds to a small ratio of blue luminosity to total luminosity. A low value of $B - V$ is the opposite. Thus moving to the right on this diagram corresponds to getting redder, and moving left corresponds to getting bluer. The value of $B - V$ corresponds to an approximate surface temperature, which is indicated on the top x axis.

The first thing you notice about the diagram is that the stars don't fall anything like randomly on it. The great majority of them fall along a single fat line, which we call the main sequence. The Sun sits right in the middle of it. The main sequence extends from stars that are bright and blue to stars that are dim and red. It covers an enormous range of absolute luminosities, from 10^{-3} $L_\odot$ to 10^3 $L_\odot$ – and that's just for nearby stars. The range is larger if we include more stars, because very bright stars are rare, and this particular HR diagram doesn't have any of the brightest ones on it.

You can also see a second prominent population, extending like a branch of the main sequence. These are stars that are red but bright. For reasons we'll see in a moment, this means they must have very large radii, and so they are called red giants.

The HR diagram we've been looking at is limited to relatively bright and nearby stars. One can extend it by adding in some observations to measure dimmer stars, as well as a selection of other more exotic stars. This HR diagram includes 22,000 stars from Hipparcos supplemented by 1,000 stars from the Gliese catalog.

[Slide 2 – extended Hipparcos HR diagram]

As before, we see that the most prominent feature is the main sequence, and the second most prominent is a branch extending out of it consistent of bright, red stars. However, we can also see some other populations start to emerge. First, there is a collection of stars that run from medium color to blue, but that are very dim – a factor of ~ 1000 dimmer than the Sun. These too fall along a rough line.

As we will see in a moment, the combination of high surface temperature and low luminosity implies that these stars must have very small radii. For this reason, we call them dwarfs. Because these stars have fairly high surface temperatures, their colors are whitish-blue. Thus, these stars are called white dwarfs.

One can also get glimpses of other types of stars in other parts of the diagram, which don't fall on either the main sequence, the red giant branch, or the white dwarf sequence. Almost all of these stars are very bright, and lie above the main sequence. These other types of stars must be very rare, since even a catalog containing 23,000 stars includes only a handful of them. We'll discuss these more exotic types of stars when we get to stellar evolution in the second half of the course.

It's worth stopping for a moment to realize that the HR diagram is rather surprising. Why should it be that stars do not occupy the full range of luminosities and temperatures continuously, and instead seem to cluster into distinct groups? Explaining the existence of these distinct groups where stars live, and their relationship to one another, is the single big theoretical problem from stellar physics. Our goal at the end of this class is to understand why the HR diagram looks like this, and what it means.

Similarly, the fact that the main sequence is a single curve means that all stars of a given luminosity have about the same color, and vice-versa. We would like to understand why that is.

B. HR Diagrams of Clusters

So far we've been looking at HR diagrams of stars that happen to be near the Sun, which are a hodgepodge of stars of many different masses, ages, and abundances of heavy elements. However, it is very instructive to instead look at the HR diagram for more homogenous populations. How do we pick a homogenous population? Fortunately, nature has provided for us. Some stars are found in clusters that are held together by the stars mutual gravity.

[Slide 3 – globular cluster M80]

The slide shows an example of a star cluster: a globular cluster known as M80, which contains several hundred thousand stars. The stars in a cluster like M80 generally formed in a single burst of star formation, and as a result they are very close to one another in age and chemical abundance. The stars in a cluster are also of course all at about the same distance from us, which means that we can compare their relative brightnesses even if we don't know exactly how far away the cluster is.

[Slide 4 – HR diagram for NGC 6397]

The slide shows the HR diagram for the star clusters NGC 6397. The axis labels here correspond to the filters used on the Hubble telescope, but the idea is the same as before: the y axis shows the magnitude in a single filter using a reverse axis scale, so it measures luminosity, with up corresponding to higher luminosity. The x axis shows the ratio of the luminosities in two different color filters, with the orientation chosen so that red is to the right and blue is to the left.

So what's different about this HR diagram as opposed to the ones we looked at earlier for nearby stars? Just like in the previous picture we see that most stars fall along a single curve – the main sequence – and that there is a secondary curve at lower luminosity and bluer color – the white dwarf sequence. In comparison to the other HR diagram, however, this main sequence is much narrower. Instead of a fat line we have a very thin one. In fact, it's even thinner than it appears in the diagram – many of the points that lie off the main sequence turn out to be binary stars that are so close to one another that the telescope couldn't separate them, and thus treated them as a single star.

The thinness of this main sequence suggests that the spread we saw in the solar neighborhood main sequence must be due to factors that are absent in the star cluster: a spread in stellar ages and a spread in chemical composition. The fact that all the stars fall along a single thin line is compelling evidence that there is a single intrinsic property of stars that varies as we move along the main sequence, and is responsible for determining a star's luminosity and temperature. The age and chemical composition can alter this slightly, causing the thin sequence to puff up a little, but basically there's one number that is going to determine everything about a star's properties. The natural candidate, of course, is the star's mass.

If you're very sharp you may have noticed something else slightly different about this HR diagram, compared to the one we looked at for the solar neighborhood. You may have noticed that, in the solar neighborhood, the white dwarf sequence does not go as far toward the blue as the bluest part of the main sequence. In this HR diagram, on the other hand, the white dwarfs go further toward the blue. This is not an accident, as becomes clear when we compare HR diagrams for different star clusters.

[Slide 5 – HR diagrams for M 67 and NCG 188]

The slide shows HR diagrams for two different clusters: M 67 and NGC 188. These are both a type of cluster called on open cluster. The scatter is mostly an observational artifact: it's due to stars that happen to be along the same line of sight as the star cluster, but aren't really members of it, and are at quite different distances. Those have been cleaned out of the HR diagram for NGC 6397, but not for these clusters.

The interesting thing to notice is the difference between the two main sequences. It seems that the dim, red end of the main sequence is about the same from one cluster to another, but the bright, blue side ends at different points in different clusters. Where the main sequence ends, it turns upward and becomes the red giant branch, and the red giant branch is at two different places in the two clusters. The place where the main sequence ends and bends upward into the red giant branch is called the main sequence turn-off.

What is going on here? What's the difference between these two clusters? The answer turns out to be their ages. M 67 formed somewhat more recently than NGC 188. If we repeat this exercise for clusters of different ages, we see that this is a general trend. In younger clusters the main sequence turn off is more toward the bright, blue side of the sequence, and it vanishes completely in the youngest clusters. In older clusters it moves to lower luminosity and redder color. That's the reason that the white dwarf sequence extended further to the blue than the main sequence in the globular cluster NGC 6397 – globular clusters are very, very old.

Thus, we have another clue: bright, blue main sequence stars disappear at a certain age. The brighter and bluer the star, the shorter the time for which it can be found. This too is something that our theory needs to be able to explain.

C. Stellar Radii on the HR Diagram

So far we have talked only about quantities we can observe directly – the observational HR diagram. Now let's see what we can infer based on our knowledge of physics. The first and most obvious thing to do with an HR diagram is to see what it tells us about stars' radii.

To remind you, radius, luminosity, and temperature are all related by the black-body formula we wrote down last time:

$$L = 4\pi R^2 \sigma T^4$$

Since the HR diagram is a plot of temperature versus luminosity, at every point in the plot we can use the temperature and luminosity to solve for the corresponding radius:

$$R = \sqrt{\frac{L}{4\pi\sigma T^4}}$$

Putting this in terms of some useful units, this is

$$\frac{R}{R_\odot} = 1.33 \left(\frac{L}{L_\odot}\right)^{1/2} \left(\frac{T}{5000\text{ K}}\right)^{-2} \tag{1}$$

Just from this formula you can figure out what a line of constant radius will look like. We've been plotting things on a logarithmic scale, so our HR diagrams have $\log L$ on the y axis and $\log T$ on the x axis. If we take the logarithm of this equation, we get

$$\log L = 4\log T + \log(4\pi\sigma) + 2\log R.$$

If R is constant, then in the $(\log T, \log L)$ plane this is just a line with a slope of 4.

[Slide 6 – theorist's HR diagram]

Note that the line does have a slope of 4, but because of the annoying astronomy convention that we plot red, lower temperature, to the right, the lines appear to have a negative slope. Comparing these lines to the main sequence, you can see that the main sequence has a somewhat steeper slope than 4 in the $(\log T, \log L)$ plane. This means that stars at the low temperature, low luminosity end have smaller radii than stars at the high luminosity, high temperature end. The range in radii is a factor of ~ 100.

Although they're not shown on this particular diagram, you can immediately see why white dwarfs and red giants have the names they do. The red giant branch extends above and to the right of the main sequence, so it goes up to several hundred $R_\odot$. The white dwarf sequence is below the main sequence, so that it hovers around 0.01 $R_\odot$. To put these numbers in perspective, $100 R_\odot = 0.47$ AU, so a star with a radius just above $200 R_\odot$ would encompass the Earth's orbit. A radius of 0.01 $R_\odot$ corresponds to 10% more than the radius of the Earth. Thus the largest red giants would swallow the Earth, while the smallest white dwarfs are about the size of the Earth.

D. Stellar Masses on the HR Diagram

The radii of stars we can measure directly off the HR diagram, but the masses are a bit trickier, since we need additional information to obtain those. As we discussed earlier, we can only get independent measurements for the masses of stars if they are members of binary systems. Fortunately, astronomers have now compiled a fairly large list of binary systems within which we can measure masses, so we can plot mass against luminosity and color.

[Slide 7 – mass-luminosity and mass-effective temperature relations]

Note that this slide only contains main sequence stars, not red giants, white dwarfs, etc. Also note that, on the plot on the right, the axes are reverse, so more massive, brighter stars are to the left.

Here's one interesting thing to take away from this plot: the luminosity-mass relation is a line just like the main sequence. In other words, all main sequence stars of a given mass have the same luminosity, and, with some error, the same radius and surface temperature. Since we already saw that the main sequence is a single curve on the HR diagram, this plot tells us something critical: *for main sequence stars, the stellar mass determines where the star falls on the main sequence.*

This is a profound statement. It means that, for a main sequence star, if you know it's mass, then you know pretty much everything about it. the properties of a star are, to very good approximation, dictated solely by its mass. Nothing else matters much. In this class we will attempt to understand why this is.

Another interesting point is that over significant ranges in mass, the mass-luminosity relation is a straight line on a log-log plot. What sort of function produces a straight line on a log-log plot? The answer is a powerlaw. To see why, write down the equation of a line in the log-log plot:

$$\log L = p \log M + c,$$

where p is the slope of the line and c is the y-intercept. It immediately follows that

$$L = cM^p.$$

Thus the luminosity is proportional to some power of the mass, and the power is equal to the slope of the line on the log-log plot. If you actually measure the slope of the data, you find that for masses in the vicinity of 1 $M_\odot$, a slope of 3.5 is a reasonable fit.

Class 3 Notes: Hydrostatic Balance and the Virial Theorem

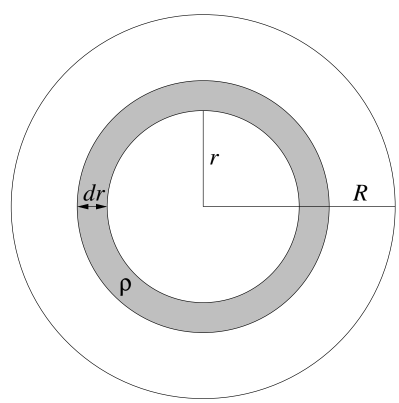
Thus far we have discussed what observations of the stars tell us about them. Now we will begin the project that will consume the next 5 weeks of the course: building a physical model for how stars work that will let us begin to make sense of those observations. This week we'll try to write down some equations that govern stars' large-scale properties and behavior, before diving into the detailed microphysics of the stellar plasma next week.

In everything we do today (and for the rest of the course) we will assume that stars are spherically symmetric. In reality stars rotate, convect, and have magnetic fields; these induce deviations from spherical symmetry. However, these deviations are small enough that, for most stars, we can ignore them to first order.

I. Hydrostatic Balance

A. The Equation of Motion

Consider a star of total mass M and radius R, and focus on a thin shell of material at a distance r from the star's center. The shell's thickness is dr, and the density of the gas within it is ρ. Thus the mass of the shell is $dm = 4\pi r^2 \rho \, dr$.



The shell is subject to two types of forces. The first is gravity. Let m be the mass of the star that is interior to radius r and note that $\rho \, dr$ is the mass of the shell per cross sectional area. The gravitational force per area on the shell is just

$$F_g = -\frac{Gm}{r^2} \rho \, dr,$$

where the minus sign indicates that the force per area is inward.

The other force per area acting on the shell is gas pressure. Of course the shell feels pressure from the gas on either side of it, and it feels a net pressure force only due to the difference in pressure on either side. This is just like the forces caused by air the room. The air pressure is pretty uniform, so that we feel equal force from all directions, and there is no net force in any particular direction. However, if there is a difference in pressure, there will be a net force.

Thus if the pressure at the base of the shell is $P(r)$ and the pressure at its top is $P(r + dr)$, the net pressure force per area that the shell feels is

$$F_p = P(r) - P(r + dr)$$

Note that the sign convention is chosen so that the force from the top of the shell (the $P(r + dr)$ term) is inward, and the force from the bottom of the shell is outward.

In the limit $dr \to 0$, it is convenient to rewrite this term in a more transparent form using the definition of the derivative:

$$\frac{dP}{dr} = \lim_{dr \to 0} \frac{P(r + dr) - P(r)}{dr}.$$

Substituting this into the pressure force per area gives

$$F_p = -\frac{dP}{dr} dr.$$

We can now write down Newton's second law, $F = ma$, for the shell. The shell mass per area is $\rho \, dr$, so we have

$$\begin{aligned} (\rho \, dr)\ddot{r} &= -\left(\frac{G\rho m}{r^2} + \frac{dP}{dr}\right) dr \\ \ddot{r} &= -\frac{Gm}{r^2} - \frac{1}{\rho}\frac{dP}{dr} \end{aligned}$$

This equation tells us how the shell accelerates in response to the forces applied to it. It is the shell's equation of motion.

B. The Dynamical Timescale

Before going any further, let's back up and ask a basic question: how much do we actually expect a shell of material within a star to be accelerating? We can approach the question by asking a related one: suppose that the pressure force and the gravitational force were very different, so that there was a substantial acceleration. On what timescale would we expect the star to change its size or other properties?

We can answer that fairly easily: if pressure were not significant, the outermost shell would free-fall inward due to the star's gravity. The characteristic speed to which it would accelerate is the characteristic free-fall speed produced by that gravity: $v_{\text{ff}} = \sqrt{2GM/R}$. The amount of time it would take to fall inward to the star's center is roughly the distance to the center divided by this speed. We define this as the star's dynamical (or mechanical) timescale: the time that would be required for the star to re-arrange itself if pressure and gravity didn't balance. It is

$$t_{\text{dyn}} = \frac{R}{\sqrt{2GM/R}} \approx \sqrt{\frac{R^3}{GM}} \approx \sqrt{\frac{1}{G\bar{\rho}}},$$

where $\bar{\rho}$ is the star's mean density. Obviously we have dropped factors of order unity at several points, and it is possible to do this calculation more precisely – in fact, we will do so later in the term when we discuss star formation.

In the meantime, however, let's just evaluate this numerically. If we plug in $R = R_\odot$ and $M = M_\odot$, we get $\bar{\rho} = 1.4$ g cm^{-3}, and $t_{\text{dyn}} = 3000$ s. There are two things about this result that might surprise you. The first is how low the Sun's density is: 1.4 g cm^{-3} is about the density of water, 1 g cm^{-3}. Thus the Sun has about the same mean density as water. The second, and more important for our current problem, is how incredibly short this time is: 3000 seconds, or a bit under an hour.

This significance of this is clear: if gravity and pressure didn't balance, the gravitational acceleration of the Sun would be sufficient to induce gravitational collapse in about an hour. Even if gravity and pressure were out of balance by 1%, collapse would still occur, just in 10 hours instead of 1. (It's 10 and not 100 because distance varies like acceleration times time squared, so a factor of 100 change in the acceleration only produces a factor of 10 change in time.) Given that the Sun is more than 10 hours old, the pressure and gravity terms in the equation of motion must balance to at better than 1%. In fact, just from the fact that the Sun is at least as old as recorded history (not to mention geological time), we can infer that the gravitational and pressure forces must cancel each other to an extraordinarily high degree of precision.

C. Hydrostatic Equilibrium

Given this result, in modeling stars we will simply make the assumption that we can drop the acceleration term in the equation of motion, and directly equate the gravitational and force terms. Thus we have

$$\frac{dP}{dr} = -\rho \frac{Gm}{r^2}.$$

This is known as the equation of hydrostatic equilibrium, since it expresses the condition that the star be in static pressure balance.

This equation expresses how much the pressure changes as we move through a given radius in the star, i.e. if we move upward 10 km, by how much will the pressure change? Sometimes it is more convenient to phrase this in terms of change per unit mass, i.e. if we move upward far enough so that an additional 0.01 $M_\odot$ of material is below us, how much does the pressure change. We can express this mathematically via the chain rule. The change in pressure per unit mass is

$$\frac{dP}{dm} = \frac{dP}{dr}\frac{dr}{dm} = \frac{dP}{dr}\left(\frac{dm}{dr}\right)^{-1} = \frac{dP}{dr}\frac{1}{4\pi r^2 \rho} = -\frac{Gm}{4\pi r^4}.$$

This is called the Lagrangian form of the equation, while the one involving dP/dr is called the Eulerian form.

In either form, since the quantity on the right hand side is always negative, the pressure must decrease as either r or m increase, so the pressure is highest at the star's center and lowest at its edge. In fact, we can exploit this to make a rough estimate for the minimum possible pressure in the center of star. We can integrate the Lagrangian form of the equation over mass to get

$$\begin{aligned}\int_0^M \frac{dP}{dm} dm &= -\int_0^M \frac{Gm}{4\pi r^4} dm \\ P(M) - P(0) &= -\int_0^M \frac{Gm}{4\pi r^4} dm\end{aligned}$$

On the left-hand side, $P(M)$ is the pressure at the star's surface and $P(0)$ is the pressure at its center. The surface pressure is tiny, so we can drop it. For the right-hand side, we know that r is always smaller than R, so $Gm/4\pi r^4$ is always larger than $Gm/4\pi R^4$. Thus we can write

$$P(0) \approx \int_0^M \frac{Gm}{4\pi r^4} dm > \int_0^M \frac{Gm}{4\pi R^4} dm = \frac{GM^2}{8\pi R^4}$$

Evaluating this numerically for the Sun gives $P_c > 4 \times 10^{14}$ dyne cm^{-2}. In comparison, 1 atmosphere of pressure is 1.0×10^6 dyne cm^{-2}, so this argument demonstrates that the pressure in the center of the Sun must exceed 10^8 atmospheres. In fact, it is several times larger than this.

II. A Digression on Lagrangian Coordinates

Before going on, it is worth pausing to think a bit about the coordinate system we made use of to derive this result, because it is one that we're going to encounter over and over again throughout the class. Intuitively, the most natural way to think about stars is in terms of Eulerian coordinates. The idea of Eulerian coordinates is simple: you pick some particular distance r from the center of the star, and ask questions like what is the pressure at this position? What is the temperature at this position? How much mass is there interior to this position? In this system, the independent coordinate is position, and everything is expressed a function of it: $P(r)$, $T(r)$, $m(r)$, etc.

However, there is an equally valid way to think about things inside a star, which goes by the name Lagrangian coordinates. The basic idea of Lagrangian coordinates is to label things not in terms of position but in terms of mass, so that mass is the independent coordinate and everything is a function of it.

This may seem counter intuitive, but it makes a lot of sense, particularly when you have something like a star where all the mass is set up in nicely ordered shells. We label each mass shell by the mass m interior to it. Thus for a star of total mass M, the shell $m = 0$ is the one at the center of the star, the one $m = M/2$ is at the point that contains half the mass of the star, and the shell $m = M$ is the outermost one. Each shell has some particular radius $r(m)$, and we can instead talk about the pressure, temperature, etc. in given mass shell: $P(m)$, $T(m)$, etc.

The great advantage of Lagrangian coordinates is that they automatically take care of a lot of bookkeeping for us when it comes to the question of advection. Suppose we are working in Eulerian coordinates, and we want to know about the change in gas temperature at a particular radius r. The change could happen in two different ways. First, the gas could stay still, and it could get hotter or colder. Second, all the gas could stay at exactly the same temperature, but it could move, so that hotter or colder gas winds up at radius r. In Eulerian coordinates the change in temperature at a given radius arises from some arbitrary combination of these two processes, and keeping track of the combination requires a lot of bookkeeping. In contrast, for Lagrangian coordinates, only the first type of change is possible. $T(m)$ can increase or decrease only if the gas really gets hotter or colder, not if it moves.

Of course the underlying physics is the same, and doesn't depend on which coordinate system we use to describe it. We can always go between the two coordinate systems by a simple change of variables. The mass interior to some radius r is

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

so

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

These relations allow us to go between derivatives with respect to one coordinate and derivatives with respect to another. For an arbitrary quantity f, the chain rule tells us that

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

However, in the vast majority of the class, it will be simpler for us to work in Lagrangian coordinates.

III. The Virial Theorem

We will next derive a volume-integrated form of the equation of hydrostatic equilibrium that will prove extremely useful for the rest of the class, and, indeed, is perhaps one of the most important results of classical statistical mechanics: the virial theorem. The first proof of a form of the virial theorem was accomplished by the German physicist Claussius in 1851, but numerous extensions and generalizations have been developed since. We will be using a particularly simple version of it, but one that is still extremely powerful.

A. Derivation

To derive the virial theorem, we will start by taking both sides of the Lagrangian equation of hydrostatic balance and multiplying by the volume $V = 4\pi r^3/3$ interior to some radius r:

$$V\,dP = -\frac{1}{3}\frac{Gm\,dm}{r}.$$

Next we integrate both sides from the center of the star to some radius r where the mass enclosed is $m(r)$ and the pressure is $P(r)$:

$$\int_{P(0)}^{P(r)} V\,dP = -\frac{1}{3}\int_0^{m(r)} \frac{Gm'\,dm'}{r'}.$$

Before going any further algebraically, we can pause to notice that the term on the right side has a clear physical meaning. Since Gm'/r' is the gravitational potential due to the material of mass m' inside radius r', the integrand $(Gm'/r')dm'$ just represents the gravitational potential energy of the shell of material of mass dm' that is immediately on top of it. Thus the integrand on the right-hand side is just the gravitational potential energy of each mass shell. When this is integrated over all the mass interior to some radius, the result is the total gravitational potential energy of the gas inside this radius. Thus we define

$$\Omega(r) = -\int_0^{m(r)} \frac{Gm'\,dm'}{r'}$$

to the gravitational binding energy of the gas inside radius r.

Turning back to the left-hand side, we can integrate by parts:

$$\int_{P(0)}^{P(r)} V\,dP = [PV]_0^r - \int_0^{V(r)} P\,dV = [PV]_r - \int_0^{V(r)} P\,dV.$$

In the second step, we dropped PV evaluated at $r' = 0$, because $V(0) = 0$. To evaluate the remaining integral, it is helpful to consider what dV means. It is the volume occupied by our thin shell of matter, i.e. $dV = 4\pi r^2\,dr$. While we could make this substitution to evaluate, it is even better to think in a Lagrangian way, and instead think about the volume occupied by a given mass. Since $dm = 4\pi r^2 \rho\,dr$, we can obviously write

$$dV = \frac{dm}{\rho},$$

and this changes the integral to

$$\int_0^{V(r)} P\,dV = \int_0^{m(r)} \frac{P}{\rho} dm.$$

Putting everything together, we arrive at our form of the virial theorem:

$$[PV]_r - \int_0^{m(r)} \frac{P}{\rho} dm = \frac{1}{3}\Omega(r).$$

If we choose to apply this theorem at the outer radius of the star, so that $r = R$, then the first term disappears because the surface pressure is negligible, and we have

$$\int_0^M \frac{P}{\rho} dm = -\frac{1}{3}\Omega,$$

where Ω is the total gravitational binding energy of the star.

This might not seem so impressive, until you remember that, for an ideal gas, you can write

$$P = \frac{\rho k_B T}{\mu m_{\rm H}} = \frac{\mathcal{R}}{\mu}\rho T,$$

where μ is the mean mass per particle in the gas, measured in units of the hydrogen mass, and $\mathcal{R} = k_B/m_{\rm H}$ is the ideal gas constant. If we substitute this into the virial theorem, we get

$$\int_0^M \frac{\mathcal{R}T}{\mu} dm = -\frac{1}{3}\Omega.$$

For a monatomic ideal gas, the internal energy per particle is $(3/2)k_B T$, so the internal energy per unit mass is $u = (3/2)\mathcal{R}T/\mu$. Substituting this in, we have

$$\begin{aligned} \int_0^M \frac{2}{3} u \, dm &= -\frac{1}{3}\Omega \\ U &= -\frac{1}{2}\Omega, \end{aligned}$$

where U is just the total internal energy of the star, i.e. the internal energy per unit mass u summed over all the mass in the star. This is a remarkable result. It tells us that the total internal energy of the star is simply $-(1/2)$ of its gravitational binding energy.

The total energy is

$$E = U + \Omega = \frac{1}{2}\Omega.$$

Note that, since $\Omega < 0$, this implies that the total energy of a star made of ideal gas is negative, which makes sense given that a star is a gravitationally bound object. Later in the course we'll see that, when the material in a star no longer acts like a classical ideal gas, the star can have an energy that is less negative than this, and thus is less strongly bound.

Incidentally, this result bears a significant resemblance to one that applies to orbits. Consider a planet of mass m, such as the Earth, in a circular orbit around a star of mass M at a distance R. The planet's orbital velocity is the Keplerian velocity

$$v = \sqrt{\frac{GM}{R}},$$

so its kinetic energy is

$$K = \frac{1}{2}mv^2 = \frac{GMm}{2R}.$$

Its potential energy is

$$\Omega = -\frac{GMm}{R},$$

so we therefore have

$$K = -\frac{1}{2}\Omega,$$

which is basically the same as the result we just derived, except with kinetic energy in place of internal energy. This is no accident: the virial theorem can be proven just as well for a system of point masses interacting with one another as we have proven it for a star, and an internal or kinetic energy that is equal to $-1/2$ of the potential energy is the generic result.

B. Application to the Sun

We'll make use of the virial theorem many times in this class, but we can make one immediate application right now: we can use the virial theorem to estimate the mean temperature inside the Sun. Let $\bar{T}$ be the Sun's mass-averaged temperature. The internal energy is therefore

$$U = \frac{3}{2}M\frac{\mathcal{R}\bar{T}}{\mu}.$$

The gravitational binding energy depends somewhat on the internal density distribution of the Sun, which we are not yet in a position to calculate, but it must be something like

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

where α is a constant of order unity that describes our ignorance of the internal density structure. Applying the virial theorem and solving, we obtain

$$\begin{aligned} \frac{3}{2}M\frac{\mathcal{R}\bar{T}}{\mu} &= \frac{1}{2}\alpha\frac{GM^2}{R} \\ \bar{T} &= \frac{\alpha}{3}\frac{\mu}{\mathcal{R}}\frac{GM}{R} \end{aligned}$$

If we plug in $M = M_\odot$, $R = R_\odot$, $\mu = 1/2$ (appropriate for a gas of pure, ionized hydrogen), and $\alpha = 3/5$ (appropriate for a uniform sphere), we obtain $\bar{T} = 2.3 \times 10^6$ K. This is quite impressively hot. It is obviously much hotter than the surface temperature of about 6000 K, so if the average temperature is more than 2 million K, the temperature in the center must be even hotter.

It is also worth pausing to note that we were able to deduce the internal temperature of Sun to within a factor of a few from nothing more than its bulk characteristics, and without any knowledge of the Sun's internal workings. This sort of trick is what makes the virial theorem so powerful!

We can also ask what the Sun's high temperature implies about the state of the matter in its interior. The ionization potential of hydrogen is 13.6 eV, and for

$T = 2 \times 10^6$ K, the thermal energy per particle is $(3/2)k_B T = 260$ eV. Thus the thermal energy per particle is much greater than the ionization potential of hydrogen. Any collision will therefore lead to an ionization, and we conclude that the bulk of the gas in the interior of a star must be nearly fully ionized.

IV. The First Law of Thermodynamics

Thus far we have written down the equation of hydrostatic balance and derived results from it. Hydrostatic balance is essentially a statement of conservation of momentum. However, there is another, equally important conservation law that all material obeys: the first law of thermodynamics, i.e. conservation of energy. Conservation of energy should be a familiar concept, and all we're going to do here is express it in a form that is appropriate for the gas that makes up a star, and that includes the types of energy that are important in a star.

As in the last class, let's consider a thin shell of mass $dm = 4\pi r^2 \rho \, dr = \rho dV$ at some radius within the star. This mass element has an internal energy per unit mass u, so the total energy of the shell is $u \, dm$. The internal energy can consist of thermal energy (i.e. heat) and chemical energy (i.e. the energy associated with changes in the chemical state of the gas, for example the transition between neutral and ionized). For now we will leave the nature of the energy unspecified, because for our argument it won't matter.

We would like to know how much the energy changes in a small amount of time. Let the change in energy over a time δt be

$$\delta E = \delta(u \, dm) = \delta u \, dm,$$

where conservation of mass implies that dm is constant, so that any change in the energy of the shell is due to changes in the energy per unit mass, not due to change in the mass. The first law of thermodynamics tells us that the change in energy of the shell must be due to heat it absorbs or emits (from radiation, from neighboring shells, or from other sources) or due to work done on it by neighboring shells. Thus we write

$$\delta u \, dm = \delta Q + \delta W$$

By itself this isn't a very profound statement, since we have not yet specified the work or the heat. Let's start with the work. Work on a gas is always $P \, \delta V$, i.e. the change in the volume of the gas multiplied by the pressure of the gas that opposes or promotes that change. The volume of our shell is dV, so the change in its volume is δdV. Thus we can write

$$\delta W = -P \, \delta dV = -P \, \delta \left(\frac{dV}{dm} dm \right) = -P \, \delta \left(\frac{1}{\rho} \right) dm.$$

There are a couple of things to say about this. First, notice the minus sign. This makes sense. If the volume increases (i.e. $\delta dV > 0$), then the shell must be expanding, and doing work on the gas around it. Thus its internal energy must decrease to pay for this work. Second, we have re-written the volume dV in a more convenient form, $(1/\rho) \, dm$.

What is the physical meaning of this? Well, ρ is the density, i.e. the mass per unit volume. Thus $1/\rho$ is the volume per unit mass. Since dm is the mass, $(1/\rho)\,dm$ just means the mass times the volume per unit mass – which of course is the volume. The reason this form is more convenient is that in the end we're going to do everything per unit mass, so it is useful to have a dm instead of a dV.

Now consider the heat absorbed or emitted, δQ. Heat can enter or leave the mass shell in two ways. First, it can be produced by chemical or nuclear reactions within the shell. Let q be the rate per unit mass of energy release by nuclear reactions. Here q has units of energy divided by mass divided by time, so for example we might say that burning hydrogen into helium releases a certain number of ergs per second per gram of fuel. Thus the amount of heat added by nuclear reactions in a time δt is $\delta Q_{\rm nuc} = q\,dm\,\delta t$.

The second way heat can enter or leave the shell is by moving down to the shell below or up to the shell above. The actual mechanism of heat flow can take various forms: radiative (i.e. photons carry energy), mechanical (i.e. hot gas moves and carries energy with it), or conductive (i.e. collisions between the atoms of a hot shell and the colder shell next to it transfer energy to the colder shell). For now we will leave the mechanism of heat transport unspecified, and return to it later on. Instead, we just let $F(m)$ be the flux of heat entering the shell from below. Similarly, the flux of heat leaving the top of the shell is $F(m + dm)$. Note that the flux has units of energy per unit time, *not* energy per unit area per unit time, like the flux we talked about last week. This is unfortunate nomenclature, but we're stuck with it.

With these definitions, we can write the heat emitted or absorbed as

$$
\begin{aligned}
\delta Q &= \left[q\,dm + F(m) - F(m+dm)\right]\delta t \\
&= \left[q\,dm + F(m) - F(m) - \frac{\partial F}{\partial m}\,dm\right]\delta t \\
&= \left(q - \frac{\partial F}{\partial m}\right) dm\,\delta t.
\end{aligned}
$$

In the second step, we used a Taylor expansion to rewrite $F(m + dm) = F(m) + (\partial F/\partial m)\,dm$.

Putting together our expressions for δW and δQ in the first law of thermodynamics, we have

$$
\begin{aligned}
\delta u\,dm + P\delta\left(\frac{1}{\rho}\right) dm &= \left(q - \frac{\partial F}{\partial m}\right) dm\,\delta t \\
\frac{du}{dt} + P\frac{d}{dt}\left(\frac{1}{\rho}\right) &= q - \frac{\partial F}{\partial m},
\end{aligned}
$$

where in the second step we divided through by $dm\,\delta t$, and wrote quantities of the form $\delta f/\delta t$ as derivatives with respect to time. This equation is the first law of thermodynamics for the gas in a star. It says that the rate at which the specific internal energy of a shell of mass in a star changes is given by minus the pressure time the rate at which the volume per unit mass of the shell changes, plus the rate at which nuclear energy is generated within it, minus any difference in the heat flux between across the shell.

This equation described conservation of energy for stellar material. We'll see what we can do with it next time.

Class 4 Notes: Energy and Chemical Balance in Stars

In the last class we introduced the idea of hydrostatic balance in stars, and showed that we could use this concept to derive crude limits on their internal properties even without constructing a detailed model. In this class we will apply the same sort of analysis to the energy and chemical balance in stars – in effect examining the principles governing the energy and chemical content of stars rather than their mechanical equilibrium. We will see that this leads to similar non-obvious conclusions. It also sets us up to begin making detailed models in the next few weeks.

I. The Energy Equation

A. Static Stars

We ended last class by writing down the first law of thermodynamics for a given shell within a star. This enables us to make some interesting statements about the total energy content of a star. First consider the simplest case of a star in equilibrium, so that each shell's volume and specific internal energy are constant in time. In this case the left-hand side of the first law of thermodynamics is exactly zero because nothing is changing with time, and we have

$$q = \frac{\partial F}{\partial m}.$$

If we integrate this equation over all the mass in a star, we have

$$\begin{aligned} \int_0^M q\,dm &= \int_0^M \frac{\partial F}{\partial m}\,dm \\ \int_0^M q\,dm &= F(M) - F(0) \end{aligned}$$

Consider the physical meaning of this equation. The left-hand side is the total rate of nuclear energy generation in the star, summing over all the star's mass. We call this quantity the nuclear luminosity $L_{\rm nuc}$ – a luminosity because it has units of energy per time; i.e., it is the total rate at which nuclear reactions in the star release energy.

When we integrate the right-hand side, we wind up with the difference between the flux passing through the last mass shell, $F(M)$, and the flux entering the star at $m = 0$, $F(0)$. The latter is obviously zero, unless there is a magical energy source at the center of the star. The former, $F(M)$ is just the energy per unit time leaving the stellar surface. Thus, $F(M)$ must be the star's total luminous output, which we see as light, and denote L.

Thus the equation we have derived simply states

$$L_{\rm nuc} = L,$$

i.e. for a state in equilibrium, the total energy leaving the stellar surface must be equal to the total rate at which nuclear reactions within the star release energy. This isn't exactly a shocking conclusion, but the machinery we used to derive it will be prove extremely useful when we consider stars that are *not* exactly in equilibrium.

B. Time-Variable Stars

The calculation we just performed can be generalized to the case of a star that is not exactly in equilibrium, so that the time derivatives are not zero. If we retain these terms and integrate over mass again, we have

$$\int_0^M \frac{du}{dt} dm + \int_0^M P \frac{d}{dt}\left(\frac{1}{\rho}\right) dm = \int_0^M q\, dm - F(M) + F(0) = L_{\rm nuc} - L.$$

For the first term on the left-hand side, since m does not depend on t, we can interchange the integral and the time derivative. Thus, we have

$$\int_0^M \frac{du}{dt} dm = \frac{d}{dt} \int_0^M u\, dm = \frac{d}{dt} U,$$

where U is the total internal energy of the star.

For the second term, it is convenient to rewrite the time derivative of $1/\rho$ in a slightly different form:

$$\frac{d}{dt}\left(\frac{1}{\rho}\right) = \frac{d}{dt}\left(\frac{dV}{dm}\right) = \frac{d}{dm}\left(\frac{dV}{dt}\right),$$

where we have used the fact that m does not depend on t to interchange the order of the derivatives. Here V is the volume of the material inside mass shell m. If this shell is at radius r, then $V = (4/3)\pi r^3$, and

$$\frac{dV}{dt} = 4\pi r^2 \frac{dr}{dt}.$$

In other words, the rate at which the volume occupied by a given mass of gas changes is equal to the surface area of its outer boundary ($4\pi r^2$) multiplied by the rate at which it expands or contracts (dr/dt).

Putting this into the integral, we have

$$\int_0^M P \frac{d}{dt}\left(\frac{1}{\rho}\right) dm = \int_0^M P \frac{d}{dm}\left(4\pi r^2 \frac{dr}{dt}\right) dm.$$

We can evaluate this integral by parts. Doing so gives

$$\int_0^M P\frac{d}{dt}\left(\frac{1}{\rho}\right)dm = \left[4\pi r^2 P\frac{dr}{dt}\right]_0^M - \int_0^M 4\pi r^2 \frac{dr}{dt}\frac{dP}{dm}dm.$$

For the first term, note that at $m = 0$, $dr/dt = 0$. This is because the innermost shell must stay at the center, unless a vacuum appears around the origin. Similarly, $P = 0$ at $m = M$, because the pressure drops to zero at the edge of the star. (Strictly speaking it is not exactly zero, but it is negligibly small.) Thus, the term in square brackets must be zero.

For the second term, we can evaluate using the equation of motion that we derived in the last class. To remind you, we derived the equation

$$\ddot{r} = -\frac{Gm}{r^2} - \frac{1}{\rho}\frac{dP}{dr}.$$

If we re-arrange this we get

$$\frac{dP}{dr} = -\frac{GM\rho}{r^2} - \rho\ddot{r}.$$

Converting to Lagrangian coordinates, we have

$$\frac{dP}{dm} = \frac{1}{4\pi r^2 \rho}\frac{dP}{dr} = -\frac{Gm}{4\pi r^4} - \frac{\ddot{r}}{4\pi r^2}$$

If we substitute this into our integral, we have

$$\int_0^M P\frac{d}{dt}\left(\frac{1}{\rho}\right)dm = \int_0^M \frac{Gm}{r^2}\dot{r}\,dm + \int_0^M \dot{r}\ddot{r}\,dm$$

The meanings of the terms are a bit clearer if we rewrite them a bit. In particular, we can write

$$\frac{Gm}{r^2}\dot{r} = -\frac{d}{dt}\left(\frac{Gm}{r}\right)$$

and

$$\dot{r}\ddot{r} = \frac{1}{2}\frac{d}{dt}(\dot{r}^2),$$

so that we have

$$\int_0^M P\frac{d}{dt}\left(\frac{1}{\rho}\right)dm = -\frac{d}{dt}\int_0^M \frac{Gm}{r}\,dm + \frac{1}{2}\frac{d}{dt}\int_0^M \dot{r}^2\,dm = \dot{\Omega} + \dot{T}.$$

As earlier, we have used the fact that mass does not depend on time to interchange the integral with the time derivative.

The first term, which I have written $\dot{\Omega}$, is clearly just minus the time derivative of the gravitational potential energy; $-Gm/r$ is the gravitational potential experienced by the shell at mass m, so the integral of $-Gm\,dm/r$ is just the total gravitational potential energy over the entire star. Similarly, the term I have labelled $\dot{T}$ is just the time derivative of the total kinetic energy of the cloud; $\dot{r}^2/2$ is the kinetic energy per unit mass of a shell, so the integral of $(1/2)\dot{r}^2\,dm$ is just the total kinetic energy of the star.

Putting it all together, we arrive at the total energy equation for the star:

$$\dot{U} + \dot{\Omega} + \dot{T} = L_{\rm nuc} - L.$$

This just represents the total energy equation for the star, and it is fairly intuitively obvious. It just says that the time rate of change of internal energy plus gravitational potential energy plus kinetic energy is equal to the rate at which nuclear reactions add energy minus the rate at which energy is radiated away from the stellar surface.

We can get something slightly more interesting if we consider a star that is expanding or contracting extremely slowly, so that its very close to hydrostatic balance. In this case we can make two simplifications: (1) we can drop the term $\dot{T}$, because the star's kinetic energy is tiny compared to its internal energy or gravitational potential energy; (2) we can use the virial theorem for hydrostatic objects, which says that $U = -\Omega/2$. In this case the energy equation reads

$$\frac{1}{2}\dot{\Omega} = L_{\rm nuc} - L.$$

This equation is decidedly non-obvious. It tells us how quickly the gravitational potential energy of the star changes in response to energy generation by nuclear reactions and energy loss by radiation.

C. The Kelvin-Helmholtz Timescale

It is useful to consider an order-of-magnitude version of the energy equation, because it tells us something important about the nature of stars. Consider how long will it take the gravitational potential energy (and thus the stellar radius) to change by a significant amount in a star without any nuclear energy generation (i.e. where $L_{\rm nuc} = 0$).

The gravitational potential energy is

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

where α is a number of order unity that depends on the density distribution within the star, and R is the star's radius. If we define t as the time required to alter the gravitational potential energy significantly, then at the order-of-magnitude level the equation we have written reads

$$\begin{aligned} -\frac{1}{t}\left(\frac{GM^2}{R}\right) &\sim -L \\ t &\sim \frac{GM^2}{RL}. \end{aligned}$$

We define the quantity $t_{\rm KH} = GM^2/(RL)$ as the Kelvin-Helmholtz timescale, named after the 19th century physicists Kelvin (of the Kelvin temperature scale) and Helmholtz, who first pointed out its importance. The meaning of $t_{\rm KH}$ is that it is the time for which a star could be powered by gravity alone without its radius changing very much. Similarly, if we have a star that is not in energy balance for some reason, so that $L_{\rm nuc} \ll L$, then $t_{\rm KH}$ tells us about the time that will be required for the star to reach energy equilibrium.

If we plug in numerical values for the Sun, we find $t_{\rm KH} = GM_\odot^2/(R_\odot L_\odot) = 30$ Myr. This number is the answer to the "Star Trek" problem of what would happen if you somehow shut off all nuclear reactions inside a star. The answer is: absolutely nothing for about 30 Myr. You could turn off all the nuclear reactions in the Sun, and unless you were observing neutrinos (which get out immediately), you wouldn't notice anything change for tens of millions of years.

In fact, before the discovery of nuclear energy, it was believed that gravity was the main process causing the Sun to shine. This played an important role in the history of science, because, if gravity really did power the Sun, it would imply that the Sun's properties could only have been constant for ~ 30 Myr. This would be a natural limit for the age of the Earth, or at least for the time for which life as we know it could have existed on Earth. This was an important argument against the Darwinian theory of evolution, which required billions of years to explain the development of life. Of course, as we'll discuss in a few minutes, in this case the biologists were right and the astronomers were wrong. In reality $L_{\rm nuc}$ is very close to L, so the real timescale for the Sun's evolution is vastly longer than 30 Myr.

II. Nuclear Energy and the Nuclear Timescale

There is one more important timescale for star's, which comes from considering how long $L_{\rm nuc} \approx L$ can be maintained. This depends a bit on nuclear chemistry, which will be our topic in a few weeks. For now, we'll simply take on faith that the main nuclear reaction that occurs in the Sun is burning hydrogen into helium. You can figure out how much energy this yields by comparing the masses of hydrogen and helium. The starting point of the reaction is 4 hydrogen nuclei and the final point is 1 helium nucleus. The mass of 1 hydrogen nucleus is 1.6726×10^{-24} g, while the mass of a helium nucleus is 6.64648×10^{-24}. The difference in mass is

$$\Delta m = 4m_{\rm H} - m_{\rm He} = 4.39 \times 10^{-26} \text{ g}.$$

If we phrase this as the change in mass per hydrogen atom we started with, this is

$$\epsilon = \frac{\Delta m}{4m_{\rm H}} = 0.0066.$$

In other words, this reaction converts 0.66% of the mass of each proton into energy. Einstein's relativity then tell us that the excess energy released per hydrogen atom by this reaction is

$$\Delta E = \epsilon m_{\rm H} c^2 = 9.9 \times 10^{-6} \text{ erg} = 6.2 \text{ MeV}.$$

We can use this to estimate how long the nuclear reaction that burns hydrogen into helium can keep a star going. The star radiates energy at a rate L, and thus the rate at which hydrogen atoms must be burned is $L/\Delta E = L/(\epsilon m_{\rm H} c^2)$. To estimate how long this can keep going, we simply divide the total number of hydrogen atoms in the star, roughly $M/m_{\rm H}$, but the rate at which they are burned. This defines the nuclear timescale

$$t_{\rm nuc} = \frac{M/m_{\rm H}}{L/(\epsilon m_{\rm H} c^2)} = \frac{\epsilon M c^2}{L}$$

(In making this estimate we have implicitly neglected the fact that not all of the Sun's mass is hydrogen, but that's a fairly small correction.)

Evaluating this for the Sun gives 100 Gyr, a staggeringly long time – almost a factor of 10 larger than the age of the universe. In fact, we'll see later in the course that the true time for which nuclear reactions can old up the Sun is about a factor of 10 smaller, because the Sun can't actually use all of its hydrogen as fuel. Nonetheless, this result demonstrates that nuclear energy is able to hold up the Sun for much, much longer than the Kelvin-Helmholtz timescale.

III. The Hierarchy of Timescales and Evolutionary Models

A. General Idea

The three timescales we've computed this week tell us a great deal about the ingredients we need to make a model of stars. Putting them in order, we have $t_{\rm nuc} \gg t_{\rm KH} \gg t_{\rm dyn}$, and this is true not just for the Sun, but for all stars. This has the important implication that, on timescales comparable to $t_{\rm nuc}$, we can assume that stars are in nearly perfect mechanical and thermal equilibrium.

This enables a great simplification in making models of stars. Our approach to making stellar models for the rest of the course will therefore proceed through a series of steps:

1. Assume that the star is in perfect mechanical equilibrium (since $t_{\rm dyn} \ll t_{\rm KH} \ll t_{\rm nuc}$), and compute the resulting luminosity.
2. Assume that the star is in perfect energy equilibrium (since $t_{\rm KH} \ll t_{\rm nuc}$), and compute the reaction rate required to provide this luminosity.
3. Evolve the chemical makeup of the star using the derived reaction rates.

For a star that is powered by hydrogen burning into helium, steps 2 and 3 are extraordinarily simple. If we want to keep track of a star's evolution, we just need to keep track of the mass of hydrogen and the mass of helium within it. Let M be the total stellar mass, and let $M_{\rm H}$ and $M_{\rm He}$ be the mass of hydrogen and helium. If we neglect the mass of other elements (a reasonable first approximation) and any change in mass due to radiation (which is tiny) and due to stellar winds (which is small), then we can write $M = M_{\rm H} + M_{\rm He} =$ constant. Steps 2 and 3 therefore are entirely embodied by the equation

$$\dot{M}_{\mathrm{He}} = -\dot{M}_{\mathrm{H}} = \frac{L}{\epsilon c^2}.$$

This is a bit of an oversimplification, since in reality we need to do this on a shell-by-shell basis, and to worry about chemical mixing between the shells. Nonetheless, it conveys the basic idea.

Of course the hard part of this is in step 1: compute the luminosity of a star that is in mechanical equilibrium. This luminosity, it turns out, will depend on nothing but the star's mass – which is what observations have already told us, if we recall the mass luminosity relation. Deriving that mass-luminosity relation from first principles is going to consume the next four weeks of the course. Once we've done that, however, what this timescale analysis tells us is that we will essentially have a full theory for stellar evolution.

B. The Chemical Evolution Equations

To make this more precise, we need to write down the equations governing changed in the composition of the star. To do this, we need to introduce some notation. A star is made of many different elements – the vast majority are hydrogen and helium, but there are others, and it turns out that they can play important roles, particularly in evolved stars. If we examine some volume of gas in a star, we can discuss its density ρ, but we could also count only the hydrogen atoms, only the helium atoms, etc., and compute the density for that species only. For convenience we number these species, so that we might write the density of hydrogen as ρ_1, the density of helium as ρ_2., etc. Depending on the level of sophistication of our model, we might also distinguish different isotopes of the same atom, so that we would count ordinary hydrogen and deuterium (which has an extra neutron) separately. It is often convenient to work with quantities other than mass densities, so we define some alternatives. The mass fraction of species i is written

$$X_i = \frac{\rho_i}{\rho}.$$

We often also want to count the number of atoms, instead of measuring their mass. To do this, we have to take into account the differences in their atomic weights. For example, helium atoms have four times the mass of hydrogen atoms, so if there are four hydrogen atoms for every helium atom, then they both have the same mass fraction. We write the *atomic mass number* for species i as an integer $\mathcal{A}_i$, which means that each atom of that species has an *atomic mass* of approximately $\mathcal{A}_i m_p$, where $m_p = 1.67 \times 10^{-24}$ is the proton mass. Hydrogen has $\mathcal{A} = 1$, so $m_{\mathrm{H}} = m_p$, and we frequently write masses in terms of m_H rather than m_p. Given this definition, it is clear that the number density of atoms is related to their mass density by

$$n_i = \frac{\rho_i}{\mathcal{A}_i m_{\mathrm{H}}},$$

and substituting this into the mass fraction X_i immediately gives

$$X_i = n_i \frac{\mathcal{A}_i}{\rho} m_{\mathrm{H}}.$$

We can describe any species in terms of its atomic mass $\mathcal{A}_i$ and also its atomic number, $\mathcal{Z}_i$. The atomic number gives the number of protons, and thus the charge of the nucleus. For example for the stable isotopes of the most common elements we have

Element	Name	$\mathcal{Z}$	$\mathcal{A}$
^{1}H	Hydrogen	1	1
^{2}H	Deuterium	1	2
^{3}He	Helium-3	2	3
^{4}He	Helium	2	4
^{12}C	Carbon	6	12
^{13}C	Carbon-13	6	13

Some nuclear reactions also involve electrons and positrons. This have $\mathcal{Z} = \pm 1$ and $\mathcal{A} = 0$ (to good approximation – we can generally neglect the mass of electrons compared to protons and neutrons).

The mass fractions X_i can be altered by nuclear reactions, which leave the total mass density fixed (to very good approximation), but convert atoms of one species into atoms of another. We can write one species as $I(\mathcal{Z}_i, \mathcal{A}_i)$, another as $J(\mathcal{Z}_j, \mathcal{A}_j)$, and so forth. In this notation, any chemical reaction is

$$I(\mathcal{Z}_i, \mathcal{A}_i) + J(\mathcal{Z}_j, \mathcal{A}_j) \leftrightarrows K(\mathcal{Z}_k, \mathcal{A}_k) + L(\mathcal{Z}_l, \mathcal{A}_l),$$

where this can obviously be extended to more elements as needed if the reaction involves more species. Chemical reactions always have to conserve mass and charge, so we have two conservation laws:

$$\begin{aligned} \mathcal{Z}_i + \mathcal{Z}_j &= \mathcal{Z}_k + \mathcal{Z}_l \\ \mathcal{A}_i + \mathcal{A}_j &= \mathcal{A}_k + \mathcal{A}_l. \end{aligned}$$

If we want to know how the star's composition changes in time, we need to know the rate at which reactions between two species occur. Computing these reactions rates is a problem we'll defer for now, but we can begin to think about them by noting that the reaction rate will always be proportional to the rate at which atoms of the two reactant species run into one another. To see what this implies, consider a given volume of space within which a chemical reaction is taking place, for example

$$^2\mathrm{H} +^1 \mathrm{H} \rightarrow^3 \mathrm{He},$$

one of the steps in the energy-generating reaction chain in the Sun. What would happen to the rate at which this reaction occurred in that volume if I were to remove half the deuterium (^{2}H)? The remaining deuterium atoms would still encounter hydrogen atoms just as often, since their number would be unchanged, so the reaction rate would just be reduced by a factor of 2. Similarly, the same argument shows that if I were, for example, double the number of hydrogen atoms, the reaction rate would increase by a factor of 2. Clearly the reaction rate must be proportional to the number of members of each reactant species in the volume. Expressing this mathematically, the reaction rate per unit volume must be proportional to $n_i n_j$, where n_i and n_j are the number density of the species i and j involved in the reaction. If there are more species involved, then we just multiply by n_k, n_l, etc. We call the constant of proportionality the reaction rate, and write it R_{ijk}, meaning the rate at which reactions between particles of species i and j occur, leading to species k.

If the reaction involves two members of the same species, the same argument applies, except that we have to be careful not to double-count. Thus rather than having the reaction rate be proportional to $n_i n_j$, it is proportional to $n_i(n_i - 1)/2 \approx n_i^2/2$, where the factor of 1/2 is to handle the double-counting problem. You can convince yourself that this is right. Suppose there are 4 people in a room: Alice, Bob, Cathy, and David (A, B, C, and D). How many distinct couples can we make? Counting them is pretty easy, since we just pick one person, then pick another different person. If we pick Alice, there are 3 possible partners: Bob, Cathy, and David. Similarly, if we pick Bob, there are three possible partners: Alice, Cathy, and David. We can write this out as

AB	AC	AD
BA	BC	BD
CA	CB	CD
DA	DB	DC

The table has $3 \times 4 = 12$ entries. It's pretty clear, however, that half of the entries in this table are duplicates: we have both BA and AB. Thus to count the number of distinct couples, we have $n_i = 4$ and $n_i(n_i - 1)/2 = 6$. The argument is the same for chemical reactions: if we were to make a list of possible collisions, there would be $n_i(n_i - 1)/2$, which for large values of n_i is approximately $n_i^2/2$. Thus the reaction rate is $n_i^2 R_{iik}/2$.

With this notation out of the way, we are now prepared to write down the equations of chemical evolution. Suppose that we have a number density n_i of species i, and that these atoms are destroyed by a reaction with species j, which leads to species k. Clearly the rate of change in the number density of species i is just given by minus the rate at which reactions occur:

$$\frac{d}{dt} n_i = -n_i n_j R_{ijk}.$$

In this case we don't divide by 2 when the reaction is species i with itself because each reaction destroys two atoms of species i, and the factors of 2 cancel. In general there are multiple possible reactions with many possible partners, and we have to sum over all the reactions that destroy members of species i. Thus

$$\frac{d}{dt}n_i = -\sum_{j,k} n_i n_j R_{ijk}.$$

We must also take into account that reactions can create members of species i. Suppose we have a reaction between species l and species k that creates members of species i. In this case the rate at which members of species i are created is

$$\frac{d}{dt}n_i = n_l n_k R_{lki}$$

if l and k are distinct, or

$$\frac{d}{dt}n_i = \frac{1}{2} n_l n_k R_{lki}$$

if l and k are the same. We can unify the notation for these two cases by writing

$$\frac{d}{dt}n_i = \frac{n_l n_k}{1+\delta_{lk}} R_{lki},$$

where δ_{lk} is simply defined to be 1 if l and k are the same, and 0 otherwise. This is nothing more than a notational convenience. Again, we need to sum over all possible reactions that can create species i:

$$\frac{d}{dt}n_i = \sum_{l,k} \frac{n_l n_k}{1+\delta_{lk}} R_{lki},$$

Finally, combining the rates of creation and destruction, we have

$$\frac{d}{dt}n_i = \sum_{l,k} \frac{n_l n_k}{1+\delta_{lk}} R_{lki} - \sum_{j,k} n_i n_j R_{ijk}.$$

If we prefer, we can also write this in terms of the composition fraction by substituting:

$$\frac{d}{dt}X_i = \rho \frac{\mathcal{A}_i}{m_\mathrm{H}} \left(\sum_{l,k} \frac{X_l}{\mathcal{A}_l} \frac{X_k}{\mathcal{A}_k} \frac{R_{lki}}{1+\delta_{lk}} - \frac{X_i}{\mathcal{A}_i} \sum_{j,k} \frac{X_j}{\mathcal{A}_j} R_{ijk} \right)$$

C. The Evolution Equations

We are now in a position to write down the basic equations governing a star's evolution. We envision a spherical star of fixed mass, whose chemical composition at some initial time is known. To figure out how its structure changes in time, our first step is to assume mechanical equilibrium, i.e. hydrostatic balance. We're already written down the equation for this:

$$\frac{dP}{dm} = -\frac{Gm}{4\pi r^4}.$$

This equation gives us a relationship between the position of each shell of mass, r, and the gradient in the pressure of the gas P. The solution tells us how the shells of gas must arrange themselves to maintain mechanical equilibrium.

The second step is to require thermal equilibrium, which means that we balance the energy being generated in the star against the energy it is radiating. We've already written down the integrated version of this as $L = L_{\rm nuc}$, but if we're going to model the shells individually, we want to use the non-integrated version we wrote down earlier:

$$\frac{dF}{dm} = q,$$

where F is the flux passing through a mass shell, and q is the heat generated within it by nuclear burning.

The final step is to figure out how it is changing chemically, which is described by the equations we have just written out:

$$\frac{d}{dt}X_i = \rho\frac{\mathcal{A}_i}{m_{\rm H}}\left(\sum_{l,k}\frac{X_l}{\mathcal{A}_l}\frac{X_k}{\mathcal{A}_k}\frac{R_{lki}}{1+\delta_{lk}} - \frac{X_i}{\mathcal{A}_i}\sum_{j,k}\frac{X_j}{\mathcal{A}_j}R_{ijk}\right)$$

for each species i.

Of course these equations are all coupled. The rates of nuclear reactions in the final step determine the rate of heat generation q in the second one, or vice versa. Similarly, the densities, which come from $r(m)$, also affect the chemical evolution rates. Thus we have a set of coupled non-linear differential equations to solve. Moreover, our set of equations is not yet complete. Most obviously, we haven't yet written down the reaction rates R_{ijk} or the way that nuclear energy generation rate q depends on them. Also, we have not specified yet how the pressure depends on density, temperature, or anything else. To solve for the evolution of a star, we will need to fill in these gaps.

Class 5 Notes: The Pressure of Stellar Material

As we discussed at the end of last class, we've written down the evolution equations, but we still need to specify how to fill in the things like pressures, reaction rates, rates of energy transfer through the star, etc. Today we're going to tackle the problem of pressure in stars, also known as the equation of state: an equation that specifies the pressure in a gas given its density and temperature. You're all familiar with the most common of these, the ideal gas law: $P = nk_B T$. While this works well under terrestrial conditions, inside a star things get a bit trickier. To derive the equation of state for a star, we will need to talk a little about the kinetic theory of gasses.

I. Why Stars Are Gasses

As a preliminary step, let's just confirm to ourselves that it is legitimate to think of material in a star is a gas, rather than a solid or a liquid. The distinguishing characteristic of a gas is that the potential energy associated with inter-particle forces at the typical particle-particle separation is small compared to the particles' thermal energy. In other words, a gas is a set of particles that are moving around at high enough speeds that the forces they exert on one another are negligible except on those rare occasions when they happen to pass extremely close to one another.

To check this for a star, consider a region of density ρ and temperature T, consisting of atoms with atomic mass $\mathcal{A}$ and atomic number $\mathcal{Z}$. The number density of the particles is $n = \rho/(\mathcal{A}m_{\mathrm{H}})$, so the typical distance between them must be

$$d = n^{-1/3} = \left(\frac{\mathcal{A}m_{\mathrm{H}}}{\rho}\right)^{1/3}.$$

The typical electromagnetic potential energy is therefore at most

$$E \simeq \frac{\mathcal{Z}^2 e^2}{d} = \mathcal{Z}^2 e^2 \left(\frac{\rho}{\mathcal{A}m_{\mathrm{H}}}\right)^{1/3},$$

where e is the electron charge. The "at most" is because this assumes that the potential energy comes from the full charge of the nuclei, neglecting any cancellation coming from electrons of opposite charge "screening" the nuclear charges.

To see how this compares to the thermal energy, i.e. to compute the ratio $E/k_B T$, ideally we would check at every point in the star, since both ρ and $k_B T$ change with position. However, we can get a rough idea of what the result is going to be if we use mean values of ρ and T. For a star of mass M and radius R, $\overline{\rho} = 3M/(4\pi R^3)$, and we proved using the virial theorem last week that

$$\overline{T} = \frac{\alpha}{3}\frac{\mu}{\mathcal{R}}\frac{GM}{R} = \frac{\alpha}{3}\frac{\mathcal{A}}{\mathcal{R}}\frac{GM}{R},$$

where α is a constant of order unity, and $\mu = \mathcal{A}$ is the mean atomic mass per particle. If the gas is fully ionized this will be lower, but the effect is not large. Substituting $\overline{\rho}$ and $\overline{T}$, and dropping constants of order unity, we find

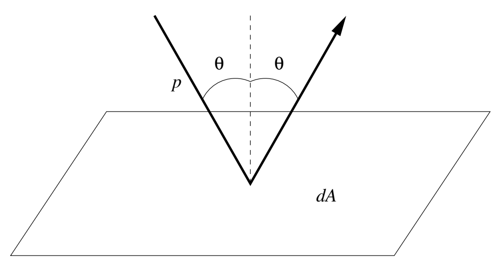
$$\frac{E}{k_B\overline{T}} \sim \frac{\mathcal{Z}^2 e^2}{G\mathcal{A}^{4/3} m_{\mathrm{H}}^{4/3} M^{2/3}} = 0.011 \frac{\mathcal{Z}^2}{\mathcal{A}^{4/3}} \left(\frac{M}{M_\odot}\right)^{-2/3}$$

Even for a pure iron star, $\mathcal{Z} = 26$ and $\mathcal{A} = 56$, we have $E/k_B\overline{T} = 0.035(M/M_\odot)^{-2/3}$. This may vary some within the star, but the general result is that $E \ll k_B T$, so something with the mass of a star is essentially always going to be a gas, unless something very strange happens (which it does in some exotic cases). In contrast, planets do not satisfy this condition. If we plug in $M = M_\oplus = 6.0 \times 10^{27}$ g and consider pure iron, we get a ratio of 167 – the center of the Earth is definitely not a gas!

II. The Kinetic Theory Model of Pressure

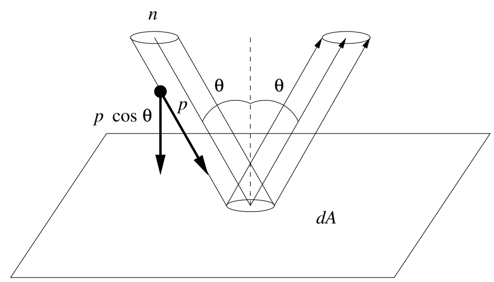
In order to compute the pressure of stellar material, we need to recall that pressure is the force exerted by a gas on a surface, such as the walls of its container, and that force is a change in momentum per unit time. In other words, the pressure is the momentum per unit time per unit area that a gas transfers to the walls of the vessel containing it. The reason there is a momentum transfer is that particles in the gas are moving around at random, and that some of them will strike the walls of the vessel, bounce off, and transfer momentum. We can compute the pressure by computing this momentum transfer.

To see what this implies, consider an immobile surface with a gas on one side of it, and focus on an area of that surface dA. First consider a single particle with momentum p approaching the surface at an angle θ relative to the normal and bouncing off it elastically.



A little geometry quickly shows that the momentum transferred to the surface is $2p\cos\theta$.

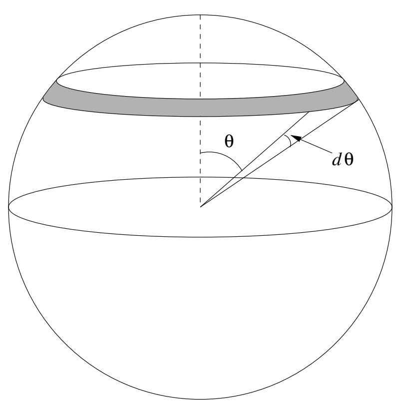
Now consider a beam of particles, all moving toward the surface at angle θ and bouncing off, and all moving with the same momentum p. Suppose the number density of particles in the beam is n, and that they are moving at velocity v (which is related to p in the usual way).



The rate at which particles strike the surface is $nv \cos\theta \, dA$. The $nv \, dA$ comes from multiplying the density of particles available by the speed at which they move by the area available to catch them. You can understand the factor of $\cos\theta$ in two equivalent ways. One is that only a fraction $\cos\theta$ of that velocity is in the direction perpendicular to the surface, and velocity parallel to the surface doesn't produce any collisions. Alternately, you can think about the projected area of the surface as seen by a particle in the beam, which is smaller than its total area by a factor $\cos\theta$. Since each collision transfers a momentum $2p\cos\theta$, the total rate at which the beam transfer momentum to the surface is

$$\frac{d^2 p_{\mathrm{surf}}}{dt \, dA} = 2nvp \cos^2\theta.$$

To generalize from the case of a beam to the case of a gas, we have to consider the fact that particles are moving in every possible direction. Continuing for the moment to imagine that all particles have the same momentum, the n be the total number density of particles, and let $dn(\theta)/d\theta$ be the number density of particles coming in at angles between θ and $\theta + d\theta$ relative to the normal. If the particle distribution is isotropic, then the fraction of particles at angle θ is just proportional to the fraction of the solid angle that lies between θ and $\theta + d\theta$.



The solid angle of the indicated strip is $2\pi \sin\theta \, d\theta$, as compared to 4π sr in total, so we must have that

$$\frac{dn(\theta)}{d\theta} = \frac{1}{2} n \sin\theta.$$

Thus the collision rate for particles coming in at angle θ is $(dn(\theta)/d\theta) v \cos\theta \, dA$, and each collision still transfers momentum $2p\cos\theta$. To get the total rate of momentum transfer we just have to multiply collision rate times momentum transfer and integrate this over all angles:

$$\begin{aligned}\frac{d^2 p_{\rm surf}}{dt\, dA} &= npv \int_0^{\pi/2} \cos^2\theta \sin\theta\, d\theta \\ &= npv \int_0^1 \cos^2\theta\, d\cos\theta \\ &= \frac{1}{3} npv\end{aligned}$$

Finally, to generalize this to a distribution of particles that aren't all moving at the same speed, we just have to integrate over their momentum distribution. We let $dn(p)/dp$ be the number of particles with momenta between p and $p + dp$. The pressure is then simply the momentum transferred to the surface per unit time per unit area, which we obtain simply by integrating over all the possible particle momenta:

$$P \equiv \frac{d^2 p_{\rm surf}}{dt\, dA} = \int_0^\infty \frac{1}{3} \frac{dn(p)}{dp} pv\, dp$$

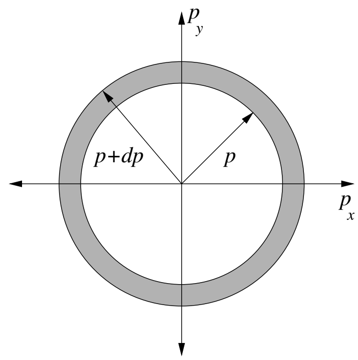
III. Types of Pressure

We have now written the pressure of a gas in terms of the momentum distribution of its particles. This required a lengthy mathematical exercise, but this was worth it because this enables us to define the pressure in all sorts of complicated situations where we can't blindly apply the ideal gas law. There are several relevant for stars that we'll consider next.

A. Re-Derivation of the Ideal Gas Law

The first step in our analysis is to use this kinetic model of pressure to re-derive the ideal gas law. This will provide a guide to how to proceed when dealing with more complicated situations. To derive the ideal gas law, we begin with a gas whose particles all have mass m, and where the particles have a Maxwell-Boltzmann velocity distribution. This is the same as the Boltzmann distribution we wrote down during the first class: the probability that a particle is in a state with energy E is proportional to $e^{-E/k_B T}$.

To derive the momentum distribution from this, it is helpful to first think about things two-dimensionally. A gas particle can have any vector momentum, $\mathbf{p}$. In two dimensions, this has two components, p_x and p_y. We can think of the momentum of a given particle as corresponding to a point in the two-dimensional plane of p_x and p_y.



We want to know the probability that a particle will be at a point (p_x, p_y) in this plane, and the Boltzmann distribution can tell us. The energy and momentum of a particle are related by

$$E = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2}{2m},$$

where p is the magnitude of the vector $\mathbf{p}$. Thus the probability of being at a point (p_x, p_y) is proportional to $e^{-(p_x^2+p_y^2)/(2mk_BT)}$. The probability that the magnitude of the momentum will fall in the ring between p and $p + dp$ is just the probability of being at a point (p_x, p_y) in the ring times the area of the ring, which is $2\pi p\, dp$. Thus in two dimensions we have

$$\frac{dn(p)}{dp} \propto 2\pi p e^{-p^2/(2mk_BT)}.$$

The three-dimensional generalization is obvious: instead of a ring of area $2\pi p\, dp$, we now have a shell of volume $4\pi p^2\, dp$. Thus in three dimensions the momentum distribution for the particles must follow

$$\frac{dn(p)}{dp} \propto 4\pi p^2 e^{-p^2/(2mk_BT)}.$$

To get the normalization constant, we just require that, when we integrate over all momenta, we get the right number of particles. Thus we say that $dn(p)/dp = k \cdot 4\pi p^2 e^{-p^2/(2mk_BT)}$ and solve for the constant k by requiring that

$$\begin{aligned} n &= 4\pi k \int_0^\infty p^2 e^{-p^2/(2mk_BT)}\, dp \\ &= 4\pi k (2mk_BT)^{3/2} \int_0^\infty q^2 e^{-q^2}\, dq \\ &= k(2\pi m k_B T)^{3/2} \end{aligned}$$

where in the second step we have made the substitution $q = p/\sqrt{2mk_BT}$, and in the third step we evaluated the integral to get $\sqrt{\pi}/4$ – the integral is fairly straightforward to do by standard tricks. This gives us k, which in turn gives us $dn(p)/dp$:

$$\frac{dn(p)}{dp} = \frac{4\pi n}{(2\pi m k_B T)^{3/2}} p^2 e^{-p^2/(2mk_BT)}$$

Given this result, computing the pressure is just a matter of plugging in and evaluating the integral:

$$\begin{aligned} P &= \int_0^\infty \frac{1}{3}\left[\frac{4n}{\pi^{1/2}(2mk_BT)^{3/2}}p^2 e^{-p^2/(2mk_BT)}\right] p\left(\frac{p}{m}\right)\,dp \\ &= \frac{4n}{3\pi^{1/2}m}(2mk_BT)\int_0^\infty q^4 e^{-q^2}\,dq \\ &= nk_BT, \end{aligned}$$

where in the last step we again evaluated the integral, this time to $3\sqrt{\pi}/8$. Thus we have successfully re-derived the ideal gas law from first principles using the kinetic theory of gasses.

B. Gasses with Multiple Species

The first complication to add to this story is what happens if we have multiple types of particles, each with a different mass. This is relevant to a gas that contains a mixture of hydrogen and helium, for example. It is also relevant in a fully ionized gas, where the ions and electrons move separately, and obviously their masses are quite different. Fortunately, the kinetic description makes the result obvious: each species follows the Boltzmann distribution, and the sum of the momentum transferred to a surface is simply the sum of the momenta transferred by the particles of each species, each of which is given by nk_BT. Thus, if we have N species present in the gas, then the total pressure is simply

$$P = \left(\sum_{i=1}^N n_i\right) k_BT.$$

We can write this equivalently in terms of the mass fraction and mass. If we let $\mathcal{A}_i m_\mathrm{H}$ be the mass per particle of species i and X_i be the fraction of the mass at a given point that comes from species i, then, as before, we have

$$n_i = \frac{X_i}{\mathcal{A}_i m_\mathrm{H}}\rho,$$

and therefore we can write the pressure as

$$P = \left(\sum_{i=1}^N \frac{X_i}{\mathcal{A}_i}\right)\rho\mathcal{R}T$$

For convenience we define

$$\frac{1}{\mu} = \sum_{i=1}^N \frac{X_i}{\mathcal{A}_i},$$

where μ is the mean mass (measured in units of hydrogen masses) per particle, so that the ideal gas law becomes

$$P = \frac{\mathcal{R}}{\mu}\rho T.$$

If we only include ions (not electrons) in the sum, then we get the pressure due to ions alone, and we write μ in this case as μ_I, for the mean mass per particle of ions. Since the Sun is mostly hydrogen and helium, it is convenient to express its composition in terms of the fraction of the mass that is hydrogen, the fraction that is helium, and the fraction that is everything else – the everything else we call metals. Note that, to an astronomer, carbon, oxygen, and neon are all metals! We define X as the hydrogen mass fraction, Y as the helium mass fraction, and Z as the metal mass fraction. For the Sun, $X = 0.707$, $Y = 0.274$, and $Z = 1 - X - Y = 0.019$.

We can write μ_I in terms of these definitions:

$$\frac{1}{\mu_I} = \frac{X}{1} + \frac{Y}{4} + \frac{Z}{\langle \mathcal{A} \rangle_{\text{metals}}},$$

where $\langle \mathcal{A} \rangle_{\text{metals}}$ is the mean atomic mass of the metals, which is about 20 in the Sun. Thus for the Sun $\mu_I = 1.29$.

We can similarly calculate the pressure due to electrons. In the outer layers of a star where it is cold there are none, but we showed using the virial theorem that in the stellar interior the gas is close to fully ionized. Thus there is one free electron per proton. If n_i is the number density of ions of species i, then the number density of electrons is

$$n_e = \sum_i \mathcal{Z}_i n_i = \frac{\rho}{m_{\text{H}}} \sum_i X_i \frac{\mathcal{Z}_i}{\mathcal{A}_i}.$$

Again, for convenience we give this sum a name:

$$\frac{1}{\mu_e} = \sum_i X_i \frac{\mathcal{Z}_i}{\mathcal{A}_i}.$$

The meaning of $1/\mu_e$ is that it is the average number of free electrons per nucleon, meaning per proton or neutron. In terms of our X, Y, and Z numbers,

$$\frac{1}{\mu_e} = X + \frac{Y}{2} + Z \left\langle \frac{\mathcal{Z}}{\mathcal{A}} \right\rangle_{\text{metals}},$$

where the term $\langle \mathcal{Z}/\mathcal{A} \rangle_{\text{metals}}$ represents the ratio of electrons (or protons) averaged over all the metal atoms.This is roughly 1/2, so to good approximation

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

since $Z = 1 - X - Y$. Thus for the Sun $\mu_e = 1.17$.

Thus the pressures of the ions and the electrons are $P_I = (\mathcal{R}/\mu_I)\rho T$ and $P_e = (\mathcal{R}/\mu_e)\rho T$, so the total pressure is $P = P_I + P_e = (\mathcal{R}/\mu)\rho T$, where

$$\frac{1}{\mu} = \frac{1}{\mu_I} + \frac{1}{\mu_e}.$$

C. Relativistic Gasses and Radiation

The rule that pressures from different gasses just add is fairly intuitive, and one could probably have guessed it without the kinetic theory. We do need the kinetic theory, however, to generalize the concept of pressure to gasses that are not ideal, classical gasses. The simplest generalization to make is to gases that are relativistic, meaning that the particles within them are moving at close to the speed of light. This occurs in some extreme stars. We will limit ourselves to considering gases in the extremely relativistic limit, where most particles have speeds very close to c. The partially relativistic case is conceptually the same, but involves a great deal more algebra.

For a relativistic gas the pressure integral is exactly the same as for a non-relativistic one. There are only two differences. The first is that the velocity v in the pressure integral is nearly c, the speed of light. The second is that energy and momentum are no longer related by $E = p^2/(2m)$. That is the relationship that applies when the particle's rest energy is much greater than its kinetic energy. In the extreme relativistic limit the opposite is true, and the particle's rest energy is negligible. For such a particle, energy and momentum are related by

$$E = pc.$$

For electrons, the transition between the two regimes occurs when $(3/2)k_BT$ becomes comparable to $(1/2)m_ec^2$, the electron rest energy. Thus, an electron gas becomes relativistic at a temperature of roughly

$$T_{\rm rel} \sim \frac{m_e c^2}{3k_B} = 2 \times 10^9 \text{ K}.$$

With these two changes, the procedure is exactly the same as for a non-relativistic gas. The momentum distribution is

$$\frac{dn(p)}{dp} = 4\pi k p^2 e^{-E/k_BT} = 4\pi k p^2 e^{-pc/k_BT},$$

and the constant k is again determined by requiring that

$$\begin{aligned} n &= 4\pi k \int_0^\infty p^2 e^{-pc/k_BT}\, dp = 8\pi k \left(\frac{k_BT}{c}\right)^3 \\ k &= \left(\frac{c}{k_BT}\right)^3 \frac{n}{8\pi} \end{aligned}$$

The pressure is

$$\begin{aligned} P &= \frac{n}{6} \frac{c^4}{(k_BT)^3} \int_0^\infty p^3 e^{-pc/k_BT}\, dp \\ &= nk_BT \end{aligned}$$

This is exactly the same as for a non-relativistic gas.

However, for relativistic gasses we have a complication which is not present for non-relativistic ones, which is that the number of particles is not necessarily fixed. Instead, when particles are moving around with an energy comparable their rest energy, collisions can create or destroy particles. Thus n and $dn(p)/dp$ are no longer fixed, and instead becomes functions of T. Our result is valid only for fixed n.

We will not solve this problem in general, but we will solve it for one particular type of relativistic gas: radiation. We can think of photons as a relativistic gas, since photons move at the speed of light and have energies much larger than their rest energy (which is zero). To compute the pressure of a photon gas, we need to know how the number density of photons and its distribution in momentum, $dn(p)/dp$, varies with the temperature T. You will see this result in your quantum mechanics or statistical mechanics class, and I will not re-derive it. The distribution is known as Planck's Law, and it states

$$\begin{aligned} \frac{dn(\nu)}{d\nu} &= \frac{8\pi\nu^2}{c^3}\frac{1}{e^{h\nu/k_BT}-1} \\ \frac{dn(p)}{dp} &= \frac{8\pi p^2}{h^3}\frac{1}{e^{pc/k_BT}-1}. \end{aligned}$$

The first form is in terms of the frequency, and the second is in terms of the momentum. The two are related by $E = pc = h\nu$. This distribution, known as Planck's Law, was first found empirically by Max Planck in 1901 and was finally understood theoretically by Satyendra Nath Bose in 1924. An interesting historical aside: Bose was a professor at the University of Dhaka in India, and when he first produced this result, no journal in Europe was willing to accept his paper. Eventually he sent the paper to Einstein, who recognized its significance and wrote a companion paper in support of Bose's. The two were then published together, giving rise to what is known today as Bose-Einstein statistics.

Given this distribution, we derive the pressure as before:

$$P = \frac{1}{3}\int_0^\infty c\frac{h\nu}{c}\frac{dn(\nu)}{d\nu}dp = \frac{1}{3}aT^4,$$

where

$$a = \frac{8\pi^5 k_B^4}{15c^3h^3} = \frac{4\sigma}{c}$$

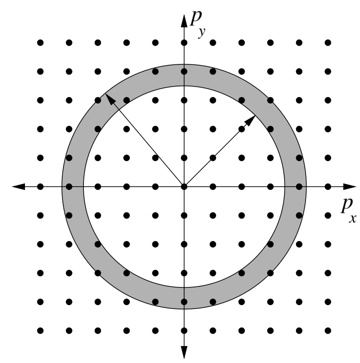
is known as the radiation constant.

You will generalize to the case of a mixed fluid of gas and radiation on your homework.

D. Degenerate Gasses

The second generalization we will consider is to gasses where quantum mechanical effects become important. These are called degenerate gasses. A full theory of degenerate gasses and their pressures is beyond what we will do in this class, but we will deal with one limiting case, and we can use that to provide a good approximation to the quantum effect.

Consider again the picture of the (p_x, p_y) plane, where we describe every particle's momentum in terms of a position in the plane. In classical mechanics, a particle can occupy any position in the (p_x, p_y) plane, but quantum mechanics tells us that in reality there actually only discrete, quantized values of p_x and p_y that particles are allowed to have – in effect there is a grid in (p_x, p_y)-space, and particles can only be found on the grid points.



Most of the time this doesn't matter, because the grid points are packed so densely that they might as well be a continuum. Particles can't really be anywhere, but they can be nearly anywhere. However, there are some situations where it does matter. In the classical picture, the probability of being at a given point is e^{-E/k_BT}, where E is the energy associated with that point. This distribution continues to apply in quantum mechanics. If T is small, then E/k_BT is a big number for most grid points, so the particles all try to crowd into the points close to the origin, where E is small. As a result, they're all trying to occupy the same few grid points. However, the Pauli exclusion principle says that no two fermions (a category of particles that includes electrons) can occupy the same quantum state. For electrons, which can be spin up or spin down, no more than two can sit at any grid point. Because the electrons can't all pack into the few central grid points, they are forced to occupy a wider range of momenta than classical mechanics would suggest they should. As a result, their pressure is much higher than you would expect based on classical mechanics.

To know when this effect is important, you need to know what the density of grid points is, since that will dictate when you start to have problems with too man electrons trying to sit at the same site. Actually calculating this rigorously is something to be left for your quantum mechanics class, but you can get the basic result from the Heisenberg uncertainty principle. The most common way of stating this is that there is an irreducible uncertainty in the product of a particle's momentum and its position:

$$\Delta x \Delta p \geq h,$$

where $h = 6.63 \times 10^{-27}$ erg s is Planck's constant. In 3D, we can write this as

$$\Delta V \Delta^3 p \geq h^3.$$

This relates the uncertainty in the volume where a particle is located to the uncertainty in its 3D momentum. An equally valid interpretation of the Heisenberg uncertainty principle is that it is tell us how tightly packed the quantum grid points are. If we have a volume of space ΔV, then the grid points for particles in that volume each occupy a space $\Delta^3 p = h^3/\Delta V$ in the (p_x, p_y, p_z)-space.

This tells us that quantum effects are going to start become important in two circumstances. One is when the temperature is low, and all the particles try to pack into the inner few gridpoints. The other is when the density of particles is high. This is because a high density means a large number of particles in a small space ΔV. However, when ΔV is small, then the quantum grid points are spaced a larger distance apart, which means there are few sites available for particles to occupy.

To apply this idea to calculating the pressure of a gas, consider the limit of a gas where the temperature approaches 0. In this case, the particles will try to crowd as close to the origin in (p_x, p_y, p_z)-space as possible. One can imagine placing the particles at the grid points. The first two electrons will go at the grid point closest to the origin, the next two and the second closest point, and so forth until all the electrons are used up. Thus the particles fill a circle of radius p_0 in the (p_x, p_y)-plane in the 2D case, or a sphere in the (p_x, p_y, p_z)-volume in the 3D case. All the grid points with momentum $p < p_0$ are occupied, and all the grid points further from the origin than p_0 will be empty. A gas of this sort is fully degenerate, meaning that the particles are packed as closely as possible.

To get the pressure in this fully degenerate state, we need to know the momentum distribution $dn(p)/dp$ – that is, we need to know how many electrons there are inside the shell from p to $p + dp$. For the fully degenerate case this is easy. If $p > p_0$, then $dn(p)/dp = 0$, because all the grid points at $p > p_0$ are empty. If $p < p_0$, then all the grid points are full, so the number of electrons is just twice the number of grid points within the shell (since there are two electrons per grid point). Since the shell has volume $4\pi p^2\, dp$, and each grid point takes up a volume $\Delta^3 p = h^3/\Delta V$, the number of electrons inside the shell is

$$N_e = 2\frac{4\pi p^2\, dp}{\Delta^3 p} = \frac{2}{h^3} 4\pi p^2\, dp \Delta V.$$

To change this to a number density, we just divide both sides by ΔV, which gives

$$\frac{dn(p)}{dp} = \frac{2}{h^3} 4\pi p^2$$

To figure out the momentum p_0 where this distribution stops, we simply set it by the condition that, when we integrate over all momenta, we get the right number of particles:

$$\begin{aligned}
n &= \int_0^{p_0} \frac{2}{h^3} 4\pi p^2 \, dp \\
&= \frac{8\pi}{3h^3} p_0^3 \\
p_0 &= \left(\frac{3h^3 n}{8\pi}\right)^{1/3}.
\end{aligned}$$

Finally, we are in a position to calculate the pressure. Suppose that all the particles have mass m. Then

$$\begin{aligned}
P &= \frac{1}{3}\int_0^{\infty} \frac{dn(p)}{dp} pv \, dp \\
&= \frac{1}{3}\int_0^{p_0} \left(\frac{2}{h^3} 4\pi p^2\right) p \left(\frac{p}{m}\right) dp \\
&= \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20m} n^{5/3} \\
&= \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20 m_e m_{\mathrm{H}}^{5/3}} \left(\frac{\rho}{\mu_e}\right)^{5/3},
\end{aligned}$$

where in the last step we have assumed that the particles are electrons, and we have inserted the electron density for a fully ionized gas. The result applies equally well to protons and neutrons, since they are fermions too, but since the degeneracy pressure varies as $1/m$, the much higher mass of these particles means that their degeneracy pressure is much lower. Thus we are generally concerned only with electrons. The combination of constants in front of the ρ term comes up often enough that it is useful to compute it. We define

$$K_1' = \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20 m_e m_{\mathrm{H}}^{5/3}} = 1.00 \times 10^{13} \text{ dyn cm}^{-2} \left(\text{g cm}^{-3}\right)^{-5/3},$$

so that $P = K_1' (\rho/\mu_e)^{5/3}$.

This is the pressure of a fully degenerate gas, and it represents a lower limit on the pressure, which is achieved at zero temperature. In reality at any finite temperature the pressure is higher than this. As a very crude approximation, we can write the electron pressure as

$$P_e = \max\left[\mathcal{R}\frac{\rho}{\mu_e} T, K_1' \left(\frac{\rho}{\mu_e}\right)^{5/3}\right],$$

i.e. the electron pressure is either the thermal pressure or the degeneracy pressure, whichever is greater. In reality the transition between the two is more smooth than this, and can be calculated quantum-mechanically. The transition between degenerate and non-degenerate occurs roughly where these two pressures are equal, which requires that

$$\frac{\rho}{\mu_e} = \left(\frac{\mathcal{R}}{K_1'}T\right)^{3/2} = 750\left(\frac{T}{10^7\ \mathrm{K}}\right)^{3/2}\ \mathrm{g\ cm^{-3}}.$$

One subtle but important thing to notice is that the degeneracy pressure, unlike the thermal pressure, does not depend on the gas temperature – a degenerate gas has essentially fixed pressure until the temperature rises high enough to make it non-degenerate. We will see that this has profound consequences for the evolution of degenerate stars. It causes some of them to explode.

E. Relativistic Degenerate Gasses

In some very dense stars, the gas is degenerate, and it is also dense enough so that the electrons have speeds that approach the speed of light. In this case we have a relativistic degenerate gas. Again, the procedure to calculate the pressure is the same, except that the velocity is now c, and the energy and momentum are related by $E = pc$. In the degenerate case, however, the change in the relationship between energy and momentum doesn't matter, because the momentum distribution is dictated by how many particles you can pack into a given volume in momentum-space, not by Boltzmann factors. Thus $dn(p)/dp$ is the same as for the non-relativistic case, and we have

$$P = \frac{1}{3}\int_0^{p_0}\left(\frac{2}{h^3}4\pi p^2\right)pc\,dp = \frac{2\pi c}{3h^3}p_0^4 = \left(\frac{3}{\pi}\right)^{1/3}\frac{hc}{8}n^{4/3} = \left(\frac{3}{\pi}\right)^{1/3}\frac{hc}{8m_{\mathrm{H}}^{4/3}}\left(\frac{\rho}{\mu_e}\right)^{4/3}.$$

Again, in the last step we have assumed that the particles in question are electrons.

As with the non-relativistic degenerate case, it is convenience to give the collection of constants a name, so we say that

$$P = K_2'(\rho/\mu_e)^{4/3},$$

where $K_2' = 1.24 \times 10^{15}$ dyn cm^{-2} (g cm^{-3})$^{-4/3}$. Again, note that tthe pressure does not depend on temperature.

The condition for a degenerate gas to be relativistic is that p_0 must be large enough so that the kinetic energy is comparable to the rest energy of the electron. Thus the gas becomes relativistic when $p_0^2/(2m_e) \sim m_e c^2$. This requires that

$$\frac{\rho}{\mu_e} = \frac{16\pi\sqrt{2}}{3}\frac{m_{\mathrm{H}} m_e^3 c^3}{h^3} = 3 \times 10^6\ \mathrm{g\ cm^{-3}}.$$

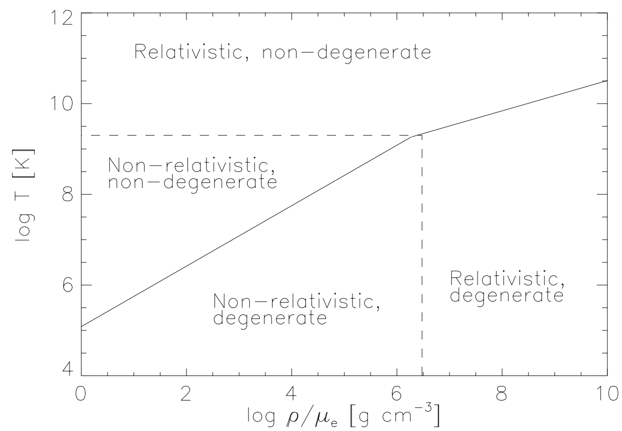
The condition for a relativistic gas to become degenerate is that the degeneracy pressure equal the gas pressure, which requires that

$$\frac{\rho}{\mu_e}\mathcal{R}T = K_2'\left(\frac{\rho}{\mu_e}\right)^{4/3}$$

$$\frac{\rho}{\mu_e} = \left(\frac{\mathcal{R}T}{K_2'}\right)^3 = 0.3\left(\frac{T}{10^7\text{ K}}\right)^3 \text{ g cm}^{-3}$$

F. Regimes of Pressure

We can summarize the four cases we have just derived for relativistic and non-relativistic, degenerate and non-degenerate gasses in a single figure, by combining the conditions for switching between the regimes.



The plot indicates where each case applies.

One thing that initially seems surprising about this plot is that it seems to suggest that fluid at a density of 1 g cm^{-3} should be degenerate unless its temperature is more than 10^5 K or so. Does this mean that water at room temperature is degenerate? The answer is no. Recall that this plot is for electrons. A gas of free electrons with the density of water and a temperature comparable to room temperature would indeed be degenerate. However, water molecules are not free electrons. The H_2O molecule has a mass of $18m_{\rm H}$, which is 3.3×10^4 electron masses. Recall that degeneracy pressure varies as $1/m$, so the degeneracy pressure of water is $33,000$ times smaller than that of electrons. If you use the mass of a water molecule to compute K_1' instead of the mass of an electron, you will find that at a density of 1 g cm^{-3}, degeneracy does not set in until the temperature drops below 4 K.

This does make an important point, however: for fully ionized gasses, it is much easier to be degenerate than you might think.

Class 6 Notes: Internal Energy and Radiative Transfer

In the last class we used the kinetic theory of gasses to understand the pressure of stellar material. The kinetic view is essential to generalizing the concept of pressure to the environments found in stars, where gas can be relativistic, degenerate, or both. The goal of today's class is to extend that kinetic picture by thinking about pressure in stars in terms of the associated energy content of the gas. This will let us understand how the energy flow in stars interacts with gas pressure, a crucial step toward building stellar models.

I. Pressure and Energy

A. The Relationship Between Pressure and Internal Energy

The fundamental object we dealt with in the last class was the distribution of particle mometa, $dn(p)/dp$, which we calculated from the Boltzmann distribution. Given this, we could compute the pressure. However, this distribution also corresponds to a specific energy content, since for particles that don't have internal energy states, internal energy is just the kinetic energy of particle motions. Given a distribution of particle momenta $dn(p)/dp$ in some volume of space, the corresponding density of energy within that volume of space is

$$e = \int_0^\infty \frac{dn(p)}{dp}\epsilon(p)\,dp,$$

where $\epsilon(p)$ is the energy of a particle with momentum p. It is often more convenient to think about the energy per unit mass than the energy per unit volume. The energy per unit mass is just e divided by the density:

$$u = \frac{1}{\rho}\int_0^\infty \frac{dn(p)}{dp}\epsilon(p)\,dp.$$

The kinetic energy of a particle with momentum p and rest mass m is

$$\epsilon(p) = mc^2\left(\sqrt{1+\frac{p^2}{m^2c^2}}-1\right).$$

This formula applies regardless of p. In the limit $p \ll mc$ (i.e. the non-relativistic case), we can Taylor expand the square root term to $1 + p^2/(2m^2c^2)$, and we recover the usual kinetic energy: $\epsilon(p) = p^2/(2m)$. In the limit $p \gg mc$ (the ultra-relativistic case), we can drop the plus 1 and the minus 1, and we get $\epsilon(p) = pc$.

Plugging the non-relativistic, non-degenerate values for $\epsilon(p)$ and $dn(p)/dp$ into the integral for u and evaluating gives

$$u = \frac{1}{\rho}\int_0^\infty \frac{4\pi n}{(2\pi m k_B T)^{3/2}} p^2 e^{-p^2/(2mk_BT)}\left(\frac{p^2}{2m}\right) dp$$

$$\begin{aligned}
&= \frac{2\pi}{m^2(2\pi m k_B T)^{3/2}} \int_0^\infty p^4 e^{-p^2/(2mk_BT)}\, dp \\
&= \frac{4}{m\pi^{1/2}} k_B T \int_0^\infty q^4 e^{-q^2}\, dq \\
&= \frac{3}{2m} k_B T \\
&= \frac{3}{2}\frac{P}{\rho},
\end{aligned}$$

where the integral over q evaluates to $3\sqrt{\pi}/8$. This is the same as the classic result that an ideal gas has an energy per particle of $(3/2)k_BT$. Since the pressure and energies are simply additive, it is clear that the result $u = (3/2)(P/\rho)$ applies even when there are multiple species present.

Applying the same procedure in the relativistic, non-degenerate limit gives

$$\begin{aligned}
u &= \frac{1}{\rho}\int_0^\infty \left(\frac{c}{k_BT}\right)^3 \frac{n}{2} p^2 e^{-pc/k_BT}(pc)\, dp \\
&= \frac{c^4}{2m(k_BT)^3}\int_0^\infty p^3 e^{-pc/k_BT}\, dp \\
&= \frac{3}{m} k_B T \\
&= 3\frac{P}{\rho}.
\end{aligned}$$

It is straightforward to show that this result applies to radiation too, by plugging $\epsilon = h\nu$ for the energy and the Planck distribution for $dn(\nu)/d\nu$. Note that this implies that the volume energy density of a thermal radiation field is

$$e_{\text{rad}} = aT^4 = 3P_{\text{rad}}.$$

For the non-relativistic, degenerate limit we have a step-function distribution that has a constant value $dn(p)/dp = (8\pi/h^3)p^2$ out to some maximum momentum $p_0 = [3h^3 n/(8\pi)]^{1/3}$, so the energy is

$$\begin{aligned}
u &= \frac{1}{\rho}\int_0^{p_0} \frac{8\pi}{h^3} p^2 \left(\frac{p^2}{2m}\right) dp \\
&= \frac{4\pi}{5m\rho h^3} p_0^5 \\
&= \left(\frac{3}{\pi}\right)^{2/3} \frac{3h^2 n^{5/3}}{40\rho m} \\
&= \frac{3}{2}\frac{P}{\rho},
\end{aligned}$$

exactly as in the non-dengerate case.

Finally, for the relativistic degenerate case we have

$$\begin{aligned} u &= \frac{1}{\rho}\int_0^{p_0} \frac{8\pi}{h^3} p^2 \left(pc\right)\, dp \\ &= \left(\frac{3}{\pi}\right)^{1/3} \frac{3}{8}\frac{hcn^{4/3}}{\rho} \\ &= 3\frac{P}{\rho} \end{aligned}$$

B. Adiabatic Processes and the Adiabatic Index

Part of the reason that internal energies are interesting to compute is because of the problem of adiabatic processes. An adiabatic process is one in which the gas is not able to exchange heat with its environment or extract it from internal sources (like nuclear burning), so any work it does must be balanced by a change in its internal energy. The classic example of this is a gas that is sealed in an insulated box, which is then compressed or allowed to expand. In many circumstances we can think of most of the gas in a star (that outside the region where nuclear burning takes place) as adiabatic. It can exchange energy with its environment via radiation, but, as we have previously shown, the radiation time is long compared to the dynamical time. Thus any process that takes place on timescale shorter than a Kelvin-Helmholtz timescale can be thought of as adiabatic.

To understand how an adiabatic gas behaves, we use the first law of thermodynamics, which we derived a few classes back:

$$\frac{du}{dt} + P\frac{d}{dt}\left(\frac{1}{\rho}\right) = q - \frac{\partial F}{\partial m} = 0,$$

where we have set the right-hand side to zero under the assumption that the gas is adiabatic, so it does not exchange heat with its environment and does not generate heat by nuclear fusion. We have just shown that for many types of gas $u = \phi P/\rho$, where ϕ is a constant that depends on the type of gas. If we make this substitution in the first law of thermodynamics, then we get

$$0 = \phi P \frac{d}{dt}\left(\frac{1}{\rho}\right) + \phi\frac{1}{\rho}\frac{d}{dt}P + P\frac{d}{dt}\left(\frac{1}{\rho}\right) = (\phi+1)P\frac{d}{dt}\left(\frac{1}{\rho}\right) + \phi\frac{1}{\rho}\frac{d}{dt}P$$

This implies that

$$\begin{aligned} \frac{dP}{dt} &= -\frac{\phi+1}{\phi}\rho P \frac{d}{dt}\left(\frac{1}{\rho}\right) \\ &= \left(\frac{\phi+1}{\phi}\right)\frac{P}{\rho}\frac{d\rho}{dt} \\ \frac{dP}{P} &= \left(\frac{\phi+1}{\phi}\right)\frac{d\rho}{\rho} \\ \ln P &= \gamma_a \ln\rho + \ln K_a \\ P &= K_a \rho^{\gamma_a}, \end{aligned}$$

where $\gamma_a = (\phi + 1)/\phi$ and K_a is a constant. Thus we have shown that, for an adiabatic gas, the pressure and density are related by a powerlaw.

The constant of integration K_a is called the adiabatic constant, and it is determined by the entropy of the gas. The exponent γ_a is called the adiabatic index, and it is a function solely of the type of gas: all monatomic ideal gasses have $\gamma_a = (3/2 + 1)/(3/2) = 5/3$, whether they are degenerate or not. All relativistic gasses have $\gamma_a = 4/3$, whether they are degenerate or not.

The adiabatic index is a very useful quantity for know for a gas, because it describes how strongly that gas resists being compressed – it specifies how rapidly the pressure rises in response to an increase in density. The larger the value of γ_a, the harder it is to compress a gas. In a few weeks, we will see that the value of the adiabatic index for material in a star has profound consequences for the star's structure. For most stars γ_a is close to 5/3 because the gas within them is non-relativistic. However, as the gas becomes more relativistic, γ_a approaches 4/3, and resistance to compression drops. When that happens the star is not long for this world. A similar effect is responsible for star formation. Under some circumstances, radiative effects cause interstellar gas clouds to act as if they had $\gamma_a = 1$, which means very weak resistance to compression. The result is that these clouds collapse, which is how new stars form.

C. The Adiabatic Index for Partially Ionized Gas

The adiabatic index is fairly straightforward for something like a pure degenerate or non-degenerate gas, but the idea can be generalized to considerably more complex gasses. One case that is of particular interest is a partially ionized gas. This will be the situation in the outer layers of a star, where the temperature falls from the mean temperature, where the gas is fully ionized, to the surface temperature, where it is fully neutral. This case is tricky because the number of free gas particles itself becomes a function of temperature, and because the potential energy associated with ionization and recombination becomes an extra energy source or sink for the gas.

Consider a gas of pure hydrogen within which the number density of neutral atoms is n_0 and the number densities of free protons and electrons are $n_p = n_e$. The number density of all atoms regardless of their ionization state is $n = n_e + n_0$. We define the ionization fraction as

$$x = \frac{n_e}{n},$$

i.e. x is the fraction of all the electrons present that are free, or, equivalently, the fraction of all the protons present that do not have attached electrons. The pressure in the gas depends on how many free particles there are:

$$P = n_e k_B T + n_p k_B T + n_0 k_B T = (1 + x) n k_B T = (1 + x)\mathcal{R}\rho T$$

Thus the pressure at fixed temperature is higher if the gas is more ionized, because there are more free particles.

The number densities of free electrons and protons are determined by the Saha equation, which we encountered in the first class. As a reminder, the Saha equation is that the number density of ions at ionization states i and $i+1$ are related by

$$\frac{n_{i+1}}{n_i} = \frac{2Z_{i+1}}{n_e Z_i} \left(\frac{2\pi m_e k_B T}{h^2}\right)^{3/2} e^{-\chi/k_B T},$$

where χ is the ionization potential and Z_i and Z_{i+1} are the partition functions of the two states. Applying this equation to the neutral and ionized states of hydrogen gives

$$\frac{n_e^2}{n_0} = \frac{2}{Z_0} \left(\frac{2\pi m_e k_B T}{h^2}\right)^{3/2} e^{-\chi/k_B T},$$

where Z_0 is the partition function of the neutral hydrogen, and we have set $Z_1 = 1$ because the ionized hydrogen has a partition function of 1. Making the substitution $n_e = xn$ and $n_0 = (1-x)n$, the equation becomes

$$\frac{x^2}{1-x} n = \frac{2}{Z_0} \left(\frac{2\pi m_e k_B T}{h^2}\right)^{3/2} e^{-\chi/k_B T}.$$

Finally, if we use the pressure relation to write $n = P/[(1+x)k_B T]$, we arrive at

$$\frac{x^2}{1-x^2} = \left(\frac{2}{h^3 Z_0}\right) \frac{(2\pi m_e)^{3/2} (k_B T)^{5/2}}{P} e^{-\chi/k_B T}.$$

This equation gives us the ionization fraction in terms of P and T.

To compute the adiabatic index, we must first know the internal energy. For a partially ionized gas this has two components. The first is the standard kinetic one, $(3/2)(P/\rho)$. However, we must also consider the ionization energy: neutral atoms have a potential energy that is lower than that of ions by an amount $\chi = 13.6$ eV. Thus the total specific internal energy including both kinetic and potential parts is

$$u = \frac{3}{2}\frac{P}{\rho} + \frac{\chi}{m_{\rm H}} x.$$

The second term says that the potential energy per unit mass associated with ionization is 13.6 eV per hydrogen mass, multiplied by the fraction of hydrogen atoms that are ionized. If none are then this term is 0, and if they all are, then the energy is 13.6 eV divided by per hydrogen atom mass.

Now we plug into the first law of thermodynamics for an adiabatic gas, $du/dt + P(d/dt)(1/\rho) = 0$:

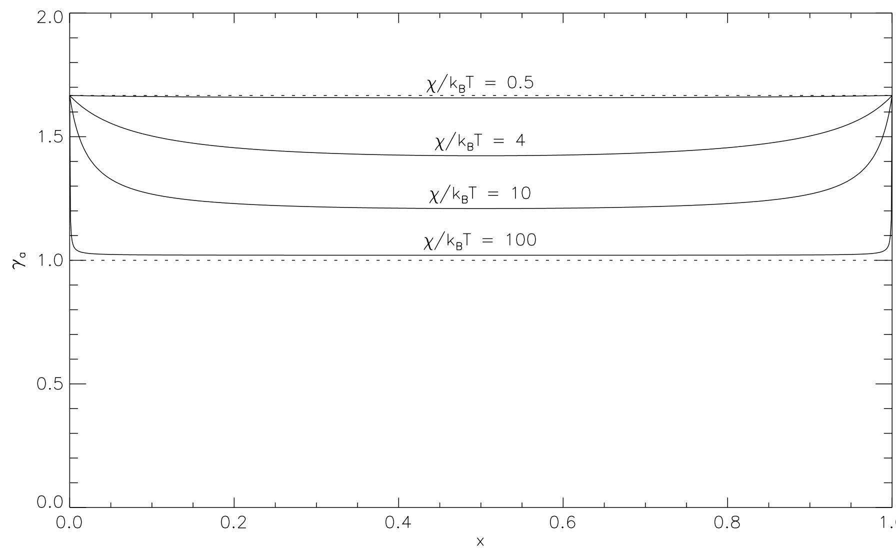
$$\begin{aligned}
\frac{3}{2}\left(\frac{1}{\rho}\right)\frac{dP}{dt} - \frac{3}{2}\frac{P}{\rho^2}\frac{d\rho}{dt} + \frac{\chi}{m_{\rm H}}\frac{\partial x}{\partial \rho}\frac{d\rho}{dt} + \frac{\chi}{m_{\rm H}}\frac{\partial x}{\partial P}\frac{dP}{dt} - \frac{P}{\rho^2}\frac{d\rho}{dt} &= 0 \\
\left[\frac{3}{2} + \chi n \left(\frac{\partial x}{\partial P}\right)\right]\frac{dP}{P} - \left[\frac{5}{2} - \frac{\chi n \rho}{P}\left(\frac{\partial x}{\partial \rho}\right)\right]\frac{d\rho}{\rho} &= 0 \\
\left[\frac{3}{2} + \frac{\chi}{k_B T}\left(\frac{P}{1+x}\right)\left(\frac{\partial x}{\partial P}\right)\right]\frac{dP}{P} - \left[\frac{5}{2} - \frac{\chi}{k_B T}\left(\frac{\rho}{1+x}\right)\left(\frac{\partial x}{\partial \rho}\right)\right]\frac{d\rho}{\rho} &= 0.
\end{aligned}$$

In the second step we multiplied by ρ/P, and in the third step we substituted $n = P/[(1+x)k_B T]$, so that everything is in terms of x and T. Note that we have pulled a small trick in that we have ignored the temperature dependence of the partition function Z_0. This is justified because it doesn't change much over the temperatures where ionization occurs.

The remained of the calculation is straightforward but algebraically tedious. One evaluates the partial derivatives using using the formula for x derived from the Saha equation, then integrates to get the dependence between P and ρ. The final result is

$$\gamma_a = \frac{5 + \left(\frac{5}{2} + \frac{\chi}{k_B T}\right)^2 x(1-x)}{3 + \left[\frac{3}{2} + \left(\frac{3}{2} + \frac{\chi}{k_B T}\right)^2\right] x(1-x)}.$$

This has the limiting behavior we would expect. For $x \to 0$ or $x \to 1$, the result approaches 5/3, the value expected for a monatomic gas. In between γ_a is lower.



The amount by which γ_a drops depends on $\chi/k_B T$. When $\chi/k_B T$ is small, γ_a doesn't change much even when $x = 0.5$. When $\chi/k_B T$ is large, however, then for any x appreciably different from 0, the gas drops to near $\gamma_a = 1$.

We can understand this result intuitively. When $\chi/k_B T$ is small, ionizing an atom doesn't take much energy compared to the thermal energy, so the fact that there is an additional energy source or sink doesn't make much difference. When it is large, however, then every ionization requires a huge amount of thermal energy, and every recombination provides a huge amount of thermal energy. The value $\gamma_a = 1$ has a special significance: $P \propto \rho^1$ is what we expect for a gas that is isothermal, meaning at constant temperature. The reason that γ_a is close to 1 for a partially ionized gas where $\chi/k_B T$ is large is that any excess energy from compression is immediately used up in ionizing the gas just a little bit more, and any work done by the expansion of the gas is immediately balanced by reducing the ionization state just a tiny bit, releasing a great deal of thermal energy. Thus partial ionization can, at certain temperatures, at like a thermostat that keeps the gas at fixed temperature as it expands or contracts.

This phenomenon in a different guise is familiar from every day life. When one heats water on a stove, the temperature rises to 100 C and the water begins to boil. While it is boiling, however, the water doesn't get any hotter. Its temperature stays at 100 C even as more and more heat is pumped into it. That is because all the extra heat is going into changing the phase of the water from liquid to gas, and the energy per molecule required to make that phase change happen is much larger than $k_B T$. The potential energy associated with the phase change causes the temperature to stay fixed until all the water boils away.

II. Radiative Transfer

Having understood how pressure and energy are related and what this implies for adiabatic gas, we now turn to the topic of how energy moves through stars. Although there are many possible mechanisms, the most ubiquitous is radiative transfer. The basic idea of radiative transfer is that the hot material inside a star reaches thermal equilibrium with its local radiation field. The hot gas cannot move, but the photons can diffuse through the gas. Thus when a slightly cooler fluid element in a star sits on top of a slightly hotter one below it, photons produced in the hot element leak into the cool one and heat it up. This is the basic idea of radiative transfer.

A. Opacity

In order to understand how radiation moves energy, we need to introduce the concepts of radiation intensity and matter opacity. First think about a beam of radiation. To describe the beam I need to specify how much energy it carries per unit area per unit time. I also need to specify its direction, give giving the solid angle into which it is aimed. Finally, the beam may contain photons of many different frequencies, and I need to give you this information about each frequency. We define the object that contains all this information as the radiation intensity I, which has units if energy per unit time per unit area per unit solid angle per unit frequency. At any point in space, the radiation intensity is a function of direction and of frequency.

Before making use of this concept, it is important to distinguish between intensity and the flux of radiation that you're used to thinking about, H, which has units of energy per unit time per unit area per unit frequency. They are related very simply: flux is just the average of intensity over all directions. To understand the difference, imagine placing a sensor inside an oven whose walls are of uniform temperature. There is radiation coming from all directions equally, so the net flux of radiation is zero – as much energy moves from the left to the right each second as move from the right to the left. However, the intensity is not zero. The sensor would report that photons were striking it all the time, in equal numbers from every direction. Formally, we can define the relationship as follows. Suppose we want to compute the flux in the z direction. This is given by the average of the intensity over direction:

$$H = \int I \cos\theta \, d\Omega = \int_0^{2\pi} \int_0^{\pi} I(\theta, \phi) \cos\theta \sin\theta \, d\theta \, d\phi.$$

Now consider aiming a beam traveling in some direction at a slab of gas. If the slab consists of partially transparent material, only some of the radiation will be absorbed. An example is shining a flashlight through misty air. The air is not fully opaque, but it is not fully transparent either, so the light is partially transmitted. The opacity of a material is a measure of its ability to absorb light. To make this formal, suppose the slab consists of material with density ρ, and that its thickness is ds. The intensity of the radiation just before it enters the slab is I, and after it comes out the other side, some of the radiation has been absorbed and the intensity is reduced by an amount dI. We define the opacity κ by

$$\begin{aligned} -\frac{dI}{I} &= \kappa\rho \, ds \\ \kappa &= -\frac{1}{\rho I}\frac{dI}{ds}. \end{aligned}$$

This definition makes intuitive sense: the larger a fraction of the radiation the slab absorbs (larger magnitude $(1/I)(dI/ds)$), the higher the opacity. The more material it takes to absorb a fixed amount of radiation (larger ρ), the smaller the opacity. Of course κ can depend on the frequency of the radiation in question, since some materials are very good at absorbing some frequencies and not good at absorbing others. We can imagine repeating this experiment with radiation beams at different frequencies and measuring the absorption for each one, and thus figuring out κ_ν, the opacity as a function of frequency ν.

If the radiation beam moves through a non-infinitesimal slab of uniform gas, we can calculate how much of it will be absorbed in terms of κ. Suppose the beam shining on the surface has intensity I_0, and the slab has a thickness s. The intensity obeys

$$\frac{dI}{ds} = -\rho\kappa I,$$

so integrating we get

$$I = I_0 e^{-\kappa\rho s},$$

where the constant of integration has been chosen so that $I = I_0$ at $s = 0$. Thus radiation moving through a uniform absorbing medium is attenuated exponentially. This exponential attenuation is why common objects appear to have sharp edges – the light getting through them falls of exponentially fast. The quantity $\kappa\rho s$ comes up all the time, so we give it a special name and symbol:

$$\tau = \kappa\rho s$$

is defined as the optical depth of a system. Equivalently, we can write the differential equation describing the absorption of radiation as

$$\frac{dI}{d\tau} = -I,$$

which has the obvious solution $I = I_0 e^{-\tau}$.

B. Emission and the Radiative Transfer Equation

If radiation were only ever absorbed, life would be simple. However, material can emit radiation as well as absorb it. The equation we've written down only contains absorption, but we can generalize it quite easily to include emission. Suppose the slab of material also emits radiation at a certain rate. We describe its emission in terms of the emission rate j_ν. It tells us how much radiation energy the gas emits per unit time per unit volume per unit solid angle per unit frequency.

Including emission in our equation describing a beam of radiation traveling through a slab, we have

$$\frac{dI}{ds} = -\kappa\rho I + j.$$

The first term represents radiation taken out of the beam by absorption, and the second represents radiation put into the beam by emission. Equivalently, we can work in terms of optical depth:

$$\frac{dI}{d\tau} = -I + S,$$

where we define $S = j/(\kappa\rho)$ to be the source function. This equation is called the equation of radiative transfer.

Thus far all we've done is make formal definitions, so we haven't learned a whole lot. However, we can realize something important if we think about a completely uniform, opaque medium in thermal equilibrium at temperature T. We've already discussed that in thermal equilibrium the photons have to follow a particular distribution, called the Planck function:

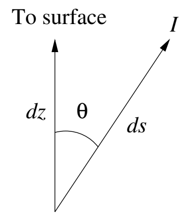
$$B(\nu, T) = \left(\frac{2h\nu^3}{c^2}\right)\frac{1}{e^{h\nu/k_BT} - 1}.$$

In such a medium the intensity clearly doesn't vary from point to point, so $dI/d\tau = 0$, and we have $I = S = B(\nu, T)$. This means that we know the radiation intensity and the source function for a uniform medium.

C. The Diffusion Approximation

The good news regarding stars is that, while the interior of a star isn't absolutely uniform, it's pretty close to it. There is just a tiny anisotropy coming from the fact that it is hotter toward the center of the star and colder toward the surface. The fact that the difference from uniformity is tiny means that we can write down a simple approximation to figure out how energy moves through the star, called the diffusion approximation. It is also sometimes called the Rosseland approximation, after its discoverer, the Norwegian astrophysicist Svein Rosseland.

To derive this approximation, we set up a coordinate system so that the z direction is toward the surface of the star. For a ray at an angle θ relative to the vertical, the distance ds along the ray is related to the vertical distance dz by $ds = dz/\cos\theta$.



Therefore the transfer equation for this particular ray reads

$$\frac{dI(z,\theta)}{ds} = \cos\theta \frac{dI(z,\theta)}{dz} = \kappa\rho[S(T) - I(z,\theta)],$$

where we have written out the dependences of I and S explicitly to remind ourselves of them: the intensity depends on depth z and on angle θ, while the source function depends only on temperature T. We can rewrite this as

$$I(z,\theta) = S(T) - \frac{\cos\theta}{\kappa\rho}\frac{dI(z,\theta)}{dz}.$$

Thus far everything we have done is exact, but now we make the Rosseland approximation. In a nearly uniform medium like the center of a star, I is nearly constant, so the term $\cos\theta/(\kappa\rho)(dI(z,\theta)/dz)$ is much smaller than the term S. Thus we can set the intensity I equal to S plus a small perturbation. Moreover, since we are dealing with material that is a blackbody, S is equal to the Planck function. Thus, we write

$$I(z,\theta) = B(T) + \epsilon I^{(1)}(z,\theta),$$

where ϵ is a number much smaller than 1. To figure out what the small perturbation should be, we can substitute this approximation back into the original equation for I:

$$\begin{aligned}
I(z,\theta) &= S(T) - \frac{\cos\theta}{\kappa\rho}\frac{dI(z,\theta)}{dz} \\
B(T) + \epsilon I^{(1)}(z,\theta) &= B(T) - \frac{\cos\theta}{\kappa\rho}\frac{d}{dz}\left[B(T) + \epsilon I^{(1)}(z,\theta)\right] \\
\epsilon I^{(1)}(z,\theta) &= -\frac{\cos\theta}{\kappa\rho}\frac{d}{dz}\left[B(T) + \epsilon I^{(1)}(z,\theta)\right] \\
&\approx -\frac{\cos\theta}{\kappa\rho}\frac{dB(T)}{dz}.
\end{aligned}$$

In the last step, we dropped the term proportional to ϵ on the right-hand side, on the ground that $B(T) \gg \epsilon I^{(1)}(z,\theta)$. Thus we arrive at our approximate form for the intensity:

$$I(z,\theta) \approx B(T) - \frac{\cos\theta}{\kappa\rho}\frac{dB(T)}{dz}.$$

From this approximate intensity, we can now compute the radiation flux:

$$\begin{aligned}
H_\nu &= \int I(z,\theta)\cos\theta\, d\Omega \\
&= \int \left[B(T) - \frac{\cos\theta}{\kappa\rho}\frac{dB(T)}{dz}\right]\cos\theta\, d\Omega \\
&= -\int \frac{\cos\theta}{\kappa\rho}\frac{dB(T)}{dz}\cos\theta\, d\Omega \\
&= -\frac{2\pi}{\kappa\rho}\frac{dB(T)}{dz}\int_0^\pi \cos^2\theta\sin\theta\, d\theta \\
&= -\frac{2\pi}{\kappa\rho}\frac{dB(T)}{dz}\int_{-1}^1 \cos^2\theta\, d(\cos\theta) \\
&= -\frac{4\pi}{3\kappa\rho}\frac{dB(T)}{dz} \\
&= -\frac{4\pi}{3\kappa\rho}\frac{\partial B(T)}{\partial T}\frac{\partial T}{\partial z}
\end{aligned}$$

This is at a particular frequency. To get the total flux over all frequencies, we just integrate:

$$H = -\frac{4\pi}{3\rho}\frac{\partial T}{\partial z}\int_0^\infty \frac{1}{\kappa}\frac{\partial B(T)}{\partial T}\, d\nu$$

We have written all the terms that do not depend on frequency outside the integral, and left inside it only those terms that do depend on frequency. Finally, we define the Rosseland mean opacity by

$$\frac{1}{\kappa_R} \equiv \frac{\int_0^\infty \kappa^{-1}\frac{\partial B(T)}{\partial T}\, d\nu}{\int_0^\infty \frac{\partial B(T)}{\partial T}\, d\nu} = \frac{\pi\int_0^\infty \kappa^{-1}\frac{\partial B(T)}{\partial T}\, d\nu}{4\sigma T^3},$$

where in the last step we plugged in $dB(T)/dT$ and evaluated the integral. With this definition, we can rewrite the flux as

$$H = -\frac{16\sigma T^3}{3\kappa_R\rho}\frac{\partial T}{\partial z}.$$

This is the flux per unit area. If we want to know the total energy passing through a given shell inside the star, we just multiply by the area:

$$F = -4\pi r^2\frac{16\sigma T^3}{3\kappa_R\rho}\frac{dT}{dr}.$$

This is an extremely powerful result. It gives us the flux of energy through a given shell in the star in terms of the gradient in the temperature and the mean opacity.

We can also invert the result and write the temperature gradient in terms of the flux:

$$\begin{aligned} \frac{dT}{dr} &= -\frac{3}{16\sigma}\frac{\kappa_R \rho}{T^3}\frac{F}{4\pi r^2} \\ \frac{dT}{dm} &= -\frac{3}{16\sigma}\frac{\kappa_R}{T^3}\frac{F}{(4\pi r^2)^2}. \end{aligned}$$

This is particularly useful in the region of a star where there is no nuclear energy generation, because in such a region we have shown that the flux is constant. Thus this equation lets us figure out the temperature versus radius in a star given the flux coming from further down, and the opacity.

D. Opacity Sources in Stars

This brings us to the final topic for this class: the opacity in stars. We need to know how opaque stellar material is. There are four main types of opacity we have to worry about:

- Electron scattering: photons can scatter off free electrons with the photon energy remaining constant, a process known as Thompson scattering. The Thomson scattering opacity can be computed from quantum mechanics, and is a simply a constant opacity per free electron. This constancy breaks down if the mean photon energy approaches the electron rest energy of 511 keV, but this is generally not the case in stars.
- Free-free absorption: a free electron in the vicinity of an ion can absorb a photon and go into a higher energy unbound state. The presence of the ion is critical to allowing absorption, because the potential between the electron and the ion serves as a repository for the excess energy.
- Bound-free absorption: this is otherwise known as ionization. When there are neutral atoms present, they can absorb photons whose energies are sufficient to ionize their electrons.
- Bound-bound absorption: this is like ionization, except that the transition is between one bound state and another bound state that is at a higher excitation. The Hα and calcium K transitions we discussed a few weeks ago in the solar atmosphere are examples of this.

Which of these sources of opacity dominates depends on the local temperature and density, and changes from one part of a star to another. In the deep interior we have already seen that the gas is almost entirely ionized due to the high temperatures there, and as a result bound-free and bound-bound absorption contribute very little – there are simply too few bound electrons around. In stellar atmospheres, on the other hand, bound-free and bound-bound absorptions dominate, because there are comparatively few free electrons. For our purposes we will mostly be concerned with stellar interiors, where electron scattering and free-free are most important.

Electron scattering is fairly easy to calculate, since it just involves the interaction of an electromagnetic wave with a single charged particle. In fact, the calculation can be done classically as long as the photon energy is much smaller than the electron rest mass. We will not do the derivation in class, and will simply quote the result. The cross-section for a single electron, called the Thomson cross-section, is

$$\sigma_T = \frac{8\pi}{3}\left(\frac{\alpha\hbar}{m_e c}\right)^2 = 6.65 \times 10^{-25} \text{ cm}^2,$$

where $\alpha \approx 1/137$ is the fine structure constant. This is independent of frequency, so the Rosseland mean opacity is simply the opacity at any frequency (except for frequencies where the photon energy approaches 511 keV.) To figure out the corresponding opacity, which is the cross-section per unit mass, we simply have to multiply this by the number of free electrons per gram of material. For pure hydrogen, there are $1/m_{\rm H}$ hydrogen atoms per gram, and thus $1/m_{\rm H}$ free electrons is the material is fully ionized. Thus for pure hydrogen we have

$$\kappa_{\rm es,0} = \frac{\sigma_T}{m_{\rm H}} = 0.40 \text{ cm}^2 \text{ g}^{-1}.$$

If the material is not pure hydrogen we get a very similar formula, and we just have to plug in the appropriate number of hydrogen masses per free electron, μ_e. Thus we have

$$\kappa_{\rm es} = \frac{\sigma_T}{\mu_e m_{\rm H}} = \frac{\kappa_{\rm es,0}}{\mu_e} = \kappa_{\rm es,0}\left(\frac{1+X}{2}\right).$$

The other opacity source to worry about in the interior of a star is free-free absorption. This process is vastly more complicated to compute, since one must consider interactions of free electrons with many types of nuclei and with photons of many frequencies, over a wide range in temperatures and densities. The calculation these days is generally done by computer. However, we have some general expectations based on simple principles.

The idea of free-free absorption is that the presence of ions enhances the opacity of the material above and beyond what would be expected with just free electrons, because the potential energy associated with the electron-ion interaction provides a repository into which to deposit energy absorbed from photons. Since the effect depends on ion-electron interactions, a higher density should increase the opacity, since it means electrons and ions are more closely packed and thus interact more. For this reason, $\kappa_{\rm ff}$ should increase with density.

At very high temperatures, on the other hand, free-free opacity should become unimportant. This is because the typical photon energy is much higher than the electron-ion potential, so having the potential energy of the electron-ion interaction available doesn't help much. Thus $\kappa_{\rm ff}$ should decline with increasing temperature, until it becomes small compared to electron-scattering opacity.

When making stellar models on a computer, one can directly use the tables of $\kappa_{\rm ff}$ as a function of ρ and T that a computer spits out. However, we can make simple analytic models and get most of the general results right using an analytic fit to the numerical data. The free-free opacity, first derived by Hendrik Kramers, is well approximated by a power law in density and temperature:

$$\kappa_{\rm ff} \approx \frac{\kappa_{\rm ff,0}}{\mu_e} \left\langle \frac{\mathcal{Z}^2}{\mathcal{A}} \right\rangle \rho[\text{g cm}^{-3}] T[\text{K}]^{-7/2} \approx \kappa_{\rm ff,0} \left(\frac{1+X}{2}\right) \left\langle \frac{\mathcal{Z}^2}{\mathcal{A}} \right\rangle \rho[\text{g cm}^{-3}] T[\text{K}]^{-7/2},$$

where $\kappa_{\rm ff,0} = 7.5 \times 10^{22}$ cm^2 g^{-1}, and the square brackets after ρ and T indicate the units in which the are to be measured when plugging into this formula. As expected, the opacity increases with density and decreases with temperature. The factor $\langle \mathcal{Z}^2/\mathcal{A} \rangle$ appears because the free-free opacity is affected by the population of ions available for interactions. Interactions are stronger with more charged ions, hence the $\mathcal{Z}^2$ in the numerator. They are weaker with more massive ions, hence the $\mathcal{A}$ in the denominator.

An opacity of this form, given as $\kappa = \text{const} \cdot \rho^a T^b$ for constant powers a and b, is known as a Kramers law opacity.

Class 7 Notes: Basics of Nuclear Fusion

In this class we continue the process of filling in the missing microphysical details that we need to make a stellar model. To recap, in the last two classes we computed the pressure of stellar material and the rate of energy transport through the star. These were two of the missing pieces we needed. The third, which we'll sketch out over the next two lectures, is the rate for nuclear reactions, and the energy that they generate.

I. Energetics

A. Energy Release

All nuclear reactions fundamentally work by converting mass into energy. (In some ways the same could be said of chemical reactions, but for those the masses involved are so tiny as to not be worth worrying about.) The masses of the reactants involved therefore determine the energy released by the reaction.

Consider a reaction between two species that produced some other species

$$\mathcal{I}(\mathcal{A}_i, \mathcal{Z}_i) + \mathcal{J}(\mathcal{A}_j, \mathcal{Z}_j) \rightarrow \mathcal{K}(\mathcal{A}_k, \mathcal{Z}_k) + \mathcal{L}(\mathcal{A}_l, \mathcal{Z}_l),$$

where as usual $\mathcal{Z}$ is the atomic number and $\mathcal{A}$ is the atomic mass number. At this point we must distinguish between atomic mass number and actual mass, so let $\mathcal{M}$ be the mass of a given species. The atomic mass number times m_{H} and the true mass are nearly identical, $\mathcal{M} \approx \mathcal{A} m_{\mathrm{H}}$, but not quite, and that small difference is the source of energy for the reaction. For the reaction we have written down, the energy released is

$$Q_{ijk} = (\mathcal{M}_i + \mathcal{M}_j - \mathcal{M}_k - \mathcal{M}_l)c^2,$$

i.e. the initial mass minus the final mass, multiplied by c^2.

To remind you, the rate per unit volume at which the reaction we have written down occurs is

$$\frac{\rho^2}{m_{\mathrm{H}}^2} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R_{ijk},$$

where R_{ijk} is the rate coefficient. If each such reaction released an energy Q_{ijk}, then the rate of nuclear energy release per unit volume is simply given by this rate, multiplied by Q_{ijk}, and summed over all possible reactions:

$$\frac{\rho^2}{m_{\mathrm{H}}^2} \sum_{ijk} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R_{ijk} Q_{ijk}.$$

The rate of nuclear energy release per unit mass is just this divided by the mass per volume ρ:

$$q_{\rm nuc} = \frac{\rho}{m_{\rm H}^2} \sum_{ijk} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R_{ijk} Q_{ijk}.$$

If the reaction produces neutrinos, they will carry away some of the energy and escape the star, and thus the amount by which the star is heated will be reduced. However this loss is small in most stars under most circumstances.

B. Binding Energy per Nucleon

A very useful way to think about the amount of energy available in nuclear reactions is to compute the binding energy per nucleon. Suppose that we start with hydrogen, which consists of one proton of mass $m_{\rm H}$ (ignoring electrons), and we define that to have zero binding energy. Since binding energy is potential energy, we can do this, since we can choose the zero of potential energy to be anywhere.

Now consider some other element, with atomic mass number $\mathcal{A}$ and actual mass $\mathcal{M}$ per atom; and consider how much energy is released in the process of making that element from hydrogen. The exact reaction processes used don't matter, just the initial and final masses. Since atomic number is conserved, we must use $\mathcal{A}$ hydrogen atoms to make the new nucleus, so the difference between the final and initial mass is $\mathcal{M} - \mathcal{A} m_{\rm H}$. We define the mass excess as this quantity multiplied by c^2:

$$\Delta M = (\mathcal{M} - \mathcal{A} m_{\rm H}) c^2.$$

This is just the difference in energy between the bound nucleus and the equal number of free protons. The name is somewhat confusing, since this is really an energy not a mass. The reason for the name is that in relativity one doesn't really need to distinguish between mass and energy. They're the same thing, just measured in different units.

A more useful quantity than this is the binding energy per nucleon, i.e. minus the mass excess divided by the number of nucleons (protons or neutrons) in the nucleus. The minus here is added so that the binding energy is positive if the nucleus is more strongly bound than the corresponding number of free nucleons. Thus we define the binding energy per nucleon as

$$-\frac{\Delta M}{\mathcal{A}} = \left(1 - \frac{\mathcal{M}}{\mathcal{A} m_{\rm H}}\right) m_{\rm H} c^2.$$

Since $\mathcal{M}$ and $\mathcal{A}$ can be determined experimentally, this quantity is fairly straightforward to measure. The results are very illuminating.

[Silde 1 – binding energy per nucleon]

This plot contains an enormous amount of information, and looking at it immediately explains a number of facts about stars and nuclear physics. To interpret this plot, recall that number of nucleons is conserved by nuclear reactions. Thus any nuclear reaction just involves taking a fixed number of nucleons and moving them to the left or right on this plot. The energy released or absorbed in the process is just the number of nucleons involved multiplied by the change in binding energy per nucleon.

The first thing to notice about this plot is that there is a maximum at ^{56}Fe – iron-56. This is the most bound nucleus. At smaller atomic masses the binding energy per nucleon generally increases with atomic number, while at larger atomic masses it decreases. This marks the divide between fusion and fission reactions. At atomic masses below 56, energy is released by increasing the atomic number, so fusion is exothermic and fission is endothermic. At atomic number above 56, energy is released by decreasing the atomic number, so fission is exothermic and fission is endothermic.

Second, notice that the rise is very sharp at small atomic number. This means that fusing hydrogen into heavier things is generally the most exothermic reaction available, and that it releases far more energy per nucleon than later stages of fusion, say helium into carbon. This has important implications for the fate of stars that exhaust their supply of hydrogen.

Third, notice that there are several local maxima and minima at small atomic number. ^{4}He is a maximum, as are ^{12}C and ^{16}O. There is a good reason that helium, carbon, and oxygen are the most common elements in the universe after hydrogen: they are local maxima of the binding energy, which means that they are the most strongly bound, stable elements in their neighborhood of atomic number. Conversely, lithium is a minimum. For this reason nuclear reactions in stars destroy lithium, and they do not produce it. Essentially all the lithium there is in the universe was made in the big bang, and stars have been destroying it ever since.

Finally, notice that these are big numbers as far as the energy yield. The scale on this plot is MeV per nucleon. In terms of more familiar units, 1 MeV per H mass corresponds to 10^{18} erg g^{-1}, or roughly 22 tons of TNT per gram of hydrogen fuel.

II. Reaction Rates

A. The Coulomb Barrier

The binding energy curve tells us the amount of energy available from nuclear reactions, but not the rates at which they occur. Given that the reaction for fusing hydrogen to helium is highly exothermic, why doesn't the reaction happen spontaneously at room temperature? The answer is the same as the reason that coal doesn't spontaneously combust at room temperature: the reaction has an activation energy, and that energy is quite high.

To understand why, consider the potential energy associated with two nuclei of charge $\mathcal{Z}_i$ and $\mathcal{Z}_j$ separated by a distance r. The Coulomb (electric) potential energy is

$$U_{\rm C} = \frac{\mathcal{Z}_i \mathcal{Z}_j e^2}{r} = \mathcal{Z}_i \mathcal{Z}_j \frac{1.4 \text{ MeV}}{r/\text{fm}},$$

where 1 fm $= 10^{-13}$ cm $= 10^{-15}$ m. Since this is positive, the force between the protons is repulsive, as it should be.

In addition to that positive energy, there is a negative energy associated with nuclear forces. The full form of the proton-proton force is complicated, but we can get an idea of its behavior by noting that, at larger ranges, it is mediated by the exchange of virtual mesons such as pions. Because these particles have mass, their range is limited by the Heisenberg uncertainty principle: they can only exist for a short time, and they only exert significant force at distances they can reach within that time. Specifically, the uncertainty principle tells us that

$$\Delta E \, \Delta t \geq \frac{\hbar}{2}$$

If the particle travels at the maximum possible speed of c, its range is roughly

$$r \sim c\Delta t \sim \frac{c\hbar}{E},$$

where E is the rest energy of the particle being exchanged. For pions, which mediate the proton-proton force, $\Delta E = 135$ MeV or 140 MeV, depending on whether they are neutral or charged. Plugging this in for ΔE gives $r \sim 1$ fm. Thus the nuclear force is negligible at distances greater than ~ 1 fm. Within that range, however, the nuclear force is dominant. Potentials arising from exchanges of massive particles like this are called Yukawa potentials, and they have the form

$$U_{\rm Y} = -g^2 \frac{e^{-r/\lambda}}{r},$$

where g is a constant and $\lambda = c\hbar/E$ is the range of the force. This is only an approximation to the true potential energy, but it is reasonably good one at large ranges.

The total potential is the sum of the Yukawa and Coulomb potentials. The functional form of this potential is something like a $1/r$ rise that is cut off by a sharp decrease at small radii. This slide shows an example for an important reaction: $^{12}\text{C} + \alpha$, which has $\mathcal{Z}_i = 6$ and $\mathcal{Z}_j = 2$.

[Slide 2 – Coulomb barrier for ^{12}C + α reaction]

For the reaction to proceed, the two particles must get close enough to one another to reach the region where the potential drops, and the force becomes attractive. If they do not, they will simply bounce off one another without reacting. This is called the Coulomb barrier, and it applies to chemical as well as nuclear reactions. The existence of the Coulomb barrier means that there is a minimum relative velocity the particles must have in order for the reaction to go, which we can calculate from the height of the Coulomb barrier. This is much like rolling a ball up a steep hill with a peak – there is a minimum velocity with which you must roll the ball if you want it to reach the top of the hill.

Suppose that the potential follows the Coulomb form until some minimum radius $r_0 \sim 1$ fm, then suddenly drops at smaller radii. The maximum potential energy is

$$U_C = \frac{\mathcal{Z}_i \mathcal{Z}_j e^2}{r_0} = \mathcal{Z}_i \mathcal{Z}_j \frac{1.4 \text{ MeV}}{r_0/\text{fm}}.$$

The minimum relative velocity of the particles is given by the condition that the kinetic energy in the center of mass frame exceed this value:

$$\frac{1}{2} \mu_{\text{red}} m_{\text{H}} v^2 \geq U_c,$$

where $\mu_{\text{red}} m_{\text{H}}$ is the reduced mass.

A more useful calculation than this is to ask what temperature the gas must have such that the typical collision is at sufficient velocity for the reaction to occur. The typical collision energy is

$$\frac{1}{2} \mu_{\text{red}} m_{\text{H}} v^2 = \frac{3}{2} k_B T,$$

so setting this equal to U_C and solving gives

$$T \geq \frac{2 \mathcal{Z}_i \mathcal{Z}_j e^2}{3 k_B r_0} = 1.1 \times 10^{10} \text{ K } \frac{\mathcal{Z}_i \mathcal{Z}_j}{r_0/\text{fm}}.$$

Thus the typical particle does not have enough energy to penetrate the Coulomb barrier until the temperature is $\sim 10^{10}$ K for proton-proton reactions, and even higher temperatures for higher atomic numbers. This is much higher than the temperatures for stars' centers than we estimated earlier in the class. You might think that it's not a problem because some particles move faster than the average, and thus are going fast enough to penetrate the Coulomb barrier. You will show on your homework that this solution doesn't work. At the temperature of $\sim 10^7$ K in the center of the Sun, if this calculation is correct then fusion should not be possible.

B. Quantum Tunneling

The resolution to this problem lies in the phenomenon of quantum tunneling. The calculation we just did is based on classical physics, and predicts that no nuclei will get within r_0 of one another unless they reach a high enough velocity to overcome the Coulomb barrier. However, in quantum mechanics there is a non-zero probability of finding the particle inside r_0 even if it does not have enough energy to break the Coulomb barrier. This phenomenon is known as tunneling, because it is like the particle takes a tunnel through the peak rather than going over it.

We can make a crude estimate of when tunneling will occur using wave-particle duality. Recall that each proton can be thought of as a wave whose wavelength is dictated by the uncertainty principle. The wavelength associated with a particle of momentum p is

$$\lambda = \frac{h}{p}.$$

This is known as the particle's de Broglie wavelength.

As a rough estimate of when quantum tunneling might allow barrier penetration, we can estimate that the two particles must be able to get within one de Broglie wavelength of one another. This in turn requires that the kinetic energy of the particles be equal to their Coulomb potential energy at a separation of one de Broglie wavelength:

$$\frac{\mathcal{Z}_i \mathcal{Z}_j e^2}{\lambda} = \frac{1}{2}\mu_{\rm red} m_{\rm H} v^2 = \frac{p^2}{2\mu_{\rm red} m_{\rm H}} = \frac{h^2}{2\mu_{\rm red} m_{\rm H} \lambda^2}$$

Solving this for λ, we find that barrier penetration should occur is the particles are able to get within a distance

$$\lambda = \frac{h^2}{2\mu_{\rm red} m_{\rm H} \mathcal{Z}_i \mathcal{Z}_j e^2}.$$

of one another.

To figure out the corresponding temperature, we can just evaluate our result from the classical problem using λ in place of r_0:

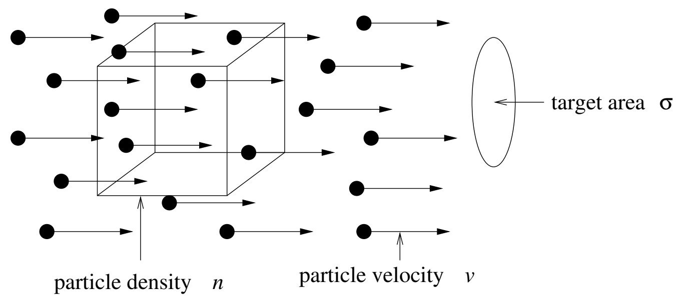
$$T \geq \frac{2\mathcal{Z}_i \mathcal{Z}_j e^2}{3k_B \lambda} = \frac{4\mathcal{Z}_i^2 \mathcal{Z}_j^2 e^4 \mu_{\rm red} m_{\rm H}}{3h^2 k_B} = 9.6 \times 10^6 \text{ K } \mathcal{Z}_i^2 \mathcal{Z}_j^2 \left(\frac{\mu_{\rm red}}{1/2}\right).$$

Thus proton-proton reactions, which have $\mathcal{Z}_i = \mathcal{Z}_j = 1$ and $\mu_{\rm red} = 1/2$, should begin to occur via quantum tunneling at a temperature of $\sim 10^7$ K, much closer to the temperatures we infer in the center of the Sun.

C. The Gamow Peak

Having seen that quantum effects are important, we will now try to perform a more rigorous calculation of the reaction rate. Consider reactions between two nuclei with number densities n_i and n_j in a gas at temperature T. In order to compute the reaction rate, we need to know the rate at which these nuclei collide with enough energy to tunnel through the Coulomb barrier. That's what we'll calculate now.

The first step is to compute the rate at which particles strike one another closely enough to interact. This is very much like calculating the pressure. We consider a particle, and we want to know how often other particles run into it. If we had a beam of particles of density n and velocity v, and the target particle had a cross-sectional area σ, the impact rate would be $n\sigma v$. Note that this formula is almost exactly like the one describing the rate at which particles strike the wall of a vessel, which we used to compute pressures.



In reality the particle in question isn't a solid target with a fixed area. We're interested in interactions that lead to reactions, which require that the collision be close enough to allow the nuclei to tunnel through the Coulumb barrier, but also require that the interaction have enough energy to make such tunneling possible. A direct bullseye at a very low energy won't lead to a reaction, so the cross-section at very low energies is basically zero. However, we can still extend the analogy of shooting a beam of particles at a target by defining the cross-section at energy E. Let $dN_{\text{reac}}(E)/dt$ be the number of reactions per time interval dt produced by shooting a beam of particles of density n at velocity v at a target nucleus. We define the cross-section $\sigma(E)$ via the relation

$$\frac{dN_{\text{reac}}(E)}{dt} = n\sigma(E)v(E).$$

Next we want to generalize from a the case of a beam to the case of a thermal gas where not all particles have the same energy. We proved a few classes ago that the momentum distribution of particles of mass m at temperature T is

$$\frac{dn}{dp} = \frac{4n}{\pi^{1/2}(2mk_BT)^{3/2}} p^2 e^{-p^2/(2mk_BT)}.$$

Since we're interested in particle energies, we'll change this to a distribution over energy instead of momentum. Since $E = p^2/(2m)$, or $p = \sqrt{2mE}$, we have

$$\frac{dn}{dE} = \frac{dn}{dp}\frac{dp}{dE} = \frac{4n}{\pi^{1/2}(2mk_BT)^{3/2}} p^2 e^{-p^2/(2mk_BT)} \cdot \sqrt{\frac{m}{2E}} = \frac{2n}{\pi^{1/2}(k_BT)^{3/2}} E^{1/2} e^{-E/k_BT}.$$

Note that this only applies to non-relativistic particles, since we used $E = p^2/(2m)$ instead of $E = pc$. However, nuclei are generally always non-relativistic, except in neutron stars.

In this case, the number of reactions dN per time interval dt that a given target nucleus undergoes is given by integrating over the possible energies of the impacting particles. In particular, the number of reactions per unit time for a particle of species i due collisions with particles of species j is

$$\frac{dN_i}{dt} = \int_0^\infty \sigma(E)v(E)\frac{dn_j}{dE}\,dE.$$

Since the velocity that matters here is the relative velocity, we have to compute it in terms of the reduced mass: $v(E) = \sqrt{2E/\mu_{\rm red} m_{\rm H}}$, where $\mu_{\rm red} m_{\rm H} = m_i m_j/(m_i + m_j)$. Finally, if we want to know the number of reactions per unit time in a given volume of gas, we just have to multiply this by the number of target nuclei per unit volume, n_i, and divide by $(1+\delta_{ij})$ to avoid double-counting. This gives

$$\frac{dn_{\rm reac}}{dt} = \frac{n_i}{(1+\delta_{ij})} \int_0^\infty \sigma(E) v(E) \frac{dn_j}{dE}\, dE.$$

Recall that we defined the rate coefficient R_{ijk} so that the reaction rate is $R_{ijk} n_i n_j$ for different species, or $R_{ijk} n_i^2/2$ for two of the same species. Thus the rate coefficient is

$$\begin{aligned} R_{ijk} &= \frac{(1+\delta_{ij})}{n_i n_j} \frac{dn_{\rm reac}}{dt} \\ &= \frac{2}{\pi^{1/2}} \frac{1}{(k_B T)^{3/2}} \int_0^\infty \sigma(E) v(E) E^{1/2} e^{-E/k_B T}\, dE \\ &= \frac{1}{(\pi \mu_{\rm red} m_{\rm H})^{1/2}} \left(\frac{2}{k_B T}\right)^{3/2} \int_0^\infty \sigma(E) E e^{-E/k_B T}\, dE \end{aligned}$$

The final remaining step is to figure out the cross-section $\sigma(E)$ at energy E. Computing this in general is quite difficult, and often laboratory measurements are required to be sure of exact values. However, we can get a rough idea of how $\sigma(E)$ varies with energy based on general quantum-mechanical principles. The first such principle is that particles should interact when they come within distances that are comparable to their de Broglie wavelengths – a higher energy particles has a smaller wavelength, and thus represents a smaller target. Thus we expect that

$$\sigma(E) \propto \lambda^2 = \frac{h^2}{p^2} \propto \frac{1}{E}.$$

The second principle is that nuclear reactions like the ones we are interested in require tunneling through the Coulomb barrier. A quantum mechanical calculation of the probability that tunneling will occur shows that it is proportional to

$$e^{-2\pi^2 U_C/E},$$

where U_C is the height of the Coulomb barrier at a distance of one de Broglie wavelength. You will see this calculation in your quantum mechanics class, and I will not go through it here. In terms of the energy, the Coulomb barrier U_C is

$$U_C = \frac{\mathcal{Z}_i \mathcal{Z}_j e^2}{\lambda} = \frac{\mathcal{Z}_i \mathcal{Z}_j e^2 p}{h} = \frac{\mathcal{Z}_i \mathcal{Z}_j e^2}{h} \sqrt{2\mu_{\rm red} m_{\rm H} E},$$

so the exponential factor is

$$2\pi^2 \frac{U_C}{E} = 2^{3/2} \pi^2 \frac{\mu_{\rm red}^{1/2} m_{\rm H}^{1/2} \mathcal{Z}_i \mathcal{Z}_j e^2}{h} E^{-1/2} \equiv b E^{-1/2},$$

where

$$b = 2^{3/2}\pi^2 \frac{\mu_{\rm red}^{1/2} m_{\rm H}^{1/2} \mathcal{Z}_i \mathcal{Z}_j e^2}{h} = 0.0013 \mu_{\rm red}^{1/2} \mathcal{Z}_i \mathcal{Z}_j \ ({\rm erg})^{1/2}.$$

Thus we also expect to have $\sigma \propto e^{-bE^{-1/2}}$. Note that the factor b depends only on the charges and masses of the nuclei involved in the reaction. It is therefore constant for any given reaction.

Combining the two factors our analysis reveals, we define

$$\sigma(E) \equiv \frac{S(E)}{E} e^{-bE^{-1/2}},$$

where $S(E)$ is, ideally, either a constant or a function that varies only very, very weakly with E. Plugging all this in, the reaction rate coefficient is

$$R_{ijk} = \frac{1}{(\pi \mu_{\rm red} m_{\rm H})^{1/2}} \left(\frac{2}{k_B T}\right)^{3/2} \int_0^\infty S(E) e^{-bE^{-1/2}} e^{-E/k_B T}\, dE.$$

It is instructive to look at the behavior of the two exponential factors, $e^{-bE^{-1/2}}$ and e^{-E/k_BT}. Clearly the first function increases as E increases, while the second one decreases as E increases. We therefore expect their product to reach a maximum at some intermediate energy. In fact, we can compute the maximum analytically, by taking the derivative and setting in equal to zero:

$$\begin{aligned}
0 &= \frac{d}{dE}\left(e^{-bE^{-1/2}} e^{-E/k_B T}\right) \\
&= \frac{d}{dE} e^{-(E/k_BT + bE^{-1/2})} \\
&= -\left(\frac{E}{k_B T} + bE^{-1/2}\right)\left(\frac{1}{k_B T} - \frac{b}{2E^{3/2}}\right) e^{-(E/k_BT + bE^{-1/2})} \\
E_0 &= \left(\frac{b k_B T}{2}\right)^{2/3} \\
&= 1.22 \left[\mathcal{Z}_i^2 \mathcal{Z}_j^2 \mu_{\rm red} \left(\frac{T}{10^6\ {\rm K}}\right)^2\right]^{1/3} \ {\rm keV},
\end{aligned}$$

where E_0 is the energy at the maximum. This maximum is known as the Gamow peak, after George Gamow, who discovered it in 1928. The plot shows the Gamow peak for proton-proton interactions at $T = 1.57 \times 10^7$ K, the Sun's central temperature.

[Slide 3 – the Gamow peak]

If we let $x = E/E_0$, then we can rewrite the reaction rate coefficient as

$$\begin{aligned}
R_{ijk} &= \frac{E_0}{(\pi\mu_{\rm red} m_{\rm H})^{1/2}} \left(\frac{2}{k_B T}\right)^{3/2} \int_0^\infty S(x) \exp\left[-\left(\frac{b^2}{4k_BT}\right)^{1/3}\left(x + \frac{2}{x^{1/2}}\right)\right] dx \\
&= \left[2^{11}\pi^5 \frac{\mathcal{Z}_i^4 \mathcal{Z}_j^4 e^8}{\mu_{\rm red} m_{\rm H} h^4 (k_B T)^5}\right]^{1/6} \int_0^\infty S(x) \exp\left[-\left(\frac{b^2}{4k_BT}\right)^{1/3}\left(x + \frac{2}{x^{1/2}}\right)\right] dx
\end{aligned}$$

To get a sense of how narrowly peaked this function is, it is helpful to evaluate the factor $[b^2/(4k_BT)]^{1/3}$ for some typical examples. If we consider proton-proton interactions (so $\mathcal{Z}_i = \mathcal{Z}_j = 1$ and $\mu_{\rm red} = 1/2$) at the Sun's central temperature of 1.57×10^7 K, then we have

$$b = 8.8 \times 10^{-4}\ (\text{erg})^{1/2} \text{ and } \left(\frac{b^2}{4k_BT}\right)^{1/3} = 4.5.$$

Evaluating the function $e^{-4.5(x+2/x^{1/2})}$ shows that for $x = 3$ (i.e. at energies three times the peak), it is a factor of 180 lower than it is at peak. For $x = 1/3$ (i.e. at energies three times below the peak), it is 35 times smaller than it is at peak. Thus the reaction rate is strongly dominated by energies near the peak, with energies different from the peak by even as little as a factor of 3 contributing negligibly.

When we are near the peak, i.e. $x \approx 1$, the reaction rate varies exponential with the quantity $[b^2/(k_BT)]^{1/3}$. This means that the reaction rate is extremely sensitive to temperature. For this reason, we often think of nuclear reactions as having a threshold temperature at which they turn on. This threshold temperature clearly increases with nuclear charge: since $b \propto \mathcal{Z}_i\mathcal{Z}_j$, and the reaction rate depends on b^2/T, we expect the temperature needed to ignite a particular reaction to vary as $\mathcal{Z}_i^2\mathcal{Z}_j^2$. Thus higher $\mathcal{Z}$ nuclei require progressively higher temperatures to fuse.

Of course we still have not assigned a value of $S(E)$ near the Gamow peak. We have only said that we expect it to be nearly constant. Its actual value depends on the reaction in question and the type of physics it involves, and must be obtained either by laboratory measurement, theoretical quantum calculation, or a combination of both. Unfortunately these values sometimes have significant uncertainties. In a star, reactions can occur at an appreciable rate at relatively low temperatures because the density is high – recall that the reaction rate per unit volume varies as n_in_j. In a laboratory, we have to work with much lower densities, and as a result the reaction rates at the temperatures found in stars are often unobservably small. Instead, we are forced to make measurements at higher temperatures and extrapolate.

D. Temperature Dependence of Reaction Rates

It is often helpful to know roughly how the reaction rate varies with temperature when one is near the ignition temperature. To find that out, we can approximately evaluate the integral in the formula for the rate coefficient. As a first step in this approximation, we neglect any variation in the $S(E)$ factor across the Gamow peak, and simply set it equal to a constant value $S(E_0)$. Thus the reaction rate coefficient is approximately

$$R_{ijk} = \frac{1}{(\pi\mu_{\rm red}m_{\rm H})^{1/2}}\left(\frac{2}{k_BT}\right)^{3/2} S(E_0)\int_0^\infty \exp\left(-\frac{E}{k_BT} - \frac{b}{E^{1/2}}\right) dE.$$

The maximum value of the integrand occurs when $E = E_0$, and is given by

$$I_{\max} \equiv \exp\left(-\frac{3E_0}{k_B T}\right) \equiv e^{-\tau},$$

where we define

$$\tau = \frac{3E_0}{k_B T} = 42.46 \left[\mathcal{Z}_i^2 \mathcal{Z}_j^2 \mu_{\rm red} \left(\frac{T}{10^6\ {\rm K}}\right)^{-1}\right]^{1/3}$$

The second step in the approximation is to approximate the exponential factor in the integral by a Gaussian of width Δ:

$$\exp\left(-\frac{E}{k_B T} - \frac{b}{E^{1/2}}\right) \approx I_{\max} \exp\left[-\left(\frac{E - E_0}{\Delta/2}\right)^2\right].$$

The width Δ is generally chosen by picking the value such that the second derivatives of the exact and approximate forms for the integrand are equal at $E = E_0$. A little algebra shows that this gives

$$\Delta = \frac{4}{\sqrt{3}} \left(E_0 k_B T\right)^{1/2}.$$

The approximation is reasonably good. The graph shown is for two protons at a temperature of 1.6×10^7 K.

[Slide 4 – Gaussian approximation to the Gamow peak]

The final step in the approximation is to change the limits of integration from 0 to ∞ to $-\infty$ to ∞. This is not a bad approximation because the vast majority of the power in the Gaussian occurs at positive energies, and if the limits are $-\infty$ to ∞, the integral can be done exactly:

$$\int_{-\infty}^{\infty} \exp\left[-\left(\frac{E - E_0}{\Delta/2}\right)^2\right] dE = \frac{\sqrt{\pi}}{2}\Delta.$$

With this approximation complete, we can write the reaction rate coefficient as

$$\begin{aligned} R_{ijk} &= \frac{1}{(\pi \mu_{\rm red} m_{\rm H})^{1/2}} \left(\frac{2}{k_B T}\right)^{3/2} S(E_0) I_{\max} \frac{\sqrt{\pi}}{2}\Delta \\ &= I_{\max} \left(\frac{2}{\mu_{\rm red} m_{\rm H}}\right)^{1/2} \frac{\Delta}{(k_B T)^{3/2}} S(E_0). \end{aligned}$$

We can rewrite this in terms of τ by substituting in for Δ and $k_B T$ in terms of τ. Doing so and simplifying a great deal produces

$$R_{ijk} = \frac{4}{3^{5/2}\pi^2} \frac{h}{\mu_{\rm red} m_{\rm H} \mathcal{Z}_i \mathcal{Z}_j e^2} S(E_0) \tau^2 e^{-\tau}.$$

All the temperature-dependence is encapsulated in the $\tau^2 e^{-\tau}$ term. The factor τ itself varies as

$$\tau \propto \frac{E_0}{T} \propto T^{-1/3}.$$

It is often useful to approximate the reaction rate as a powerlaw in T, i.e. to set $R_{ijk} \propto T^\nu$ for some power ν. Obviously the relationship is not a powerlaw in general, since there is an exponential in τ. However, we can approximate the behavior as a powerlaw if we are in the vicinity of a particular temperature T_0, near which $\tau = \tau_0 (T/T_0)^{-1/3}$. To understand what this entails, recall that a powerlaw is just a straight line in a log-log plot. In effect, fitting to a powerlaw is just the same as computing the slope at some point in the log-log plot. Thus we have

$$\nu = \frac{d \ln R_{ijk}}{d \ln T}$$

Since $R_{ijk} \propto \tau^2 e^{-\tau}$,

$$\ln R_{ijk} = 2 \ln \tau - \tau + \text{const} = -\frac{2}{3} \ln T - \tau_0 \left(\frac{T}{T_0}\right)^{-1/3} + \text{const}$$

Taking the derivative:

$$\begin{aligned}
\nu = \frac{d \ln R_{ijk}}{d \ln T} &= -\frac{2}{3} - \tau_0 T_0^{1/3} \frac{d}{d \ln T} T^{-1/3} \\
&= -\frac{2}{3} - \tau_0 T_0^{1/3} T \frac{d}{dT} T^{-1/3} \\
&= -\frac{2}{3} + \frac{\tau_0 T_0^{1/3}}{3 T^{1/3}} \\
&= \frac{\tau}{3} - \frac{2}{3}
\end{aligned}$$

This lets us approximate the behavior of R_{ijk} as a powerlaw:

$$R_{ijk} = R_{0,ijk} T^{(\tau-2)/3}.$$

We will use this in the next class to evaluated several of the important reactions inside stars. Given such a powerlaw fit, we can come up with an equivalent one for the rate of nuclear energy generation per unit mass when the gas temperature is near the ignition temperature for a given reaction:

$$q_{\text{nuc}} = \frac{\rho}{m_{\text{H}}^2} \sum_{ijk} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R_{ijk} Q_{ijk} = \rho \sum_{ijk} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} q_{0,ijk} T^{p_{ijk}},$$

where $q_{0,ijk}$ and p_{ijk} are constants for a given reaction, i.e. they do not depend on gas density, element abundances, or gas temperature, as long as the temperature is near the ignition temperature.

E. Resonances and Screening

The simple model we have just worked out is reasonably good for many reactions of importance in stars, but it omits a number of complications, two of which we will discuss briefly.

First, the assumption that $S(E)$ varies weakly with E over the Gamow peak is not always valid. The most common way for the assumption to fail is if there is a resonance, which means that the energy of the collision corresponds closely to the energy of an excited state of the final product nucleus. If this happens, the cross section increases dramatically in a narrow range of energies, and $S(E)$ becomes sharply peaked. While none of the reactions involved in hydrogen burning in main sequence stars are resonant, some of the important reactions that occur in more evolved stars are. Resonances can enhance the reaction rate by orders of magnitude compared to what our our simple model would suggest.

A second complication is screening. Our calculation of the Coulomb barrier was based on the potential of two nuclei of charge $\mathcal{Z}_i$ and $\mathcal{Z}_j$ interacting with one another. However, this ignores the presence of electrons. For neutral atoms, the electric potential drops to zero for distances greater than a few angstroms, because the nucleus is surrounded by a cloud of electrons of equal and opposite charge. From a point outside the cloud, the net charge seen is zero, because the electronic and nuclear charges cancel – the electrons screen the nucleus. This is why neutral atoms do not violently repel one another.

In the fully ionized plasma inside a star electrons are not bound to atoms, and they float about freely. However, they are still attracted to the positively charged nuclei, and thus they tend to cluster around them, partly screening them. This effect reduces the Coulomb barrier. Screening is strongest at lower temperatures, since when $k_B T$ is smaller compared to the electric potential energy, electrons tend to to cluster more tightly around nuclei. This effect can enhance reaction rates for turning H into He by $\sim 10 - 50\%$ compared to the results of our naive calculation.

Class 8 Notes: Nuclear Chemistry in Stars

In the last class we discussed the physical process of nuclear fusion, and saw how rate coefficients for nuclear reactions are calculated. With this understanding in place, it is possible to examine to study which reactions chains are actually important in generating energy and driving the evolution of stars. Those reaction chains and their properties will be the topic of this class.

I. Characterizing Reactions

To begin, it is useful to review what we know about nuclear reactions and what we want to know. A first principle of importance is that reaction rates are very, very sensitive to temperature, so that they can go very rapidly from negligible to huge. As a result, there is usually only a narrow range of temperatures where a reaction can occur for a period of time comparable to t_{KH}, the thermal timescale of the star. At lower temperatures the reaction would produce negligible energy, and at higher temperatures it would rapidly consume all the available fuel in a time much less than t_{KH}. Because the temperature windows where reactions occur at moderate rates are narrow, it is usually (though not always) the case that, in a given region within a star, there is only one reaction (or reaction chain) that occurs over an extended period in a given part of a star.

By figuring out at what temperature such moderate reaction rates set in, we can assign a characteristic ignition temperature T_i to a given reaction or reaction chain. This ignition temperature is one of the basic things we want to know about a nuclear reaction. We would also like to know the reaction rate coefficient

$$R \approx \frac{4}{3^{5/2}\pi^2} \frac{h}{\mu_{\mathrm{red}} m_{\mathrm{H}} \mathcal{Z}_i \mathcal{Z}_j e^2} S(E_0) \tau^2 e^{-\tau},$$

where

$$\tau = \frac{3E_0}{k_B T} = 42.46 \left[\mathcal{Z}_i^2 \mathcal{Z}_j^2 \mu_{\mathrm{red}} \left(\frac{T}{10^6 \mathrm{~K}} \right)^{-1} \right]^{1/3}.$$

The corresponding rate at which the reaction generates energy per unit mass

$$q = \frac{\rho}{m_{\mathrm{H}}^2} \left(\frac{1}{1+\delta_{ij}} \right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R Q.$$

Finally, the we can approximate the reaction coefficient at some particular temperature T as a powerlaw

$$R \approx R_0 T^{\nu},$$

where the index $\nu = (\tau - 2)/3$. The energy generation rate can similarly be written $q = q_0 \rho T^{\nu}$, where the factor of ρ appears because the reaction rate varies as rate $\propto \rho^2 R$, and the energy per unit mass varies as rate$/\rho$.

II. Major Fusion Reactions

With this framework established, we can apply it to the major fusion reactions of importance in stars. A quick note on notation: to help us keep track of charges, we will write symbols for nuclei as ${}^{\mathcal{A}}_{\mathcal{Z}}\mathrm{C}$, where $\mathcal{A}$ is the atomic weight, $\mathcal{Z}$ is the atomic number, and C is the symbol for that element. This is slightly redundant, since a chemical symbol C uniquely identifies the atomic number $\mathcal{Z}$. However, writing out the numbers explicitly makes it easier to keep track of the charges, and to assure ourselves that they balance on each side of a reaction.

A. The $p-p$ Chain

1. Reaction Path

The most important mechanism for generating power in the Sun is known as the $p-p$ chain, for proton proton chain. It is not surprising that the reaction involves protons, i.e. hydrogen nuclei. These are by far the most abundance nuclei in main sequence stars, and, since the strength of the Coulomb barrier scales as $\mathcal{Z}_i\mathcal{Z}_j$, it is also the reaction with the lowest Coulomb barrier. Thus it begins at the lowest temperature.

[Slide 1 – binding energy per nucleon]

Before going into the details of the reaction, it is useful to re-examine the chart of binding energy per nucleon. Clearly the first big peak is at helium-4, so that is where we expect the reaction to go. However, getting there is not so easy, because all the stable nuclei shown in the chart except ${}^1_1\mathrm{H}$ contain neutrons. The reason is that neutrons are required to provide enough nuclear force to hold a nucleus together against the Coulomb repulsion of the protons. Thus the most obvious first step for a reaction involving two hydrogen nuclei doesn't work. We can't easily do

$$ {}^1_1\mathrm{H} + {}^1_1\mathrm{H} \rightarrow {}^2_2\mathrm{He} $$

because ${}^2_2\mathrm{He}$ is not a stable nucleus. Any ${}^2_2\mathrm{He}$ made in such a manner almost immediately disintegrates into two protons, producing no net energy release.

Thus for a reaction to generate energy, one must find a way to bypass ${}^2_2\mathrm{He}$ and jump to a stable state. One possible solution to this problem was discovered by Hans Bethe in 1939: during the very brief period that ${}^2_2\mathrm{He}$ lives, a weak nuclear reaction can occur that converts one of the protons into a neutron plus a positron plus a neutrino. The positron and neutrino, which do not feel the strong nuclear force, immediately escape from the nucleus, leaving behind a proton plus a neutron. The proton plus neutron do constitute a stable nucleus: deuterium. The net reaction is exothermic, and the excess energy mostly goes into the recoil of the deuterium and positron. This excess energy is then turned into heat when the nuclei collide with other particle. The final reaction is

$$ {}^1_1\mathrm{H} + {}^1_1\mathrm{H} \rightarrow {}^2_1\mathrm{D} + e^+ + \nu_e $$

The electron neutrino, ν_e, escapes the star immediately, while the positron very quickly encounters an electron and annihilates, producing gamma rays which are then absorbed and converted into heat:

$$e^+ + e^- \rightarrow 2\gamma,$$

where γ is the symbol for photon. As we'll discuss further in a moment, the proton-neutron conversion is very unlikely because it relies on the weak force, so the reaction coefficient for this reaction is very small compared to others in the chain. In terms of our earlier notation, $S(E_0)$ is very small for this reaction.

The next step in the chain is an encounter between the deuterium nucleus and another proton, producing helium:

$$^2_1\mathrm{D} + \, ^1_1\mathrm{H} \rightarrow \, ^3_2\mathrm{He} + \gamma.$$

This reaction goes very quickly compared to the first step, because the Coulomb barrier is the same (deuterium and ordinary hydrogen both have $\mathcal{Z} = 1$), but no weak nuclear forces are required.

For the last part of the chain, there are three possibilities, known as the pp I, pp II, and pp III branches. Branch I involves an encounter between two $^3_2$He nuclei produced in the previous step:

$$^3_2\mathrm{He} + \, ^3_2\mathrm{He} \rightarrow \, ^4_2\mathrm{He} + 2\, ^1_1\mathrm{H}.$$

This reaction has a Coulomb barrier four times higher than the first one, but, since it does not require a weak nuclear interaction, it actually proceeds faster than the first step. At this point the reaction stops, because $^4_2$He is stable, and there is no route from there to a more massive nucleus that is accessible at the temperatures of $\sim 10^7$ K where hydrogen burning occurs.

Branch II involves an encounter between the $^3_2$He and a pre-existing $^4_2$He nucleus to make beryllium, followed by capture of an electron to convert the beryllium to lithium, followed by capture of one more proton and fission of the resulting nucleus:

$$\begin{aligned} ^3_2\mathrm{He} + \, ^4_2\mathrm{He} &\rightarrow \, ^7_4\mathrm{Be} + \gamma \\ ^7_4\mathrm{Be} + e^- &\rightarrow \, ^7_3\mathrm{Li} + \nu_e \\ ^7_3\mathrm{Li} + ^1_1\mathrm{H} &\rightarrow 2\, ^4_2\mathrm{He}. \end{aligned}$$

Finally, branch III involves production of beryllium-7 just like the first step of the pp II branch, but then an encounter between that and another proton to produce boron. The boron then decays to beryllium via positron emission, and finally ends at beryllium-8, which spontaneously splits:

$$^3_2\mathrm{He} + \, ^4_2\mathrm{He} \rightarrow \, ^7_4\mathrm{Be} + \gamma$$

$$\begin{aligned}
{}^7_4\mathrm{Be} + {}^1_1\mathrm{H} &\rightarrow {}^8_5\mathrm{B} + \gamma \\
{}^8_5\mathrm{B} &\rightarrow {}^8_4\mathrm{Be} + e^+ + \nu_e \\
{}^8_4\mathrm{Be} &\rightarrow 2\,{}^4_2\mathrm{He}.
\end{aligned}$$

As before, the positron produced in the third step immediately encounters an electron and annihilates into gamma rays.

Which of these chains is most important depends on the local density, temperature, and chemical composition. Obviously pp II and pp III are more likely when there is more ${}^4_2$He around, since they require it. In Sun, pp I is 69% of all reactions, pp II is 31%, and pp III is 0.1%.

2. Energetics and Rates

The net reactions associated with these chains can be written:

$$4\,{}^1_1\mathrm{H} \rightarrow {}^4_2\mathrm{He} + \text{photons, neutrinos, and light particles},$$

where the exact number of photons, neutrinos, and light particles depends on which branch is taken. The total energy release is given by subtracting the mass of He-4 from the mass of 4 protons, and is given by

$$\Delta E = (4m_{\mathrm{H}} - m_{\mathrm{He}})c^2 = 26.73 \text{ MeV}.$$

The actual amount of energy that goes into heating up the gas depends on the amount of energy carried away by neutrinos, which escape the star. This is different for each branch, because each branch involves production of a different number of neutrinos with different energies. The neutrino loss is 2.0% for pp I, 4.0% for pp II, and 28.3% for pp III.

In any of the pp branch, the first step, which requires spontaneous conversion of a proton into a neutron, is by far the slowest, and thus the rate at which it occurs controls the rate for the entire chain. For this reason, we can calculate the rate coefficient simply by knowing the properties of this reaction. The reaction begins to occur at an ignition temperature that is roughly equal $T_i = 4 \times 10^6$ K. The Sun's central temperature $T_0 \approx 1.57 \times 10^7$ K, which gives

$$\tau = \frac{3E_0}{k_B T} = 42.46 \left[\mathcal{Z}_i^2 \mathcal{Z}_j^2 \mu_{\mathrm{red}} \left(\frac{T}{10^6 \text{ K}} \right)^{-1} \right]^{1/3} = 13.5.$$

The reaction rate varies as temperature to roughly the 4th power. Measuring the value of $S(E_0)$ for this reaction lets us compute the rate coefficient. If we do not make the powerlaw approximation and just plug into

$$R \approx \frac{4}{3^{5/2}\pi^2} \frac{h}{\mu_{\mathrm{red}} m_{\mathrm{H}} \mathcal{Z}_i \mathcal{Z}_j e^2} S(E_0) \tau^2 e^{-\tau},$$

we get

$$R \approx 6.34 \times 10^{-37} \left(\frac{T}{10^6 \text{ K}} \right)^{-2/3} \exp\left[-\frac{33.8}{(T/10^6 \text{ K})^{1/3}} \right] \text{ cm}^3/\text{ s}.$$

If we multiply this by the number density of protons, we get an estimate for the rate of reactions that a single proton undergoes. The inverse of this is the lifetime of a proton – the amount of time it takes for it to react with another proton and begin the reaction chain that will turn it into helium. Thus the proton lifetime is

$$t = \frac{1}{n_p R} = \frac{m_\mathrm{H}}{\rho X R} = \frac{8.3 \times 10^4 \text{ yr}}{X} \left(\frac{1 \text{ g cm}^{-3}}{\rho}\right) \left(\frac{T}{10^6 \text{ K}}\right)^{2/3} \exp\left[\frac{33.8}{(T/10^6 \text{ K})^{1/3}}\right].$$

Plugging in a density of 100 g cm^{-3} and a temperature of 1.5×10^7 K, the result is a bit over 10^9 yr. Thus the typical proton in the center of the Sun requires $> 10^9$ yr to undergo fusion. Averaging over a larger volume of the Sun, which has a lower density and temperature, makes the timescale even longer.

Finally, combining the reaction rate coefficient R with an energy release of $Q = 13.4$ MeV per reaction (since 26.73 MeV is what we get when we use 4 protons, and each pp reaction only uses 2), the corresponding energy generation rate is

$$\begin{aligned} q &\approx \frac{\rho}{m_\mathrm{H}^2} \left(\frac{1}{1+\delta_{ij}}\right) \frac{X_i X_j}{\mathcal{A}_i \mathcal{A}_j} R Q \\ &= 2.4 \times 10^6 X^2 \left(\frac{\rho}{1 \text{ g cm}^{-3}}\right) \left(\frac{T}{10^6 \text{ K}}\right)^{-2/3} \exp\left[-\frac{33.8}{(T/10^6 \text{ K})^{1/3}}\right] \text{ erg g}^{-1} \text{ s}^{-1}. \end{aligned}$$

If we do want to make a powerlaw fit, the index is

$$\nu = \frac{\tau - 2}{3} \approx 4.$$

B. The CNO Cycle

1. Reaction Path

The $p - p$ chain faces a relatively small Coulomb barrier, since the rate-limiting step has $\mathcal{Z} = 1$ for both reactants. However, it is slow because it requires spontaneous proton-neutron conversion within the short time that two protons are close to one another in a violently unstable configuration. There is another possible route to turning hydrogen into helium-4 which has a different tradeoff: a larger Coulomb barrier, but no need for a weak reaction in a short period.

This second route is called the CNO cycle, and was discovered independently by Hans Bethe and Carl-Friedrich von Weizsäcker in 1938. It relies on the fact that cabron, nitrogen, and oxygen are fairly abundant in the universe, and are present in a star even before it starts nuclear burning. They can act as catalysts in a proton fusion reaction. The reaction chain is

$${}^{12}_{6}\mathrm{C} + {}^{1}_{1}\mathrm{H} \rightarrow {}^{13}_{7}\mathrm{N} + \gamma$$

$$
\begin{aligned}
{}^{13}_{7}\mathrm{N} &\rightarrow {}^{13}_{6}\mathrm{C} + e^{+} + \nu_e \\
{}^{13}_{6}\mathrm{C} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{14}_{7}\mathrm{N} + \gamma \\
{}^{14}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{15}_{8}\mathrm{O} + \gamma \\
{}^{15}_{8}\mathrm{O} &\rightarrow {}^{15}_{7}\mathrm{N} + e^{+} + \nu_e \\
{}^{15}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{12}_{6}\mathrm{C} + {}^{4}_{2}\mathrm{He}
\end{aligned}
$$

Alternately, the chain can be:

$$
\begin{aligned}
{}^{14}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{15}_{8}\mathrm{O} + \gamma \\
{}^{15}_{8}\mathrm{O} &\rightarrow {}^{15}_{7}\mathrm{N} + e^{+} + \nu_e \\
{}^{15}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{16}_{8}\mathrm{O} + \gamma \\
{}^{16}_{8}\mathrm{O} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{17}_{9}\mathrm{F} + \gamma \\
{}^{17}_{9}\mathrm{F} &\rightarrow {}^{17}_{8}\mathrm{O} + e^{+} + \nu_e \\
{}^{17}_{8}\mathrm{O} + {}^{1}_{1}\mathrm{H} &\rightarrow {}^{14}_{7}\mathrm{N} + {}^{4}_{2}\mathrm{He}
\end{aligned}
$$

The first route is generally the more important one, by a large factor.

2. Energetics and Rates

Note that both of these chains have the property that it neither creates nor destroys any carbon or nitrogen nuclei. One starts with ${}^{12}_{6}\mathrm{C}$ and ends with it, or starts with ${}^{14}_{7}\mathrm{N}$ and ends with it. Thus the net reaction is exactly the same as for the $p - p$ chain:

$$
4\,{}^{1}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + \text{photons, neutrinos, and light particles,}
$$

In this sense, the carbon or nitrogen acts as a catalyst. They enable the reaction to take place, but are not themselves consumed or created by it. Since the net reaction is the same as for $p - p$, the net energy release is also the same, except for slightly different neutrino losses. For the CNO cycle, $Q \approx 25$ MeV once the neutrino losses are factored in, as opposed to 27 MeV for $p - p$.

In each of these reaction chains, it makes sense to distinguish between reactions that involve creation of a positron e^{+} and reactions that do not. The former are called β decays, and they rely on the weak nuclear force. However, they are much faster than the first step of the $p - p$ chain, because the take place in nuclei that are stable except for the weak reaction they undergo. Thus there is no need for precisely timing the reaction with the period when two protons are in close proximity.

At the temperature found in the Sun, however, the rate-limiting step is not the β decays, but the need to overcome strong Coulomb barriers. The ignition temperature is about 1.5×10^{7} K, about the Sun's central temperature. Analysis of the full reaction rate is tricky because which step is the rate-limiting one depends on the relative abundances of C, N, O, and the other catalysts, which are in turn determined by the reaction cycle itself. Once things reach equilibrium, however, it turns out that the step

$$ {}^{14}_{7}\mathrm{N} + {}^{1}_{1}\mathrm{H} \;\rightarrow\; {}^{15}_{8}\mathrm{O} + \gamma $$

is the rate-limiting one. This step appears in both cycles.

Plugging $\mathcal{Z}_i = 7$, $\mathcal{Z}_j = 1$, and $\mu_{\rm red} = (14)(1)/(14{+}1) = 0.93$ into our equation for the temperature-dependence gives

$$ \tau = \frac{3E_0}{k_B T} = 42.46 \left[\mathcal{Z}_i^2 \mathcal{Z}_j^2 \mu_{\rm red} \left(\frac{T}{10^6 \text{ K}} \right)^{-1} \right]^{1/3} = 152 \left(\frac{T}{10^6 \text{ K}} \right)^{-1/3} $$

Using the laboratory measurement for $S(E_0)$ for this reaction, the rate coefficient is

$$ R = 8.6 \times 10^{-19} \left(\frac{T}{10^6 \text{ K}} \right)^{-2/3} \exp\left[-\frac{152}{(T/10^6 \text{ K})^{1/3}} \right] \text{ cm}^3/\text{ s}, $$

and the corresponding energy generation rate is

$$ q = 8.7{\times}10^{27} X X_{\rm CNO} \left(\frac{\rho}{1 \text{ g cm}^{-3}} \right) \left(\frac{T}{10^6 \text{ K}} \right)^{-2/3} \exp\left[-\frac{152}{(T/10^6 \text{ K})^{1/3}} \right] \text{ erg g}^{-1} \text{ s}^{-1}, $$

where $X_{\rm CNO}$ is the total mass fraction of carbon, nitrogen, and oxygen. This is roughly $Z/2$, where Z is the total mass fraction of metals.

It is informative to evaluate q for the $p-p$ chain and for the CNO cycle using values appropriate to the center of the Sun: $\rho \approx 100$ g cm^{-3}, $T \approx 1.5 \times 10^7$ K, $X = 0.71$, $Z = 0.02$. This gives

$$ \begin{aligned} q_{p-p} &= 82 \text{ erg g}^{-1} \text{ s}^{-1} \\ q_{\rm CNO} &= 6.4 \text{ erg g}^{-1} \text{ s}^{-1} \end{aligned} $$

Thus the $p-p$ chain dominates in the Sun by about a factor of 10. However, it is important to notice that, because it has 152 instead of 33.8 in the exponential, the CNO cycle is much more temperature-sensitive than the $p-p$ chain. If we assign a powerlaw approximation, the index is

$$ \nu = \frac{\tau - 2}{3} = 20. $$

Thus stars a bit more massive than the Sun, which we will see have higher central temperatures, the CNO cycle dominates. In stars smaller than the Sun, the CNO cycle is completely irrelevant.

This also brings out a general feature of all the nuclear reactions we will consider: the temperature-sensitivity is determined by τ, and τ in turn depends on the charges of the nuclei involved, $\mathcal{Z}$, because it is determined by the Coulomb barrier. The stronger the nuclear charge, the stronger the Coulomb barrier, and thus the higher the ignition temperature and the more temperature-sensitive the reaction becomes. We have already seen that the CNO cycle produces energy as a rate that varies as T^{20}, and the temperature-sensitivity only gets stronger as we march up the periodic table.

C. The Triple-α Process

1. Reaction Path

The transition from hydrogen to helium-4 takes us to the first big peak in the binding energy per nucleon curve, and moves us from a reaction where the Coulomb barrier to the first step is 1 ($\mathcal{Z}_i = \mathcal{Z}_j = 1$) to one where it is 4 ($\mathcal{Z}_i = \mathcal{Z}_j = 2$). In fact, we'll see in a moment that it is even worse than that. As a result, the temperature required to burn any further up the periodic table is significantly higher than that required to burn hydrogen. Hydrogen burning tends to keep stars' central temperatures about constant while hydrogen remains available, so helium burning is generally not initiated until all the hydrogen has been exhausted, which means that the star must leave the main sequence. Thus all the reactions we will talk about for the remainder of the class take place in post-main-sequence stars.

[Slide 1 – binding energy per nucleon]

In the core of such a star, essentially all the hydrogen will have been converted to helium-4. Looking at the curve of binding energy per nucleon, the next peak is clearly carbon-12. However, we again see a problem in getting there. The most obvious reaction to start is

$$^4_2\mathrm{He} + {}^4_2\mathrm{He} \leftrightarrows {}^8_4\mathrm{Be}.$$

However, the $^8_4$Be nucleus is violently unstable, and disintegrates in about 3×10^{-16} s. Nor can we get out of the problem by hoping for a weak reaction to convert a proton into a neutron, because there is no stable nucleus with an atomic mass number $\mathcal{A} = 8$.

Thus we need to jump past atomic number 8 in order to burn He. The solution to this problem was found by Edwin Salpeter in 1952. If the density and temperature get high enough, it may be possible for the $^8_4$Be nucleus to collide with another $^4_2$He nucleus before it decays. Then it will undergo the reaction

$$^8_4\mathrm{Be} + {}^4_2\mathrm{He} \rightarrow {}^{12}_6\mathrm{C} + \gamma,$$

and arrive at carbon-12, which is stable and is another peak of binding energy per nucleon. This reaction is known as the triple-α process, because it effectively involves a three-way collision between three helium-4 nuclei, which are also known as α particles. It is not a true three-way collision, because some extra time for the third collision is provided by the lifetime of the $^8_4$Be nucleus, but it is nearly so.

In addition to the short-lived beryllium state, another factor that helps this reaction go is the existence of a resonance in the carbon-12 nucleus that coincides closely in energy with that produced by colliding another helium nucleus with beryllium-8. This greatly enhances the rate at which the second step in the reaction chain takes place.

In environments where a significant amount of carbon builds up and the temperature is high, carbon will occasionally capture an additional helium nucleus and jump to the next peak in the binding energy curve, oxygen-16:

$$^{12}_{6}\mathrm{C} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{16}_{8}\mathrm{O} + \gamma.$$

Thus stars in which the triple-α process takes place wind up containing a mixture of carbon and oxygen, with the exact ratio depending on their age, density, and temperature. Further He captures are also possible, but become increasingly unlikely as one moves up in atomic number due to the increasing Coulomb barrier.

2. Energetics and Rates

The first step in the triple-α process is actually endothermic, although only mildly so. The mildly endothermic nature of the reaction is important. The fact that it is endothermic is the reason that the $^{8}_{4}\mathrm{Be}$ nucleus is unstable: it can spontaneously split back into two helium nuclei. That is is only mildly endothermic (it requires 92 keV) means that collision between nuclei moving around at the thermal speed are always producing some of it, so there is always a small amount of beryllium-8 present.

The second step, converting beryllium-8 to carbon, is exothermic. The net energy released can be calculated by comparing the mass of the carbon-12 nucleus to that of three helium-4 nuclei:

$$Q = (3m_{\mathrm{He}} - m_{\mathrm{C}})c^2 = 7.28 \text{ MeV}.$$

The capture of a fourth He nucleus leading to oxygen-16 yields another 7.16 MeV.

To compute the reaction rate and its temperature-dependence, one can assume that there is always a small amount of $^{8}_{4}\mathrm{Be}$ by equating the creation and destruction rates – a process that we will not go through, but which yields an amount of $^{8}_{4}\mathrm{Be}$ that is roughly independent of temperature. It does not depend on temperature because the limiting factor in how much beryllium is present is the very rapid spontaneous decay of the beryllium nucleus, not the Coulomb barrier to creating it. Calculations show that the beryllium fraction is ~ 1 part in 10^{10}.

All the temperature-dependence in the reaction rate comes in the next step that of converting beryllium-8 to carbon-12. As discussed a moment ago, the reaction process for creating carbon-12 depends on a resonance. We will not go through the details of how to calculate a resonant reaction in class, but we can sketch it out briefly in order to understand the temperature-dependence of the reaction. Recall from last class that the reaction rate is proportional to

$$R \propto \int_0^\infty \sigma(E) E e^{-E/k_B T} \, dE.$$

For a non-resonant reaction, we evaluated this by using a calculation of quantum tunneling to estimate $\sigma(E)$. For a resonant reaction, however, the process is much simpler: when there is a dominant resonance, essentially all reactions take place at energies very close to the energy required to hit the resonance. For this reason we can treat the factor $Ee^{-E/k_B T}$ as nearly constant over the resonance, and take it out of the integral, yielding

$$R \propto E_R e^{-E_R/k_B T} \int_0^\infty \sigma(E) \, dE,$$

where E_R is the energy that the incoming particle must have in order to hit the resonance. Then if we let

$$\tau_R = \frac{E_R}{k_B T},$$

we have

$$R \propto e^{-\tau_R} \int_0^\infty \sigma(E) \, dE.$$

As with the non-resonant case, all the temperature-dependence is encapsulated in the parameter τ_R, which varies as T^{-1}.

The second step in the triple-α process relies on a resonance that is at an energy $E_R = 379.5$ keV above the energy of the beryllium-8 nucleus, so that is the energy an incoming particle must have to trigger the resonance. (Note that the state in question has an energy 7.95 MeV above the ground state of carbon-12, but the relevant question is the difference between that energy and the energy of the beryllium-8 nucleus, which is much smaller.)

$$\tau_R = \frac{379.5\,\mathrm{keV}}{k_B T} = 44.0 \left(\frac{10^8\,\mathrm{K}}{T} \right).$$

This is normalized to 10^8 K, which is about the ignition temperature for this reaction.

To go further in computing the reaction rate, we must recall that triple-α effectively requires a three-way collision. For a single particle, we said that the rate at which it encounters other particles is proportional to n. We can also view this as a probability: the probability of a collision per unit time is proportional to n. For a three-body process, we need to ask about the probability of two of them striking simultaneously or nearly so (within the

10^{-16} s lifetime of the $^8_4$Be nucleus. The rate at which such double-collisions occurs is proportional to the probability of one collision times the probability of another: n^2v^2. Thus we expect a collision rate that varies as n^2. We will not walk through putting this in terms of a rate coefficient, but a straightforward generalization of our existing calculation shows that the reaction rate per unit volume varies as Rn^3, while the kinetic part rate coefficient itself varies as T^{-3} – it is T^{-3} instead of the usual $T^{-3/2}$ because the collision rate varies as n^2 rather than n.

Putting this together, we expect the rate coefficient to vary with temperature as

$$R \propto T^{-3} \int_0^\infty \sigma(E) E e^{-E/k_B T} \, dE \propto \tau_R^3 e^{-\tau_R} \int_0^\infty \sigma(E) \, dE.$$

It will vary with density as ρ^2. Putting in the measured cross sections, the final result for the energy generation rate is

$$q = 5.1 \times 10^8 Y^3 \left(\frac{\rho}{1\,\mathrm{g\ cm^{-3}}}\right)^2 \left(\frac{T}{10^8\,\mathrm{K}}\right)^{-3} \exp\left(-\frac{44}{T/10^8\,\mathrm{K}}\right)\ \mathrm{erg\ g^{-1}\ s^{-1}}.$$

To get a powerlaw fit for the temperature-dependence, we can proceed exactly as we did in the non-resonant case. The slope in a log-log plot of R vs. T is

$$\begin{aligned}
\nu = \frac{d\ln R}{d\ln T} &= \frac{d}{d\ln T}\left(3\ln\tau_R - \tau_R + \mathrm{const}\right) \\
&= \frac{d}{d\ln T}\left[-3\ln T - \tau_{R,0}\left(\frac{T}{T_0}\right)^{-1}\right] \\
&= -3 - \tau_{R,0} T_0 T \frac{d}{dT}\left(\frac{1}{T}\right) \\
&= -3 + \frac{\tau_{R,0} T_0}{T} \\
&= \tau_R - 3 \\
&\approx 41
\end{aligned}$$

Thus we see that the triple-α reaction is extraordinarily temperature-sensitive.

D. Carbon and Oxygen Burning

At even higher temperatures, the Coulomb barrier for oxygen and carbon can be overcome, creating yet heavier nuclei. Many reaction paths are possible. Carbon can burn to produce

$$^{12}_{6}\mathrm{C} + {}^{12}_{6}\mathrm{C} \rightarrow \begin{cases} ^{16}_{8}\mathrm{O} + 2\,{}^{4}_{2}\mathrm{He} \\ ^{20}_{10}\mathrm{Ne} + {}^{4}_{2}\mathrm{He} \\ ^{23}_{11}\mathrm{Na} + {}^{1}_{1}\mathrm{H} \\ ^{23}_{12}\mathrm{Mg} + n \\ ^{24}_{12}\mathrm{Mg} + \gamma \end{cases}$$

Oxygen burning produces similar reactions, which I will not write down – they are in your textbook. The final outcome of this sort of burning is generally $^{32}_{16}$Si, silicon.

These reactions begin at temperatures of about 6×10^8 K. Given the number of possible pathways and reactions involved, there isn't a single simple formula for the energy generation rate. Calculations of reactions of this sort are generally done by a computer, which keeps track of the densities of many different nuclei and calculates all the possible reactions and the energies they yield.

E. Silicon to Iron

At still higher temperatures, around 3×10^9 K, the typical photon becomes energetic enough that it can disrupt nuclei, knocking pieces off them in a process known as photodisintegration. The chemical balance in the star is then determined by a competition between this process and reactions between nuclei. However, as we might expect, the net effect is to drive the chemical balance ever further toward the most stable nucleus, iron. Once the temperature around 3×10^9 K, more and more nuclei begin to convert to $^{56}_{26}$Fe, and its close neighbors cobalt and nickel. Things stay in this state until the temperature is greater than about 7×10^9 K, at which point photons have enough energy to destroy even iron, and the entire process reverses: all elements are converted back into its constituent protons and neutrons, and photons reign supreme.

III. The r and s Processes

We have already seen how elements up through iron are built, but we have not yet mentioned how even heavier elements can be created. The answer is that they are not made in stars under normal circumstances, because when the only forces at work at electromagnetism and nuclear forces, it is never energetically favorable to create such elements in any significant number. Creating such elements requires the intervention of another force: gravity.

When stars are in the process of being crushed by gravity, right before they explode as supernovae (which we will discuss toward the end of the course), gravity drives a process that converts most of the protons to neutrons. This creates a neutron-rich environment unlike any found at earlier stages of stellar evolution, when the lack of neutrons was often the rate-limiting step.

In a neutron-rich environment, it becomes possible to create heavy nuclei via the absorption of neutrons by existing nuclei. Since the neutrons are neutral, there is no Coulomb barrier to overcome, and the reaction proceeds as quickly as the neutron supply allows. Reactions look like this:

$$\begin{aligned} I_0(\mathcal{A}, \mathcal{Z}) + n &\rightarrow I_1(\mathcal{A}+1, \mathcal{Z}) \\ I_1(\mathcal{A}+1, \mathcal{Z}) + n &\rightarrow I_2(\mathcal{A}+2, \mathcal{Z}) \\ I_2(\mathcal{A}+2, \mathcal{Z}) + n &\rightarrow I_3(\mathcal{A}+3, \mathcal{Z}) \\ &\text{etc.} \end{aligned}$$

Neutron captures therefore increase $\mathcal{A}$ at constant $\mathcal{Z}$.

This continues until it produces a nucleus that is unstable and undergoes β decay, converting one of the neutrons back into a proton:

$$I_N(\mathcal{A}+N,\mathcal{Z}) \rightarrow J(\mathcal{A}+N,\mathcal{Z}+1)+e^-+\overline{\nu}.$$

(The bar over the ν indicates that this is an anti-neutrino.) If the new element produced in this way is stable, it will begin neutron capturing again. If not, it will keep undergoing β decays until it becomes stable:

$$\begin{aligned} J(\mathcal{A}+N,\mathcal{Z}+1) &\rightarrow K(\mathcal{A}+N,\mathcal{Z}+2)+e^-+\overline{\nu} \\ K(\mathcal{A}+N,\mathcal{Z}+2) &\rightarrow L(\mathcal{A}+N,\mathcal{Z}+3)+e^-+\overline{\nu} \\ &\text{etc.} \end{aligned}$$

β decays therefore increase $\mathcal{Z}$ at constant $\mathcal{A}$.

These processes together lead to the build-up of elements heavier than iron. The chain stops if at any point it reaches a nucleus that is stable against β decay, and is also not able to capture neutrons because neutron capture is endothermic for it, i.e. adding a neutron makes the nucleus less rather than more bound.

The rate of neutron capture depends on the local density and temperature, while the rate of β decay does not. As a result, either can be the rate-limiting step in the build-up, depending on the local environment and the element in question. Elements that are build up by reaction chains in which β decays occur faster are called r process, for rapid. Elements where β decays are slower are called s process, for slow. Knowing which process produces which elements requires knowing the stability, binding energy, and β decay lifetimes of the various elements, which must be determined experimentally.

Astronomy 112: The Physics of Stars

Class 9 Notes: Polytropes

With our discussion of nuclear reaction rates last time, we have mostly completed our survey of the microphysical properties of stellar matter – its pressure, how energy flows through it, and how it generates energy from nuclear reactions. For the next few weeks we will be using those microphysical models to begin to make our first models of stars.

I. The Stellar Structure Equations

To begin we will collect the various equations we have developed thus far to describe the behavior of material in stars. As always, we consider a shell of material of mass dm and thickness dr, which is at a distance r from the center of the star and has a mass m interior to it. The shell has density ρ, pressure P, temperature T, and opacity κ. The radiation flux passing through it is F, and the shell generates energy via nuclear reactions at a rate per unit mass q. We will write down the equations describing this shell, under the assumption that the star is in both hydrostatic and thermal equilibrium, so we can drop all time derivatives describing change in position, energy, etc. Since we are assuming equilibrium, for now we will also assume that the composition is fixed, so that we know X, Y, Z, and any other quantities we need that describe the chemical makeup of the gas.

The first equation is just the definition of the density for the shell, which says that $\rho = dm/dV$. Writing this in Eulerian or Lagrangian form (i.e. with either radius or mass as the independent variable), we have

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

The second equation is the equation of hydrostatic balance, which we can also write in either Eulerian or Lagrangian form:

$$\frac{dP}{dr} = -\frac{Gm}{r^2}\rho \qquad\qquad \frac{dP}{dm} = -\frac{Gm}{4\pi r^4}.$$

This just equates the change in gradient in pressure with the force of gravity. The third equation is the equation describing how the temperature changes with position within a star as a result of radiative diffusion:

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F}{4\pi r^2} \qquad\qquad \frac{dT}{dm} = -\frac{3}{4ac}\frac{\kappa}{T^3}\frac{F}{(4\pi r^2)^2}.$$

Finally, we have the equation of energy conservation, which for material in equilibrium just equates the change in energy flux across a shell to the rate at which nuclear reactions generate power within it:

$$\frac{dF}{dr} = 4\pi r^2 \rho q \qquad\qquad \frac{dF}{dm} = q.$$

These are four coupled, non-linear ordinary differential equations. As we discussed a few weeks ago, by themselves they are not a complete system, because by themselves they contain more unknowns than equations. The unknowns appearing are ρ, T, P, F, κ, and q. If we adopt Eulerian coordinates, then m is also an unknown, and r is the independent variable. For Lagrangian coordinates, r is also an unknown, and m is the independent variable. Thus we have seven unknowns, but only four equations. We therefore need three more equations, and that is what we have spent the last few weeks providing.

The pressure depends on density and temperature via

$$P = \frac{\mathcal{R}}{\mu_I}\rho T + P_e + \frac{1}{3}aT^4.$$

The first term is the ion pressure, which we have written assuming that ions are non-degenerate, which they are except in neutron stars. The last term is the radiation pressure. The middle term, the electron pressure, takes a form that depends on whether the electrons are degenerate or not, but which is a known function of ρ and T.

The opacity and nuclear energy generate rate, we have seen, are in general quite complicated functions. However, we have also seen that they can be approximated reasonably well as powerlaws:

$$\kappa = \kappa_0 \rho^a T^b \qquad\qquad q = q_0 \rho^m T^n.$$

Regardless of whether we make the powerlaw approximation or not, we now know how to compute κ and q from ρ and T.

Thus we have written down three more equations involving the unknowns. We are therefore up to seven equations for our seven unknowns, which is sufficient to fully specify the system. The only thing missing is boundary conditions, since differential equations produce constants of integration that must be determined by boundary conditions. Since there are four differential equations, we need four boundary conditions. Three of them are obvious. In Lagrangian coordinates, we have $r = 0$ and $F = 0$ at $m = 0$, and $P = 0$ at $m = M$. In words, the first condition says that the innermost mass element must reside at radius $r = 0$, and it must have zero flux ($F = 0$) entering it from below. The third condition says that the pressure falls to zero at the boundary of the star, $m = M$.

The fourth condition is slightly more complicated. The simplest approach is to set $T = 0$ at the star's surface. This is actually not a terrible approximation, since the temperature at the surface is very low compared to that in the interior. A better approach is to specify the relation between flux and temperature at the stellar surface as $F = 4\pi r^2 \sigma T^4$. An even better approach is to make a detailed model of a stellar atmosphere and figure out how the flux through it depends on its temperature and pressure, and use that as a boundary condition for the stellar model.

The set of seven equations and four boundary conditions we have written down now fully specifies the structure of a star. Solving those equations, however, is another matter entirely. There is no general method for solving sets of coupled non-linear differential-algrebraic equations subject to boundary conditions specified at two points. There is every reason to believe that such equations cannot, in general, be solved in closed form.

Today the standard approach is to hand the problem to a computer. A computer can integrate the equations and find solutions to any desired level of accuracy, and this problem is sufficiently simple that the calculations will run on an ordinary desktop machine in a matter of seconds. However, in the days when people first approached these problems, there were no such things as computers. Instead, people were forced to come up with analytic approximations, and it turns out that one can understand a great deal about the behavior of stars using such approximations. (It turns out that a significant fraction of being a good physicist consists in the ability to come up with good approximations for intractable differential equations.)

II. Polytropes

A. Definition and Motivation

The first two stellar structure equations, describing the definition of density and hydrostatic equilibrium, are linked to the second two only via the relationship between pressure and temperature. If we can write the pressure in terms of the density alone, without reference to the temperature, then we can separate these two equations from the others and solve them by themselves. Solving two differential equations (plus one algebraic equation relating P and ρ is much easier than solving seven equations.

As a first step in this strategy, we can combine the first two first-order ODEs into a single second-order ODE. To do so, we start with the equation of hydrostatic equilibrium and multiply by r^2/ρ to obtain

$$\frac{r^2}{\rho}\frac{dP}{dr} = -Gm.$$

Next we differentiate both sides:

$$\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dP}{dr}\right) = -G\frac{dm}{dr}.$$

Finally, we substitute for dm/dr using the definition of density, $dm/dr = 4\pi r^2 \rho$. Doing so we obtain

$$\frac{1}{r^2}\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dP}{dr}\right) = -4\pi G\rho.$$

This is just another form of the equation of hydrostatic balance, this time with the definition of density explicitly substituted in.

Thus far everything we have done is exact. Now we make our approximation. We approximate that the pressure and density are related by a powerlaw

$$P = K_P \rho^{\gamma_P}.$$

Equations of state of this sort are called polytropes. For historical reasons, it is common to define

$$\gamma_P = 1 + \frac{1}{n} \text{ or } n = \frac{1}{\gamma_P - 1}$$

where we call n the polytropic index.

Before going any further, it is important to consider whether an equation of state like this is at all sensible. Why should a star ever obey such an equation of state? The answer to this question becomes clearer if we recall that, for an adiabatic gas, the equation of state reads

$$P = K_a \rho^{\gamma_a},$$

where K_a is the adiabatic constant and adiabatic index. It is important to understand the difference between this relation and the polytropic relation. The polytropic relation describes how the pressure changes with density inside as one moves through a star, while the adiabatic equation of state describes how a given gas shell would respond to being compressed. The constant K_a depends on the entropy of the gas in a given shell, so different shells in a star can have different values of K_a. If different shells have different K_a values, then as I move through the star the pressure will not vary as $P \propto \rho^\gamma$, because different shells will have different constants of proportionality. Thus a star can be described by a polytropic relation only if K_a is the same for every shell.

While this condition might seem far-fetched, it is actually satisfied under a wide range of circumstances. One circumstance when it is satisfied is if a star is dominated by the pressure of degenerate electrons, since in that case we proved a few classes back that $P_e = K_1'(\rho/\mu_e)^{5/3}$ for a non-relativistic gas, or $P_e = K_2'(\rho/\mu_e)^{4/3}$. The proportionality constants K_1' and K_2' depend only on constants of nature like h, c, the electron mass, and the proton mass, and thus do not vary within a star. For such stars, $\gamma_P = 5/3$ or $4/3$, corresponding to $n = 1.5$ or $n = 3$, for the non-relativistic and highly-relativistic cases, respectively.

Another situation where K_a is constant is if a star is convective. As we will discuss in a week or so, under some circumstances the material in a star can be subject to an instability that causes it to move around in such a way as to enforce that the entropy is constant. In a region undergoing convection, K_a does not vary from shell to shell, and a polytropic equation of state is applicable. Significant fractions of the mass in the Sun and similar stars are subject to convection, and for those parts of stars, a polytropic equation of state applies well, and is a good approximation for the star as a whole. In low mass stars where the gas is non-relativistic and radiation pressure is not significant, $\gamma_a = 5/3$, so constant K_a means that $\gamma_P = 5/3$ as well, and $n = 1.5$.

Because they apply in a broad range of situations, polytropic models turn out to be extremely useful.

B. The Lane-Emden Equation

Having motivated the choice $P = K_P \rho^{\gamma_P} = K_P \rho^{(n+1)/n}$ for an equation of state, we can proceed to substitute it into the equation of hydrostatic balance. Doing so, after a little bit of simplification we find

$$\frac{(n+1)K_P}{4\pi G n} \frac{1}{r^2} \frac{d}{dr} \left(\frac{r^2}{\rho^{(n-1)/n}} \frac{d\rho}{dr} \right) = -\rho.$$

As a second-order equation, this ODE requires two boundary conditions. At the surface $r = R$, the pressure $P(R) = 0$, and since $P \propto \rho^{\gamma_P}$, one of our boundary conditions could be $\rho(R) = 0$. However, usually one instead sets the density at the center, $\rho(0) = \rho_c$, and then R is the radius at which ρ first goes to zero. For the second boundary condition, recall that the original equation of hydrostatic equilibrium read $dP/dr = -\rho G m/r^2$. Unless the density becomes infinite in the center, m must vary as $m \propto \rho_c r^3$ near the center of the star, so m/r^2 must vary as r. Thus dP/dr goes to zero at $r = 0$, and, for our polytropic equation of state, $d\rho/dr$ must therefore approach zero in the center as well. This is our second boundary condition: $d\rho/dr = 0$ at $r = 0$.

Note that our equation now involves three constants: K_P and n, which come from the equation of state, and R, which sets the total radius of the star. Since these are the only constants that appear (other than physical ones), these must fix the solution. In other words, for a given polytrope with a given choice of K_P, n, and R, there is a single unique density profile $\rho(r)$ which is in hydrostatic equilibrium.

In fact, it is even simpler than that, as becomes clear if we make a change of variables. Let ρ_c be the density in the center of the star, and let us define the new variable Θ by

$$\frac{\rho}{\rho_c} = \Theta^n.$$

Note that Θ is a dimensionless number, and that it runs from 1 at the center of the star to 0 at the edge of the star. With this change of variable, we can re-write the equation of hydrostatic balance as

$$\left[\frac{(n+1)K_P}{4\pi G \rho_c^{(n-1)/n}} \right] \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{d\Theta}{dr} \right) = -\Theta^n.$$

The quantity in square brackets has units of length squared, and this suggests a second change of variables. We let

$$\alpha^2 = \left[\frac{(n+1)K_P}{4\pi G \rho_c^{(n-1)/n}} \right]$$

and we then set

$$\xi = \frac{r}{\alpha}$$

so that ξ is also a dimensionless number. With this substitution, the equation becomes

$$\frac{1}{\xi^2} \frac{d}{d\xi} \left(\xi^2 \frac{d\Theta}{d\xi} \right) = -\Theta^n.$$

The two central boundary conditions are now $\Theta = 1$ and $d\Theta/d\xi = 0$ at $\xi = 0$. The equation we have just derived is called the Lane-Emden Equation. Clearly the only constant in the equation is n.

Before moving on to solve the equation, it is worth mentioning the procedure used to derive it, which is a very general and powerful one. This procedure is called non-dimensionalization. The basic idea is to take an equation relating physical quantities, like the hydrostatic balance equation we started with, and re-write all the physical quantities as dimensionless numbers times their characteristic values. For example, we re-wrote the density as $\rho = \rho_c \Theta^n$. We re-wrote the length as $r = \alpha\xi$. The advantage of this approach is that it allows us to factor out all the dimensional quantities and leave behind only a pure mathematical equation, and in the process we often discover that the underlying problem does not depend on the dimensional quantities. In this case, it was far from obvious that the structure of a polytrope didn't depend on R, ρ_c, K_P, etc. After all, those quantities appear in the hydrostatic balance equation. However, the non-dimensionalization procedure shows that they just act as multipliers on an underlying solution whose behavior depends only on n. Tricks like this come up all the time in the study of differential equations, and are well worth remembering if you plan to think about them at any point in the future.

With that aside out of the way, consider the Lane-Emden equation itself. To get a sense of how it behaves, we can solve it for some chosen values of n. First consider $n = 0$, corresponding to the limit $\gamma_P \to \infty$. In this case the equation is

$$\frac{d}{d\xi}\left(\xi^2 \frac{d\Theta}{d\xi}\right) = -\xi^2,$$

which we can integrate to obtain

$$\xi^2 \frac{d\Theta}{d\xi} = -\frac{\xi^3}{3} + C,$$

where C is a constant of integration. Bringing the ξ^2 to the other side and integrating again gives

$$\Theta = -\frac{\xi^2}{6} - \frac{C}{\xi} + D.$$

Applying the boundary conditions that $\Theta = 1$ and $d\Theta/d\xi = 0$ at $\xi = 0$, we immediately see that we must choose $C = 0$ and $D = 1$; so the solution is therefore

$$\Theta = 1 - \frac{\xi^2}{6}.$$

This is a function that decreases monotonically at $\xi > 0$, and reaches 0 at $\xi = \xi_1 = \sqrt{6}$. Analytic solutions also exist for $n = 1$ and $n = 5$.

Numerically integrating the equation for other values of n shows that, for $n < 5$, this is generic behavior: Θ decreases monotonically with ξ and reaches zero at some finite value ξ_1. As we will see in a moment, it is particularly useful to know ξ_1 and $-\xi_1^2(d\Theta/d\xi)_{\xi_1}$, and these can be obtained trivially from the numerical solution. Some reference values are

n	ξ_1	$-\xi_1^2(d\Theta/d\xi)_{\xi_1}$
1.0	3.14	3.14
1.5	3.65	2.71
2.0	4.35	2.41
2.5	5.36	2.19
3.0	6.90	2.02

The radius at which ξ reaches zero is clearly the radius of the star, so

$$R = \alpha\xi_1.$$

Similarly, given a solution $\Theta(\xi)$, we can also compute the mass of the star:

$$\begin{aligned} M &= \int_0^R 4\pi r^2 \rho \, dr \\ &= 4\pi\alpha^3\rho_c \int_0^{\xi_1} \xi^2 \Theta^n \, d\xi \\ &= -4\pi\alpha^3\rho_c \int_0^{\xi_1} \frac{d}{d\xi}\left(\xi^2 \frac{d\Theta}{d\xi}\right) d\xi \\ &= -4\pi\alpha^3\rho_c\xi_1^2 \left(\frac{d\Theta}{d\xi}\right)_{\xi_1}. \end{aligned}$$

In the third step we used the equation of hydrostatic balance to replace $\xi^2\Theta^n$.

From a polytropic model, we can derive a bunch of other useful numbers and relationships. As one example, it is often convenient to know how centrally concentrated a star is, i.e. how much larger its central density is than its mean density. We define this quantity as

$$\begin{aligned} D_n &\equiv \frac{\rho_c}{\overline{\rho}} \\ &= \rho_c \frac{4\pi R^3}{3M} \\ &= \frac{4\pi}{3}\rho_c\left(\alpha\xi_1\right)^3 \left[-4\pi\alpha^3\rho_c\xi_1^2 \left(\frac{d\Theta}{d\xi}\right)_{\xi_1}\right]^{-1} \\ &= -\left[\frac{3}{\xi_1}\left(\frac{d\Theta}{d\xi}\right)_{\xi_1}\right]^{-1} \end{aligned}$$

A second useful relationship is between mass and radius. We start by expressing the central density ρ_c in terms of the other constants in the problem and our length scale α:

$$\rho_c = \left[\frac{(n+1)K_P}{4\pi G\alpha^2}\right]^{n/(n-1)}.$$

Next we substitute this into the equation for the mass:

$$M = -4\pi\alpha^3\left[\frac{(n+1)K_P}{4\pi G\alpha^2}\right]^{n/(n-1)}\xi_1^2\left(\frac{d\Theta}{d\xi}\right)_{\xi_1}.$$

Finally, we substitute $\alpha = R/\xi_1$. Making the substitution and re-arranging, we arrive at

$$\left[\frac{GM}{-\xi_1^2(d\Theta/d\xi)_{\xi_1}}\right]^{n-1}\left(\frac{R}{\xi_1}\right)^{3-n} = \frac{[(n+1)K_P]^n}{4\pi G}$$

Thus mass and radius are related by $M \sim R^{(n-3)/(n-1)}$.

A third useful expression is for the central pressure. From the equation of state we have $P_c = K_P\rho_c^{(n+1)/n}$. We can then use the mass-radius relation to solve for K_P and then substitute it into the equation of state, which gives

$$P_c = \frac{(4\pi G)^{1/n}}{n+1}\left[\frac{GM}{-\xi_1^2(d\Theta/d\xi)_{\xi_1}}\right]^{(n-1)/n}\left(\frac{R}{\xi_1}\right)^{(3-n)/n}\rho_c^{(n+1)/n}.$$

Then, using the centrally concentrated measure, D_n, to elliminate R, we get

$$P_c = (4\pi)^{1/3}B_nGM^{2/3}\rho_c^{4/3}$$

where

$$B_n = -\left[(n+1)\left(\xi_1^2\left(\frac{d\theta}{d\xi}\right)_{\xi_1}\right)^{2/3}\right]^{-1}$$

which is relatively independent of n. As we will see, all realistic models have $n = 1.5$ to $n = 3$. Thus we expect a nearly universal relation among the central pressure, central density and mass of stars.

C. The Chandrasekhar Mass and Relativistic Gasses

Consider what this analysis of polytropes implies for stars where the pressure is dominated by electron degeneracy pressure. White dwarf stars are examples of such stars. To see why, consider what their position on the HR diagram and their masses tells us about them. Observations of white dwarfs in binary systems imply that they have masses comparable to the mass of the Sun. On the other hand, their extremely low luminosities, combined with surface temperatures that are not very different from that of the Sun, implies that they must have very small radii: $r \sim r_E = 6\times 10^8$ cm. The corresponding mean density is $\rho \sim 2\times 10^6$ g cm^{-3}. In contrast, we showed a few classes ago that electrons become degenerate at a density above $\rho/\mu_e \approx 750(T/10^7\,\mathrm{K})^{3/2}$ g cm^{-3}. Thus, unless white dwarf interiors are hotter than $\sim 10^9$ K, which they do not appear to be, the gas must be degenerate.

As long as the gas is non-relativistic, this implies that its pressure is $P = K_1'(\rho/\mu_e)^{5/3}$, where K_1' is a constant that depends only on fundamental constants. Thus a white dwarf is a polytrope, and $\gamma_a = \gamma_P = 5/3$ implies that $n = 3/2$. For a polytrope of index n, we have just shown that mass and radius are related by

$$R \propto M^{(n-1)/(n-3)} = M^{-1/3}.$$

Thus more massive white dwarfs have smaller radii, with the radius falling as $M^{-1/3}$. It is important to realize that this statement is not just true of a particular white dwarf, it is true of *all* white dwarfs. The constant of proportionality between M and R depends only on K_P and G, and for a degenerate gas $K_P = K_1'/\mu_e^{5/3}$. For a given composition (fixed μ_e), this is a universal constant of nature, since K_1' depends only on fundamental constants. Thus all white dwarfs in the universe follow a common mass-radius relation that is imposed by quantum physics.

However, this cannot continue forever, with ever more massive white dwarfs having smaller and smaller radii. The thing that breaks is that, as M increases and R shrinks, eventually the electrons in the white dwarf are forced by the Pauli exclusion principle to occupy higher and higher momenta. As a result, the gas becomes relativistic. The mean density varies as

$$\overline{\rho} \propto \frac{M}{R^3} \propto M^2,$$

and we showed a few weeks that a degenerate gas becomes relativistic when its density exceeds

$$\frac{\rho}{\mu_e} \approx 3 \times 10^6 \text{ g cm}^{-3}.$$

We got a density of 2×10^6 g cm^{-3} for a mass of $M_\odot$ and a radius of r_E, so it doesn't take much of an increase in mass to push the star above the relativistic limit.

For a relativistic electron gas, we have shown that $P = K_2'(\rho/\mu_e)^{4/3}$, where, as with K_1', the quantity K_2' depends only on fundamental constants. Since such a gas has $\gamma_a = 4/3$ and fixed K_a, it can be described by a polytropic equation of state with index $n = 1/(\gamma_P - 1) = 3$.

The mass-radius relation has an interesting behavior near $n = 3$. We showed that it can be written

$$\left[\frac{GM}{-\xi_1^2(d\Theta/d\xi)_{\xi_1}}\right]^{n-1} \left(\frac{R}{\xi_1}\right)^{3-n} = \frac{[(n+1)K_P]^n}{4\pi G},$$

or equivalently

$$M = -\frac{1}{(4\pi)^{1/(n-1)}} \xi_1^{(n+1)/(n-1)} \left(\frac{d\Theta}{d\xi}\right)_{\xi_1} \frac{[(n+1)K_P]^{n/(n-1)}}{G^{n/(n-1)}} R^{(n-3)/(n-1)}.$$

Notice that, for $n = 3$, the R dependence disappears, and we find

$$M = -\frac{4}{\sqrt{\pi}}\xi_1^2 \left(\frac{d\Theta}{d\xi}\right)_{\xi_1} \left(\frac{K_P}{G}\right)^{3/2}$$

A relativistic electron gas has

$$K_P = \frac{K_2'}{\mu_e^{4/3}} = \left(\frac{3}{\pi}\right)^{1/3} \frac{hc}{8(\mu_e m_\mathrm{H})^{4/3}},$$

and plugging this into the mass we have just calculated gives

$$M = M_\mathrm{ch} = \left(\frac{3}{32}\right)^{1/2} \frac{1}{\pi} \left[-\xi_1^2 \left(\frac{d\Theta}{d\xi}\right)_{\xi_1}\right] \left(\frac{hc}{G}\right)^{3/2} \frac{1}{(m_\mathrm{H}\mu_e)^2} = \frac{5.83}{\mu_e^2} M_\odot.$$

Everything on the right-hand side except μ_e is a constant, which means that for a highly relativistic electron gas, there is only a single possible mass which can be in hydrostatic equilibrium. If we have a gas that is depleted of hydrogen, so $X = 0$, then $\mu_e \approx 2$, and we have $M_\mathrm{ch} = 1.46\ M_\odot$.

This quantity is called the Chandresekhar mass, after Subrahmanyan Chandresekhar, who first derived it. He did the calculation while on his first trip out of India, to start graduate school at Cambridge... at age 20... which goes to prove that the rest of us are idiots....

To understand the significance of this mass, consider what happens when a star has a mass near it. If the mass is slightly smaller than M_ch, the star can respond by puffing out a little and adopting a larger radius. This reduces the density and makes the gas slightly less relativistic, so that n decreases slightly and the star can be in equilibrium.

On the other hand, if $M > M_\mathrm{ch}$, the gas is already relativistic, and there is no adjustment possible. If a Chandresekhar mass white dwarf gets just a little bit more massive, there simply is no equilibrium state that the star can possibly reach. Instead, it is forced to collapse on a dynamical timescale. The result is invariably a massive explosion. For a white dwarf composed mainly of carbon and oxygen, the carbon and oxygen undergo rapid nuclear burning to elements near the iron peak, and the result is essentially that a nuclear bomb goes off with a solar mass worth of fuel. The resulting explosion, known as a type Ia supernova, can easily outshine an entire galaxy.

D. Very Massive Stars

The Chandrasekhar limit is one way that a star can get into trouble if it is supported by a relativistic gas. It is, however, not the only way. Massive stars are also close to being $n = 3$ polytropes, but instead of being supported by a relativistic gas of electrons, they are supported by a relativistic gas of photons – i.e. by radiation pressure.

Massive stars don't explode while they are on the main sequence, but they still suffer from instabilities associated with being close to the instability line of $n = 3$. In massive stars, this instability tends to manifest as rapid mass loss and ejection of gas. The mechanism is not fully understood today, but observations make it clear that massive stars do suffer from instabilities.

Perhaps the most spectacular example is the star η Carinae, a massive star that is reasonably close to Earth. In 1843, η Carinae (which is visible with the naked eye, despite being > 2 kpc away), suddenly brightened, briefly becoming the second-brightest star in the night sky. Subsequently it dimmed again, although it remains naked-eye visible. In modern times, observations using the Hubble Telescope show that η Carinae is surrounded by a nebula of gas, called the Homunculus Nebula, which was presumably ejected in the 1843 eruption.

[Slide 1 – η Carinae and the Homunculus Nebula]

The physical mechanism behind the explosion is still not very well understood, and it is an active area of research. Nonetheless, it seems clear that it is connected with the fact that η Carinae is supported primarily by radiation pressure, which puts it dangerously close to the $n = 3$ line of instability. We'll discuss the nature of this instability more in a week or so.

Class 10 Notes: Applications and Extensions of Polytropes

In the last class we saw that polytropes are a simple approximation to a full solution to the stellar structure equations, which, despite their simplicity, yield important physical insight. This is particularly true for certain types of stars, such as white dwarfs and very low mass stars. In today's class we will explore further extensions and applications of simple polytropic models, which we can in turn use to generate our first realistic models of typical main sequence stars.

I. The Binding Energy of Polytropes

We will begin by showing that any realistic model must have $n < 5$. Of course we already determined that solutions to the Lane-Emden equation reach $\Theta = 0$ at finite ξ only for $n < 5$, which already hints that $n \geq 5$ is a problem. Nonetheless, we have not shown that, as a matter of physical principle, models of this sort are unacceptable for stars. To demonstrate this, we will prove a generally useful result about the energy content of polytropic stars.

As a preliminary to this, we will write down the polytropic relation

$$P = K_P \rho^{(n+1)/n}$$

in a slightly different form. Consider a polytropic star, and imagine moving down within it to the point where the pressure is larger by a small amount dP. The corresponding change in density $d\rho$ obeys

$$dP = K_P \frac{n+1}{n} \rho^{1/n} d\rho.$$

Similarly, since

$$\frac{P}{\rho} = K_P \rho^{1/n},$$

it follows that

$$d\left(\frac{P}{\rho}\right) = \frac{K_P}{n} \rho^{(1-n)/n} d\rho = \frac{1}{n+1} \left(\frac{dP}{\rho}\right)$$

We can regard this relation as telling us how much the ratio of pressure to density changes when we move through a star by an amount such that the pressure alone changes by dP.

With this out of the way, we now compute the gravitational potential energy of a polytropic star of total mass M and radius R. For convenience, since we'll be changing variables many times, we will write limits of integration as s for surface or c for center, indicating that, whatever variable we're using, it is to be evaluated at the surface or the center. The potential energy is

$$
\begin{aligned}
\Omega &= -\int_c^s \frac{Gm}{r}\,dm \\
&= -\frac{1}{2}\int_c^s \frac{G}{r}\,d(m^2) \\
&= -\left[\frac{Gm^2}{2r}\right]_c^s - \frac{1}{2}\int_c^s \frac{Gm^2}{r^2}\,dr \\
&= -\frac{GM^2}{2R} - \frac{1}{2}\int_c^s \frac{Gm^2}{r^2}\,dr
\end{aligned}
$$

In the third step, we integrated by parts.

Next, we use the equation of hydrostatic balance, $dP/dr = -Gm\rho/r^2$, or $dP = -(Gm/r)(\rho/r)\,dr$ to replace Gm/r in the integral, and then make use of the result we just derived:

$$
\begin{aligned}
\Omega &= -\frac{GM^2}{2R} + \frac{1}{2}\int_c^s m\frac{dP}{\rho} \\
&= -\frac{GM^2}{2R} + \frac{n+1}{2}\int_c^s m\,d\left(\frac{P}{\rho}\right).
\end{aligned}
$$

Now we integrate by parts one more time:

$$
\begin{aligned}
\Omega &= -\frac{GM^2}{2R} + \left[\frac{n+1}{2}m\frac{P}{\rho}\right]_c^s - \frac{n+1}{2}\int_c^s \frac{P}{\rho}dm \\
&= -\frac{GM^2}{2R} - \frac{n+1}{2}\int_c^s \frac{P}{\rho}4\pi r^2\rho\,dr \\
&= -\frac{GM^2}{2R} - \frac{n+1}{2}\int_c^s P\frac{4\pi}{3}\,d(r^3).
\end{aligned}
$$

In the second step, the term in brackets vanishes because $m = 0$ at the center and $P/\rho = 0$ at the surface.

Finally, one last integration by parts, followed by another application of hydrostatic balance:

$$
\begin{aligned}
\Omega &= -\frac{GM^2}{2R} - \left[\frac{n+1}{2}\frac{4\pi}{3}Pr^3\right]_c^s + \frac{n+1}{6}\int_c^s 4\pi r^3\,dP \\
&= -\frac{GM^2}{2R} - \frac{n+1}{6}\int_c^s 4\pi r^3\frac{Gm}{r^2}\rho\,dr \\
&= -\frac{GM^2}{2R} - \frac{n+1}{6}\int_c^s \frac{Gm}{r}\,dm
\end{aligned}
$$

As before, the term in brackets vanished in the second step because $r = 0$ at the center and $P = 0$ at the surface.

Notice that we have gotten back to where we started: the integral on the right hand side is just $-\Omega$. Thus we have shown that

$$\Omega = -\frac{GM^2}{2R} + \frac{n+1}{6}\Omega$$
$$\Omega = -\left(\frac{3}{5-n}\right)\frac{GM^2}{R}.$$

Given this result, one can get the total energy simply by applying the virial theorem. If the star consists of monotonic ideal gas, we have

$$E = \Omega + U = \frac{1}{2}\Omega = -\left(\frac{3}{10-2n}\right)\frac{GM^2}{R}.$$

This is a handy formula to know, but it also shows us why only polytropic models with $n < 5$ are acceptable models for stars. At $n = 5$, the total energy of the polytrope becomes infinite, and for $n > 5$ the energy $E > 0$, meaning the object is unbound. Thus only models with $n < 5$ represent gravitationally bound objects with finite binding energy.

II. Radiation Pressure and the Eddington Limit

Polytropes are very useful, but, since they separate the hydrostatic balance of a star (as embodied by the first two stellar structure equations) from its energy balance (as embodied by the second two), they have limitations. In particular, they cannot by themselves tell us anything about how energy flows through a star, and thus about stars' luminosities. To put this into our models, we must re-insert the temperature-dependence. This brings back into the picture the third structure equation:

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F}{4\pi r^2} \qquad \frac{dT}{dm} = -\frac{3}{4ac}\frac{\kappa}{T^3}\frac{F}{(4\pi r^2)^2}.$$

We care about the temperature is because of its relationship with the pressure, and in particular with the pressure of radiation. That is because the gas pressure follows $P_{\mathrm{gas}} \propto T$, but radiation pressure has a much steeper dependence: $P_{\mathrm{rad}} = aT^4/3$. Thus at sufficiently high temperatures radiation pressure always dominates. You calculated when radiation pressure begins to dominate on your last homework.

The strong dependence of P_{rad} on T has important consequences for stellar structure. To see this, it is helpful to rewrite our last equation in terms of radiation pressure rather than temperature:

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F}{4\pi r^2}$$
$$\frac{4}{3}aT^3\frac{dT}{dr} = -\frac{\kappa\rho}{c}\frac{F}{4\pi r^2}$$
$$\frac{dP_{\mathrm{rad}}}{dr} = -\frac{\kappa\rho}{c}\frac{F}{4\pi r^2}$$

Repeating the same trick using the Lagrangian form gives

$$\frac{dP_{\rm rad}}{dm} = -\frac{\kappa}{c}\frac{F}{(4\pi r^2)^2}.$$

Now consider what this implies for hydrostatic balance. Since $P = P_{\rm gas} + P_{\rm rad}$, the equation of hydrostatic balance can be written as

$$\begin{aligned} \frac{dP_{\rm rad}}{dr} + \frac{dP_{\rm gas}}{dr} &= -\rho\frac{Gm}{r^2} \\ \frac{dP_{\rm rad}}{dr} &= -\rho\frac{Gm}{r^2} - \frac{dP_{\rm gas}}{dr} \end{aligned}$$

Since density and temperature always fall with radius within a star, $dP_{\rm gas}/dr$ is always negative, so the term $-dP_{\rm gas}/dr > 0$. Thus we have

$$\begin{aligned} \frac{dP_{\rm rad}}{dr} &> -\rho\frac{Gm}{r^2} \\ -\frac{\kappa\rho}{c}\frac{F}{4\pi r^2} &> -\rho\frac{Gm}{r^2} \\ F &< \frac{4\pi c G m}{\kappa} = 3.2\times 10^4\left(\frac{M}{M_\odot}\right)\left(\frac{0.4\ {\rm cm}^2\ {\rm g}^{-1}}{\kappa}\right) L_\odot, \end{aligned}$$

where we have normalized κ to the electron scattering opacity because that is usually the smallest possible opacity in a star, producing the maximum possible F. This result is known as the Eddington limit, named after Arthur Eddington, who first derived it. In his honor, the quantity on the right-hand side is called the Eddington luminosity:

$$L_{\rm Edd} = \frac{4\pi c G m}{\kappa}.$$

This result represents a fundamental limit on the luminosity of any object in hydrostatic equilibrium. It applies to stars, but it applies equally well to any other type of astronomical system, and the Eddington limit is important for black holes, entire galaxies, and in many other contexts.

This also lets us derive a useful relation describing how the ratio of radiation pressure to total pressure varies within a star. We have already shown that

$$\begin{aligned} \frac{dP_{\rm rad}}{dr} &= -\frac{\kappa\rho}{c}\frac{F}{4\pi r^2} \\ \frac{dP}{dr} &= -\rho\frac{Gm}{r^2}, \end{aligned}$$

and taking the ratio of these two equations gives

$$\frac{dP_{\rm rad}}{dP} = \frac{\kappa F}{4\pi c G m} = \frac{F}{L_{\rm Edd}}.$$

The Eddington limit also tells us something about nuclear energy generation at the center of a star, where $m = 0$ and $F(m) = 0$. We can Taylor expand F around $m = 0$, so $F \approx m(dF/dm)$. Inserting this into our inequality gives

$$\frac{dF}{dm} = q_c < \frac{4\pi c G}{\kappa},$$

where we have noted that the final stellar structure equation asserts that $dF/dm = q$. Thus we have derived an upper limit on the rate of nuclear energy generation q_c in the center of a star.

There is an important subtlety here, which is important to notice. Our equation relating dT/dr and F is calculated assuming that F comes solely from radiation, i.e. that there are no other sources of energy transport in stars. This means that the F that appears in the limit includes only the radiative heat flow F, not any other sources of heat transport – and we will see next week that there is potentially another type of energy transport that is important in some stars. Although the radiative heat flow always obeys the limit we have derived, the total heat flow need not. This is obviously not much a limitation at the surface of a star, since there radiation is the only means by which energy can move around.

III. The Eddington Model

Considering the importance of radiation pressure leads to an important model, essentially the first quasi-realistic model for stellar structure. It was first derived by Arthur Eddington, and is therefore known as the Eddington model.

The first step in this model is the equation we just derived for the relation between

$$\frac{dP_{\rm rad}}{dP} = \frac{\kappa F}{4\pi c G m} = \frac{F}{L_{\rm Edd}}.$$

Eddington then assumes that the right hand side is constant. This might seem like a crazy assumption, but it turns out to be reasonably good. You will show that it holds for massive stars on your homework. For low mass stars, recall that we showed that burning on the $p-p$ chain occurs at a rate that depends on temperature as roughly $q \propto \rho T^4$. Since $dF/dm = q$, this means that we expect $F/m \sim T^4$ as long as nuclear burning mostly takes place near the center of the star where ρ doesn't vary much. On the other hand, free-free opacity varies with $\kappa \propto T^{-3.5}$. Thus $\kappa F/m$ depends relatively weakly on temperature. This is why the Eddington approximation works reasonably well.

Given the assumption that the right-hand side is constant, we can integrate both sides to get

$$P_{\rm rad} = \frac{F}{L_{\rm Edd}} P,$$

or, rewriting this in terms of $\beta = P_{\rm gas}/P$,

$$\beta = 1 - \frac{F}{L_{\rm Edd}}.$$

Thus this model implies that β is constant throughout the star.

It is easy to see that this implies that the star must be an $n = 3$ polytrope. To see why, note that

$$P = \frac{P_{\rm rad}}{1-\beta} = \frac{aT^4}{3(1-\beta)}$$
$$P = \frac{P_{\rm gas}}{\beta} = \frac{\mathcal{R}}{\beta\mu}\rho T$$

Using the first equation to solve for T gives

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

and inserting this into the second equation gives

$$P = \frac{\mathcal{R}}{\beta\mu}\rho\left[\frac{3}{a}(1-\beta)P\right]^{1/4}$$
$$P = \left[\frac{3\mathcal{R}^4(1-\beta)}{a\mu^4\beta^4}\right]^{1/3}\rho^{4/3}.$$

Thus if β is constant throughout the star, then $P \propto \rho^{4/3}$, and the star is a polytrope with $\gamma_P = 4/3$, corresponding to an index $n = 1/(\gamma_P - 1) = 3$.

Recall that we showed in the last class that for $n = 3$ polytropes there exists only a single possible mass for a given K_P:

$$M = -\frac{4}{\sqrt{\pi}}\xi_1^2\left(\frac{d\Theta}{d\xi}\right)_{\xi_1}\left(\frac{K_P}{G}\right)^{3/2} = 4.56\left(\frac{K_P}{G}\right)^{3/2}.$$

Thus the value of K_P uniquely determines the value of M. Since K_P in our model is determined by β, this means that the value of β uniquely determines the value of M. Inserting the value of K_P we just obtained into the relationship between M and K_P gives

$$M = -\frac{4}{\sqrt{\pi}}\xi_1^2\left(\frac{d\Theta}{d\xi}\right)_{\xi_1}\left[\frac{3\mathcal{R}^4(1-\beta)}{aG^3\mu^4\beta^4}\right]^{1/2}$$
$$= \frac{18.1\,M_\odot}{\mu^2}\left(\frac{1-\beta}{\beta^4}\right)^{1/2}.$$

This gives us M in terms of β and μ. Alternately, we can re-arrange to get an equation for β in terms of M:

$$0 = -1+\beta+\left(\frac{aG^3}{3\mathcal{R}^4}\right)\frac{\pi}{16\xi_1^4(d\Theta/d\xi)_{\xi_1}^2}\mu^4M^2\beta^4$$
$$= -1+\beta+0.003\left(\frac{M}{M_\odot}\right)^2\mu^4\beta^4$$
$$= -1+\beta+0.0004\left(\frac{M}{M_\odot}\right)^2\left(\frac{\mu}{0.61}\right)^4\beta^4$$

This is known as the Eddington quartic.

The final trick is to insert this into the expression we derived earlier for the heat flow: $\beta = 1 - F/L_{\rm Edd}$. If we go to the surface of the star, then the flow F is just the star's total luminosity L. Thus we have

$$\begin{aligned} L &= (1-\beta) L_{\rm Edd} \\ &= \left(\frac{aG^3}{3\mathcal{R}^4}\right) \frac{\pi}{16\xi_1^4 (d\Theta/d\xi)_{\xi_1}^2} \mu^2 \beta^4 M^2 \left(\frac{4\pi c G}{\kappa_s} M\right) \\ &= \frac{\pi^2}{12\xi_1^4 (d\Theta/d\xi)_{\xi_1}^2} \frac{acG^4}{\mathcal{R}^4 \kappa_s} \mu^4 \beta^4 M^3 \\ &= 5.5\,\beta^4 \left(\frac{\mu}{0.61}\right)^4 \left(\frac{1\text{ cm}^2\text{ g}^{-1}}{\kappa_s}\right) \left(\frac{M}{M_\odot}\right)^3 L_\odot, \end{aligned}$$

where κ_s is the value of κ at the stellar surface. We have thus derived, for the first time, a theoretical mass luminosity relation.

We can get some idea of how this mass-luminosity relation behaves by solving the Eddington quartic in the limits of high and low masses. First consider stars with masses $\sim M_\odot$ or less. For these stars, the term

$$0.0004 \left(\frac{M}{M_\odot}\right)^2 \left(\frac{\mu}{0.61}\right)^4 \beta^4$$

is very small. Thus the solution is near $\beta = 1$. If κ_s doesn't vary much between stars, then at these masses we therefore have $L \propto M^3$.

On the other hand, consider very massive stars, those with $M \sim 100\ M_\odot$ or more. In this case the coefficient of β^4 is large. We can very roughly approximate the solution in that case by dropping the β term, which gives

$$0.0004 \left(\frac{M}{M_\odot}\right)^2 \left(\frac{\mu}{0.61}\right)^4 \beta^4 \approx 1 \qquad \Longrightarrow \qquad \beta^4 \propto M^{-2}.$$

The approximation is rough because, even for $M = 100\ M_\odot$, the coefficient of the β^4 term is only 4.

Nonetheless, plugging this into the mass-luminosity relation gives $L \propto \beta^4 M^3 \propto M$. Thus we expect that for very massive stars the mass-luminosity relation should flatten and approach $L \propto M$. Again, this expectation is assuming constant surface opacity κ_s, which is an ok approximation, but not a great one, since the surface temperature varies significantly between low and high mass stars, and the opacity therefore varies as well.

This rough trend that at low masses $L \propto M^3$ (it's actually a bit closer to 3.5 in reality), flattening to $L \propto M$ at high masses, is actually seen in the observations. Thus this model at least roughly reproduces reality.

IV. Cowling Models

Eddington models are the simplest class of models that produce a mass-luminosity relation, but other simple models are possible as well. It is simply a matter of what assumptions one is willing to make – it is a tradeoff between accuracy of result and ease of calculation, with Eddington models at one extreme and full computer-based models at the other.

An intermediate step is the Cowling model. The basic assumptions of the Cowling model are: (1) assume that all the nuclear reactions in a star occur in a small core, which we can approximate to be a point, and (2) adopt a Kramers Law form for the opacity.

The first assumption implies that

$$\frac{dF}{dm} = q = 0$$

outside the core. Thus the heat flow through the star is constant, and must match the star's total luminosity, since that is just F evaluated at the stellar surface. Thus we can set $F = L =$ constant.

Next we consider the temperature equation, rewritten in terms of radiation pressure as in the Eddington model:

$$\frac{dP_{\rm rad}}{dr} = -\frac{\kappa\rho}{c}\frac{F}{4\pi r^2}.$$

Inserting $F = L$ and the Kramers Law form $\kappa = \kappa_0 \rho^a T^b$ for κ, we obtain

$$\frac{dP_{\rm rad}}{dr} = -\frac{\kappa_0 L}{c}\frac{\rho^{a+1}T^b}{4\pi r^2}.$$

This plus the equation of hydrostatic balance, the definition of the density, and the equation of state complete specification of the model:

$$\begin{aligned} \frac{dP}{dr} &= -\rho\frac{Gm}{r^2} \\ \frac{dm}{dr} &= 4\pi r^2 \rho \\ P &= P_{\rm rad} + \frac{\mathcal{R}}{\mu}\rho T. \end{aligned}$$

As it is this model would still need to be numerically integrated for a general choice of a and b. However, we can solve analytically (mostly) in the special case $a = 0$ and $b = 0$, i.e. an opacity that does not depend on density or temperature. Electron scattering opacity has this form, and in very massive stars it dominates, so this assumption is not terrible for very massive stars.

For the case $a = b = 0$, the equation for the radiation pressure becomes

$$\frac{dP_{\rm rad}}{dr} = -\frac{\kappa L \rho}{4\pi r^2 c}.$$

The basic idea of the Cowling model is to recast the equations by eliminating the density, and re-arrange so that the radiation pressure is the independent variable. This turns out to yield something analytically tractable.

If we write the equation of hydrostatic balance as

$$\frac{dP_{\rm gas}}{dr} + \frac{dP_{\rm rad}}{dr} = -\rho\frac{Gm}{r^2}$$

and divide both sides by $dP_{\rm rad}/dr$, we get

$$\frac{dP_{\rm gas}}{dP_{\rm rad}} + 1 = \frac{4\pi cG}{\kappa L}m.$$

Next we take d/dr of both sides, giving

$$\begin{aligned}
\frac{d^2P_{\rm gas}}{dP_{\rm rad}^2}\frac{dP_{\rm rad}}{dr} &= \frac{4\pi cG}{\kappa L}\frac{dm}{dr} \\
\frac{d^2P_{\rm gas}}{dP_{\rm rad}^2}\left(-\frac{\kappa L\rho}{4\pi r^2 c}\right) &= \frac{4\pi cG}{\kappa L}\left(4\pi r^2\rho\right) \\
\frac{d^2P_{\rm gas}}{dP_{\rm rad}^2} &= -\left(\frac{64\pi^3c^2G}{\kappa^2L^2}\right)r^4.
\end{aligned}$$

This is a transformed version of the hydrostatic balance equation, using $P_{\rm rad}$ as the independent variable.

Next, the equation of radiative diffusion must be similarly transformed. The equation reads

$$\frac{dP_{\rm rad}}{dr} = -\frac{\kappa L\rho}{4\pi r^2 c}.$$

Since we want $P_{\rm rad}$ as the independent variable, we flip both sides and then re-arrange

$$\begin{aligned}
\frac{dr}{dP_{\rm rad}} &= -\frac{4\pi r^2 c}{\kappa L\rho} \\
-\frac{1}{r^2}\frac{dr}{dP_{\rm rad}} &= \frac{4\pi c}{\kappa L\rho} \\
\frac{d}{dP_{\rm rad}}\left(\frac{1}{r}\right) &= \frac{4\pi c}{\kappa L\rho}.
\end{aligned}$$

Since the transformed equation of hydrostatic balance involves $P_{\rm gas}$ and not ρ, we similarly want to eliminate ρ from this equation. We note that the relationship between radiation pressure and temperature can be rewritten

$$T = \left(\frac{3}{a}P_{\rm rad}\right)^{1/4},$$

and inserting this into the ideal gas law for the gas pressure gives

$$\begin{aligned}
P_{\rm gas} &= \frac{\mathcal{R}}{\mu}\rho\left(\frac{3}{a}P_{\rm rad}\right)^{1/4} \\
\rho &= \frac{\mu}{\mathcal{R}}\left(\frac{a}{3}\right)^{1/4}P_{\rm gas}P_{\rm rad}^{-1/4}
\end{aligned}$$

Substituting this for ρ, we obtain the transformed version of the radiation diffusion equation:

$$\frac{d}{dP_{\rm rad}}\left(\frac{1}{r}\right) = \frac{4\pi c\mathcal{R}}{\mu\kappa L}\left(\frac{3}{a}\right)^{1/4} P_{\rm gas}^{-1} P_{\rm rad}^{1/4}.$$

We have now reduced the problem to a pair of coupled, non-linear ODEs. These must still be solved numerically, but the solution is considerably easier than for the full set of equations, and can be made even more so by an appropriate non-dimensionalization. We will not solve them in class, however.

Class 11 Notes: Stellar Instabilities

Before moving on to make more sophisticated models of stars than those discussed in the last week, we need to check that our models are stable. What this means is that we need to check not only that we have found solutions to the equations of stellar structure that are steady in time (since we dropped the time derivatives), we need to check that those solutions have the property that small perturbations, which are always present, tend to damp out and return the system back to the equilibrium we have identified. This is what is mean by stability. In contrast, instability occurs when any small deviation from an equilibrium solution tends to drive the system further and further away from it.

The classic example of an unstable system is a pencil standing on its point. If one could get the pencil to balance completely perfectly, it would be in equilibrium. However, any small perturbation that causes the pencil to tilt slightly will grow, and the pencil will fall over. We need to make sure that our solutions to the stellar structure equations aren't like a pencil standing on end, or, if they are, to understand why and what that implies.

I. Stability of Nuclear Burning

We begin our consideration of stability in stars by examining thermal stability. That is, we recall that $t_{\text{dyn}} \ll t_{\text{KH}} \ll t_{\text{nuc}}$, and for now we neglect instabilities that occur on the dynamical timescale. We assume that the star is in dynamical equilibrium, and we do not worry about the stability of our solution to the equation of hydrostatic balance. Instead, we worry about the stability of our solutions to the equations describing energy generation and transport.

A. Non-Degenerate Ideal Gas with Radiation

One way to approach this problem is to consider a star as a whole and apply the virial theorem. Consider a star that is supported by a combination of ideal, non-degenerate, non-relativistic gas pressure and radiation pressure, so that

$$P = \frac{\mathcal{R}}{\mu}\rho T + \frac{1}{3}aT^4.$$

In terms of internal energy, recall that the specific internal energies of gas and radiation are given by

$$u_{\text{gas}} = \frac{3}{2}\frac{\mathcal{R}}{\mu}T = \frac{3}{2}\frac{P_{\text{gas}}}{\rho} \qquad\qquad u_{\text{rad}} = \frac{aT^4}{\rho} = 3\frac{P_{\text{rad}}}{\rho}.$$

The virial theorem tells us that pressure and gravitational binding energy are related by

$$\Omega = -3\int_0^M \frac{P}{\rho}\,dm,$$

so plugging in the pressure gives

$$\Omega = -3 \int_0^M \left(\frac{2}{3} u_{\rm gas} + \frac{1}{3} u_{\rm rad} \right) \, dm = -(2U_{\rm gas} + U_{\rm rad}) \, ,$$

where $U_{\rm gas}$ and $U_{\rm rad}$ are the total internal energies of gas and radiation in the star. Thus

$$U_{\rm gas} = -\frac{1}{2}(\Omega + U_{\rm rad})$$

and the total energy is

$$E = \Omega + U_{\rm rad} + U_{\rm gas} = -U_{\rm gas}.$$

Thus the total energy of the star is $-U_{\rm gas}$. What does this imply about the mean temperature in the star? Recall that if $\overline{T}$ is the mass-averaged temperature in the star, the gas internal energy is

$$U_{\rm gas} = \frac{3}{2} M \frac{\mathcal{R}}{\mu} \overline{T}.$$

Conservation of energy for a star therefore requires that

$$L_{\rm nuc} - L = \frac{dE}{dt} = -\frac{3}{2} M \frac{\mathcal{R}}{\mu} \frac{d\overline{T}}{dt},$$

where $L_{\rm nuc}$ is the total rate of nuclear energy generation in a star, and L is its total luminosity. This assumes M and μ are constant over the time we are considering.

In thermal equilibrium $L_{\rm nuc} = L$ and the left-hand side vanishes. To investigate stability, consider what would happen if $L_{\rm nuc}$ and L were slightly different, so that $L_{\rm nuc} - L = \delta L \neq 0$. In this case the mean temperature would change according to

$$\frac{d\overline{T}}{dt} = -\frac{2}{3} \frac{\mu}{\mathcal{R}} \frac{\delta L}{M}.$$

Thus if $\delta L > 0$, meaning that $L_{\rm nuc} > L$, then $d\overline{T}/dt < 0$, and the temperature decreases. Since, as we have seen, $L_{\rm nuc}$ is a strongly increasing function of T, this means that $L_{\rm nuc}$ will in turn decrease, and δL will decrease too. Conversely, if $\delta L < 0$, then the temperature will increase, $L_{\rm nuc}$ will rise, and δL will increase. Thus an imbalance in one direction creates a restoring force in the opposite direction. This is the hallmark of a stable system. Stars supported by non-degenerate ideal gas plus radiation pressure are therefore thermally stable.

It is worth pausing to note that this result is actually somewhat counterintuitive, and it arises because gravity is a strange force. If $\delta L > 0$, this means that the star is generating more energy than it is radiating. If one considers an ordinary object that is producing more heat than it is radiating, one expects it to heat up – when one starts a fire, the fireplace gets hot because it is producing more heat than it is radiating. A star, however, does exactly the opposite. If it produces more heat than it radiates, it actually gets colder, not hotter. This is a generic feature of systems that are held together by gravity: adding energy to a self-gravitating system makes it colder, not hotter, exactly the opposite of our experience in everyday life. The reason this happens is that gravity is an attractive force, and this causes self-gravitating systems to have a negative specific heat: adding heat makes them colder. Systems we're used to deal with in everyday life do not have strong long-range attractive forces, and as a result they have positive specific heat. Adding heat makes them hotter.

B. Degenerate Ideal Gas with Radiation

Now let us extend this analysis to a degenerate ideal gas. For a non-relativistic degenerate gas, recall that we showed that the internal energy is related to pressure and density exactly as for a non-degenerate gas: $u_{\mathrm{gas}} = 3/2\, P_{\mathrm{gas}}/\rho$, where now P_{gas} is the degeneracy pressure. Consequently, our calculation of the total energy and the effects or radiation using the virial theorem is unchanged:

$$E = -U_{\mathrm{gas}} \qquad \Longrightarrow \qquad L_{\mathrm{nuc}} - L = -\frac{dU_{\mathrm{gas}}}{dt}.$$

The difference that degeneracy makes is that now U_{gas} does not depend on the gas temperature, because P_{gas} does not depend on temperature for a degenerate gas. Consequently, if L_{nuc} and L are out of balance the star can expand or contract (since P and ρ can change), but this does not cause the temperature to change. The temperature instead will respond only to the local rate of energy generation.

This is an unstable situation. Suppose there is a fluctuation in which $L_{\mathrm{nuc}} > L$. The star will expand, but the temperature will not drop as a result; instead, it will rise, responding to the increase in the local rate of energy generation. As a result L_{nuc} will increase rather than decrease, pushing the star further out of balance, consuming nuclear fuel even faster. This is called a thermonuclear runaway, and it leads to a phenomenon called novae that occur on the surfaces of white dwarf stars. It is also important for the evolution of red giant stars.

The runaway ends once the temperature becomes high enough that the star is no longer degenerate. Once degeneracy ends, the temperature no longer increases for $L_{\mathrm{nuc}} > L$. Instead, it decreases, as in a non-degnerate star, and the situation is stabilized. This is called lifting the degeneracy.

One can write down the condition for instability somewhat more rigorously by considering the center of the star. Recall that polytropes obey

$$P_c = (4\pi)^{1/3} B_n G M^{2/3} \rho_c^{4/3},$$

where P_c and ρ_c are the central pressure and density, and B_n is a number that depends on the polytropic index n, but only very weakly. We therefore expect a relation of this sort to hold approximately even in stars that are not perfect polytropes. This relation is only true in hydrostatic equilibrium, but we are assuming hydrostatic equilibrium for now, since we are only concerned with instabilities on a Kelvin-Helmholtz timescale, not a dynamical timescale.

Now consider a perturbation that causes the central density to change by an amount $d\rho_c$ over a time dt. The corresponding change in pressure is given by

$$\frac{dP_c}{dt} = \left[(4\pi)^{1/3} B_n G M^{2/3}\right] \frac{4}{3} \rho_c^{1/3} \frac{d\rho_c}{dt}.$$

Dividing this equation by the previous one gives

$$\frac{dP_c}{P_c} = \frac{4}{3} \frac{d\rho_c}{\rho_c}.$$

The pressure and density are related by the equation of state $P(\rho, T)$. We can write a general equation of state near some particular density and pressure as

$$P = P_0 \rho^a T^b,$$

where a and b are numbers that depend on the type of gas. A non-degenerate gas has $a = 1$ and $b = 1$, a degenerate non-relativistic gas has $a = 5/3$ and $b = 0$, and a degenerate relativistic gas has $a = 4/3$, $b = 0$.

Again, let us consider perturbing the pressure by a small amount dP. The density and temperature respond according to

$$dP = P_0 \left(a\rho^{a-1} T^b \, d\rho + b\rho^a T^{b-1} \, dT \right).$$

Dividing this by the equation of state gives

$$\frac{dP}{P} = a\frac{d\rho}{\rho} + b\frac{dT}{T}.$$

If we apply this relation at the center of the star and substitute in our result based on hydrostatic balance for dP_c/P_c, we get

$$\begin{aligned} \frac{4}{3} \frac{d\rho_c}{\rho_c} &= a\frac{d\rho_c}{\rho_c} + b\frac{dT_c}{T_c} \\ b\frac{dT_c}{T_c} &= \left(\frac{4}{3} - a\right) \frac{d\rho_c}{\rho_c}. \end{aligned}$$

Consider what this implies for various types of gas. For a non-degenerate gas, $a = 1$ and $b = 1$, the coefficients on both sides are positive, and this means that an increase in density causes an increase in temperature. This means that contraction of the star, which raises ρ_c, also raises the temperature. This increases the rate of nuclear burning, raising the pressure and causing the star to re-expand. Conversely, expansion of the star lowers ρ_c and thus also lowers T_c. This reduces the rate of nuclear burning and causes the star to stop expanding.

For a degenerate gas, on the other hand, $a \geq 4/3$ (depending on how relativistic the gas is) and $b \ll 1$ (reaching 0 exactly for a perfectly degenerate gas). Thus the coefficients on the left and right sides have opposite signs. As a result, expansion of the star ($d\rho_c < 0$) raises the temperature ($dT_c > 0$), and the rate of nuclear burning increases. This pushes the star to expand even further, and leads to an unstable runaway that ends only once the gas is hot enough to drive a back below 4/3.

C. Thin Shell Instability

The two cases we have considered thus far are for entire stars. However, it sometimes occurs that nuclear burning takes place not in an entire star, but in a thin shell within it. This often happens in evolved stars that have used up their hydrogen fuel. The center of the star fills with ash supported by degeneracy pressure that the star is too cool to burn further, but on top of this ash layer there is still fuel left and burning continues. In this case the burning is generally confined to a thin shell on top of the degenerate ash core.

Consider such a burning shell of mass dm, temperature T, density ρ, outer radius $r_{\rm sh}$, and inner radius r_0, which is taken to be fixed, as will be the case for a shell supported by a degenerate ash core. The thickness is dr, which is much less than the radius of the star, R.

The star is in hydrostatic equilibrium, so the pressure in the shell is determined by the equation of hydrostatic balance:

$$\frac{dP}{dm} = -\frac{Gm}{4\pi r^4}.$$

Thus the pressure in the shell is given by

$$P_{\rm sh} = -\int_{m_{\rm sh}}^{M} \frac{Gm}{4\pi r^4}\, dm,$$

where $m_{\rm sh}$ is the mass interior to the shell.

Now consider perturbing the shell by changing its pressure by an amount δP. We would like to know the corresponding amount δr by which the outer radius $r_{\rm sh}$ of the shell changes. In hydrostatic equilibrium the star should behave homologously, meaning that the radius of ever shell simply rises in proportion to the amount by which the outer edge of the perturbed shell expands. The outer edge of the perturbed shell expands by a factor $1 + \delta r/r_{\rm sh}$, so a shell of gas that was at radius r moves to radius $r(1 + \delta r/r_{\rm sh})$ after the perturbation.

Since hydrostatic balance still holds, the new pressure in the shell is

$$P_{\rm sh} + \delta P \approx -\int_{m_{\rm sh}}^{M} \frac{Gm}{4\pi[r(1+\delta r/r_{\rm sh})]^4}\, dm = -\left(1 + \frac{\delta r}{r_{\rm sh}}\right)^{-4} \int_{m_{\rm sh}}^{M} \frac{Gm}{4\pi r^4}\, dm.$$

For small perturbations, $\delta r \ll r_{\rm sh}$, we can use a Taylor expansion to approximate the term in parentheses, and drop higher-order terms on the grounds that they are small:

$$
\begin{aligned}
P_{\rm sh} + \delta P &= -\left(1 - 4\frac{\delta r}{r_{\rm sh}}\right)\int_{m_{\rm sh}}^{M} \frac{Gm}{4\pi r^4}\, dm' \\
\delta P &= 4\frac{\delta r}{r_{\rm sh}}\int_{m_{\rm sh}}^{M} \frac{Gm}{4\pi r^4}\, dm \\
\frac{\delta P}{P_{\rm sh}} &= -4\frac{\delta r}{r_{\rm sh}}.
\end{aligned}
$$

This procedure of Taylor expanding in the small perturbation and dropping the higher-order terms is known as linearization, and it is one of the most powerful techniques available for analyzing differential equations.

Given this result, we can figure out how the density in the shell responds to the perturbation. The shell density is the ratio of its mass to its volume:

$$\rho = \frac{dm}{4\pi r_{\rm sh}^2\, dr}.$$

We want to know the amount $\delta\rho$ by which the density changes. After the perturbation, the new thickness of the shell is $dr + \delta r$, so the new density is

$$\rho + \delta\rho = \frac{dm}{4\pi r_{\rm sh}^2 (dr + \delta r)} = \frac{dm}{4\pi r_{\rm sh}^2 dr}\left(1 + \frac{\delta r}{dr}\right)^{-1} = \rho\left(1 + \frac{\delta r}{dr}\right)^{-1}.$$

If we now do the same trick of Taylor-expanding the $(1 + \delta r/dr)$ term, we have

$$\rho + \delta\rho \approx \rho\left(1 - \frac{\delta r}{dr}\right) \qquad\Longrightarrow\qquad \frac{\delta\rho}{\rho} = -\frac{\delta r}{dr} = -\frac{\delta r}{r_{\rm sh}}\frac{r_{\rm sh}}{dr}.$$

Combining this result for the perturbed density with the one for the perturbed pressure, we have

$$\frac{dP}{P_{\rm sh}} = 4\frac{dr}{r_{\rm sh}}\frac{d\rho}{\rho}.$$

Before going on, we can pause to understand the physical significance of this equation. Notice that, for a thin shell, $dr/r_{\rm sh}$ is a very small number. This means that, for a given fractional change in density $\delta\rho/\rho$, the fractional change in pressure dP/P, is much smaller. This makes intuitive sense. Suppose we have a thin shell and we double its thickness. The shell's volume goes up by a factor of two, so its density drops by a factor of two. On the other hand, the pressure is dictated by the weight of the overlying material, which depends on its mass and radius. The mass hasn't changed, and the radius of that material has only changed by a trivial amount, because expanding a thin shell by a factor of two doesn't move anything very far. Consequently, even though the density has changed by a factor of two, the pressure changes very little.

Finally, to obtain the change in temperature, we again assume a powerlaw equation of state $P \propto \rho^a T^b$, so that

$$\frac{dP}{P} = a\frac{d\rho}{\rho} + b\frac{dT}{T}.$$

Plugging the result we obtained for dP/P from hydrostatic equilibrium into this formula based on the equation of state gives

$$\left(4\frac{dr}{r_{\rm sh}} - a\right)\frac{d\rho}{\rho} = b\frac{dT}{T}$$

Instability occurs when the two coefficients have opposite signs, since this means that expansion of the star, which lowers ρ_c, increases the temperature, driving more nuclear burning and more expansion. Since pressure never decreases with increasing temperature, b is never negative. The coefficient on the left, however, can be: a is always positive (1 for an ideal gas, 4/3 for a degenerate relativistic gas, 5/3 for a degenerate non-relativistic gas), and $dr/r_{\rm sh}$ can be arbitrarily small for a sufficiently thin shell. This means that thin shells are always unstable to thermonuclear runaway.

As with the case of a degenerate star, this runaway cannot continue indefinitely. For a degenerate star, stability returns when the temperature becomes high enough to lift the degeneracy. For a thin shell, stability returns when the pressure in the shell becomes large enough to expand it to the point where it is no longer thin, and $dr/r_{\rm sh}$ becomes larger than $a/4$.

II. Global Dynamical Stability

We have now completed our discussion of thermal instabilities, those that take place over a Kelvin-Helmholtz timescale within stars that are in hydrostatic equilibrium. We now turn to the question of dynamical instabilities, those that involve departures from hydrostatic balance. Again, we can take advantage of the fact that $t_{\rm KH} \gg t_{\rm dyn}$. Since $t_{\rm KH}$ is the time required for processes involving heat or radiation, in modeling dynamical instabilities on timescales $t_{\rm dyn}$ we can generally treat stars as adiabatic, meaning that there is negligible heat exchange. Today we will only consider global instabilities, those involving the entire star. We will leave local instabilities for the next class.

A. Stability Against Homologous Perturbations

The general theory of hydrodynamic stability is a complex one, but we can obtain the basic results by considering one particular type of perturbation: a homologous perturbation. A homologous perturbation is one in which we expand or contract the star uniformly, so that every shell expands or contracts by the same factor.

To see how a star responds to a homologous perturbation, we begin by recalling the equation of motion for a shell of material in the star, which we wrote down early in the class. The equation is

$$\ddot{r} = -\frac{Gm}{r^2} - \frac{1}{\rho}\frac{dP}{dr}.$$

The first term on the right is the gravitational force, and the second is the pressure force. The standard equation of hydrostatic balance amounts to setting $\ddot{r} = 0$ in this equation. For convenience we will multiply both sides by $dm = 4\pi r^2 \rho\, dr$, which gives

$$dm\,\ddot{r} = -\frac{Gm}{r^2}\,dm - 4\pi r^2\,dP,$$

where dP is the change in pressure across the shell.

Now consider a star that is initially in hydrostatic equilibrium, meaning that $\ddot{r} = 0$ for every shell, and add a homologous perturbation in which we expand or contract the star by a factor $1 + \epsilon$. That is, a shell that was previously at radius r_0 is moved to radius $r_0(1 + \delta r/r_0)$. Expansion corresponds to $\delta r/r_0 > 0$, and contraction to $\delta r/r_0 < 0$. We assume that the perturbation is small, so $|\delta r/r_0| \ll 1$. This perturbation will also cause the pressure everywhere within the star to change by some factor. We let P_0 be the pressure before the perturbation, and we write the new pressure as $P_0(1+\delta P/P_0)$. We expect $|\delta P/P_0| \ll 1$ as well.

The unperturbed configuration satisfies

$$0 = -\frac{Gm}{r_0^2}\,dm - 4\pi r_0^2\,dP_0.$$

Inserting the perturbed radius and pressure into the equation of motion gives

$$dm\frac{d^2}{dt^2}\left[r_0\left(1+\frac{\delta r}{r_0}\right)\right] = -\frac{Gm}{[r_0(1+\delta r/r_0)]^2}\,dm - 4\pi\left[r_0\left(1+\frac{\delta r}{r_0}\right)\right]^2 d\left[P_0\left(1+\frac{\delta P}{P_0}\right)\right]$$

We now Taylor expand and keep only terms that are linear in $\delta r/r_0$ and $\delta P/P_0$. To remind you, the first term in the Taylor expansion of a polynomial is

$$(1+x)^n = 1 + nx + O(x^2).$$

Doing the expansion and plugging in gives

$$\begin{aligned} dm\,\ddot{\delta r} &= -\left(1 - 2\frac{\delta r}{r_0}\right)\frac{Gm}{r_0^2}\,dm - \left(1 + 2\frac{\delta r}{r_0} + \frac{\delta P}{P_0}\right)4\pi r_0^2\,dP_0 \\ dm\,\ddot{\delta r} &= 2\frac{Gm}{r_0^3}\delta r\,dm - 4\pi\left(2\frac{\delta r}{r_0} + \frac{\delta P}{P_0}\right)r_0^2\,dP_0, \end{aligned}$$

where in the second step we cancelled the terms $-Gm\,dm/r_0^2$ and $4\pi r_0^2\,dP_0$, since they add up to zero.

This equation describes how the perturbation δr varies in time. To make progress, however, we must know how δP is related to δr, and this is where we make use of the fact that, over short timescales, the star behaves adiabatically. Recall that for an adiabatic gas, we have

$$P = K_a \rho^{\gamma_a},$$

where γ_a is the adiabatic index, which depends only on the microphysical properties of the gas (i.e. is it degenerate or not, relativistic or not). Suppose that the perturbation causes the density to change from its original value ρ_0 to a new value $\rho_0(1+\delta\rho/\rho_0)$. Since the gas is adiabatic, the perturbed pressure and density must satisfy the same adiabatic equation of state as the unperturbed values, so

$$\begin{aligned} P_0 \left(1 + \frac{\delta P}{P_0}\right) &= K_a \left[\rho_0 \left(1 + \frac{\delta\rho}{\rho_0}\right)\right]^{\gamma_a} \\ &\approx K_a \rho_0^{\gamma_a} \left(1 + \gamma_a \frac{\delta\rho}{\rho_0}\right) \\ \delta P &= K_a \rho_0^{\gamma_a} \gamma_a \frac{\delta\rho}{\rho_0} \\ \frac{\delta P}{P_0} &= \gamma_a \frac{\delta\rho}{\rho_0}. \end{aligned}$$

The last step is to relate the change in density $\delta\rho$ to the change in radius δr. The mass of a shell is

$$dm = 4\pi r^2 \rho \, dr.$$

Homologous expansion or contraction involves changing r_0 to $r_0(1+\delta r/r_0)$, changing dr_0 to $dr_0(1 + \delta r/r_0)$, and ρ to $\rho(1 + \delta\rho/\rho_0)$, while leaving the shell mass dm unchanged. Thus

$$dm = 4\pi \left[r_0 \left(1 + \frac{\delta r}{r_0}\right)\right]^2 \rho_0 \left(1 + \frac{\delta\rho}{\rho_0}\right) dr_0 \left(1 + \frac{\delta r}{r_0}\right) = 4\pi r_0^2 \rho_0 \, dr_0 \left(1 + 3\frac{\delta r}{r_0} + \frac{\delta\rho}{\rho_0}\right),$$

where, again, we have linearized and dropped higher-order terms in the perturbations. However, we know that $dm = 4\pi r_0^2 \rho_0 \, dr_0$ exactly, so all the terms in the parentheses except the 1 must vanish. Thus we have

$$\frac{\delta\rho}{\rho_0} = -3\frac{\delta r}{r_0}$$

Combining this with the relationship we derived from the adiabatic equation of state shows that

$$\frac{\delta P}{P_0} = -3\gamma_a \frac{\delta r}{r_0},$$

and substituting this into the perturbed equation of motion gives

$$\begin{aligned} dm \, \ddot{\delta r} &= 2\frac{Gm}{r_0^3} \delta r \, dm - 4\pi r_0^2 \left(2\frac{\delta r}{r_0} - 3\gamma_a \frac{\delta r}{r_0}\right) dP_0 \\ &= \left[2\frac{Gm}{r_0^2} \, dm - 4\pi r_0^2 \, dP_0 (2 - 3\gamma_a)\right] \frac{\delta r}{r_0}. \end{aligned}$$

Recall that the unperturbed configuration satisfies

$$\frac{Gm}{r_0^2}\,dm = -4\pi r_0^2\,dP_0.$$

Making this substitution for $-4\pi r_0^2\,dP_0$ gives

$$\ddot{\delta r} = -(3\gamma_a - 4)\frac{Gm}{r_0^3}\delta r.$$

This is the differential equation describing a harmonic oscillator, and it has the trivial solution

$$\delta r = Ae^{i\omega t},$$

where

$$\omega = \pm\sqrt{(3\gamma_a - 4)\frac{Gm}{r_0^3}}.$$

For a non-relativistic ideal gas, $\gamma_a = 5/3$, so the term inside the square root is positive and ω is real. This means that $i\omega t$ is a pure imaginary number, so δr varies sinusoidally in time – the response of the star to the perturbation is to oscillate at a constant amplitude A. This is a stable behavior.

On the other hand, suppose we have a gas that is not a non-relativistic ideal gas, and has a different value of the adiabatic index. If $\gamma_a < 4/3$, then the term inside the square root is negative, and ω is an imaginary number. In this case the term $i\omega t$ that appears in the numerator of the exponential is a *real* number, which can be positive or negative depending on whether we take the positive or negative square root – both are valid solutions. A negative real number corresponds to a perturbation that decays exponentially in time, which is stable.

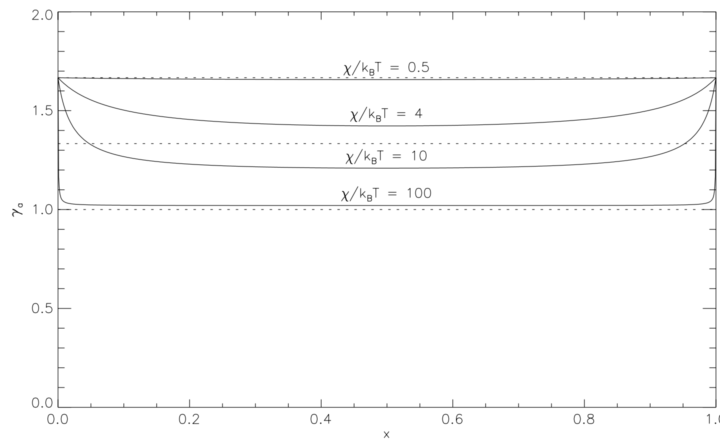
On the other hand, a positive real number for $i\omega$ corresponds to a solution for δr that grows exponentially in time. This is an instability, since it means that a small perturbation will grow to arbitrary size, or at least to the size where our analysis in the limit of small δr breaks down. The characteristic time required for this growth is just $1/(i\omega)$. Note that $1/(i\omega) \sim 1/\sqrt{G\rho} \sim t_{\rm dyn}$. Thus if $\gamma_a < 4/3$, the star will be unstable on dynamical timescale.

B. Applications of Instability

The limit that a star becomes unstable for $\gamma_a < 4/3$ has consequences in a number of circumstances. We have already explored one: a star that is dominated by the pressure of a relativistic gas (of either non-degenerate or degenerate electrons or of photons), approaches $\gamma_a = 4/3$. This causes stars that approach this limit to become unstable.

Another situation where a star can approach $\gamma_a = 4/3$ is when ionization-type processes become important. Recall that we showed toward the beginning of the class that a partially ionized gas can have γ_a below 4/3. The dotted lines show

$\gamma_a = 1, 4/3$, and $5/3$. Clearly for $\chi/k_B T$ large enough, $\gamma_a < 4/3$ over a significant range of ionization fractions.



The centers of stars are fully ionized and their surfaces are fully neutral, so only the small parts of stars where the gas is transitioning between these are extremes are subject to ionization instability. This causes them to oscillate, but the amount of mass involved is small, and it is trapped between two stable regions. Thus the consequences are small.

However, other ionization-like mechanisms that operate at high temperatures can also produce $\gamma_a < 4/3$, and these can have more severe consequences. One such example is the photodisintegration of iron nuclei and conversion of photons into electron-positron pairs at temperatures above several times 10^9 K. Both of these process are ionization-like in the sense that they use increases in thermal energy to create new particles rather than to make the existing particles move faster. Thus doing work on a gas in this condition does not cause its temperature to increase by any significant amount, and in doing work the temperature of the gas does not decrease much – everything is buffered by creation and destruction of particles. This is the hallmark of a gas with small γ_a.

Unlike hydrogen ionization, these process can take place at the centers of stars and can involve a significant amount of mass. This analysis suggests that, if they do take place, they compromise the stability of the star as a whole. Indeed, this is exactly what we think happens to initiate supernovae in massive stars: the core becomes hot enough that photodisintegration and/or pair creation push the adiabatic below 4/3, initiating a dynamical instability and collapse.

III. Opacity-Driven Instabilities

We will end this class by examining one more type of instability that can occur on stars: instability driven by variations in opacity. Although we will not be able to develop an analytic theory of how these work at the level of this class, we can understand their general behavior. Given the vital importance of this sort of instability for all of astronomy (as we will see), it is worth understanding how it works.

A. The κ Mechanism

The mechanism for instabilities based on opacity, called the κ mechanism since opacity is written with a κ, was worked out the mid-20th century. However, the basic idea for opacity-driven instabilities was suggested by Eddington, based on analogy with a steam valve.

Suppose there is a layer in a star that has the property that its opacity increases as it is compressed. If such a region is compressed, the increase in opacity will reduce the flow of heat through it, trapping more heat in the stellar interior. The layer acts like a valve that is closed. Closing the valve and trapping heat will raise the pressure interior to the opaque layer, causing it to expand. This expansion will decrease the opacity, opening the valve and letting the trapped heat out. This reduces the pressure in the stellar interior, reversing the expansion and letting the layer fall back. This raises its density and opacity, starting a new cycle.

Clearly this mechanism only operates if the opacity increases with density. However, this is generally not the case. Free-free opacity obeys $\kappa \propto \rho T^{-7/2}$, so

$$\frac{d\kappa}{\kappa} = \frac{d\rho}{\rho} - \frac{7}{2}\frac{dT}{T}.$$

For an adiabatic ideal gas, we have seen that

$$\frac{dP}{P} = \gamma_a \frac{d\rho}{\rho},$$

and since $P \propto \rho T$, we also know that

$$\frac{dP}{P} = \frac{d\rho}{\rho} + \frac{dT}{T}.$$

Thus for an adiabatic ideal gas, we have

$$\frac{dT}{T} = (\gamma_a - 1)\frac{d\rho}{\rho}.$$

Plugging this in, we see that for adiabatic gas, the opacity change $d\kappa$ associated with a small change in density $d\rho$ is

$$\frac{d\kappa}{\kappa} = \left(\frac{9 - 7\gamma_a}{2}\right)\frac{d\rho}{\rho}$$

Thus $d\kappa$ and $d\rho$ have opposite signs unless $\gamma_a < 9/7 = 1.29$. Since the star is unstable only if $d\kappa$ and $d\rho$ have the same sign, i.e. increasing density increases opacity, this means that instability occurs only for $\gamma_a < 9/7$. (In fact the real condition is a bit more complicated than this, but this gives the basic idea.)

Gasses composed of relativistic and non-relativistic, degenerate and non-degenerate particles all have $\gamma_a > 4/3$, so the instability does not operate throughout most of the star. However, we have just been reminded that γ_a can be small in the partially ionized zones of a star. In these regions of the star, this instability does operate, and these regions act like a piston, driving pulsations into the rest of the star.

Whether these pulsations actually do anything significant depends on how large the instability zone is, where it is located in the star, and how luminous the star is. There are two main instability zones, one associated with hydrogen ionization and one with helium ionization.

If the star is too hot, the ionization zones are located very close to the stellar surface, and thus they occur in a region where the density is low. This makes the piston ineffective, because it is driven by too little mass to excite motions in the rest of the star. Conversely, if a star is too cool, the ionization zones are deep in the star. The overlying layers of the star, which we will see next class are convective, then damp out the motions, and again nothing happens. Thus instability is possible only in a certain range of surface temperatures.

Moreover, since the instability is ultimately driven by the star's radiation, so the strength with which it is driven depends on the star's luminosity. The instability does not operate if the luminosity is too low. It turns out that it for this reason it does not generally operate in main sequence stars, because those which are luminous enough to meet the minimum luminosity condition are too hot at their surfaces, and those with cool enough surfaces are not sufficiently luminous.

Post-main sequence stars, however, can be unstable to the κ mechanism, and this causes them to pulsate.

B. Stellar Pulsation and Variable Stars

The κ mechanism can cause instability in stars in several different parts of the HR diagram. The regions of instability are generally characterized by a minimum luminosity and a narrow range of surface temperatures, and thus are called instability strips, since they appear as vertical strips in the diagram. The most famous of the variable stars classes is the Cepheids.

[Slide 1 – the instability strip]

Variable stars are important because the period of the oscillation depends on the luminosity of the star – which is not surprising, since the luminosity determines how hard the instability is driven. This relation was first discovered empirically in 1908 by Henrietta Swan Leavitt, and has now been understood from first principles.

[Slide 2 – light curve of SU Cygni]

[Slide 3 – Cepheid in M100 seen by HST]

[Slide 4 – M100 Cepheid vs. time]

The Cepheid period-luminosity relation is important in astronomy because it provides a distance indicator. Since one can compute the luminosity from the star's observed period, one can determine its distance by comparing the observed heat flow to the luminosity. Cepheids are bright enough to be seen in other galaxies, and thus can be used to determine the distance to those galaxies. This technique was first used on a large number of galaxies by Edwin Hubble, leading to the discovery of the expansion of the universe.

Class 12 Notes: Convection in Stars

In the last class we discussed a large number of instabilities, but we haven't yet discussed the most important one in most stars: convective instability. That will be the subject of today's class.

Convection is a process in which heat is transported by the motion of fluid elements. One common example where convection occurs is when one heats water on a stove. Initially the water is still, and heat is transported by conduction through it. However, as the water at the bottom of the pot gets hotter, eventually the water starts to churn. Hot water from the bottom of the pot rises and transports heat upwards, while cold water at the top falls. This process is called convection. Convection is also important in planetary atmospheres, in the liquid interiors of giant planets and in the liquid iron-rich cores of terrestrial planets.

I. Convective Stability and Instability

A. The Adiabatic Temperature Gradient

As with the other dynamical instabilities we have studied, since $t_{\mathrm{KH}} \gg t_{\mathrm{dyn}}$ in a star, we make the assumption that the gas behaves adiabatically on short timescales. To see what this implies, consider what happens in a convective region: parcels of gas at one radius within a star move to a different radius, and, under our assumption, they remain adiabatic while doing so. We also assume that the parcels of gas are also at the same pressure as their neighbors, because any difference in pressure will be lead to compression or expansion until the pressure balances.

Let us consider how the temperature of such a gas parcel changes as it rises. Assuming that the gas is ideal, it obeys $P = (\mathcal{R}/\mu)\rho T$. If we move the gas parcel upward a small distance dr, then the change in pressure is given by

$$dP = \frac{\mathcal{R}}{\mu}\left(\rho \frac{dT}{dr} + T\frac{d\rho}{dr}\right) dr = \left(\frac{P}{T}\frac{dT}{dr} + \frac{P}{\rho}\frac{d\rho}{dr}\right) dr,$$

where we have assumed that the composition is uniform, so μ is constant. If the gas is adiabatic, however, we also know that $P = K_a \rho^{\gamma_a}$, so we must also have

$$dP = K_a \gamma_a \rho^{\gamma_a - 1}\frac{d\rho}{dr}\, dr = \gamma_a \frac{P}{\rho}\frac{d\rho}{dr}\, dr.$$

Combining the two expressions for dP, we have

$$\gamma_a \frac{P}{\rho}\frac{d\rho}{dr} = \frac{P}{T}\frac{dT}{dr} + \frac{P}{\rho}\frac{d\rho}{dr}$$

$$\left(\frac{dT}{dr}\right)_{\rm ad} = (\gamma_a - 1)\frac{T}{P}\frac{P}{\rho}\frac{d\rho}{dr}$$
$$= \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{T}{P}\frac{dP}{dr},$$

where in the last step we substituted for $d\rho/dr$ using the adiabatic equation of state. This value of dT/dr is known as the adiabatic temperature gradient.

We can also express $(dT/dr)_{\rm ad}$ using the equation of hydrostatic balance

$$\frac{dP}{dr} = -\frac{Gm}{r^2}\rho$$

and the ideal gas law $P = (\mathcal{R}/\mu)\rho T$. (Note that we can use the equation of hydrostatic balance because we assume that the pressure of the rising fluid element is the same as the pressure of its neighbors, which are in hydrostatic balance.) Plugging in for P and dP/dr gives

$$\left(\frac{dT}{dr}\right)_{\rm ad} = -\left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{\mu}{\mathcal{R}}\frac{Gm}{r^2} = -\left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{\mu}{\mathcal{R}}g,$$

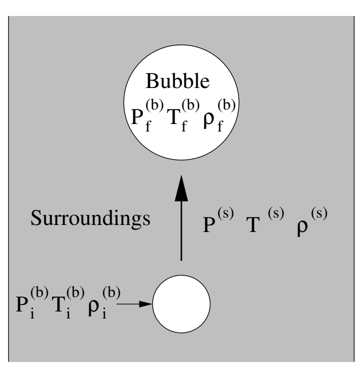
where $g = Gm/r^2$ is the local acceleration of gravity in the star.

An alternative form of the adiabatic temperature gradient is to give it in terms of a logarithmic derivative of P with respect to T. Dividing both sides by dT/dr gives

$$\frac{\gamma_a}{\gamma_a - 1} = \frac{T}{P}\frac{dP}{dT} = \left(\frac{d\ln P}{d\ln T}\right)_{\rm ad}.$$

If a star has a temperature gradient equal to the adiabatic temperature gradient, then as a parcel of fluid rises or falls, its temperature changes in exactly the same way as the background temperature. Since the moving parcel of fluid has the same temperature and pressure as its new surroundings, it necessarily has the same density, and thus nothing is really changed by the motion. Thus we expect the adiabatic temperature gradient to tell us something interesting about convection, since convection in stars whose temperature gradients are equal to $(dT/dr)_{\rm ad}$ doesn't actually do anything.

B. The Brunt-Väisälä Frequency



Now consider what happens to the rising bubble of gas if the temperature gradient in the star is not equal to $(dT/dr)_{\rm ad}$. We start with a bubble of gas that is at the same pressure, density, and temperature as its surroundings, and we perturb it upward by a distance dr. It stays at the same pressure as its new surroundings, but it is at a different temperature, and therefore a different density. If the bubble is initially at a density $\rho_i^{(b)}$, after it rises a distance dr its new density is $\rho_f^{(b)}$, where

$$\rho_f^{(b)} = \rho_i^{(b)} + \frac{d\rho^{(b)}}{dr}\,dr.$$

Since the bubble is adiabatic, we know that

$$\frac{dP^{(b)}}{dr} = \gamma_a \frac{P_i^{(b)}}{\rho_i^{(b)}}\frac{d\rho^{(b)}}{dr},$$

where $P_i^{(b)}$ is the initial pressure in the bubble. Thus the new density is

$$\rho_f^{(b)} = \rho_i^{(b)} + \frac{\rho_i^{(b)}}{\gamma_a P_i^{(b)}}\frac{dP^{(b)}}{dr}\,dr.$$

Similarly, the initial density of the surrounding gas is $\rho_i^{(s)}$, and the density of the surrounding gas a distance dr higher is

$$\rho_f^{(s)} = \rho_i^{(s)} + \frac{d\rho^{(s)}}{dr}\,dr.$$

The surrounding gas is not adiabatic, so we cannot substitute in terms of dP/dr here.

Now that we have computed the difference in density between the bubble and the surrounding gas, consider what this implies about the forces on that bubble. The bubble feels two forces: gravity, and buoyancy force. The gravitational force per unit volume on the displaced bubble is

$$f_{\rm g} = -\rho_f^{(b)} g = -\left(\rho_i^{(b)} + \frac{\rho_i^{(b)}}{\gamma_a P^{(b)}}\frac{dP^{(b)}}{dr}\,dr\right) g.$$

The buoyancy force is just the difference in pressure between its top and its bottom, and is given by Archimedes principle: the buoyancy force on an object is equal to the weight of the material it displaces. The density of the material displaced is $\rho_f^{(s)}$, so the buoyancy force per unit volume is

$$f_{\rm b} = \rho_f^{(s)} g = \left(\rho_i^{(s)} + \frac{d\rho^{(s)}}{dr}\,dr\right) g.$$

Adding the gravity and buoyancy forces gives the net force,

$$f_{\rm net} = \left(\frac{d\rho^{(s)}}{dr} - \frac{\rho_i^{(b)}}{\gamma_a P_i^{(b)}} \frac{dP^{(b)}}{dr} \right) g\, dr = \left(\frac{1}{\rho}\frac{d\rho}{dr} - \frac{1}{\gamma_a P}\frac{dP}{dr} \right) \rho g\, dr,$$

where we have dropped the subscripts because everything in the second equation refers to the surroundings, since $P^{(b)} = P^{(s)}$ and $\rho_i^{(b)} = \rho_i^{(s)}$. The term in parentheses is usually denoted by the letter A:

$$A = \frac{1}{\rho}\frac{d\rho}{dr} - \frac{1}{\gamma_a P}\frac{dP}{dr}.$$

Since we have computed the net force, we can write down the equation of motion for the bubble. Since $f_{\rm net}$ is the force per unit volume, and ρ is the mass per unit volume, Newton's second law tells us that the displacement of the bubble dr obeys

$$\frac{d^2}{dt^2}(dr) = \frac{f_{\rm net}}{\rho} = Ag\, dr$$

This is the equation of motion for a harmonic oscillator, and it has the usual solution:

$$dr = Ce^{iNt},$$

with

$$N = \pm\sqrt{-Ag} = \sqrt{\left(\frac{1}{\gamma_a P}\frac{dP}{dr} - \frac{1}{\rho}\frac{d\rho}{dr} \right) g}.$$

The quantity N is the frequency of oscillation, and is known as the Brunt-Väisäla frequency.

As we found before when considering homologous perturbations, the behavior of the solution depends on whether the term inside square root is positive or negative, corresponding to a real or imaginary value for N. If N is real, the solutions are oscillations, and the system is stable. If N is imaginary, then the solutions corresponding to an exponentially decaying and an exponentially growing mode, and the system is unstable.

Convective instability corresponds to the case when N is imaginary. Physically, we can understand this fairly easily. If $A < 0$, then the differential equation for dr looks like a harmonic oscillator, in the sense that the force $-Ag\, dr$ is opposite to the displacement. It therefore constitutes a restoring force, which pushes the system back to stability. The value of A in turn is determined by the balance between gravity and buoyancy, with $A < 0$ corresponding to the case where gravity is stronger. As a result we get a real value for N, and any displaced fluid element just oscillates, bobbing up and down like a buoy in the ocean.

If $A > 0$, the net force is in the same direction as the displacement. Physically, what is going on is that a blob of fluid rises and expands because it is at higher pressure than its surroundings. Although gravity wants to pull it back down, its high pressure makes it expand so much that it experiences a large buoyancy force that is stronger than gravity. The net force is therefore upward, and the bubble accelerates further up. This is an unstable situation, hence N is imaginary.

This physical interpretation makes sense if we examine the terms inside the square root, and recall that dP/dr and $d\rho/dr$ are both negative. If dP/dr is very big (in absolute value), then the system is unstable. This is because the value of dP/dr determines how much the rising bubble expands, and thus how large the buoyancy force is. If $d\rho/dr$ is very large (in absolute value), the system is stable. That is because $d\rho/dr$ measures how much denser the rising bubble is than its new surroundings, and thus how strongly gravity wants to pull it down.

Thus, we have derived the condition for stability against convection: $A < 0$. Convection does not occur for $A < 0$, and it does for $A > 0$.

C. Convective Stability and the Adiabatic Temperature Gradient

We have now determined a condition for stability in terms of the gradients of P and ρ, but it is helpful to instead phrase things in terms of temperature, because this allows us to see how convective stability relates to the adiabatic temperature gradient we derived a moment ago.

We use the ideal gas law $P = (\mathcal{R}/\mu)\rho T$, which we showed earlier gives

$$\begin{aligned} \frac{dP}{dr} &= \frac{P}{T}\frac{dT}{dr} + \frac{P}{\rho}\frac{d\rho}{dr} \\ \frac{d\rho}{dr} &= \frac{\rho}{P}\frac{dP}{dr} - \frac{\rho}{T}\frac{dT}{dr} \end{aligned}$$

for a gas of uniform composition. Substituting for $d\rho/dr$ in A gives

$$\begin{aligned} A &= \frac{1}{\rho}\left[\frac{\rho}{P}\frac{dP}{dr} - \frac{\rho}{T}\frac{dT}{dr}\right] - \frac{1}{\gamma_a P}\frac{dP}{dr} \\ &= \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{1}{P}\frac{dP}{dr} - \frac{1}{T}\frac{dT}{dr} \end{aligned}$$

The stability condition is $A < 0$, so a system is stable against convection if

$$\begin{aligned} 0 &> \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{1}{P}\frac{dP}{dr} - \frac{1}{T}\frac{dT}{dr} \\ \frac{dT}{dr} &> \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{T}{P}\frac{dP}{dr} \end{aligned}$$

The right-hand side, however, is just the adiabatic temperature gradient. Thus the criterion for no convection is that

$$\frac{dT}{dr} > \left(\frac{dT}{dr}\right)_{\rm ad}.$$

The signs get a little confusing here. Recall that dT/dr is negative – temperature falls as one moves upward through a star. Thus this equation means that a system is stable to convection as long as the true temperature gradient is less negative than the adiabatic one. To avoid confusion, it is common to take the absolute value of both sides, which, since both sides are negative, gives

$$\left|\frac{dT}{dr}\right| < \left|\frac{dT}{dr}\right|_{\mathrm{ad}}.$$

Thus a system is stable against convection as long as the actual temperature gradient is shallower than the adiabatic temperature gradient. Equivalently, the condition for convect stability can be written as

$$\frac{d\ln P}{d\ln T} > \frac{\gamma_a}{\gamma_a - 1}.$$

A star within which the temperature gradient is steeper than the adiabatic temperature gradient is said to be super-adiabatic. What we have shown is that superadiabatic temperature gradients are convectively unstable.

Finally, it is important to point out that this analysis is for regions of a star dominated by gas pressure. It can be extended to include radiation pressure in a fairly straightforward manner, and this extension can be important in the centers of massive or evolved star where radiation pressure is important. The general result is that, if radiation pressure is important, convection is more likely.

II. Effects of Convection

Now that we have determined when convection should occurs, we turn to the question of how it affects stars.

A. Locations of Convection

As a first step toward this, let us consider where convection is likely to occur in a star. To do this, it is helpful to write down the temperature gradient that is produced by radiation alone, and compare it to the adiabatic value. If there is no convection, then the temperature gradient is given by the equation we have already derived:

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F_{\mathrm{rad}}}{4\pi r^2},$$

where we have added the subscript rad on F to emphasize that this is the flow carried by radiation, which need not match the total flow if convection is occuring. Note, here we define F as the *heat flow*, which is the *heat flux* times $4\pi r^2$. The convective stability condition that $dT/dr > (dT/dr)_{\mathrm{ad}}$ therefore implies that

$$-\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F_{\mathrm{rad}}}{4\pi r^2} > -\left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{\mu}{\mathcal{R}}g$$

$$\left(\frac{\gamma_a}{\gamma_a - 1}\right)\frac{3\mathcal{R}}{4ac\mu g}\frac{\kappa\rho}{T^3}\frac{F_{\mathrm{rad}}}{4\pi r^2} < 1.$$

Convection begins if this inequality is violated. In practice, it is violated in three situations.

First, if the stellar opacity (κ) is large, the inequality is violated. Physically, this occurs because a large opacity means that the temperature gradient must become steep to carry the same heat flow. The star responds by developing a steeper temperature gradient until it becomes so steep that it exceeds the adiabatic gradient, at which point convection starts. Since κ generally increases with decreasing temperature, this situation occurs most commonly in the cooler outer parts of stars than in their cores.

Second, in the ionization zones in a star, γ_a can become small due to ionization effects. This also makes the left-hand side large. Due to this effect, we expect the ionization zones in stars to be highly convective. Again, this occurs fairly near the stellar surface, since the deep interior is fully ionized.

Third, if the energy generation rate in the star is very sharp function of temperature, then F rises rapidly as r approaches 0 inside a star. This large heat flow at a small radius leads to violation of the inequality. This happens only in the center of the star, and only if the nuclear reactions are very sensitive to temperature, e.g. the CNO cycle or the triple-α process.

In the Sun, the since the $p - p$ chain dominates, the third type of convective instability doesn't occur. The center of the Sun is convectively stable. In the outer part of the Sun, the first and second types of convective instability do occur, so the outer part of the Sun is convective. In less massive stars, the gas is cooler, and the first and second types of convection occur over ever-larger fractions of the star, working their way down toward the center. At $\sim 0.3\ M_\odot$ the star is fully convective.

In the opposite direction, as one moves to stars more massive and hotter than the Sun the convection zone at the top of the star disappears, while one driven by the strong temperature-dependence of the CNO cycle appears at the base of the star and covers more and more of its mass as the stellar mass increases.

B. Convective Energy Transport

In regions where convection does occur, it can transport energy in addition to radiative diffusion. This will modify the stellar structure equations, since the equation for dT/dr is derived based on the assumption that transport is entirely by radiation. We therefore need to understand how much energy is carried by convection.

To determine that, we can return to our picture of convection as a bubble of material rising through a star, being driven by buoyancy force that dominates over gravity. The bubble rises adiabatically until it spreads out and mixes with the surrounding material, delivering its heat. This process of hot bubbles rising and then mixing with their surroundings is what carries the heat flux in a convective star, and it is that flux we want to calculate.

In this picture, the star has some temperature gradient dT/dr, which is more negative than the adiabatic gradient $(dT/dr)_{\rm ad}$. Thus when the bubble rises a distance dr, the gas surrounding it has decreased in temperature by an amount

$$dT^{(s)} = \frac{dT}{dr}\,dr$$

The bubble, on the other hand, is adiabatic until the point where it stalls and mixes with its environment. Therefore after it rises a distance dr, its temperature changes by an amount

$$dT^{(b)} = \left(\frac{dT}{dr}\right)_{\rm ad} dr.$$

The difference in temperature between the bubble and its surroundings is therefore

$$\delta T = dT^{(b)} - dT^{(s)} = \left[\left(\frac{dT}{dr}\right)_{\rm ad} - \frac{dT}{dr}\right] dr \equiv \delta\left(\frac{dT}{dr}\right) dr.$$

The quantity $\delta(dT/dr)$ that we have defined is a measure of how superadiabatic the gas is. At $\delta(dT/dr) = 0$ the temperature gradient is adiabatic and convection shuts off.

Now suppose that a hot, rising bubble travels a distance ℓ before it fully mixes with the surrounding gas and gives up its thermal energy. As the bubble mixes, the amount of energy per unit bubble volume that it transfers to its surroundings is

$$\delta q = \rho c_P\,\delta T = \rho c_P\,\delta\left(\frac{dT}{dr}\right)\ell,$$

where c_P is the specific heat capacity of the gas at constant pressure. For an ideal monatomic gas, $c_P = (5/2)(\mathcal{R}/\mu)$, but we leave the expression as c_P because in convective zones where ionization is important one must use a value of c_P that accounts for ionization energy.

This is the heat per unit volume carried by one bubble. If we want to know the heat flow associated with the collective motion of all the rising bubbles in the star, we must multiply by the average speed with which the bubbles move and the area through which they move:

$$F_c = \rho c_P\,\delta\left(\frac{dT}{dr}\right)\ell\overline{v}_c(4\pi r^2).$$

This expression gives the convective heat flow in the star, which must be added to the radiative flow to find the total.

The remaining steps are to evaluate ℓ and $\overline{v}_c$, the characteristic distance that bubbles get before dissolving, and the characteristic velocity with which they rise. Unfortunately at this point we lack a "spherically symmetric" theory of convection that would tell us with certainty. Instead, we are forced to use an empirical approximation called Mixing Length theory. This is right at the order of magnitude level, and mostly tells us what we need to know; but it is definitely not complete. Getting a better understanding of how convection really works is a major challenge in 3D time-dependent fluid dynamics.

The first approximation of Mixing Length theory is to guess that the typical distance that a convective bubble travels before breaking up is set by the condition that the pressure change significantly, so that the bubble must expand significantly to stay in pressure balance. As long as the bubble expands by a small amount, it should survive, but once it has to roughly double its volume, it should break up.

To make this definite, we use the equation of hydrostatic balance:

$$\frac{dP}{dr} = -\frac{Gm}{r^2}\rho = -\rho g,$$

where $g = Gm/r^2$ is the local gravitational acceleration, which we have defined for convenience. We are interested in the distance dr that one must travel before the change in pressure dP is of order P, i.e.

$$1 \sim \frac{dP}{P} = \frac{1}{P}\frac{dP}{dr}\,dr = -\frac{\rho g}{P}\,dr$$

Thus we expect a change in the pressure of order unity when $dr \sim P/(\rho g)$. We define this quantity as the pressure scale height,

$$H_P = \frac{P}{\rho g},$$

and the first basic assumption of Mixing Length theory is that $\ell \sim H_P$. To make it formal, we write

$$\ell = \alpha H_P = \alpha\frac{P}{\rho g} = \alpha\frac{\mathcal{R}}{\mu}\frac{T}{g}$$

where α is a dimensionless fudge factor of order unity that represents our ignorance.

The second thing we need to approximate is the velocity of the convective bubbles, $\overline{v}_c$. To estimate this, recall our equation of motion for the bubble, which we used in deriving the Brunt-Väisälä frequency:

$$\frac{d^2}{dt^2}(dr) = Ag\,dr,$$

where the quantity A is given by

$$A = \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{1}{P}\frac{dP}{dr} - \frac{1}{T}\frac{dT}{dr} = \frac{1}{T}\left[\left(\frac{dT}{dr}\right)_{\rm ad} - \frac{dT}{dr}\right] = \frac{1}{T}\delta\left(\frac{dT}{dr}\right).$$

Thus the equation of motion can be written

$$\frac{d^2}{dt^2}(dr) = \frac{g}{T}\delta\left(\frac{dT}{dr}\right) dr.$$

The quantity on the right-hand side is the acceleration of the bubble.

Since the acceleration increases as dr does, we make the simple assumption that the characteristic acceleration is given by its value at the point halfway between the final and initial points, so

$$a = \frac{g}{T}\delta\left(\frac{dT}{dr}\right)\frac{\ell}{2}.$$

For such a uniform acceleration we can use the old first-term physics standby formula $v_f^2 = v_i^2 + 2a\,\Delta x$. Since the initial velocity is $v_i = 0$, and $\Delta x = \ell$ is the distance traveled, the final velocity is

$$v_f = (2a\ell)^{1/2} = \left[\frac{g}{T}\delta\left(\frac{dT}{dr}\right)\right]^{1/2}\ell = \left[\frac{g}{T}\delta\left(\frac{dT}{dr}\right)\right]^{1/2}\alpha\frac{\mathcal{R}}{\mu}\frac{T}{g} = \alpha\frac{\mathcal{R}}{\mu}\left[\frac{T}{g}\delta\left(\frac{dT}{dr}\right)\right]^{1/2}.$$

We are after the mean velocity, which must be somewhere between 0 and v_f, so we again insert another parameter to represent our ignorance. We set

$$\overline{v}_c = \alpha\frac{\mathcal{R}}{\mu}\left[\beta\frac{T}{g}\delta\left(\frac{dT}{dr}\right)\right]^{1/2}.$$

where β is another dimensionless number of order unity.

Now we're ready to plug in. The convective heat flow is

$$\begin{aligned}
F_c &= 4\pi r^2\rho c_P\,\delta\left(\frac{dT}{dr}\right)\ell\overline{v}_c \\
&= 4\pi r^2\rho c_P\,\delta\left(\frac{dT}{dr}\right)\left(\alpha\frac{\mathcal{R}}{\mu}\frac{T}{g}\right)\left\{\alpha\frac{\mathcal{R}}{\mu}\left[\beta\frac{T}{g}\delta\left(\frac{dT}{dr}\right)\right]^{1/2}\right\} \\
&= 4\pi r^2\rho c_P\left(\frac{\mathcal{R}}{\mu}\right)^2\left(\frac{T}{g}\right)^{3/2}\alpha^2\beta^{1/2}\left[\delta\left(\frac{dT}{dr}\right)\right]^{3/2}.
\end{aligned}$$

We have therefore succeeded in calculating the convective heat flow in terms of the local properties of the star and our two fudge factors α and β.

C. Deviation from Adiabaticity in Convective Regions

The first thing to check based on this result is how superadiabatic the temperature gradient can get. Note, Mixing Length theory assumes that F_c increases as $\delta(dT/dr)$ does, i.e. as the star becomes more and more superadiabatic. We therefore expect this to be strong feedback effect that stops the temperature gradient becoming too steep. For example, if $\delta(dT/dr)$ then F_c increases but then $|dT/dr|$ decreases, which make $\delta(dT/dr)$ decrease.

To make this argument quantitative, suppose that none of the heat flow is carried by radiation; only convection occurs. This is the limit that gives the maximum possible temperature gradient, since any additional heat flow due to radiation on top of convection will only smooth things out further.

If we are outside the part of the star where nuclear burning occurs, then F is simply the total stellar luminosity L, and under our assumption that there is no radiative flow, this means that $F_c = L$. Plugging this into the formula for F_c and solving for $\delta(dT/dr)$ gives

$$\delta\left(\frac{dT}{dr}\right) = \left[\frac{1}{\alpha^2\beta^{1/2}}\left(\frac{\mu}{\mathcal{R}}\right)^2 \frac{L}{4\pi r^2}\frac{1}{\rho c_P}\left(\frac{g}{T}\right)^{3/2}\right]^{2/3}.$$

This represents the difference between the true temperature gradient and the adiabatic temperature gradient. We want to know what fraction of the adiabatic temperature gradient this is, so we divide by $|dT/dr|_{\mathrm{ad}} = g/C_P$, which is true for a monatomic, ideal gas. This gives

$$\frac{\delta\,(dT/dr)}{|dT/dr|_{\mathrm{ad}}} = \alpha^{-4/3}\beta^{-1/3}\left(\frac{\mu}{\mathcal{R}}\right)^{4/3}\left(\frac{L}{4\pi r^2}\right)^{2/3} C_P^{1/3}\rho^{-2/3}T^{-1}.$$

We can evaluate this directly by plugging in, but it is more instructive to examine the physical meaning of this expression. To the order of magnitude level, $r \sim R$ and $\rho \sim M/R^3$. If we are dealing with an ideal gas, then $c_P \sim \mathcal{R}/\mu$. Finally, recall that the virial theorem implies that the mean temperature $T \sim (\mu/\mathcal{R})(GM/R)$ Plugging this in, and dropping factors of order unity,

$$\begin{aligned}
\frac{\delta\,(dT/dr)}{|dT/dr|_{\mathrm{ad}}} &\sim \left(\frac{\mu}{\mathcal{R}}\right)^{4/3}\left(\frac{L}{R^2}\right)^{2/3}\left(\frac{\mathcal{R}}{\mu}\right)^{1/3}\left(\frac{R^3}{M}\right)^{2/3}\left(\frac{\mathcal{R}}{\mu}\frac{R}{GM}\right) \\
&= \left(\frac{L^2R^5}{G^3M^5}\right)^{1/3} \\
&= \left[\left(\frac{RL}{GM^2}\right)^2\left(\frac{R^3}{GM}\right)\right]^{1/3} \\
&= \left(\frac{t_{\mathrm{dyn}}}{t_{\mathrm{KH}}}\right)^{2/3}
\end{aligned}$$

Thus the physical meaning of this expression is that the deviation from adiabaticity is of order the ratio of the dynamical to the KH timescale, to the 2/3 power. It makes sense that the deviation from adiabaticity should involve this ratio. Convection is a dynamical instability, where the speeds of motion are set by the forces of buoyancy and gravity. Thus it should be able to transport heat on a dynamical timescale. Effects trying to produce a large temperature gradient, like radiation, operate on a KH timescale. Thus the amount by which convection dominates is determined by the ratio of these timescales.

Numerically, recall that for the Sun $t_{\text{dyn}} \sim 3000$ s and $t_{\text{KH}} \sim 30$ Myr. Thus

$$\frac{\delta\,(dT/dr)}{|dT/dr|_{\text{ad}}} \sim \left(\frac{3000\text{ s}}{30\text{ Myr}}\right)^{2/3} \sim 10^{-8}.$$

Thus the deviation from adiabaticity is *extremely* small. To good approximation, we can therefore assume that the true temperature gradient is equal to the adiabatic temperature gradient anywhere in a star that convection is taking place. Only in sophisticated numerical models do we even need to worry about setting exact values of α and β.

The one exception to this is near stellar surfaces, where ρ and T have values that are much lower than their average values throughout the star. Since the relative deviation from adiabaticity varies as $1/(\rho^{2/3}T)$, it can be significantly larger near the stellar surfaces and can approach order unity.

D. Implications for Stellar Structure

Convection has important implications for stellar structure, since it provides a heat transport mechanism that can sometimes be more important than radiation. One effect is that, where it operates, convection guarantees that a star is close to adiabatic, which means that the star is a polytrope in the convection zone – entropy is constant, so K_a is the same for every shell in the convective zone. If gas pressure dominates, then $\gamma_a = \gamma_P = 5/3$, corresponding to an $n = 1.5$ polytrope. If radiation dominates, then, as you will show on your homework, $\gamma_a = \gamma_P = 4/3$, and the star is an $n = 3$ polytrope. Intermediate radiation pressure strengths give intermediate values of n. Since stars are often convective over only parts of their interiors, this does not make the entire star a polytrope, however.

Perhaps more important, convection limits the polytropic index that is possible anywhere within a star, or at least anywhere that convection is capable of forcing the temperature gradient to be no larger than the adiabatic one. A value of $\gamma_P = 5/3$ corresponds to a star where $dT/dr = (dT/dr)_{\text{ad}}$, and larger values of γ_P correspond to steeper dT/dr. Thus the condition imposed by convection that $|dT/dr| \le |dT/dr|_{\text{ad}}$ is equivalent to requiring that $\gamma_P \le 5/3$. We have already seen that $\gamma_P > 4/3$ is required for stability. Thus we have shown that no stellar model is stable except for those with $4/3 < \gamma_P \le 5/3$, corresponding to $1.5 \le n < 3$. Higher values of n are unstable to dynamical collapse, and lower values of n are unstable to convection.

Convection means that we must also replace one of our stellar structure equations, since now dT/dr will be the radiative value we have been using only up to the point where convection begins. At that point dT/dr will be equal to $(dT/dr)_{\text{ad}}$.

Thus the stellar structure equation for the temperature changes to

$$\begin{aligned}\frac{dT}{dr} &= \max\left[\left(\frac{dT}{dr}\right)_{\rm rad}, \left(\frac{dT}{dr}\right)_{\rm ad}\right] \\ &= -\min\left[\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{F}{4\pi r^2}, \left(\frac{\gamma_a - 1}{\gamma_a}\right)\frac{\mu}{\mathcal{R}}\frac{Gm}{r^2}\right].\end{aligned}$$

The first line is a maximum rather than a minimum because $(dT/dr)_{\rm rad}$ and $(dT/dr)_{\rm ad}$ are both negative, meaning that taking the maximum is equivalent to selecting whichever one has a smaller absolute value. The second line is a minimum because we have factored out the minus sign.

Astronomy 112: The Physics of Stars

Class 13 Notes: Schematics of the Evolution of Stellar Cores

We're now done with our discussion of physical processes in stars, and we are ready to begin the last phase of the class: applying those principles to develop a theoretical model for how stars behave and evolve. We can get a very good general idea of this by focusing on the center of a star, and considering the different physical process that can take place in it. That is the topic for today.

I. The $(\log T, \log \rho)$ Plane

To begin studying the center of a star, imagine that it has a density ρ and a temperature T. We can describe any possible stellar center in terms of these two numbers plus the composition, which doesn't vary much for main sequence stars. This suggests that we can gain a great deal of insight into the behavior of stellar cores by drawing a graph of ρ vs. T and coloring in the regions on the graph where various processes occur. Since ρ and T both cover very large ranges, it is more convenient to take the logarithm and plot in the $(\log T, \log \rho)$ plane, which is what we will now do. In passing, I'll mention that we're going to calculate things quite approximately. One can do these calculations more precisely, and the results of the more precise calculations are shown in the figures in the textbook.

A. The Pressure

First consider the pressure and the equation of state. We have already seen that there are four possible regimes for gas pressure: any combination of degenerate and non-degenerate, and relativistic and non-relativistic. There is also radiation pressure. We would like to draw approximate lines in the $(\log \rho, \log T)$ plane delineating where each type of pressure is dominant, since that will determine part of the behavior of the gas in those regions. To do this, we will ask where various types of pressure are equal.

First let's collect all the types of pressure we have to worry about. We have non-degenerate gas pressure (which is the same for relativistic or non-relativistic gases),

$$P_{\text{gas}} = \frac{\mathcal{R}}{\mu}\rho T,$$

pressure for a non-releativstic degenerate gas

$$P_{\text{deg,NR}} = K_1' \left(\frac{\rho}{\mu_e}\right)^{5/3},$$

pressure for an ultra-relativistic degenerate gas

$$P_{\text{deg,UR}} = K_2' \left(\frac{\rho}{\mu_e}\right)^{4/3},$$

and radiation pressure

$$P_{\rm rad} = \frac{1}{3} a T^4.$$

First we take the log of all the pressure equations:

$$\begin{aligned}
\log P_{\rm gas} &= \log\rho + \log T + \log\frac{\mathcal{R}}{\mu} \\
\log P_{\rm deg,NR} &= \frac{5}{3}\log\rho + \log K_1' - \frac{5}{3}\log\mu_e \\
\log P_{\rm deg,UR} &= \frac{4}{3}\log\rho + \log K_2' - \frac{4}{3}\log\mu_e \\
\log P_{\rm rad} &= 4\log T + \log\frac{a}{3},
\end{aligned}$$

For numerical evaluations we can assume standard Solar composition, $\mu = 0.61$ and $\mu_e = 1.17$.

We want to find the regions where each of these pressures is dominant, so we will look for the lines where the various pressures are equal. These divide the regions where one dominates from the regions where another dominates. Since there are 4 pressures there are 6 possible lines of equality, but in reality only 4 of them are relevant, because the other equalities either never occur, or occur only in regions of ρ and T not relevant to stars.

First, consider the line where $P_{\rm gas} = P_{\rm rad}$. Equating the two pressure and re-arranging, we get

$$\begin{aligned}
4\log T + \log\frac{a}{3} &= \log\rho + \log T + \log\frac{\mathcal{R}}{\mu} \\
\log\rho &= 3\log T + \log\frac{a}{3} - \log\frac{\mathcal{R}}{\mu}.
\end{aligned}$$

This is clearly the equation of a line with a slope of 3 in $\log\rho$ vs. $\log T$.

Next consider the line where $P_{\rm gas} = P_{\rm deg,NR}$. Using the same procedure, we have

$$\begin{aligned}
\log\rho + \log T + \log\frac{\mathcal{R}}{\mu} &= \frac{5}{3}\log\rho + \log K_1' - \frac{5}{3}\log\mu_e \\
\log\rho &= \frac{3}{2}\log T + \frac{3}{2}\log\frac{\mathcal{R}}{\mu} - \frac{3}{2}\log K_1' + \frac{5}{2}\log\mu_e.
\end{aligned}$$

This is a line of slope 3/2.

The line of equality between ideal gas pressure and degenerate relativistic gas pressure is

$$\begin{aligned}
\log\rho + \log T + \log\frac{\mathcal{R}}{\mu} &= \frac{4}{3}\log\rho + \log K_2' - \frac{4}{3}\log\mu_e \\
\log\rho &= 3\log T + 3\log\frac{\mathcal{R}}{\mu} - 3\log K_2' + 4\log\mu_e.
\end{aligned}$$

This is another slope 3 line.

Equating relativistic and non-relativistic degeneracy pressures gives the line where degenerate gas becomes relativistic. This is

$$\begin{aligned} \frac{5}{3}\log\rho + \log K_1' + \frac{5}{3}\log\mu_e &= \frac{4}{3}\log\rho + \log K_2' + \frac{4}{3}\log\mu_e \\ \log\rho &= 3\log K_2' - 3\log K_1' - \log\mu_e. \end{aligned}$$

This is a line of slope 0.

To visualize the implications of this, it is helpful to see these 4 lines on a graph, keeping in mind that we only draw the line for relativistic degenerate gas above the line separating relativistic from non-relativistic, and we only draw the line for non-relativistic degenerate gas below this line.

[Slide 1 – types of pressure in the $(\log T, \log\rho)$ plane]

The plot lets us identify which sources of pressure are dominant in which parts of the $(\log T, \log\rho)$ plane. Ideal gas pressure dominates in a strip down the center, which includes the properties found at the center of the Sun, $\log\rho \sim 2$ and $\log T \sim 7$. Increasing the density at fixed temperature makes the gas degenerate, first non-relativistically and then relativistically. Increasing the temperature at fixed density eventually leads to the radiation pressure-dominated regime.

Of course the transitions between the different regimes are smooth and continuous, not sharp as we have drawn them. The purpose of drawing them this way is to give some sense of where in parameter space we have to worry about different effects.

B. Nuclear Reactions

Now that we know what types of pressure occur in different regions, the next thing to add to our plot is regions of nuclear burning. Recall that nuclear reaction rates are extremely temperature sensitive, so the reaction rate generally increases quite dramatically once one is past a certain threshold temperature.

To get a sense of when nuclear burning of a particular type becomes important, it is useful to ask when the energy generation rate passes some minimum value at which it is significant. As a rough estimate of what it means to be significant, we can require that nuclear burning be competitive with Kelvin-Helmholtz contraction as a source of energy – if not, then the burning rate is insufficient to hold up the star.

We can estimate the required reaction rate very roughly as follows. In the absence of nuclear burning, the star will contract on a KH timescale, which means that the gas will heat up by a factor of order unity in a time $t \sim t_{\mathrm{KH}}$. In order for nuclear reactions to be significant they would also need to be able to change the gas temperature by a factor of order unity (in the absence of radiative losses that keep everything in equilibrium) on that timescale.

For an ideal gas the energy per unit mass is

$$u_{\rm gas} = \frac{3}{2}\frac{P_{\rm gas}}{\rho} = \frac{3}{2}\frac{\mathcal{R}}{\mu}T.$$

Thus if the rate of nuclear energy generation is q, the timescale required for nuclear reactions to change the gas temperature significantly is

$$t \sim \frac{u_{\rm gas}}{q} = \frac{3}{2}\frac{\mathcal{R}}{\mu}\frac{T}{q}.$$

If we want this to be comparable to the KH timescale, then we require that q have a minimum value of

$$q_{\rm min} \sim \frac{\mathcal{R}}{\mu}\frac{T}{t_{\rm KH}},$$

where we have dropped factors of order unity. Of course both T and the KH timescale vary as the type of star and the conditions in its center change. However, to get a very rough estimate we can just plug in typical Solar values, which are $T \sim 10^7$, $t_{\rm KH} \sim 10$ Myr. Doing so, we find that $q_{\rm min} \sim 10$ erg g^{-1} s^{-1} to the nearest factor of 10.

We can use our formulae for nuclear burning to see when what density and temperature give a burning rate of about this value. The three energy generation rates we wrote down are for the *pp* chain, the CNO cycle, and the 3α reaction, and those are

$$\begin{aligned}
q_{pp} &\simeq 2.4\times 10^6 X^2 \left(\frac{\rho}{1\text{ g cm}^{-3}}\right)\left(\frac{T}{10^6\text{ K}}\right)^{-2/3}\exp\left[-\frac{33.8}{(T/10^6\text{ K})^{1/3}}\right]\\
q_{\rm CNO} &\simeq 8.7\times 10^{27} X X_{\rm CNO}\left(\frac{\rho}{1\text{ g cm}^{-3}}\right)\left(\frac{T}{10^6\text{ K}}\right)^{-2/3}\exp\left[-\frac{152}{(T/10^6\text{ K})^{1/3}}\right]\\
q_{3\alpha} &\simeq 5.1\times 10^8 Y^3\left(\frac{\rho}{1\text{ g cm}^{-3}}\right)^2\left(\frac{T}{10^8\text{ K}}\right)^{-3}\exp\left(-\frac{44}{T/10^8\text{ K}}\right),
\end{aligned}$$

where everything here is in units of erg g^{-1} s^{-1}.

If we set $q = q_{\rm min}$ for each of these reactions, we can solve for ρ in terms of T. As before, it is convenient to take the log of both sides before solving. The result is

$$\begin{aligned}
\log\rho_{pp} &= 14.7 T_6^{-1/3} + \frac{2}{3}\log T_6 - 6.4 - 2\log X + \log q_{\rm min}\\
\log\rho_{\rm CNO} &= 66.0 T_6^{-1/3} + \frac{2}{3}\log T_6 - 27.9 - \log X - \log X_{\rm CNO} + \log q_{\rm min}\\
\log\rho_{3\alpha} &= 9.55 T_8^{-1} + \frac{2}{3}\log T_8 - 4.35 - \frac{3}{2}\log Y + \frac{1}{2}\log q_{\rm min},
\end{aligned}$$

where we have used the abbreviation $T_n = T/(10^n\text{ K})$. These are clearly not straight lines in the $(\log T, \log\rho)$ plane. There is a linear part, which comes from the $(2/3)\log T$ terms, but there is a far more important exponential part, coming from the $T^{1/3}$ and T^{-1} terms, which look like exponentials in the $(\log T, \log\rho)$ plane.

We can add these curves using $X = 0.71$, $X_{\rm CNO} = 0.01$ for main sequence stars and $Y = 1$ for helium-burning stars. The plot also shows curves for other nuclear reactions involving higher $\mathcal{Z}$ nuclei, which can be computed in exactly the same manner. For *pp* and CNO reactions, since they both involve hydrogen burning, the plot shows whichever reaction requires the lower threshold density to produce energy at a rate $q_{\rm min}$ – that's the one that will start first.

[Slide 2 – nuclear reactions in the $(\log T, \log\rho)$ plane]

It is important to notice that the nuclear reaction curves are quite close to vertical lines, particularly for those involving high $\mathcal{Z}$ nuclei. This is a manifestation of the extreme sensitivity of the reaction rates to temperature. Recall that we often approximate rates of nuclear energy generation as powerlaws

$$q = q_0 \rho^\mu T^\nu,$$

where $\mu = 1$ for two-body interactions and $\mu = 2$ for three-body reactions like 3α. The value of ν is ~ 4 for the *pp* chain, ~ 20 for CNO, ~ 41 for 3α, and increases even further at higher $\mathcal{Z}$.

To see what this implies about the shape of the curves we have just drawn, we can set $q = q_{\rm min}$, take the logarithm of both sides of the powerlaw approximation and then re-arrange to solve for $\log\rho$:

$$\begin{aligned} \log q_{\rm min} &= \log q_0 + \mu \log\rho + \nu \log T \\ \log\rho &= -\frac{\nu}{\mu}\log T - \frac{1}{\mu}\log\frac{q_{\rm min}}{q_0} \end{aligned}$$

This is clearly the equation of a line with a slope of $-\nu/\mu$. Now recall that $\mu = 1$ or 2 (usually 1), and that ν is a big number ranging from 4 for the *pp* chain up to many tens for higher $\mathcal{Z}$ reactions. Thus we expect the line where a nuclear reaction becomes important to look like a line with a large, negative slope. Of course it's not exactly a line, since the powerlaw approximation is only an approximation. Nonetheless, this does show why the nuclear reaction lines are so steep.

C. Instability Regions

A third thing to add to our plot is regions of instability. We have seen that stars become dynamically unstable if the gas ever reaches a condition where $\gamma_a < 4/3$, and that $\gamma_a = 4/3$ is only marginally stable, and can lead to instability. Thus the question is: where in the $(\log T, \log\rho)$ plane do we expect the conditions to be such that $\gamma_a < 4/3$.

Two answers are obvious: since relativistic gasses have $\gamma_a = 4/3$, the gas asymptotically approaches 4/3 as we move into either the relativistic degeneracy region or the radiation pressure-dominated region. Moving into these regions never causes γ_a to drop below 4/3, but we should expect some sort of instability to set in for any star whose core ventures too far into these regions.

There are also two other zones of instability to add to the plot. As we discussed earlier in the class, if the temperature exceeds $\sim 6 \times 10^9$ K, photons acquire enough energy to photodisintegrate iron nuclei, reducing them to He and reversing nucleosynthesis. This is an ionization-like process, in the sense that it creates conditions in which changes in the density or pressure do not produce much change in the temperature. Energy provided by compressing the gas and doing work on it instead goes into photodisintegrating more nuclei. If the gas expands and does work, the energy is provided by converting He nuclei back into Fe. In either case, the temperature doesn't change much, so γ_a is near 1.

For this reason, if the temperature in the core of a star reaches $\sim 6 \times 10^9$ K, the star should become dynamically unstable, and should collapse or explode. We therefore add a line to our diagram at this temperature.

The second zone of instability has to do with pair production. If photons have enough energy, they can spontaneously create electron-positron pairs when they interact with other particles. The reaction is something of the form

$$\gamma + e^- \rightarrow e^- + e^- + e^+ + \gamma,$$

where the photon on the right side has at least 1.02 MeV less energy than the one on the right, since the rest mass of the electron-positron pair is 1.02 MeV. The temperature where this reaction starts to happen is dictated by requirement that photons have enough energy to start making pairs. The typical photon energy is $\sim k_B T$, so the reaction starts when

$$k_B T \sim 1.02 \text{ MeV} \qquad \Longrightarrow \qquad T \sim 10^{10} \text{ K}.$$

In fact the temperature required is a bit lower than this, because there are always some photons with energies higher than the mean, and these can start making pairs at lower temperatures.

The effect of this on the equation of state is somewhat complicated, because it depends on the rate at which pairs are produced (which depends on the density of particles with which photons can interact), and on what fraction of the total pressure is provided by the photons as opposed to the gas. However, we can see that this is also an ionization-type process that leads to lower γ_a. The basic physical reason is the same as for ionization: when the gas can change the number of particles present, it acquires a big reservoir of energy. If the gas is compressed and some work is done on it, instead of heating up, the gas can simply increase the number of particles. If the gas expands and does work, it can get the energy by decreasing the number of particles rather than by cooling off. Thus the temperature becomes relatively insensitive to the pressure or density, and γ_a becomes small.

[Slide 3 – instability regions]

Computing the effective value of γ_a when pair creation is underway is a complex problem, and not one that we will solve in this class. The instability region is bounded by a fairly narrow region in temperature and density. If the temperature is too low photons cannot create pairs, and if it is too high then creating a pair does not require a lot of energy, and thus has little effect. If the density is too high, then gas pressure is large compared to radiation pressure, and the loss of energy from photons creating pairs doesn't make much difference. Thus the region of instability is characterized by a maximum density and a maximum and minimum temperature.

II. Stars in the $(\log T, \log \rho)$ Plane

A. Mass Lines

Now that we have established the physical processes that dominate the pressure, the nuclear reactions, and the stability or lack thereof in each part of the $(\log T, \log \rho)$ plane, let us now consider where the centers of actual stars fall in this plane.

In placing stars on this diagram, we can take advantage of the powerful and general relationship we derived between the central pressure and central density of a star. For a polytrope of index n, we showed that

$$P_c = (4\pi)^{1/3} B_n G M^{2/3} \rho_c^{4/3}.$$

Real stars aren't exactly polytropes except in certain special cases, but we have shown that their structures are generally bounded between $n = 1.5$ and $n = 3$ polytropes, depending on the strength of convection and the amount of pressure provided by radiation. For an $n = 1.5$ polytrope, $B_n = 0.206$, and for one with $n = 3$, $B_n = 0.157$. That these values are so close suggests that this equation should apply in general, with only a slight dependence of the coefficient on the internal structure of the star. For this reason, we can simply adopt an approximate value $B_n \simeq 0.2$, and expect that it won't be too far off for most stars.

In order to translate this relationship between P_c and ρ_c into our $(\log T, \log \rho)$, plane, we need to compute the central temperature T_c from ρ_c and P_c. This in turn requires that we use the equation of state. There are several equations of state on our diagram, but we really only need to worry about two: ideal gas and non-relativistic degenerate gas. That is because stars that get too far into one of the other two regimes, either relativistic degenerate gas or radiation pressure, become unstable.

For an ideal gas, we have

$$P_c = \frac{\mathcal{R}}{\mu} \rho_c T_c,$$

so combining this with the central pressure-density relation, we have

$$\frac{\mathcal{R}}{\mu} \rho_c T_c = (4\pi)^{1/3} B_n G M^{2/3} \rho_c^{4/3}$$

$$\rho_c = \frac{1}{4\pi B_n^3}\left(\frac{\mathcal{R}}{\mu}\right)^3 \frac{1}{G^3M^2}T_c^3$$

$$\log\rho_c = 3\log T_c - 2\log M - 3\log G - \log(4\pi B_n^3) + 3\log\frac{\mathcal{R}}{\mu}.$$

Thus in the case of an ideal gas, the relationship between central density and temperature is simply a line of slope 3. The y intercept of the line depends on the star's mass M, so that stars of different masses simply lie along a set of parallel lines.

If the mass is below the Chandresekhar mass, we also have to consider the possibility that the star could be degenerate. If it is, the equation of state is instead

$$P_c = K_1'\left(\frac{\rho_c}{\mu_e}\right)^{5/3}.$$

Repeating the same trick of combining this with the polytropic pressure-density relation, we have

$$K_1'\left(\frac{\rho_c}{\mu_e}\right)^{5/3} = (4\pi)^{1/3}B_nGM^{2/3}\rho_c^{4/3}$$

$$\rho_c = 4\pi B_n^3G^3M^2K_1'^{-3}\mu_e^5$$

$$\log\rho_c = \log(4\pi B_n^3) - 3\log K_1' + 5\log\mu_e + 3\log G + 2\log M$$

This is just a horizontal line, at a value that depends on the star's mass M. We can add lines for stars of various masses to our diagram.

[Slide 4 – mass lines on the $(\log T, \log\rho)$ plane]

Note that the 100 $M_\odot$ gas is really in the regime where radiation is important, but so we should really compute its central temperature taking that into account, but we're going to ignore that complication, because it doesn't move the line by that much, and it doesn't change anything essential.

B. Evolutionary Path

We can interpret these lines as evolutionary tracks for stars. As long as the star's mass remains fixed, it is constrained to spend its entire life somewhere on the line associated with its mass – it simply moves from one point on the line to another.

We can understand the great majority of stellar evolution simply by looking at this diagram. Stars form out of gas clouds that are much less dense and much colder than than the center of a star. Thus all stars begin their lives at the bottom left corner of the diagram. Since the place where they begin is in the ideal gas region, the ideal gas virial theorem applies, and we have

$$\frac{1}{2}\dot{\Omega} = -\dot{U} = L_{\rm nuc} - L \qquad U = \frac{3}{2}\frac{\mathcal{R}}{\mu}M\overline{T} \qquad \Omega = -\alpha\frac{GM}{R}.$$

Since $L_{\rm nuc} = 0$ at this point, Ω must decrease and U must increase. Since the mass is fixed, this means R decreases (and thus ρ_c increases), and $\overline{T}$ increases. Thus the star must move up and to the right in the $(\log T, \log \rho)$ plane. It will do so following its particular mass track.

The speed with which a star moves is dictated by the source of its energy. Before the star crosses any of the nuclear burning lines, it has no significant internal energy source, so it must evolve on the timescale dictated by gravitational power, the KH timescale. Thus stars move up and right on their tracks on a KH timescale.

As they move in this manner, stars eventually encounter the hydrogen burning track. Lower mass stars encounter it on the part that corresponds to the *pp* chain, while higher mass ones encounter it along the part that corresponds to the CNO cycle – the breakpoint is slightly above the mass of the Sun.

Once a star reaches the hydrogen burning line, burning begins and the star stops contracting. It happily sits at that point for a time dictated by its amount of nuclear fuel, the nuclear burning timescale. This depends on the mass of the star, as we will see soon.

After this long pause, the star keeps moving. It is still in the ideal gas part of the plot, and, since it doesn't have a power source, it must keep contracting and losing energy, moving further up and to the right on its mass track. Evolution again takes place on a KH timescale.

At this point, stars begin to diverge onto different paths based on their masses.

1. Low Mass Stars

   For the lowest mass stars, represented on the plot by the 0.1 $M_\odot$ line, the next significant line they encounter is the transition from ideal gas to non-relativistic degenerate gas. Once a star hits that line, the pressure becomes independent of the temperature. The star continues to radiate, however. Since the pressure doesn't change, the star can't contract, and thus it must pay for this radiation out of its thermal energy instead of its gravitational potential energy. As a result, instead of heating up by radiating, a degenerate star cools by radiating.

   Thus the star stops moving up and to the right on its mass track, and instead begins to move to the left at fixed ρ_c. It gets colder and colder, but the pressure and density don't depend on temperature any more, so it just sits there. As time passes the star dims, since its radius is fixed and its temperature is dropping, but it will happily continue slowly inching to the left for the entire age of the universe.

   Since the star burned H to He, but never got hot enough to ignite He, it is composed of helium. This type of star is known as a helium white dwarf.

2. Medium Mass Stars

As we move to a slightly higher mass, comparable to that of the Sun (represented by the $M_\odot$ line on the plot), what changes is that the star hits the 3α helium burning line before it reaches the degeneracy line. Thus the star ignites He. At this point things stall for another nuclear timescale, but this time the nuclear timescale is computed for helium rather than hydrogen burning.

We evaluate this nuclear timescale in the same way as for hydrogen burning:

$$t_{\rm nuc} = \frac{\epsilon M c^2}{L}.$$

For the 3α reaction, $\epsilon = 6.7 \times 10^{-4}$, roughly an order of magnitude lower than for hydrogen burning. This means that $t_{\rm nuc}$ is shorter for helium burning that for hydrogen burning.

It is even shorter than that because stars at this point are also significantly more luminous than they are on the main sequence. We can see why using the simple Eddington model, in which the luminosity of the star is given by

$$\frac{L}{L_\odot} = \frac{4\pi c G M_\odot}{\kappa_s L_\odot} 0.003 \mu^4 \beta^4 \left(\frac{M}{M_\odot}\right)^3,$$

where β is given by the Eddington quartic

$$0.003 \left(\frac{M}{M_\odot}\right)^2 \mu^4 \beta^4 + \beta - 1 = 0.$$

For low mass stars, the first term is negligible, and so we have $\beta \simeq 1$ independent of μ. Thus the luminosity simply scales as μ^4.

Consider how μ changes as the star evolves. As a reminder,

$$\begin{aligned}
\frac{1}{\mu} &= \frac{1}{\mu_I} + \frac{1}{\mu_e} \\
\frac{1}{\mu_I} &\approx X + \frac{1}{4} Y + \frac{1 - X - Y}{\langle \mathcal{A} \rangle} \\
\frac{1}{\mu_e} &\approx \frac{1}{2}(1 + X).
\end{aligned}$$

For Solar composition, $X = 0.707$ and $Y = 0.274$, we found $\mu = 0.61$. If we take all the hydrogen in the star and turn it into helium, we instead have $X = 0$ and $Y = 0.98$, and plugging in gives $\mu = 1.34$. Thus μ increases by a factor of $1.34/0.61$, and the luminosity increases by a factor of $(1.34/0.61)^4 = 23$. The surface opacity κ_s also decreases, further increasing the luminosity.

Physically, μ increases from two effects. First, converting hydrogen to helium increases the mean mass per ion of the most abundant species from 1 to 4. Second, each conversion of 4 hydrogen into 1 helium involves the conversion of 2 protons to neutrons, and, in order to conserve charge, the creation of two positrons. These two positrons annihilate with electrons, so for each He that is created, 2 electrons are destroyed. This reduces the number of electrons, further increasing the mean mass per particle.

The increase in μ means that the gas pressure is smaller at fixed temperature, which in turn means that a higher temperature is required to hold up the star against gravity. A higher temperature leads to a larger temperature gradient, which increase the rate of energy transport through the star and thus, ultimately, the stellar luminosity.

The result of this increase in luminosity along with the decrease in ϵ for 3α compared to hydrogen burning is that the nuclear timescale for helium burning is ~ 3 orders of magnitude smaller than the corresponding nuclear lifetime for hydrogen burning. Thus while the Sun will take ~ 10 Gyr to evolve off the main sequence of hydrogen burning stars, its lifetime as a helium burning star will be closer to 10 Myr.

Once the available helium is used up, the core will consist mostly of carbon and oxygen that do not burn. The star is still in the ideal gas part of the diagram, so it will being KH contraction again, shining by gravity while becoming denser and hotter. The KH timescale for this contraction phase is also reduced compared to the previous step of contraction, because the luminosity increases due to the increase in mean molecular weight.

For a star like the Sun, the next significant line it encounters is the one for degeneracy. After that point, its evolution is like that of lower mass stars: it ceases contracting and instead begins to cool at constant radius and central density, dimming as it does. The star ends it life as a carbon-oxygen white dwarf.

3. High Mass Stars

If we increase the mass a bit more, to larger than the Chadrasekhar mass, then no transition to a degenerate state is possible. The track of such a star is indicated by the 10 $M_\odot$ line on the plot. Geometrically, it is easy to see what is going on. The line between ideal gas and degenerate relativistic gas has a slope of 3, the same as the slope of the constant mass tracks in the ideal gas region. Since two lines of the same slope will never intersect unless they are identical, the constant mass track will never hit the degenerate gas region unless it does so in the non-relativistic area, where the slope is 3/2 rather than 2.

Since a star of this mass cannot become degenerate, it instead reaches the next nuclear burning line. First it burns carbon, then it contracts some more, then it burns oxygen, contracts some more, and then burns silicon. Each of these burning phases has a nuclear timescale that is shorter than the last, for three reasons. The first two are the same reasons that helium burning was shorter than hydrogen burning. First, as one moves to higher $\mathcal{Z}$, the elements are already closer to the peak of binding energy per nucleon, so ϵ is decreasing. Second, as the mean atomic weight $\mathcal{A}$ increases, the star must burn brighter and brighter to maintain hydrostatic balance.

The third reason is that, as one moves up the periodic table toward the iron peak, more and more of the energy of nuclear burning comes out in the form of neutrinos rather than photons, and the neutrinos escape the star and take their energy with them. By the time the star is burning silicon into iron, the star is roughly 1 million times brighter in neutrinos than it is in photons. This means that the nuclear reaction rate must be 1 million times faster to keep up, and the nuclear timescale is correspondingly shortened. The net outcome is that the silicon burning phase lasts only ~ 18 days.

Once the silicon is burned to iron, the star has no choice but to continue moving up and to the right on its mass evolution track. There is finally hits the photodisintegration instability strip. At that point the core switches to a $\gamma_a = 4/3$ equation of state, becomes dynamically unstable, and the story is over. The star will either collapse to a black hole or explode as a supernova, leaving behind a neutron star.

4. Very High Mass Stars

   An even more massive star can reach an instability strip even sooner, before it begins oxygen burning. If the star is massive enough, its track intersects the instability region associated with pair creation. At this point the star becomes dynamically unstable, and it will either collapse into a black hole or explode as a supernova, as in the case for a star that reaches the photodisintegration instability region.

   Pair instability supernovae are different than those that occur in less massive stars, in that they occur before the star has fused the elements in its core up to the iron peak. Some nuclear fusion can happen as the star explodes, but in general pair instability supernovae produce much more oxygen and carbon than ordinary core collapse supernovae.

5. Mass Loss

   The final thing to mention in this lecture is a major note of caution. This story is useful, and it is generally correct as an outline. However, it omits one major factor: mass loss. All the evolutionary tracks we have just discussed assume that stars are constrained to move along lines of constant mass. While this is roughly correct for stars on the main sequence, as we will see next week, it is not correct for post-main sequence stars. Instead, such stars can lose large fractions of their mass via a variety of processes we will discuss.

   The effect of mass loss is to allow stars, once they are past the main sequence, to slide upward from a higher mass evolutionary track to a lower mass one.

Thus the condition for going supernova is not that the star's initial mass exceed the Chandrasekhar mass, because even a star that starts well above $M_{\text{Ch}} = 1.4\ M_{\odot}$ may be able to lose enough mass to get down to M_{Ch} by the time it is approaching the degenerate region. The boundary between stars that turn into white dwarfs and stars that end their lives in supernovae turns out to be an initial mass of roughly 8 $M_{\odot}$, rather than 1.4 $M_{\odot}$.

For this reason, a complete theory of stellar evolution must include a model for mass loss as well.

Class 14 Notes: The Main Sequence

In the last class we drew a diagram that summarized the basic evolutionary path of stars, as seen from their centers. In this class we will focus on one the evolutionary phase where stars spend most of their lives: the main sequence. Our goal is to demonstrate that we can understand the main sequence qualitatively in terms of our simple stellar evolution model, and then to examine some detailed numerical results.

I. Homology and Scalings on the Main Sequence

A. The Non-Dimensional Structure Equations

We saw in the last class that stars are born at low ρ_c and T_c, and they evolve along tracks of constant mass to higher central density and pressure. This evolution takes place on a KH timescale, and ends when the stars' cores intersect the hydrogen burning line. At that point the stars ignite hydrogen and burn for a time $t_{\rm nuc} \gg t_{\rm KH}$. The hydrogen-burning nuclear timescale is the longest time for any evolutionary stage, because

$$t_{\rm nuc} = \frac{\epsilon M c^2}{L},$$

and the hydrogen burning stage is the one with the largest value of ϵ and the smallest value of L.

When we look at a population of stars that are at many different ages, and thus at many random points in their lives, we expect the number of stars we see in a given population to be proportional to the fraction of its life that a star spends as a member of that population. In other words, if one evolutionary phase lasts 1 million times longer than a second phase, we would expect to see roughly 1 million times as many stars in the first evolutionary phase than we do in the second. Since the main sequence is the most heavily populated part of the HR diagram and the hydrogen nuclear burning phase is the longest evolutionary phase, it seems natural to assume that main sequence stars are burning hydrogen.

To check this hypothesis, we need to investigate whether stars that have reached the hydrogen burning line in the $(\log T, \log \rho)$ plane have the right properties to be main sequence stars. In particular, do they have the right mass luminosity relation; and do they have the right relationship between luminosity and surface temperature?

To answer this question, we'll write down our stellar structure equations, in a simplified form to make life easy. We will neglect convection and radiation pressure, use a constant opacity, and use a powerlaw approximation for the rate of nuclear burning. The goal is to show that we can roughly reproduce what is observed, although not get every detail right. With these assumptions, the complete set of equations is

$$\begin{aligned}
\frac{dP}{dm} &= -\frac{Gm}{4\pi r^4} \\
\frac{dr}{dm} &= \frac{1}{4\pi r^2 \rho} \\
\frac{dT}{dm} &= -\frac{3}{4ac}\frac{\kappa}{T^3}\frac{F}{(4\pi r^2)^2} \\
\frac{dF}{dm} &= q_0 \rho T^\nu \\
P &= \frac{\mathcal{R}}{\mu}\rho T
\end{aligned}$$

The unknowns are $r(m)$, $P(m)$, $T(m)$, $F(m)$, and $\rho(m)$, and there are 5 equations, so the system is fully specified.

We don't have to solve the equations exactly to get out the basic behavior. Instead, we can figure out many scalings with some simple dimensional arguments. To do this, we will deploy the same technique of non-dimensionalizing the equations that we used so effectively with polytropes. We begin by defining a dimensionless mass variable

$$x = \frac{m}{M},$$

and then defining dimensionless versions of all the other variables:

$$\begin{aligned}
r &= f_1(x) R_* \\
P &= f_2(x) P_* \\
\rho &= f_3(x) \rho_* \\
T &= f_4(x) T_* \\
F &= f_5(x) F_*,
\end{aligned}$$

where M is the total mass of the star and R_*, P_*, ρ_*, T_*, and F_* are values of the radius, pressure, density, temperature and heat flow that we have not yet specified.

Thus far all we have done is define a new set of variables. We will now substitute this new set of variables into the equation of hydrostatic balance:

$$\frac{dP}{dm} = -\frac{Gm}{4\pi r^4} \qquad \longrightarrow \qquad \frac{P_*}{M}\frac{df_2}{dx} = -\frac{GMx}{4\pi R_*^4 f_1^4}.$$

We now exercise our freedom to define P_*. We define it by

$$P_* = \frac{GM^2}{R_*^4},$$

and with this choice the equation of hydrostatic balance reduces to

$$\frac{df_2}{dx} = -\frac{x}{4\pi f_1^4}.$$

This is the non-dimensional version of the equation.

We can non-dimensionalize the other equations in a similar fashion, in each case exercising our freedom to choose one of the starred quantities. For the dr/dm equation, we choose

$$\rho_* = \frac{M}{R_*^3},$$

which gives us the non-dimensional equation

$$\frac{df_1}{dx} = \frac{1}{4\pi f_1^2 f_3}.$$

For the other equations, the definitions and non-dimensionalized versions are

$$\begin{aligned} T_* &= \frac{\mu P_*}{\mathcal{R}\rho_*} & f_2 &= f_3 f_4 \\ F_* &= \frac{ac}{\kappa}\frac{T_*^4 R_*^4}{M} & \frac{df_4}{dx} &= -\frac{3f_5}{4f_4^3(4\pi f_1^2)^2} \\ F_* &= q_0 \rho_* T_*^\nu M & \frac{df_5}{dx} &= f_3 f_4^\nu. \end{aligned}$$

You might be suspicious that we defined F_* twice, which we can't do. The trick is that we have yet not chosen R_*. Thus we can use our last choice to define R_* in such a way as to make the second equation here true. Once we do so, have have defined all the starred quantities, and non-dimensionalized all the equations.

What is the point of this? The trick is that the non-dimensional equations for $f_1 - f_5$ now depend only on dimensionless numbers, and not on the stellar mass. Any dependence of the solution on mass *must* enter only through the starred quantities. Another way of putting it is that these equations have the property that they are homologous – one can solve for $f_1 - f_5$, and then scale that solution to an arbitrary mass by picking a different value of M. In a sense, these equations say that all stars (for which our four assumptions above are valid) have the same structure.

Of course the only reason we were able to obtain non-dimensionalized equations of this form and demonstrate homology is due to the simplifying assumptions we made – neglect of radiation pressure, neglect of convection, adopting a constant κ, and using a powerlaw form for the nuclear energy generation rate. These complications are the basic reasons that stars do not actually all have the same structure independent of mass. Nonetheless, the first and last of these assumptions are reasonably good for low mass stars (though not for massive stars). The assumption of constant κ isn't strictly necessary, as you will demonstrate on your homework. The most questionable assumption is our neglect of convection.

B. Mass Scalings

With that aside out of the way, we can proceed to use the homologous equations to deduce the dependence of all quantities on mass. Combining the equations for ρ_*, P_* and T_* gives

$$T_* = \frac{\mu}{\mathcal{R}} \frac{GM}{R_*}.$$

Notice that we have already proven essentially this result using the virial theorem.

Inserting T_* into the equation for F_* gives

$$F_* = \frac{ac}{\kappa} \frac{R_*^4}{M} \left(\frac{\mu}{\mathcal{R}} \frac{GM}{R_*} \right)^4 = \frac{ac}{\kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 M^3$$

Since this relation applies at any value of x, it must apply at $x = 1$, i.e. at the surface of the star. Since at the stellar surface $L = F = F_* f_5(1)$, it immediately follows that

$$L \propto \frac{ac}{\kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 M^3.$$

Thus the luminosity varies as M^3. Notice that this is independent of any of the other starred quantities – we have derived the dependence of L on the mass alone. Also notice that this result is basically the same as we get from Eddington's model with $\beta = 1$ (i.e. our assumption of no radiation pressure) – which makes sense, since Eddington's model is a polytrope, and therefore homologous, and also has constant κ. Thus we couldn't possibly find anything else.

We can now push further and deduce the mass scalings of other quantities as well. We have

$$F_* = q_0 \rho_* T_*^\nu M = \frac{ac}{\kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 M^3 \qquad \Longrightarrow \qquad \rho_* = \frac{ac}{q_0 \kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 \frac{M^2}{T_*^\nu}.$$

Substituting for ρ_* and T_* gives

$$\frac{M}{R_*^3} = \frac{ac}{q_0 \kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 M^2 \left(\frac{\mu P_*}{\mathcal{R} \rho_*} \right)^{-\nu}$$

Finally, substituting for P_* and ρ_* again gives

$$\begin{aligned}
\frac{M}{R_*^3} &= \frac{ac}{q_0 \kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^4 M^2 \left(\frac{\mu}{\mathcal{R}} \right)^{-\nu} \left(\frac{GM^2}{R_*^4} \frac{R_*^3}{M} \right)^{-\nu} \\
\frac{M}{R_*^3} &= \frac{ac}{q_0 \kappa} \left(\frac{\mu G}{\mathcal{R}} \right)^{4-\nu} M^{2-\nu} R_*^\nu \\
R_* &= \left[\frac{q_0 \kappa}{ac} \left(\frac{\mathcal{R}}{\mu G} \right)^{4-\nu} \right]^{1/(\nu+3)} M^{(\nu-1)/(\nu+3)}
\end{aligned}$$

Thus we expect the stellar radius to scale with mass in a way that depends on how the nuclear reactions scale with temperature. If we have a star that burns hydrogen mainly via the *pp* chain, then $\nu \approx 4$, and we obtain $R \propto M^{3/7}$. For a more massive star that burns mainly via the CNO cycle, we have $\nu \approx 20$, and we instead obtain $R \propto M^{19/23}$, a nearly linear relationship. Thus we expect the radius to increase with mass as $M^{3/7}$ at small masses, increasing in steepness to a nearly linear relationship at larger masses.

For the density, we have

$$\rho_* = \frac{M}{R_*^3} = \left[\frac{q_0\kappa}{ac}\left(\frac{\mathcal{R}}{\mu G}\right)^{4-\nu}\right]^{-3/(\nu+3)} M^{2(3-\nu)/(3+\nu)}.$$

For *pp* chain stars, this gives $\rho_* \propto M^{-2/7}$, and for CNO cycle stars it gives $\rho_* \propto M^{-34/23}$, which is nearly -1.5. Thus the density always decreases with increasing stellar mass, but does so fairly slowly for *pp* chain stars (-0.29 power) and quite rapidly for CNO cycle stars (-1.5 power). This is an important and often under-appreciated point in stellar structure: more massive stars are actually much less dense than less massive ones. Very massive stars are quite puffy and diffuse.

C. The Observed Main Sequence

Finally, we can get out the scaling that we really care about: luminosity versus temperature. This is what will determine the shape of the observed main sequence, and we had better make sure that what we get out of the theoretical model agrees reasonably well with what we actually observe. If not, the hypothesis that the main sequence is made up of stars whose cores are stalled on the hydrogen burning line will not be valid.

The effective temperature is related to the radius and luminosity by

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

However we have just shown that

$$L \propto M^3 \qquad \text{and} \qquad R \propto M^{(\nu-1)/(\nu+3)}.$$

Inverting the first relation and substituting it into the second, we have

$$M \propto L^{1/3} \qquad \Longrightarrow \qquad R \propto \left(L^{1/3}\right)^{(\nu-1)/(\nu+3)} \propto L^{(\nu-1)/[3(\nu+3)]}.$$

Now plugging this into the relationship between L and $T_{\rm eff}$, we give

$$\begin{aligned}
\frac{L}{\left[L^{(\nu-1)/[3(\nu+3)]}\right]^2} &\propto T_{\rm eff}^4 \\
L^{1-2(\nu-1)/[3(\nu+3)]} &\propto T_{\rm eff}^4 \\
\left[1 - \frac{2(\nu-1)}{3(\nu+3)}\right]\log L &= 4\log T_{\rm eff} + \text{constant} \\
\log L &= 4\left[1 - \frac{2(\nu-1)}{3(\nu+3)}\right]^{-1}\log T_{\rm eff} + \text{constant}.
\end{aligned}$$

We have therefore derived an equation for the slope that the main sequence should have in the HR diagram, which shows $\log L$ vs. $\log T_{\rm eff}$.

Plugging in $\nu = 4$ for *pp* chain stars and $\nu = 20$ for CNO cycle stars, we obtain

$$\begin{aligned} \log L &= 5.6 \log T_{\rm eff} + \text{constant} \qquad (pp \text{ chain}) \\ \log L &= 8.9 \log T_{\rm eff} + \text{constant} \qquad (\text{CNO cycle}) \end{aligned}$$

The values compare reasonably well with the observed slopes of the lower and upper main sequence on the HR diagram.

We can also explain other features of observed HR diagrams with this simple model. As we noted when we discussed star clusters of different ages, more massive, luminous stars leave the main sequence before lower mass stars. In the picture we have now developed, leaving the main sequence corresponds to moving past the hydrogen ignition point in the $(\log T, \log \rho)$ plane. The time spent at that point is given roughly by the nuclear timescale,

$$t_{\rm nuc} = \frac{\epsilon M c^2}{L} \propto M^{-2}$$

This means that the nuclear timescale should decrease with stellar mass as M^{-2}, so more massive stars have shorter nuclear timescales and leave the main sequence first. This is exactly what we observe.

We can also use this analysis to estimate the behavior of the upper and lower ends of the main sequence, by scaling from the Sun. The temperature in the star scales as

$$T_* \propto \frac{M}{R_*} \propto \frac{M}{M^{(\nu-1)/(\nu+3)}} \propto M^{4/(\nu+3)}.$$

For *pp* chain stars, this means that $T_* \propto M^{4/7}$. This applies throughout the star, including in the center.

The Sun has a central temperature of $T_{c,\odot} \simeq 1.5 \times 10^7$ K, so we expect that

$$\frac{T_c}{T_{c,\odot}} = \left(\frac{M}{M_\odot}\right)^{4/7}.$$

Thus lower mass stars than the Sun have lower central temperatures. However, the temperature cannot decline indefinitely without interfering with the star's ability to generate energy. If $T_c \lesssim 4 \times 10^6$ K, then the star will not be able to burn hydrogen, and it cannot ignite. The mass at which this limit is reached is

$$M = M_\odot \left(\frac{4 \times 10^6 \text{ K}}{1.5 \times 10^7 \text{ K}}\right)^{7/4} = 0.1 M_\odot.$$

The corresponding luminosity for a star at this limit is

$$L = L_\odot \left(\frac{M}{M_\odot}\right)^3 = 10^{-3} L_\odot.$$

Thus the dimmest stars in existence should have a luminosity about 1/1000th that of the Sun.

Of course this argument is a bit of a cheat: the way we got the homology argument in the first place is by assuming a particular scaling for the nuclear reaction rate with temperature ν, and, if the temperature gets too low, ν will change. Thus the limit is a bit more complicated. Nonetheless, this argument gives a reasonable estimate for the minimum mass of an object that can burn hydrogen. Lower mass objects never ignite hydrogen, and instead end up being supported by degeneracy pressure. This objects are called brown dwarfs, and, even though they form like stars, they end up having structures more like that of the planet Jupiter.

In the opposite direction, we can use a similar scaling argument to deduce when the mass of a star becomes large enough for it to approach the Eddington limit. The star's luminosity scales like $L \propto M^3$, and the Eddington luminosity is

$$L_{\rm Edd} = \frac{4\pi c G M}{\kappa} \propto M.$$

Thus

$$\frac{L}{L_{\rm Edd}} \propto M^2.$$

Again scaling from the Sun, we find

$$\frac{L/L_{\rm Edd}}{L_\odot/L_{\rm Edd,\odot}} = \left(\frac{M}{M_\odot}\right)^2.$$

Thus the mass at which the luminosity reaches the Eddington luminosity is

$$M = M_\odot \left(\frac{1}{L_\odot/L_{\rm Edd,\odot}}\right)^{1/2} = M_\odot \left(\frac{4\pi c G M_\odot}{\kappa L_\odot}\right)^{1/2} = 114 M_\odot.$$

The corresponding luminosity is

$$L = L_\odot \left(\frac{M}{M_\odot}\right)^3 = 1.5 \times 10^6 L_\odot.$$

Thus the most luminous stars in existence should be more than a million times as bright as the Sun.

Of course this calculation too is a bit of a cheat – the equations we used to derive the scalings on which this relationship is based assume negligible radiation pressure, which is obviously not the case in a star that is approaching the Eddington limit. Nonetheless, this again gives us a rough idea of where we should start crossing over into stars that are supported mostly by radiation. Since we have seen that these stars tend to have stability problems, we do not expect to find a lot of stars of this mass, and thus we expect this to represent a rough upper limit to the main sequence.

D. Convective Stars

The calculation we just performed is for stars where energy is transported radiatively. However, we said last week that there are parts of stars where instead the energy is transported convectively. We can repeat this homology analysis for the case of convective stars, and derive the main sequence mass-luminosity and temperature-luminosity relation we expect in that case to compare to our radiative results. Since real stars are usually convective over part but not all of their interiors, the real behavior should lie in between these extremes.

We showed that convection should bring the temperature gradient within a star extremely close to the adiabatic temperature gradient, so we can to very good approximation say that the gas is adiabatic, and the star is therefore a polytrope with $\gamma_P = \gamma_a = \gamma$. This implies that

$$P = K_a \rho^\gamma$$

and

$$T = \frac{\mu}{\mathcal{R}} \frac{P}{\rho} = \frac{\mu}{\mathcal{R}} K_a \rho^{\gamma-1} = \frac{\mu}{\mathcal{R}} K_a \left(\frac{P}{K_a}\right)^{(\gamma-1)/\gamma} = \frac{\mu}{\mathcal{R}} K_a^{1/\gamma} P^{(\gamma-1)/\gamma}$$

For an ideal non-relativistic gas, $\gamma = 5/3$. These equations replace the ideal gas law and the radiation diffusion equation in our set of stellar structure equations. The other equations are unchanged:

$$\begin{aligned} \frac{dP}{dm} &= -\frac{Gm}{4\pi r^4} \\ \frac{dr}{dm} &= \frac{1}{4\pi r^2 \rho} \\ \frac{dF}{dm} &= q_0 \rho T^\nu \end{aligned}$$

The procedure for non-dimensionalizing these equations is essentially the same as in the radiative case. We make the same changes of variables and in all equations except the temperature one, and get the set of non-dimensional equations

$$\begin{aligned} P_* &= \frac{GM^2}{R_*^4} & \frac{df_2}{dx} &= -\frac{x}{4\pi f_1^4} \\ \rho_* &= \frac{M}{R_*^3} & \frac{df_1}{dx} &= \frac{1}{4\pi f_1^2 f_3} \\ P_* &= K_a \rho_*^\gamma & f_2 &= f_3^\gamma \\ T_* &= \frac{\mu}{\mathcal{R}} K_a^{1/\gamma} P_*^{(\gamma-1)/\gamma} & f_4 &= f_2^{(\gamma-1)/\gamma} \\ F_* &= q_0 \rho_* T_*^\nu M & \frac{df_5}{dx} &= f_3 f_4^\nu. \end{aligned}$$

To get the mass-luminosity and effective-temperature luminosity relations in this case, we have the same basic problem as in the radiative case: given this set of 5 algebraic equations in 5 variables, we must manipulate them to get M by itself on the right-hand side. We're only interested in the proportionalities, so for convenience we can take the logarithm of everything and forget about the constants:

$$\begin{aligned}
\log P_* &= 2\log M - 4\log R_* + \text{constant} \\
\log \rho_* &= \log M - 3\log R_* \\
\log P_* &= \gamma \log \rho_* + \text{constant} \\
\log T_* &= \left(\frac{\gamma-1}{\gamma}\right)\log P_* + \text{constant} \\
\log F_* &= \log \rho_* + \nu \log T_* + \log M + \text{constant.}
\end{aligned}$$

We now have the same algebra problem as in the radiative case, which is to rearrange this set of 5 equations so that everything is given in terms of M. This is not hard, since in this logarithmic form the equations are linear, and this just represents a set of 5 linear equations in 5 unknowns. To solve, we begin by equating the two expressions for $\log P_*$ and substituting for ρ_*:

$$\begin{aligned}
2\log M - 4\log R_* &= \gamma \log \rho_* + \text{constant} \\
2\log M - 4\log R_* &= \gamma(\log M - 3\log R_*) + \text{constant} \\
\log R_* &= \left(\frac{\gamma-2}{3\gamma-4}\right)\log M + \text{constant.}
\end{aligned}$$

Thus we now have $\log R_*$ in terms of $\log M$ alone. We now substitute this into the equations for $\log P_*$, $\log \rho_*$, and $\log T_*$ to solve for them in terms of M:

$$\begin{aligned}
\log P_* &= \left(\frac{2\gamma}{3\gamma-4}\right)\log M + \text{constant} \\
\log \rho_* &= \left(\frac{2}{3\gamma-4}\right)\log M + \text{constant} \\
\log T_* &= \left(\frac{2(\gamma-1)}{3\gamma-4}\right)\log M + \text{constant.}
\end{aligned}$$

We then substitute this into the equation for $\log F_* = \log L + \text{constant}$:

$$\log L = \left(\frac{\gamma(2\nu+3) - 2(\nu+1)}{3\gamma-4}\right)\log M + \text{constant}$$

This gives the mass luminosity relation. For $\gamma = 5/3$ and $\nu = 4$, the constant is $25/3 = 8.33$, to the luminosity varies very steeply with mass.

The final step is to turn this into a luminosity temperature relation, using

$$L = 4\pi R^2 \sigma T_{\text{eff}}^4 \qquad \Longrightarrow \qquad \log L = 2\log R + 4\log T_{\text{eff}} + \text{constant.}$$

We therefore need $\log R$ in terms of $\log L$. To get it, we use our expressions for both $\log R_*$ and $\log L_*$ in terms of $\log M$, which give:

$$\log R_* = \left(\frac{\gamma - 2}{\gamma(2\nu + 3) - 2(\nu + 1)}\right) \log L + \text{constant}.$$

Plugging this into the expression for $\log R$ in the luminosity-effective temperature relationship and solving finally gives

$$\log L = \left(\frac{4(\gamma(2\nu + 3) - 2(\nu + 1))}{2\nu(\gamma - 1) + \gamma + 2}\right) \log T_{\text{eff}} + \text{constant}.$$

This is the luminosity-effective temperature relationship for a polytropic star with arbitrary γ. For a fully convective star composed of non-relativistic ideal gas, $\gamma = 5/3$, and for $p - p$ chain burning, $\nu = 4$. Plugging in these values gives a value of 3.7 for the coefficient. This is shallower than the value of 5.6 we found for radiative stars with constant κ. As you will show on your homework, the value for radiative stars where κ is the free-free opacity differs slightly from these.

II. Numerical Results on the ZAMS

We have now pushed as far as we are going to analytically, and the time has come to bring out the computers. We have written down all the necessary equations, and they can be solved by modern computers quite easily. We will not discuss the necessary algorithms – that is a main topic of a graduate stellar structure class. Instead, we will simply review the important results. For today we will focus on stars that have not yet processed a significant amount of hydrogen into helium. Stars of this sort are said to be on the zero age main sequence, or ZAMS for short. We will talk about evolution of stars before and after the ZAMS next week.

A. Mass-Luminosity-Effective Temperature Relations

The most basic output of the numerical codes is a prediction for the luminosity and effective temperature of a star of a given mass and composition. The figures show the results of one particular set of numerical stellar models that is freely downloadable on the web. These are called the Geneva models, since the research group that produced them is centered at Geneva Observatory.

[Slides 1 and 2 – mass-luminosity and luminosity-effective temperature relations]

The basic behavior is essentially as we predicted from our simple models (the no convection model does better). The luminosity scales as mass to roughly the third power at low masses – slightly steeper due both to the effects of convection and the varying opacity. At higher masses the dependence flattens out, approaching $L \propto M$ at the very highest masses, The most massive stars have luminosities of a bit more than $10^6\ L_\odot$, while the lowest mass ones are below $10^{-2}\ L_\odot$ – these tracks only go down to $0.4\ M_\odot$, so they don't probe what are really the absolute smallest stars.

Similarly, the plot of $\log L$ vs. $\log T_{\text{eff}}$ has a slope of $\sim 5-6$ for intermediate mass stars, with values of the effective temperature ranging from a few thousand Kelvin to several tens of thousands.

B. Interior Structure and Convection

Another basic output of the numerical models is a prediction for the internal structure of stars, meaning the run of density, mass, temperature, pressure, etc. versus radius. The plot shows an example for the Sun. We see that stars are very centrally concentrated – the density and pressure fall to $\sim 1\%$ of their central value by a radius of 50% of the total stellar radius. Thus one may reasonably think of the Sun and other stars as consisting of a compact, dense core, surrounded by a fluffy, diffuse envelope.

[Slide 3 – Solar properties versus radius]

Another useful plot is one that compares the structures of stars of different masses.

[Slide 4 – convection and structure versus mass]

This type of plot is a little complicated, because it packs in a lot of information, but it is very useful. Here's how to read it: the x axis is the mass of star we're examining. The y axis indicates position within the star, using Lagrangian coordinates. Thus the core of the star is at the bottom, $m/M = 0$, and the surface is at the top, $m/M = 1$.

Within this coordinate system, one can draw lines like the ones in the figure labelled with $0.5R$. This curve shows, for a star of mass M on the x axis, what fraction of the mass (m/M) is within a radius that is 50% of the total stellar radius. As the plot shows, this mass fraction (m/M) is usually significantly larger than 0.5, and is often well above 0.8. This is telling us that stars are quite centrally concentrated: the inner half of the radius contains the great majority of the mass.

Similarly the curves labelled with $0.5L$ and $0.9L$ indicate where in the star the luminosity is being generated. The line $0.9L$ indicates the mass that is responsible for generating 90% of the total power. This is quite close to the center, $m/M \sim 0.1-0.2$, particularly for massive stars. This is because nuclear burning is very temperature-sensitive, particularly for the CNO cycle, so most burning happens near the center where the temperature is highest.

Finally, the regions with the circles shows where convection occurs in the star. In low mass stars, convection occurs near the surface, where it is driven by a combination of partial ionziation (which produces γ_a near 1) and low temperature (which produces large opacity). In massive stars, convection occurs near the center, where the high temperature sensitivity of the CNO cycle causes the heat flow to change very rapidly with radius.

Notice that stars below about 0.3 $M_\odot$ are fully convective – and thus are well described by $n = 1.5$ polytropes. The amount of convective mass also increases with mass in very massive stars, as rapid nuclear burning causes a larger and larger temperature gradient. In contrast, the Sun lies near a minimum in terms of the fraction of its mass that is convective. Only a few percent of the Sun's mass lies in the convection zone, although it is a considerably larger fraction of the Sun's radius.

C. Main Sequence Lifetime

Knowing where the star is convective is particularly important for a number of reasons. Perhaps the biggest one is that the size of the convective zone influences the amount of fuel available to the star, and thus the main sequence lifetime.

Within a convection zone, the churning of mass up and down tends to mix the gas, and homogenize the composition – convection acts like a giant paint mixer. In stars with convective cores, the stellar paint mixer has the important effect of dragging fresh hydrogen down into the nuclear burning core, supplying additional fuel. In contrast, in stars like the Sun that are not convective in their cores, the supply of hydrogen is limited to what was present at a given radius when the star formed – there is no effective way to bring in extra hydrogen from further up in the star.

We previously estimated the nuclear timescale as

$$t_{\mathrm{nuc}} = \frac{\epsilon M c^2}{L} = 100 \frac{M/M_\odot}{L/L_\odot} \text{ Gyr}$$

where $\epsilon = 6.6 \times 10^{-3}$ for hydrogen burning. However, now we understand a subtlety in this estimate that we did not before: only gas within the nuclear burning region, or within a convective zone that includes it, is available as fuel. Thus the mass that enters this timescale estimate should only be that fraction of the mass that is available for burning.

This has important implications for the Sun. Unmodified, our estimate would suggest that the Sun should have a 100 Gyr lifetime on the main sequence. However, we have mentioned several times that this is an overestimate, and that the true answer is closer to 10 Gyr. The reason for this is that the Sun is not convective over the vast majority of its mass, so only mass within the nuclear burning region should be counted in computing t_{nuc}. Examining the structure plot, we see that for a 1 $M_\odot$ star, half the nuclear energy is produced within the central 10% of the mass, so the mass available for nuclear burning is $\sim 0.1\ M_\odot$, not $\sim 1\ M_\odot$.. This is why the Sun's main sequence lifetime is roughly 10 Gyr, not 100 Gyr.

We will return to the question of main sequence lifetimes, and what happens to stars after they have exhausted their hydrogen supplies, next week.

Class 15 Notes: Stars Before the Main Sequence

In the last class we discussed the structure of hydrogen-burning stars, which are the ones that constitute the main sequence. This is the phase in which stars spend the majority of their lives. Starting in this class and for the remainder of the term, we will discuss stellar evolution to either side of the main sequence. Today's topic is the nature of young stars, those that have not yet reached the main sequence. The next three classes will then discuss the diverse fates of stars after they leave the main sequence.

I. Star formation

A. Molecular Clouds

Any discussion of the early evolution of stars must begin with the question of how stars come to be in the first place. The topic of star formation is a vast and active field of research, and there are still numerous unanswered questions about it. However, we can sketch enough of the rough outline to get a basic picture of what must happen. In some sense, this will give us an idea of the initial conditions for a calculation of stellar evolution.

Stars form out of the interstellar medium, a diffuse gas (mean number density $n \sim 1$ cm^{-3}) that fills the space between the stars. Most of this gas is atomic or ionized hydrogen with low density, but in certain places it collects into giant clouds. For reasons we will discuss in a few moments, these clouds are the places where stars form.

In these clouds the density is much higher, ~ 100 cm^{-3}, and the gas is predominantly in the form of molecular hydrogen (H_2). These clouds are typically 10^4 - 10^6 $M_\odot$ in mass, but they occupy a tiny volume of the galaxy, because they are so much denser than the gas around them.

[Slide 1 – molecular clouds in the galaxy M33]

We detect these clouds mostly by the emission from the CO molecules within them. CO molecules can rotate, and their rotation is quantized. Molecules that are rotating with 1 quantum of angular momentum can spontaneously emit a photon and stop rotating (giving their angular momentum to the photon). These photons have energies of 4.8×10^{-4} eV, so the corresponding frequency is

$$\nu = \frac{E}{h} = 115 \text{ GHz}.$$

This is in the radio part of the spectrum, and these photons can penetrate the Earth's atmosphere and be detected by radio telescopes, which is how maps like the one I just showed can be made. In addition to this molecular line, there are many more that we can use, involving both different transitions of CO and of other molecules – thousands have been detected.

These clouds are extremely cold, typically around 10 K, mainly because the CO molecules are very efficient at radiating away energy. The clouds are also very dusty, and the dust makes them opaque in the optical. We can see this very clearly by comparing images of a galaxy in optical and CO emission – the places where there are clouds show up as dark dust lanes in the optical, because the dust absorbs all the optical light.

[Slide 2 – CO and optical images of M51]

If we zoom in to look at a single one of these clouds, we see that they are messy, complicated blobs of gas with complex structures. These complex structures are caused by the fact that the gas is moving around turbulently at speeds of several kilometers per second. To make matters even more complicated the clouds are also magnetized, and the motion of the gas is controlled by a combination of gravity, gas pressure, and magnetic forces.

[Slides 3 and 4 – the Pipe nebula and the Perseus Cloud]

We know that stars form inside these molecular clouds because we can see them if we look in the right way. The dusty gas is opaque in the optical, but dust absorbs infrared light less than optical light. As a result, if we look in infrared we can see through the dust. This is only possible from space, since the Earth's atmosphere is both opaque and blindingly bright in the infrared, but the Spitzer telescope makes it possible. Infrared images of these dark clouds reveals that they are filled with young stars.

[Slide 5 – the W5 region in optical and IR]

B. Jeans instability

So why do stars form in these cold, dense clouds, and seemingly only in them? The basic answer is Jeans instability, a phenomenon first identified by Sir James Jeans in 1902. The Jeans instability can be analyzed in many ways, but we will do so with the aid of the virial theorem.

Consider a uniform gas cloud of mass M and radius R. The density is $\rho = 3M/(4\pi R^3)$, and the density and temperature are very low, so the gas is non-degenerate and non-relativistic. Therefore its pressure is given by the ideal gas lows:

$$P = \frac{\mathcal{R}}{\mu}\rho T.$$

A subtle point here is that the value of μ for an interstellar cloud is different than it is for a star, because in a star the gas is fully ionized, while in a molecular cloud it is neutral, and the hydrogen is all in the form of H_2. This configuration has $\mu = 2.33$.

If this (spherically-symmetric) cloud is in hydrostatic equilibrium, it must satisfy the virial theorem, so the cloud must have

$$U_{\text{gas}} = -\frac{1}{2}\Omega.$$

The gravitational potential energy is

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

where α is our standard fudge factor of order unity that depends on the internal density structure. The internal energy is

$$U_{\text{gas}} = \frac{3}{2}\int \frac{P}{\rho}\, dm = \frac{3}{2}\frac{\mathcal{R}}{\mu}\int T\, dm = \frac{3}{2}\frac{\mathcal{R}}{\mu} M\overline{T},$$

where $\overline{T}$ is the mean temperature. This is something of an approximation as far as the coefficient of 3/2. In general a diatomic molecule like H_2 should have 5/2 instead. The only reason we keep 3/2 is due to an odd quantum mechanical effect: the levels of H_2 are quantized, and it turns out that the lowest lying ones are not excited at temperatures as low as 10 K. Thus the gas acts to first approximation like it is monatomic. In any event, the exact value of the coefficient is not essential to our argument.

Plugging Ω and U_{gas} into the virial theorem, we obtain

$$\alpha \frac{GM^2}{2R} = \frac{3\mathcal{R}}{2\mu} M\overline{T}.$$

It is convenient to rewrite this using density instead of radius as the variable, so we substitute in $R = (3M/4\pi\rho)^{1/3}$. With this substitution and some rearrangement, the virial theorem implies that

$$M = \frac{9}{2\sqrt{\pi\alpha^3}} \left(\frac{\mathcal{R}}{\mu G}\right)^{3/2} \sqrt{\frac{\overline{T}^3}{\rho}}.$$

Thus far we have a result that looks very much like the one we derived for stars: there is a relationship between the mass, the mean temperature, and the radius or density. In fact, this equation applies equally well to stars and interstellar gas clouds. The trick comes in realizing that stars and cloud respond very differently if you perturb them.

Consider compressing a star or cloud, so that ρ increases slightly. The mass is fixed, so in a star the gas responds by heating up a little – $\overline{T}$ rises so that the term on the right hand side remains constant. A gas cloud tries to do the same thing, but it encounters a big problem: the molecules out of which it is made are very, very good at radiating energy, and they have a particular, low temperature they want to be. Unlike a star, where it takes the energy a long time to get out because the gas is very opaque, gas clouds are transparent to the radio waves emitted by the molecules. Thus the cloud heats up slightly, but rapid radiation by molecules forces its temperature back down to their preferred equilibrium temperature immediately. One way of putting this is that for a star, $t_{\rm dyn} \ll t_{\rm KH}$, but for a molecular cloud exactly the opposite is true: $t_{\rm dyn} \gg t_{\rm KH}$.

This spells doom for the cloud, because now it cannot satisfy the virial theorem, and thus it cannot be in hydrostatic equilibrium. Instead, $U_{\rm gas} \propto P/\rho$ is too small compared to Ω. This means that the force of gravity compressing the cloud is stronger than the pressure force trying to hold it up. The cloud therefore collapses some. This further increases Ω, while leaving $U_{\rm gas}$ fixed because the molecules stubbornly keep $\overline{T}$ the same. The cloud thus falls even further out of balance, and goes into a runway collapse. This is the Jeans instability. The process ends only when the gas forms an opaque structure for which $t_{\rm dyn} < t_{\rm KH}$ – that is a newborn star.

As a result of this phenomenon, given the temperature at which the molecules like to remain, one can define a maximum mass cloud that can avoid collapsing due to Jeans instability. This is known as the Bonnor-Ebert mass, and its value is

$$M_{\rm BE} = 1.18 \left(\frac{\mathcal{R}}{\mu G}\right)^{3/2} \sqrt{\frac{\overline{T}^3}{\rho}} = \frac{4.03 \times 10^{34}}{\mu^2} \sqrt{\frac{\overline{T}^3}{n}} \text{ cgs units.}$$

The factor of 1.18 comes from self-consistently solving for the structure of the cloud, thereby determining the coefficient α. We also used $\rho = \mu m_{\rm H} n$. An important property of $M_{\rm BE}$ is that it is smallest in clouds with low temperature and high densities. In other words, regions that are dense and cold, like molecular clouds, have very small maximum masses that can be supported, while warmer, more diffuse regions have much larger masses.

Let's put some numbers on this. First think about a region of atomic gas. These typically have number densities of $n \sim 1$ cm^{-3}, $\mu = 1.67$ (because the gas is not ionized), and temperatures of $T = 8000$ K. Plugging in this numbers, we get a maximum mass $M = 5 \times 10^6\ M_\odot$ – in other words, huge clouds can be held up by pressure. On the other hand, let's try this for the interior of a molecular cloud, where the number density can be $n = 10^3$ cm^{-3} and the temperature $T = 10$ K. These numbers give $M = 4\ M_\odot$.

This leads to two conclusions. First, it explains why stars form in molecular clouds: they are much, much too massive to be stable against self-gravity given their temperatures and densities. They have no choice but to collapse, whereas lower density, warmer atomic regions won't. Second, the characteristic mass scale set by this instability in the densest regions where stars form suggests an explanation for why the typical star is comparable to the Sun in mass, and not a million times more or less massive. The mass of the Sun is about the characteristic mass at which things are prone to going into collapse because they can no longer support themselves!

C. Cores

Let's consider one of these collapsing blobs of gas and try to understand what will happen to it. Ordinarily these things come in clusters, but occasionally we can see one in isolation, and the most spectacular example is probably the object known as B68.

[Slide 6 – the core B68]

This particular object was first seen by William Herschel (the discoverer of Uranus) in the 1700s. When he saw it, in optical of course, he remarked "My God, there is a hole in the skies!" He attributed this to the inevitable decay of the cosmos caused by the Fall, and thought that it was a place where the stars had burned out. Today of course we know that this blob of gas is in fact the genesis of a new star, and that the only reason it appears dark is because the dust mixed with the gas is blocking out the background light.

We refer to objects like B68, which have masses of $\sim M_\odot$ and radii of ~ 0.1 pc, as cores. In the case of B68, we don't see a star in the center when we look in infrared, which indicates that this core has not yet collapsed to form a star at its center. However, we can work out how objects like this collapse.

Suppose at first that we neglect pressure support, and ask how long it will take before the gas at the edge of an unstable core collapses into the center. Consider a spherical core in which the mass interior to a radius r is m, and consider the shell of material of mass dm that starts at rest at radius r_0.

Since the mass interior to r_0 is m, the initial gravitational potential energy of the shell is

$$E_{\mathrm{g},0} = -\frac{Gm\,dm}{r_0}.$$

If we come back and look some time later, when the shell has fallen inward to radius r, its new potential energy is

$$E_\mathrm{g} = -\frac{Gm\,dm}{r}.$$

The kinetic energy of the shell is

$$E_\mathrm{k} = \frac{1}{2}\,dm\left(\frac{dr}{dt}\right)^2$$

Since no work is being done on the shell other than by gravity (since we have neglected pressure forces), conservation of energy requires

$$\begin{aligned} E_\mathrm{g} + E_\mathrm{k} &= E_{\mathrm{g},0} \\ -\frac{Gm\,dm}{r} + \frac{1}{2}\,dm\left(\frac{dr}{dt}\right)^2 &= -\frac{Gm\,dm}{r_0} \\ \frac{dr}{dt} &= -\sqrt{\frac{2Gm}{r_0}}\left(\frac{r_0}{r} - 1\right)^{1/2} \end{aligned}$$

To figure out when a given shell reaches the protostar at $r = 0$, we integrate from the time $t' = 0$ when the shell is at r_0 to the time $t' = t$ when it reaches protostar at the center of the core:

$$-\int_0^t \sqrt{\frac{2Gm}{r_0}}\, dt' = \int_{r_0}^{r} \left(\frac{r_0}{r'} - 1\right)^{-1/2} dr'$$

The integral on the LHS is trivial, since $\sqrt{2Gm}$ doesn't change with time, and the integral on the RHS can be done via the trigonometric substitution $r' = r_0 \cos^2 \xi$:

$$\begin{aligned}
-\sqrt{\frac{2Gm}{r_0}} t &= \int_{r_0}^{r} \left(\frac{r_0}{r'} - 1\right)^{-1/2} dr' \\
&= -2r_0 \int_0^{\pi/2} \left(\frac{1}{\cos^{-2}\xi - 1}\right)^{1/2} \cos\xi \sin\xi \, d\xi \\
&= -2r_0 \int_0^{\pi/2} \cos^2 \xi \, d\xi \\
&= -r_0 \left(\xi + \frac{1}{2}\sin 2\xi\right)\Big|_0^{\pi/2} \\
&= -r_0 \frac{\pi}{2}
\end{aligned}$$

Solving, we find that the time when a shell reaches the star is

$$t = \frac{\pi}{2}\sqrt{\frac{r_0^3}{2Gm}}$$

If the mean density interior to r_0 is ρ, then $m = (4/3)\pi r_0^3 \rho$, and we get

$$t = \sqrt{\frac{3\pi}{32 G\rho}} \equiv t_{\rm ff},$$

which defines the free-fall time $t_{\rm ff}$, the time required for an object to collapse when it is affected only by its own gravity. Note that $t_{\rm ff}$ is just the dynamical time multiplied by a constant of order unity.

For cores that form stars, typical densities are $n = 10^5$ cm^{-3} and mean molecular masses are $\mu = 2.33$; so $\rho = \mu m_{\rm H} n = 4 \times 10^{-19}$ g cm^{-3}. Plugging this in gives $t_{\rm ff} = 10^5$ yr – this is how long it would take a core to collapse if it were affected only by gravity. Of course there really is some pressure which opposes the collapse, and a more thorough analysis that includes the pressure shows that it increases the collapse time by a factor of a few. Nonetheless, what this shows is that, once a core forms, in a few hundred thousand years it must undergo collapse.

II. Protostars

Now that we have understood something about how the star formation process begins, let us turn our attention to the objects that are created by it: protostars. As we have already mentioned, a protostar first appears when the gas gets sufficiently dense that it becomes opaque, and $t_{\rm KH}$ becomes longer than $t_{\rm dyn}$. At this point it becomes possible to satisfy the virial theorem, and a hydrostatic object forms, accumulating the gas that rains down on it from its parent collapsing core.

A. Accretion Luminosity

Protostars in this configuration can be extremely bright – not because they shine from nuclear fusion like main sequence stars, but because of the material raining down onto their surfaces. To get a sense of how this works, a protostar of mass M and radius R, accreting at a rate $\dot{M}$.

The material falling onto the star started out a long distance away, which we can approximate as being infinitely far away. Its energy when it starts is zero, and conservation of energy dictates that, right before it hits the stellar surface, its kinetic and potential energy add up to zero. If we consider a blob of infalling material of mass dm, this means that

$$0 = \Omega + K = -\frac{GM}{R}\,dm + \frac{1}{2}\,dm\,v^2.$$

Thus its velocity right before it hits the stellar surface is

$$v_{\rm ff} = \sqrt{\frac{2GM}{R}},$$

which is called the free-fall velocity. For $M = M_\odot$ and $R = R_\odot$, $v_{\rm ff} = 620$ km s^{-1} – the gas is moving fast!

When the gas hits the stellar surface, it comes to a stop, and its kinetic energy drops to zero. This energy must then go into other forms. Some of it goes into internal energy: the gas heats up and its chemical state changes from molecular to ionized. The rest goes into radiation that escapes from the star, and which we can observe.

We can fairly easily establish that the fraction of the energy that goes into dissociating the molecules and then ionizing the atoms can't be very significant. Dissociating a hydrogen molecule requires 4.5 eV, and ionizing a hydrogen atom requires 13.6 eV, so for each hydrogen atom that falls onto the star,

$$\chi = 13.6 \text{ eV} + \frac{4.5 \text{ eV}}{2} = 15.9 \text{ eV}$$

go into dissociating and ionizing it. In contrast, the atom arrives at 620 km s^{-1}, so its kinetic energy is

$$K = \frac{1}{2} m_{\rm H} v_{\rm ff}^2 = 2.0 \text{ keV}.$$

Thus, less than 1% of the energy is used up in dissociating and ionizing the gas.

The rest goes into heat and radiation. Figuring out exactly how much goes into each is a complicated problem that wasn't really solved until the 1980s and 1990s, but the answer turns out to be that it is about half and half. Thus, to good approximation, half the kinetic energy of the infalling gas comes out as radiation.

To see what this implies about the luminosity, consider that, for an accretion rate $\dot{M}$, an amount of mass $dm = \dot{M}\,dt$ must arrive over a time dt. In this amount of time, the amount of energy radiated is

$$dE = \frac{1}{2}\left(\frac{1}{2}v_{\rm ff}^2\,dm\right) = \frac{GM}{2R}\,dm,$$

where we have assumed that exactly half the energy comes out as radiation, and we have neglected the 1% correction due to energy lost to ionization and dissociation. To get the luminosity, we divide both sides by the time dt over which the energy is emitted, which gives

$$L = \frac{dE}{dt} = \frac{GM\dot{M}}{2R}.$$

On your homework you will use this result to do a somewhat more sophisticated calculation of what sort of luminosity something like the proto-Sun should put out, but we can make a simple estimate now. Recall that we said that the collapse of a protostellar core takes a few hundred thousand years. To accumulate the mass of the Sun in this time, the accretion rate must be roughly $\dot{M} \sim 10^{-5}$ $M_\odot$ yr^{-1}. Plugging this in, along with $M = M_\odot$ and $R = 2R_\odot$ (since the radius of a protostar is generally bigger than that of a pre-main sequence star), gives $L = 100$ $L_\odot$. Thus a proto-Sun would be roughly 100 times as bright as the same star on the main sequence.

B. Hayashi Contraction

In addition to the radiation emitted by infalling material as it strikes the stellar surface, the star itself also radiates. However, since the protostar is initially not hot enough to burn hydrogen, it has no internal source of nuclear energy to balance out this radiation, and it is forced to contract on a Kelvin-Helmholtz timescale. (It can burn deuterium, but this all gets used up on a timescale well under the KH timescale.)

This contracting state represents the "initial condition" for a calculation of stellar evolution. In terms of the $(\log T, \log \rho)$ plane describing the center of the star, we already know what this configuration looks like: the star lies somewhere on the low T, low ρ side of its mass track, and it moves toward the hydrogen burning line on a KH timescale. We would also like to know what it looks like on the HR diagram, since this is what we can actually observe. Therefore we want to understand the movement of the star in the $(\log T_{\rm eff}, \log L)$ plane.

To figure this out, we can approximate the protostellar interior as a polytrope with

$$P = K_P \rho^{(n+1)/n},$$

or

$$\log P = \log K_P + \left(\frac{n+1}{n}\right)\log\rho.$$

Recalling way back to the discussion of polytropes, the polytropic constant K_P is related to the mass and radius of the star by

$$K_P \propto M^{(n-1)/n}R^{(3-n)/n} \Longrightarrow \log K_P = \left(\frac{n-1}{n}\right)\log M + \left(\frac{3-n}{n}\right)\log R + \text{constant},$$

so we have

$$\log P = \left(\frac{n-1}{n}\right)\log M + \left(\frac{3-n}{n}\right)\log R + \left(\frac{n+1}{n}\right)\log\rho + \text{constant}.$$

Now consider the photosphere of the star, at radius R, where it radiates away its energy into space. If the density at the photosphere is ρ_R, then hydrostatic balance requires that

$$\frac{dP}{dr} = -\rho_R\frac{GM}{R^2} \qquad \Longrightarrow \qquad P_R = \frac{GM}{R^2}\int_R^\infty \rho\, dr,$$

where P_R is the pressure at the photosphere and we have assumed that GM/R^2 is constant across the photosphere, which is a reasonable approximation since the photosphere is a very thin layer. The photosphere is the place where the optical depth τ drops to a value below ~ 1. Thus we know that at the photosphere

$$\kappa\int_R^\infty \rho\, dr \approx 1,$$

where we are also approximating that κ is constant at the photosphere. Putting this together, we have

$$P_R \approx \frac{GM}{R^2\kappa} \qquad \Longrightarrow \qquad \log P_R = \log M - 2\log R - \log\kappa + \text{constant}.$$

For simplicity we will approximate κ as a powerlaw of the form $\kappa = \kappa_0\rho_R T_{\text{eff}}^b$, where T_{eff} is the star's effective temperature, i.e. the temperature at its photosphere. Free-free opacity is $b = -3.5$. Plugging this approximation in gives

$$\log P_R = \log M - 2\log R - \log\rho_R - b\log T_{\text{eff}} + \text{constant}.$$

Finally, we know that the ideal gas law applies at the stellar photosphere, so we have

$$\log P_R = \log\rho_R + \log T_{\text{eff}} + \text{constant},$$

and we have the standard relationship between luminosity and temperature

$$\log L = 4\log T_{\text{eff}} + 2\log R + \text{constant}$$

We now have four equations

$$\begin{aligned}
\log P_R &= \left(\frac{n-1}{n}\right)\log M + \left(\frac{3-n}{n}\right)\log R + \left(\frac{n+1}{n}\right)\log \rho_R + \text{constant} \\
\log P_R &= \log M - 2\log R - \log \rho_R - b\log T_{\text{eff}} + \text{constant} \\
\log P_R &= \log \rho_R + \log T_{\text{eff}} + \text{constant} \\
\log L &= 4\log T_{\text{eff}} + 2\log R + \text{constant}.
\end{aligned}$$

in the four unknowns $\log T_{\text{eff}}$, $\log L$, $\log \rho_R$, and $\log P_R$. Solving these equations (and skipping over the tedious algebra), we obtain

$$\log L = \left(\frac{9-2n+b}{2-n}\right)\log T_{\text{eff}} - \left(\frac{2n-1}{2-n}\right)\log M + \text{constant}.$$

Thus to figure out the slope of a young star's track in the HR diagram, we need only specify n and b. Many young stars are fully convective due to their high opacities, so $n = 1.5$ is usually a good approximation, so that just leaves b. For free-free opacity $b = -3.5$, but we must recall that a young star is initially quite cold, about 4000 K. This makes its opacity very different from that of main sequence stars. In main sequence stars, the opacity is mostly free-free or, at high temperatures, electron scattering. At the low temperatures of protostars, however, there are too few free electrons for either of this to be significant, and instead the main opacity source is bound-bound. One species in particular dominates: H^-, that is hydrogen with two electrons rather than one.

The H^- opacity is very different than the opacities we're used to, in that it strongly *increases* rather than decreases, with temperature. That is because higher temperatures produce more free electrons via the ionization of metal atoms with low ionization potentials, which in turn can combine with hydrogen to make more H^-. Once the temperature passes several thousand K, H^- ions start falling apart and the opacity decreases again, but in the crucial temperature regime where protostars find themselves, opacity increases extremely strongly with temperature: $\kappa_{H^-} \propto \rho T^4$ is a reasonable approximation, giving $b = 4$.

Plugging in $n = 1.5$ and $b = 4$, we get

$$\log L = 20\log T_{\text{eff}} - 4\log M + \text{constant}.$$

Thus the slope is 20, extremely large. Stars in this phase of contraction therefore make a nearly vertical track in the HR diagram. This is called the Hayashi track. Stars of different masses have Hayashi tracks that are slightly offset from one another due to the $4\log M$ term, but they are all vertical.

Contraction along the Hayashi track ends once the star contracts and heats up enough for H^- opacity not to dominate, so that b is no longer a large positive number. Once b becomes 0 or smaller, as the opacity changes over to other sources, the track flattens, and the star contracts toward the main sequence at roughly fixed luminosity but increasing temperature. This is known as a Heyney track.

[Slide 7 – protostellar models from Palla & Stahler, showing Hayashi and Heyney tracks]

Only stars with masses $\sim M_\odot$ or less have Hayashi phases. More massive stars are "born" hot enough so that they are already too warm to be dominated by H^-.

Class 16 Notes: Post-Main Sequence Evolution of Low Mass Stars

Our topic for today is the first of three classes about post-main sequence evolution. Today the topic is low mass stars.

I. Leaving the Main Sequence

A. Main Sequence Lifetime

Stars remain on the main sequence as long as their hydrogen fuel lasts. While on the main sequence, their properties do not change much, but they do change some due to the gradual conversion of H into He. As a result of this conversion, the hydrogen mass fraction X decreases, while the helium mass fraction Y increases. As a result, the mean atomic weight changes. Recall that

$$\begin{aligned} \frac{1}{\mu_I} &\approx X + \frac{1}{4}Y + \frac{1-X-Y}{\langle \mathcal{A} \rangle} \\ \frac{1}{\mu_e} &\approx \frac{1}{2}(1+X) \\ \frac{1}{\mu} &= \frac{1}{\mu_I} + \frac{1}{\mu_e}. \end{aligned}$$

Insterstellar gas out of which stars form has roughly $X = 0.74$ and $Y = 0.24$ (in contrast to $X = 0.707$ and $Y = 0.274$ in the Sun, which has processed some of its H into He), which gives $\mu = 0.61$. In contrast, once all the H has been turned into He, $X = 0$ and $Y = 0.98$, which gives $\mu = 1.34$.

In stars like the Sun that are radiative in their cores, the changes occur shell by shell, so different shells have different compositions depending on their rate of burning. In stars that are convective in their cores, convection homogenizes the composition of the different shells, so the entire convective zone has a uniform composition.

As we mentioned briefly a couple classes ago, this difference between convective and non-convective stars affects how long it takes stars to leave the main sequence. Stars with convection in their cores do not leave the main sequence until they have converted all the mass in the convective region to He. As the convective zone fills more and more of the star, the main sequence lifetime therefore approaches the naively computed nuclear timescale $t_{\mathrm{nuc}} = \epsilon M c^2/L$.

In contrast, stars with radiative cores, like the Sun, leave the main sequence once the material in the very center where nuclear burning occurs is converted to He. This makes their lifetimes shorter than t_{nuc}, with the minimum of $t_{\mathrm{ms}}/t_{\mathrm{nuc}}$

occurring near 1 $M_\odot$, since that is where stellar cores are least convective. This general expectation agrees quite well with numerical results.

[Slide 1 – main sequence and nuclear lifetimes]

A general complication to this story is mass loss, which, for massive stars, can be significant even while they are on the main sequence. The mass loss mechanism is only generally understood. All stars have winds of gas leaving their surfaces, and these winds become more intense for more massive stars. These numerical results include a very approximate treatment of mass loss, but on the main sequence it is only significant for stars bigger than several tens of $M_\odot$.

B. Luminosity Evolution

Regardless of convection, the increase in μ results in an increase in luminosity. One can estimate this effect roughly using an Eddington model. The Eddington quartic is

$$0.003 \left(\frac{M}{M_\odot}\right)^2 \mu^4 \beta^4 = 1 - \beta = \frac{L}{L_{Edd}}$$

and so the luminosity of a star in the Eddington model is

$$L = 0.003 \frac{4\pi c G M_\odot}{\kappa_s} \mu^4 \beta^4 \left(\frac{M}{M_\odot}\right)^3.$$

Thus the luminosity at fixed mass is proportional to $(\mu\beta)^4$.

For a low mass star, the first term in the Eddington quartic is negligible, so $\beta \approx 1$ independent of μ, and thus $L \propto \mu^4$. For very massive stars the first term in the Eddington quartic dominates, which means that $\mu\beta \approx$ constant, so L stays constant. However, this will apply only to very, very massive stars. Thus in general we expect L to increase with μ, with the largest increases at low masses and smaller increases at high masses.

If an entire star were converted from H to He, this would suggest that its luminosity should go up by a factor of $(1.34/0.6)^4 = 25$ at low masses. Of course the entire star isn't converted into He except in fully convective stars that are uniform throughout, and stars are fully convective only below roughly 0.3 $M_\odot$. These stars have main sequence lifetimes larger than the age of the universe, so none have ever fully converted into He. In more massive stars that have reached the end of the main sequence, μ increases to 1.34 in their cores, but not elsewhere, so the mean value of μ and the luminosity increase by a smaller amount.

This simple understanding is also in good agreement with the results of numerical calculations. What is a bit less easy to understand analytically, but also happens, is that stars radii swell, reducing their effective temperatures. The swelling is greatest for the most massive stars, so, although they do not move very far in L, they move a considerable distance in T_{eff}.

[Slide 2 – HR diagram tracks]

II. The Red Giant Phase

A. The Schönberg-Chandrasekhar Limit

As stars reach the end of their main sequence lives, they accumulate a core of helium that is inert, in the sense that it is too cold for 3α reactions to take place, so the core generates no energy. The consequences of this become clear if we examine the two stellar structure equations that describe energy generation and transport:

$$\begin{aligned} \frac{dF}{dm} &= q \\ \frac{dT}{dm} &= -\frac{3}{4ac}\frac{\kappa}{T^3}\frac{F}{(4\pi r^2)^2}. \end{aligned}$$

Strictly we should write down the possibility for convective as well as radiative transport in the second equation, but we will see in a moment that is not necessary.

If there is no nuclear energy generation in the He core, then $q = 0$, which means $dF/dm = 0$ in the core. Thus the heat flow through the core must be constant, and, since there is no heat flow emerging from $m = 0$, this means that the core must have $F = 0$. It immediately follows that $dT/dm = 0$ in the core as well – that is, the core is isothermal. This is why we do not need to worry about convection: since $dT/dr = 0$, the temperature gradient is definitely sub-adiabatic. Thus if there was any convection going on in the core, it shuts off once the nuclear reactions stop due to lack of fuel.

The star as a whole is not necessarily pushed out of thermal equilibrium by this process because nuclear burning can continue in the material above the core that still has hydrogen in it. This can be enough to power the star. However, as this material depletes its hydrogen, it too becomes inert, adding to the mass of the helium core. Thus the core grows to be a larger and larger fraction of the star as time passes.

We can show that this configuration of a growing isothermal core cannot continue indefinitely, and, indeed, must end well before the entire star is converted to He. This point was first realized by Schönberg and Chandrasekhar in 1942, and in their honor is known as the Schönberg-Chandrasekhar limit. There are several ways to demonstrate the result, but the most straightforward is using the virial theorem.

We will apply the virial theorem to the isothermal core. It requires that

$$P_s V_c - \int_0^{M_c} \frac{P}{\rho}\, dm = \frac{1}{3}\Omega_c,$$

where V_c is the volume of the core, M_c is its mass, and Ω_c is its binding energy. The term P_s is the pressure at the surface of the core, and it is non-zero. This is somewhat different than when applying the virial theorem to the star as a whole:

normally when we do so, we drop the surface term on the grounds that the surface pressure of a star is zero. In this case, however, the core is buried deep inside the star, so we cannot assume that the pressure on its surface is zero.

Evaluating both the integral and the term on the right hand side is easy. For the term on the right hand side, we will just use our standard approximation $\Omega_c = -\alpha G M_c^2/R_c$, where α is a constant of order unity that depends on the core's internal structure. For the integral, because the core is isothermal and of uniform composition, with a temperature T_c and a mean atomic mass $\mu_c = 1.34$, appropriate for pure helium. Assuming the core is non-degenerate (more on this in a bit), we have $P/\rho = (\mathcal{R}/\mu_c)T_c$, which is constant, so

$$P_s V_c - \frac{\mathcal{R}}{\mu_c} T_c M_c = -\frac{1}{3}\alpha \frac{G M_c^2}{R_c}.$$

Re-arranging this equation, we can get an expression for the surface pressure:

$$P_s = \frac{3}{4\pi}\frac{\mathcal{R}T_c}{\mu_c}\frac{M_c}{R_c^3} - \frac{\alpha G}{4\pi}\frac{M_c^2}{R_c^4},$$

where we have replaced the core volume with $V_c = 4\pi R_c^3/3$.

An interesting feature of this expression is that, for fixed M_c and T_c, the pressure P_s reaches a maximum at a particular value of R_c. We can find the maximum in the usual way, by differentiating P_s with respect to R_c and solving:

$$\begin{aligned} 0 = \frac{dP_s}{dR_c} &= -\frac{9}{4\pi}\frac{\mathcal{R}T_c}{\mu_c}\frac{M_c}{R_c^4} + \frac{\alpha G}{\pi}\frac{M_c^2}{R_c^5} \\ R_c &= \frac{4\alpha G}{9\mathcal{R}}\frac{M_c \mu_c}{T_c}. \end{aligned}$$

Plugging this in, the maximum pressure is

$$P_{s,\max} = \frac{3^7 \mathcal{R}^4}{2^{10}\pi\alpha^3 G^3}\frac{T_c^4}{\mu_c^4 M_c^2}.$$

The physical meaning of this maximum is as follows: if one has a core of fixed mass and temperature, and exerts a certain pressure on its surface, it will pick a radius such that it is in equilibrium with the applied surface pressure. At low surface pressure R_c is big. In such a configuration self-gravity, represented by the term $\alpha G M_c^2/(4\pi R_c^4)$ in the equation for P_s, is unimportant compared to internal thermal pressure, represented by the term $3\mathcal{R}T_c M_c/(4\pi\mu_c R_c^3)$. As the external pressure is increased, the radius shrinks, and the thermal pressure of the core goes up as R_c^{-3}.

However, if the pressure is increased enough, the self-gravity of the core is no longer unimportant. As self-gravity grows in importance, one has to decrease the radius more and more quickly to keep up with an increase in surface pressure, because more and more of the pressure of the core goes into holding itself up against self-gravity, rather than opposing the external pressure. Eventually one reaches a critical radius where the core is exerting as much pressure on its surface as it can. Any further increase in the external pressure shrinks it further, and self-gravity gets stronger faster than the internal pressure grows.

We can estimate the pressure exerted on the surface of the helium core by the rest of the star (the envelope). To calculate this, we note that the envelope must obey the equation of hydrostatic equilibrium, and that we can integrate this from the surface of the isothermal core to the surface of the star:

$$\frac{dP}{dm} = -\frac{Gm}{4\pi r^4} \qquad \Longrightarrow \qquad \int_{P_s}^{0} dP = -P_s = -\int_{M_c}^{M} \frac{Gm}{4\pi r^4}\, dm.$$

Approximate this as

$$\frac{Gm}{4\pi r^4} \approx \frac{Gm}{4\pi R^4}.$$

Plugging this into the integral gives

$$P_s \approx \int_{M_c}^{M} \frac{Gm}{4\pi R^4}\, dm = \frac{G}{4\pi R^4} \int_{M_c}^{M} m\, dm = \frac{G}{8\pi R^4}\left(M^2 - M_c^2\right) \approx \frac{GM^2}{8\pi R^4}.$$

Combining this with our previous result gives a rough condition that the star must satisfy if it is to remain in hydrostatic equilibrium:

$$\frac{GM^2}{8\pi R^4} \approx \frac{3^7 \mathcal{R}^4}{2^{10}\pi\alpha^3 G^3} \frac{T_c^4}{\mu_c^4 M_c^2}$$

To see when this is likely to be violated, consider the gas just above the surface of the isothermal core. The temperature and pressure must change continuously across the core edge, so the envelope pressure and temperature there obey $T_{\rm env} = T_c$ and $P_{\rm env} = P_s$. Applying the ideal gas law to the envelope we have

$$T_{\rm env} = T_c = \frac{P_s \mu_{\rm env}}{\mathcal{R}\rho_{\rm env}},$$

where $\mu_{\rm env}$ and $\rho_{\rm env}$ are the mean molecular weight and density just above the envelope. The maximum temperature occurs when P_s is at its maximum value, and substituting in $P_s = P_{s,\rm max}$ gives

$$\begin{aligned} T_c &= \frac{\mu_{\rm env}}{\mathcal{R}\rho_{\rm env}} \left(\frac{3^7 \mathcal{R}^4}{2^{10}\pi\alpha^3 G^3} \frac{T_c^4}{\mu_c^4 M_c^2}\right) \\ T_c^3 &= \frac{2^{10}\pi\alpha^3 G^3}{3^7 \mathcal{R}^3} \frac{\mu_c^4 M_c^2 \rho_{\rm env}}{\mu_{\rm env}} \end{aligned}$$

As an extremely rough estimate we can also take $\rho_{\rm env} \sim 3M/(4\pi R^3)$, and plugging this in gives

$$T_c^3 \approx \frac{2^8 \alpha^3 G^3}{3^6 \mathcal{R}^3} \frac{\mu_c^4 M_c^2 M}{\mu_{\rm env} R^3}.$$

Thus we have now estimated T_c terms of the properties of the star. Plugging this into our condition for stability gives

$$\begin{aligned}\frac{GM^2}{8\pi R^4} &\gtrsim \frac{3^7\mathcal{R}^4}{2^{10}\pi\alpha^3G^3}\frac{1}{\mu_c^4M_c^2}\left(\frac{2^8\alpha^3G^3}{3^6\mathcal{R}^3}\frac{\mu_c^4M_c^2M}{\mu_{\rm env}R^3}\right)^{4/3} \\ \frac{M_c}{M} &\lesssim \sqrt{\frac{27}{2048\alpha^3}}\left(\frac{\mu_{\rm env}}{\mu_c}\right)^2\end{aligned}$$

Doing the analysis more carefully rather than using crude approximations, the coefficient turns out to be 0.37:

$$\frac{M_c}{M} \leq 0.37\left(\frac{\mu_{\rm env}}{\mu_c}\right)^2.$$

Since $\mu_{\rm env} < \mu_c$, this implies that the core can only reach some relatively small fraction of the star's total mass before hydrostatic equilibrium becomes impossible. Using $\mu_{\rm env} = 0.6$ and $\mu_c = 1.3$, the limit is that $M_c \lesssim 0.1M$. Once a star reaches this limit, the core must collapse.

This limit applies to stars that are bigger than about 2 $M_\odot$. For smaller stars, the gas in the He core becomes partially degenerate before the star reaches the Schönberg-Chandrasekhar limit. Since in a degenerate gas the pressure does not depend on the temperature, the pressure can exceed the result we got assuming isothermal gas. This allows the core to remain in hydrostatic equilibrium up to higher fractions of the star's mass.

B. The Sub-Giant and Red Giant Branches

Collapse of the core causes it to cease being isothermal, because it provides a new source of power: gravity. The collapse therefore allows hydrostatic equilibrium to be restored, but only at the price that the core shrinks on a Kelvin-Helmholtz timescale.

The core also heats up due to collapse, and this in turn heats up the gas around it where there is still hydrogen present. This accelerates the burning rate in the shell above the helium core. Moreover, it does so in an unstable way. The increase in temperature is driven by the KH contraction of the core, which is not sensitive to the rate of nuclear burning because none of the burning goes on in the collapsing core. Thus the burning rate will accelerate past the requirements of thermal equilibrium, and $L_{\rm nuc} > L$.

Consulting the virial theorem, we can understand what this implies must happen. Recall that we have shown several times that for stars with negligible radiation pressure support,

$$L_{\rm nuc} - L = \frac{dE}{dt} = \frac{1}{2}\frac{d\Omega}{dt} = -\frac{dU}{dt}.$$

Since $L_{\rm nuc} > L$, the left hand side is positive, and we conclude that Ω must increase and U must decrease. The potential energy $-\Omega \propto GM^2/R$, and the thermal energy $U \propto M\overline{T}$. Since the mass is fixed, the only way for Ω to increase is if R gets larger (since this brings Ω closer to zero), and the only way for U to decrease is for the mean temperature $\overline{T}$ to decrease.

Thus the unstable increase in nuclear burning causes the radius of the star to expand, while its mean temperature drops. In the HR diagram, this manifests as a drop in T_{eff}. As a result, the star moves to the right in the HR diagram. The phase is called the sub-giant branch.

[Slide 2 – HR diagram tracks]

In low mass stars the migration is slow, because the core is restrained from outright collapse by degeneracy pressure. In more massive stars the migration is rapid, since the core collapses on a KH timescale. For this reason we only see fairly low mass stars on the sub-giant branch. More massive stars cross it too rapidly for us to have any chance of finding one.

There is a limit to how red a star can get, which we encountered last time during our discussion of protostars: the Hayashi limit. As a post-main sequence star moves to the right in the HR diagram, it eventually bumps up against the ~ 4000 K limit imposed by H^- opacity. Since it can no longer deal with having $L_{\text{nuc}} > L$ by getting any colder at its surface, it instead has to increase its radius. This allows the internal temperature and the gravitational binding energy to drop, complying with energy conservation, and it also increases the luminosity, decreasing the difference between L_{nuc} and L. This phase of evolution is known as the red giant phase, and stars that are at low temperature and high and rising luminosity are called red giants.

[Slide 2 – HR diagram tracks]

Red giants also display an interesting phenomenon called dredge-up. The high opacity of the low-temperature envelope of the red giant guarantees that it will be convectively unstable, and the convective zone reaches all the way down to where the region where nuclear burning has taken place. It therefore drags up material that has been burned, changing the visible composition of the stellar surface. Nuclear burning destroys lithium (as some of you showed on your homework) and increases the abundance of C and N, and in red giants we can observe these altered compositions.

III. The Helium Burning Phase

We showed earlier that the temperature of the isothermal core of the star is given approximately by

$$T_c^3 \approx \frac{2^{10}\pi\alpha^3 G^3}{3^7\mathcal{R}^3}\frac{\mu_c^4 M_c^2 \rho_{\text{env}}}{\mu_{\text{env}}}.$$

As the star ascends the red giant branch, ρ_{env} is dropping, but at the same time M_c is rising as more and more mass is added to the core, and its 2nd power-dependence beats the first power dependence on the dropping ρ_{env}. Thus the core heats up with time. Once it violates the Schönberg-Chandresekhar limit, and it becomes powered by gravitational contraction, it heats up even more. Thus the core is always getting hotter during the red giant phase. What happens next depends on the mass of the star.

A. Stars $1.8 - 10\ M_\odot$

First consider fairly massive stars, which turn out to be those larger than 1.8 $M_\odot$. In such stars, the core temperature eventually reaches $\sim 10^8$ K, which is sufficient for He burning via the 3α process. At this point He burning provides a new source of energy in the core, which halts its contraction. Burning of hydrogen continues in the shell around the He core, but, since it is no longer being driven out of equilibrium by the contraction of the He core, it slows down. This allows the star to cease expanding and instead begin to contract, and the star's luminosity to decrease. The result is that the star comes back down from the red giant branch, and moves down and to the left on the HR diagram – higher effective temperature, lower luminosity.

[Slide 2 – HR diagram tracks]

After a short period the luminosity stabilizes, and since $L_{\mathrm{nuc}} < L$, the star responds by having its envelope contract. That contraction leaves the luminosity unchanged, but moves the star to higher effective temperature. The motion is roughly horizontal in the HR diagram, so this is known as the horizontal branch – it is shown by points 7-9 . The duration of this phase is roughly 10^8 yr, set by the amount of energy that is produced by a combination of He burning in the core and H burning in the shell. It ends when the core has been entirely transformed into C and O.

B. Stars $1 - 1.8\ M_\odot$

For stars from $1 - 1.8\ M_\odot$, the helium core becomes degenerate before it violates the Schönberg-Chandrasekhar limit. This does not stop it from heating up, but it does change what happens once the He ignites. Recall our discussion of runaway nuclear burning instability. In a degenerate gas, the pressure and density are not connected to the temperature. As a result, once a nuclear reaction starts it heats up the gas, but does not cause a corresponding expansion that pushes the temperature back down. This tends to cause the reaction rate to increase, leading to a runaway. This is exactly what happens in the He core of a low mass star. Once helium burning starts, it runs away, in a process called the helium flash.

The helium flash ends once the nuclear reactions generate enough energy to lift the degeneracy in the core, leading it to undergo rapid expansion. This only takes a few seconds. Thereafter, the envelope responds in a way that is essentially the opposite of what happens due to core collapse in the red giant phase: it contracts and heats up. The star therefore moves down off the red giant branch and across into the horizontal branch much like a more massive star, but it does so rapidly and violently, on a KH timescale rather than something like a nuclear timescale.

C. Stars Below 1 $M_\odot$

For an even smaller star, the core never heats up enough to reach He ignition, even once much of the core mass has been converted to He. In this case the remainder of the envelope is lost through processes that are not completely understood, and what is left is a degenerate helium core. This core then sits there and cools indefinitely. This is a helium white dwarf. Stars in this mass range therefore skip the AGB and PN phases we will discuss in a moment, and go directly to white dwarfs.

IV. The AGB and PN Phases

The He burning phase ends when the core has been completely converted to carbon and oxygen. At that point, what happens is essentially a repeat of the red giant phase. The core begins to contract, driving out-of-equilibrium He burning on its surface. This forces the envelope to expand, so the star moves back to the right, and lower temperature, on the HR diagram. Once the temperature drops to ~ 4000 K at the surface, the star is up against the Hayashi limit, and the envelope cannot cool any further. Instead, the star's radius expands, leading its luminosity to rise as well. The result is that the star climbs another giant branch, this one called the asymptotic giant branch, or AGB for short.

As in the red giant phase, the cool envelope becomes convective, and this convection drags up to the surface material that has been processed by nuclear burning. This is called second dredge-up, and it manifests in an increase in the helium and nitrogen abundances at the surface.

While the core is contracting and the envelope is expanding, the hydrogen burning shell goes out as its temperature drops. However, contraction of the core halts once it becomes supported by degeneracy pressure. At that point the hydrogen shell reignites, and this leads to a series of unstable thermal pulses. Thermal pulses work in the following cycle. As hydrogen burns, it produces helium, which sinks into a thin layer below the hydrogen burning shell. This layer has no source of energy, so it contracts and heats up. Once it gets hot enough, it ignites, and, as we showed a few weeks ago, nuclear burning in a thin shell is also unstable, because the shell can't expand fast enough to keep its temperature from rising. Thus all the accumulated He burns explosively, driving the core of the star to expand and cool, just like in the helium flash. This expansion also extinguishes the hydrogen burning. Once the He is gone, however, the cycle can resume again.

This chain of reactions and explosive burning has two other noteworthy effects. First, it temporarily produces neutron-rich environments, which synthesize elements heavier than iron via the s process. Second, it briefly churns up carbon from the core and convects it to the surface. The result is that carbon appears in significant quantities on the stellar surface, producing what is known as a carbon star. This process is called third dredge-up.

AGB stars also have significant stellar winds, which drive large amounts of mass loss from them. The details are not at all understood, but observationally we know that mass loss rates can reach $\sim 10^{-4}\ M_\odot\ \mathrm{yr}^{-1}$. The mechanism responsible for carrying the winds is likely radiation pressure, which is very significant in these stars due to their high luminosities. These winds carry lots of carbon with them, which condenses as the gas moves away from the stars and produces carbonaceous dust grains in interstellar space. The winds also reduce the total mass of the star significantly. As a result all stars with initial masses below roughly $8\ M_\odot$ end up with cores that are below the Chandrasekhar limit.

The winds eventually remove enough mass from the envelope that all nuclear burning there ceases, and the star finally goes out. However, the core remains very hot, and, once enough mass is removed, it is directly exposed and shines out the escaping gas. The high energy photons produced by the hot core surface are sufficient to ionize this gas, and the entire ejected shell of material lights up like a Christmas tree. This object is known as a planetary nebula. (Even though it has nothing to do with planets, the people who named it didn't know that at the time, and through a very low resolution telescope they look vaguely planetary.)

PN are some of the most visually spectacular objects in the sky, due to the variety of colors produced by the ionized gas, and the complex shapes whose origins we do not understand.

[Slides 3-5 – a gallery of PN]

V. White Dwarfs

The final state once the gas finishes escaping is a degenerate core of carbon and oxygen with a typical mass of $\sim 0.6\ M_\odot$. Lower mass stars that cannot ignite helium end up with masses of $\sim 0.2 - 0.4\ M_\odot$. We can understand the final evolution of these stars with a simple model. The center of the star consists of a degenerate electron gas. However, the pressure must go to zero at the stellar surface, so at some radius the pressure and density must begin to drop, and the gas ceases to be degenerate. Thus the star consists of a degenerate core containing most of the mass, and a non-degenerate envelope on top of it. Within the degenerate part, thermal conductivity is extremely high, so the gas is essentially isothermal – it turns out that a degenerate material acts much like a metal, and conducts very well.

In the non-degenerate part of the star, the standard equations of hydrostatic balance and radiative diffusion apply:

$$\begin{aligned} \frac{dP}{dr} &= -\rho \frac{GM}{r^2} \\ \frac{dT}{dr} &= -\frac{3}{4ac}\frac{\kappa\rho}{T^3}\frac{L}{4\pi r^2}. \end{aligned}$$

Note that we have M and not m in the numerator of the hydrostatic balance equation because we're approximating that all of the star's mass is in the inner, degenerate part. We also approximate that all the energy lost from the star comes from the inner, degenerate part, so $F = L =$ constant in the non-degenerate layer. Finally, note that the energy conservation equation $dF/dm = q$ does not apply, because we are not assuming that the star is in thermal equilibrium – indeed, it cannot be without a source of nuclear energy.

We assume that the opacity in the non-degenerate part of the star is a Kramer's opacity

$$\kappa = \kappa_0 \rho T^{-7/2} = \frac{\kappa_0 \mu}{\mathcal{R}} P T^{-9/2},$$

where we have used the ideal gas law to set $\rho = (\mu/\mathcal{R})(P/T)$. Substituting this into the radiative diffusion equation gives

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{1}{T^3}\left(\frac{\kappa_0\mu}{\mathcal{R}} P T^{-9/2}\right)\rho\frac{L}{4\pi r^2} = -\frac{3\kappa_0\mu}{16\pi ac\mathcal{R}}\frac{P\rho}{T^{15/2}}\frac{L}{r^2}.$$

If we now divide by the equation of hydrostatic balance, we obtain

$$\begin{aligned}\frac{dT}{dP} &= \frac{3\kappa_0\mu}{16\pi ac\mathcal{R}G}\frac{P}{T^{15/2}}\frac{L}{M} \\ P\,dP &= \frac{16\pi ac\mathcal{R}G}{3\kappa_0\mu}\frac{M}{L}T^{15/2}\,dT.\end{aligned}$$

We can integrate from the surface, where $P = 0$ and $T = 0$ to good approximation, inward, and obtain the relationship between pressure and temperature

$$\int_0^P P'\,dP' = \frac{16\pi ac\mathcal{R}G}{3\kappa_0\mu}\frac{M}{L}\int_0^T T'^{15/2}\,dT' \quad\Longrightarrow\quad P = \left(\frac{64\pi ac\mathcal{R}G}{51\kappa_0\mu}\right)^{1/2}\left(\frac{M}{L}\right)^{1/2}T^{17/4}.$$

Using the ideal gas law $\rho = (\mu/\mathcal{R})(P/T)$ again, we can turn this into

$$\rho = \left(\frac{64\pi ac\mu G}{51\kappa_0\mathcal{R}}\right)^{1/2}\left(\frac{M}{L}\right)^{1/2}T^{13/4}.$$

This relationship between density and temperature must hold everywhere in the ideal gas region, and so we can apply it at the boundary between that region and the degenerate region. The pressure in the non-degenerate region is just

$$P_{\text{nd}} = \frac{\mathcal{R}}{\mu_e}\rho T,$$

where we've used $\mu = \mu_e$ because the electron pressure completely dominates. Just on the other side of the boundary, in the degenerate region, the pressure is

$$P_{\text{d}} = K_1'\left(\frac{\rho}{\mu_e}\right)^{5/3}.$$

Pressure, density, and temperature must change continuously across the boundary, so the ρ that appears in these two expressions is the same. Moreover, since the core is isothermal, $T = T_c$, where T_c is the core temperature. Finally, since the pressures must match across the boundary, we have

$$\begin{aligned}
\frac{\mathcal{R}}{\mu_e}\rho T_c &= K_1'\left(\frac{\rho}{\mu_e}\right)^{5/3} \\
T &= \frac{K_1'}{\mathcal{R}\mu_e^{2/3}}\rho^{2/3} \\
&= \frac{K_1'}{\mathcal{R}\mu_e^{2/3}}\left[\left(\frac{64\pi ac\mu G}{51\kappa_0\mathcal{R}}\right)^{1/2}\left(\frac{M}{L}\right)^{1/2} T^{13/4}\right]^{2/3} \\
\frac{L}{M} &= \frac{64\pi acGK_1'^3\mu}{51\mathcal{R}^4\kappa_0\mu_e^2}T_c^{7/2}.
\end{aligned}$$

We have therefore derived the luminosity of a white dwarf in terms of the temperature of its degenerate core. Plugging in typical values gives

$$\frac{L/L_\odot}{M/M_\odot} \approx 6.8\times 10^{-3}\left(\frac{T_c}{10^7\ \mathrm{K}}\right)^{7/2}.$$

We can use this relation to infer how long white dwarfs will shine brightly enough for us to see them. The internal energy of the white dwarf is just the thermal energy of the gas. Since the electrons are degenerate they cannot lose energy – there are no lower energy states available for them to occupy. The ions, however, are not degenerate, and they can cool off. Since the ions are a non-degenerate ideal gas, their internal energy is

$$U_I = \frac{3}{2}\frac{\mathcal{R}}{\mu_I}MT_c,$$

and conservation of energy requires that

$$L = -\frac{dU_I}{dt} = -\frac{3}{2}\frac{\mathcal{R}}{\mu_I}M\frac{dT_c}{dt}.$$

It is convenient to recast this relation in terms of the luminosity. Using our temperature-luminosity relationship we have

$$\begin{aligned}
T_c &= \left(\frac{51\mathcal{R}^4\kappa_0\mu_e^2}{64\pi acGK_1'^3\mu}\frac{L}{M}\right)^{2/7} \\
\frac{dT_c}{dt} &= \frac{2}{7}\left(\frac{51\mathcal{R}^4\kappa_0\mu_e^2}{64\pi acGK_1'^3\mu}\frac{1}{M}\right)^{2/7} L^{-5/7}\frac{dL}{dt}
\end{aligned}$$

Plugging this into the equation for L gives

$$L = -\frac{3}{7}\frac{\mathcal{R}^{15/7}}{\mu_I}M^{5/7}\left(\frac{51\kappa_0\mu_e^2}{64\pi acGK_1'^3\mu}\right)^{2/7} L^{-5/7}\frac{dL}{dt}.$$

Separating the variables and integrating from an initial luminosity L_0 to a luminosity L at some later time, we have

$$
\begin{aligned}
\int_{L_0}^{L} L'^{-12/7}\, dL' &= -\frac{7}{3}\frac{\mu_I}{\mathcal{R}^{15/7}} M^{-5/7} \left(\frac{51\kappa_0\mu_e^2}{64\pi ac G K_1'^3 \mu}\right)^{-2/7} \int_0^t dt' \\
-\frac{7}{5}\left(L^{-5/7} - L_0^{-5/7}\right) &= -\frac{7}{3}\frac{\mu_I}{\mathcal{R}^{15/7}} M^{-5/7} \left(\frac{51\kappa_0\mu_e^2}{64\pi ac G K_1'^3 \mu}\right)^{-2/7} t \\
L &= L_0 \left[1 + \frac{5}{3}\frac{\mu_I}{\mathcal{R}^{15/7}} \left(\frac{L_0}{M}\right)^{5/7} \left(\frac{51\kappa_0\mu_e^2}{64\pi ac G K_1'^3 \mu}\right)^{-2/7} t\right]^{-7/5}
\end{aligned}
$$

For long times t, we can drop the $+1$, and we find that $L \propto t^{-7/5}$. Since the white dwarf birthrate in the galaxy is about constant, this immediately yields an important theoretical prediction. The number of white dwarfs we see with a given luminosity should be proportional to the amount of time they spend with that luminosity, which we have just shown varies as $t \propto L^{-5/7}$. Thus luminous white dwarfs should be rare because they cool quickly, while dimmer ones should be more common because they cool more slowly, and the ratio of the number of white dwarfs with luminosity L_1 to the number with luminosity L_2 should vary as $(L_1/L_2)^{-5/7}$. Observations confirm this result.

We can also define a characteristic cooling time t_{cool} as the time it takes a white dwarf's luminosity to change significantly. This is simply the time required for the second term in parentheses to become of order unity, which is

$$
t_{\text{cool}} \approx \frac{3\mathcal{R}^{15/7}}{5\mu_I} \left(\frac{51\kappa_0\mu_e^2}{64\pi ac G K_1'^3 \mu}\right)^{2/7} \left(\frac{M}{L_0}\right)^{5/7} \approx 2.5 \times 10^6 \left(\frac{M/M_\odot}{L/L_\odot}\right)^{5/7} \text{ yr.}
$$

Thus we conclude that white dwarfs with luminosities of $L \sim 10^4 L_\odot$, typical of the planetary nebula phase, should last only a few thousand years, while those with much lower luminosities $\sim L_\odot$ can remain at that brightness for of order a million years.

Astronomy 112: The Physics of Stars

Class 17 Notes: Core Collapse Supernovae

Our topic today is the post-main sequence evolution of massive stars, culminating in their deaths via supernova explosion. Supernovae of this type are called core collapse supernovae, to distinguish them from supernovae that occur due to accretion onto a white dwarf that pushes it above the Chandresekhar limit.

Before beginning, I will warn you that we are now entering into an area of active research where there are still very significant uncertainties. What I will tell you is the best of our understanding today, but significant parts of it may well turn out to be wrong. I will try to highlight the areas where what I say is least certain, and I will point out a couple of places where statements asserted quite confidently in your textbook have turned out to be incorrect.

I. Post-main sequence evolution

A. Mass Loss

One important effect that distinguishes the evolution of massive stars from that of lower mass stars is the importance of mass loss, both on the main sequence and thereafter. Low mass stars do not experience significant mass loss before the AGB phase, but massive stars, as we have already seen, can lose mass while still on the main sequence, and can lose even more after they leave it.

Like other aspects of stellar mass loss, the exact mechanisms are not understood. Very massive stars, those above 85 $M_\odot$ or so, lose mass in a rapid and unstable manner. We have already encountered one star like this: η Carinae. This is an example of a type of star called a luminous blue variable, or LBV.

[Slide 1 – η Car]

As these processes reduce the star's mass, its atmosphere becomes less and less dominated by hydrogen, eventually reaching $X \approx 0.1$ or even less. We see these stars are somewhat lower mass (but still very massive) stars whose atmospheres are dominated by helium rather than hydrogen. These are called Wolf-Rayet stars, and they are effectively the bare cores of massive stars. Stars from $10 - 85$ $M_\odot$ skip the LBV phase and go directly to the Wolf-Rayet phase.

Stars become WR's while they are still on the main sequence, i.e. burning hydrogen in their centers. Stars in this case are called WN stars, because they are Wolf-Rayet stars that show large amounts of nitrogen on their surfaces. The nitrogen is the product of CNO cycle burning, which produces an equilibrium level of nitrogen above the amount that the star began its life with.

WR stars continue to lose mass rapidly, often producing spectacular nebulae that look like planetary nebulae. They shine for the same reason: the expelled gas is exposed to the high energy radiation of the star, and it floresces in response.

[Slides 2, 3 – WR nebula NGC 2359, WR 124]

Mass loss continues after the star exhausts H and begins burning He – at this point the surface composition changes and we begin to see signs of 3α burning. These are WC stars. The continuing mass loss removes the enhanced nitrogen from the CNO cycle, and convection brings to the surface the result of 3α burning, which is mostly carbon. Very rarely, we see WR stars where the carbon is being blown off, and the surface is dominated by oxygen.

The mass loss can be quite dramatic – 100 $M_\odot$ stars are thought to get down to nearly 30 $M_\odot$ by the time they evolve off the main sequence.

B. Movement on the HR diagram

While these stars show dramatic mass loss, their luminosities do not evolve all that much as they age. That is for the reason we mentioned last time in the context of low mass stars' luminosity evolution: the role of radiation pressure. The luminosity varies as $L \propto \mu^4\beta^4$, and β is in turn given by the Eddington quartic:

$$0.003 \left(\frac{M}{M_\odot}\right)^2 \mu^4\beta^4 = 1 - \beta = \frac{L}{L_{Edd}}$$

For very massive stars, the first term is dominant, so $\mu\beta$ is roughly constant, and L is too. This is simply a reflection of the fact that very massive stars are largely supported by radiation pressure. As a result, their luminosity is equal to the Eddington luminosity, which depends only on total mass, not on composition.

[Slide 4 – Meynet & Maeder tracks]

This non-evolution of the luminosity continues to apply even after these stars leave the main sequence. As the stars develop inert ash cores and burning shells like lower mass stars, they cannot increase in luminosity, but they can increase in radius and go to lower effective temperature. The net effect is that they move along nearly horizontal tracks on the HR diagram. The slide shows the latest Geneva models.

As you can see, the luminosities increase less and less for stars of higher and higher masses, and instead they evolve at constant luminosity. Thus massive stars never have a red giant phase, since that would require an increase in luminosity.

C. Internal structure: the onion model

The internal structure of a massive star near the end of its lifetime comes to resemble an onion. In the center is an ash core, with the type of ash depending on the star's evolutionary state. At first it is helium, then carbon, etc., until at last the core is composed of iron. The temperature is high enough that the core is never degenerate until the last stages in the star's life, when it consists of iron.

Above the core is a burning shell where the next lowest $\mathcal{Z}$ element in the burning chain burns. Thus above an iron core is a silicon burning layer. As one moves farther outward in the star, one encounters the next burning shell where a lower $\mathcal{Z}$ element burns, and so forth until one reaches the hydrogen burning layer and the hydrogen envelope, if any is left, above it.

[Slides 5, 6 – schematic representation and numerical computation from Heger et al. of onion structure]

The onion structure grows until the star develops an iron core. Since iron is at the peak of the binding energy curve, it cannot be further burned. A star with an iron core and an onion structure around it is known as a supernova progenitor.

II. Supernovae

A. Evolution of the Core

Now consider what happens in the core of a supernova progenitor. The iron core is much like the helium core that we discussed in the context of lower mass stars: it has no nuclear reactions, so it becomes isothermal. If it gets to be more than roughly 10% of the stellar mass, it will exceed the Schönberg-Chandrasekhar limit and begin contracting dynamically. If it becomes degenerate, degeneracy pressure can slow the collapse, but if the core exceeds the Chandrasekhar mass of 1.4 $M_\odot$, electron degeneracy pressure cannot hold it up and then the core must contract.

Contraction creates two instabilities. First, at the high pressures found in the core, heavy nuclei can undergo reactions of the form

$$I(\mathcal{A},\mathcal{Z}) + e^- \rightarrow J(\mathcal{A},\mathcal{Z}-1) + \nu_e,$$

i.e. nucleus I captures a free electron, which converts one of its protons into a neutron. We'll discuss why these reactions happen in a few moments. Reactions of this sort create an instability because removing electrons reduces the number of electrons, and thus the degeneracy pressure. The loss of pressure accelerates collapse, raising the pressure again and driving the reaction to happen even faster.

Second, since the gas is degenerate, its pressure is unrelated to its temperature. As it collapses, its temperature rises, but this does not halt the collapse because it doesn't raise the pressure. Once the temperature exceeds about $6-7\times10^9$ K, photons are able to start photodisintegrating iron via the reaction

$$^{56}\mathrm{Fe} + 100\ \mathrm{MeV} \rightarrow 13\,{}^4\mathrm{He} + 4n.$$

As the equation indicates, the reaction is highly endothermic, absorbing about 100 MeV from the radiation field, or about 2 MeV per nucleon, each time it happens. In effect, all the energy that was released by burning from He to Fe is now given back. The loss of thermal energy also accelerates collapse, which leads the core to contract more, which accelerates the reaction, etc.

This is an ionization-like process, which serves to keep $\gamma_a < 4/3$, in the unstable regime where collapse cannot be halted. Once the collapse proceeds far enough, an even more endothermic reaction can take place when photons begin to disintegrate helium nuclei:

$$^4\mathrm{He} + 27\ \mathrm{MeV} \ \rightarrow \ 2p + 2n.$$

This reaction absorbs 27 MeV each time it happens, or $6 - 7$ MeV per nucleon.

The disintegration of He creates a population of free neutrons and protons. Normally free neutrons spontaneously decay into a proton plus an electron plus a neutrino:

$$n \ \rightarrow \ p + e^- + \overline{\nu}_e.$$

The free neutron lifetime is 614 seconds, and the reaction is exothermic (as it must be, since it is spontaneous). The energy released can be determined just by the difference in mass between a proton, $m_p = 1.67262 \times 10^{-24}$ g, and a neutron, $m_n = 1.67493 \times 10^{-24}$ g:

$$\Delta E = (m_n - m_p)c^2 = 1.3\ \mathrm{MeV}.$$

However, conditions in the core are very different from those in free space. The electrons are highly relativistically degenerate. Consider what this means energetically. Back at the beginning of the class, we showed that, for a population of degenerate electrons, they occupy all quantum states up to a maximum momentum

$$p_0 = \left(\frac{3h^3 n}{8\pi}\right)^{1/3},$$

where n is the number density of electrons. If a new electron were to be created by the decay of a neutron, it would have to go into an unoccupied quantum state, and the first available state has a momentum a just above p_0. The corresponding energy is

$$E_0 = pc = \left(\frac{3h^3 n}{8\pi}\right)^{1/3} c$$

in the limit where the electrons are highly relativistic. If we compare this to the energy ΔE that is released by neutron decay, we find that E_0 becomes equal to ΔE when the number density of electrons becomes

$$n = \frac{8\pi}{3}\left(\frac{\Delta E}{ch}\right)^3 = 9.6 \times 10^{30}\ \mathrm{cm}^{-3}$$

If we have one electron per two nucleons (i.e., $1/\mu_e = 1/2$), the average for elements heavier than hydrogen, the corresponding mass density is

$$\rho = n m_{\mathrm{H}} \mu_e = 3.2 \times 10^7\ \mathrm{g\ cm}^{-3}.$$

Once the density exceeds this value, it is no longer energetically possible for free neutrons to undergo spontaneous decay. Instead, the opposite is true, and the reverse reaction

$$p + e^- \rightarrow n + \nu_e$$

begins to occur spontaneously. Each such reaction requires 1.3 MeV of energy, and even further reduces the degeneracy pressure of the electrons. Electron capture by heavy elements earlier on in the collapse occurs for very similar reasons.

The collapse is only halted once another source of pressure becomes available: at sufficiently high density, the neutrons become degenerate. The structure of this degenerate neutron matter is not well understood, and is a subject of active research, but the bottom line of what we understand seems to be that the collapse is halted once the density reaches around 10^{15} g cm^{-3}. The radius of the core at this point is about 40 km, although as the neutron star cools off it eventually shrinks to about 10 km. The density of 10^{15} g cm^{-3} in the core is roughly the density of an atomic nucleus, so the core at this point is a giant atomic nucleus, several km is diameter, with the mass of the Sun.

B. Explosion Mechanism and Energy Budget

All of these processes occur in the core on dynamical timescales. The initial iron core is of order $5,000 - 10,000$ km in radius, and the mass is of order a Chandrasekhar mass, about 1.5 $M_\odot$, so the dynamical time is

$$t_{\rm dyn} \sim \frac{1}{\sqrt{G\rho}} \sim 1 \text{ second.}$$

Thus the core collapses on a timescale that is tiny compared to the dynamical time of the star as a whole – the outer envelope of the star just sits there while the core collapses.

The collapse of the iron core causes the material above it to begin falling, and the exact sequence of events thereafter is somewhat unclear. Your book gives the impression that this is a solved problem, but your book is wrong on this point. Exactly how supernovae work is far from clear. Nonetheless, we can give a rough outline and make some general statements.

First of all, we can figure out the energy budget. Ultimately what drives everything is the release of gravitational potential energy by the collapse of the iron core. It is this sudden energy release that explodes the star. The core has an initial mass of $M_c \approx 1.5\ M_\odot$, and an initial radius $R_c \approx 10^4$ km. The final neutron core has a comparable mass and a radius of $R_{nc} \approx 20$ km. Thus the amount of energy released is

$$\Delta E_{\rm grav} \approx -GM_c^2 \left(\frac{1}{R_c} - \frac{1}{R_{nc}}\right) \approx \frac{GM_c^2}{R_{nc}} \approx 3 \times 10^{53} \text{ erg.}$$

Of this, the amount that is used to convert the protons and electrons to neutrons is a small fraction. Each conversion (including the photodisintegration) ultimately uses up about 7 MeV, so the total nuclear energy absorption is

$$\Delta E_{\text{nuc}} = 7\ \text{MeV}\frac{M_c}{m_{\text{H}}} \approx 2 \times 10^{52}\ \text{erg} \approx \frac{\Delta E_{\text{grav}}}{15}.$$

Thus only $\sim$ 10% percent of the energy is used up in converting protons to neutrons. The rest is available to power an explosion.

Similarly, some of the energy is required to eject the stellar envelope. The binding energy of the envelope to the core is roughly

$$\Delta E_{\text{bind}} = \frac{GM_c(M - M_c)}{R_c} \approx 5 \times 10^{51}\ \text{erg} \approx \frac{\Delta E_{\text{grav}}}{60}.$$

Thus only a few percent of the available energy is required to unbind the envelope.

The remaining energy is available to give the envelope a large velocity, to produce radiation, and to drive nuclear reactions in the envelope. We don't have a good first-principles theory capable of telling us how this energy is divided up, but we can infer from observations.

The observed speed of the ejecta is around 10,000 km s^{-1}, so the energy required to power this is

$$\Delta E_{\text{kin}} = \frac{1}{2}(M - M_c)v^2 \approx 10^{51}\ \text{erg} \approx \frac{\Delta E_{\text{grav}}}{300}.$$

Finally, the observed amount that is released as light is comparable to that released in kinetic energy:

$$\Delta E_{\text{rad}} \approx 10^{51}\ \text{erg} \approx \frac{\Delta E_{\text{grav}}}{300}.$$

Both of these constitute only about 1% of the total power.

So where does the rest of the energy go? The answer is that it is radiated away too, but as neutrinos rather than photons. The neutrinos (produced when the protons in the core are converted into neutrons) don't escape immediately, but they do eventually escape, and they carry away the great majority of the energy with them.

Understanding the mechanism by which the energy released in the core is transferred into the envelope of the star is one of the major problems in astrophysics today. We have a general outline of what must happen, but really solving the problem is at the forefront of numerical simulation science.

Here's what we know: as long as the collapsing core has a pressure set by relativistic electrons, its adiabatic index is $\gamma_a = 4/3$. As it approaches nuclear density and more of the electrons and protons convert to neutrons, it initially experiences an attractive nuclear force that pulls it together, and this has the effect of pushing γ_a even lower, toward 1, and accelerating the collapse. Once the densities get even higher, though, the strong nuclear force becomes repulsive, and γ_a increases to a value $\gg 4/3$.

This is sufficient to halt collapse of the core, and from the perspective of the material falling on top of it, it is as if the core suddenly converted from pressureless foam ($\gamma_a < 4/3$) to hard rubber ($\gamma_a > 4/3$). The infall is therefore halted suddenly, and all the kinetic energy of the infalling material is converted to thermal energy. This thermal energy raises the pressure, which then causes the material above the neutron core to re-expand – it "bounces". The bounce launches a shock wave out into the envelope.

The bounce by itself does not appear to be sufficient to explode the star. The shock wave launched by the bounce stalls out before it reaches the stellar surface. However, at the same time all of this is going on, the core is radiating neutrinos like crazy. Every proton that is converted into a neutron leads to emission of a neutrino, and the collapsing star is sufficiently dense that the neutrinos cannot escape. Instead, they deposit their energy inside the star above the core, further heating the material there and raising its pressure.

The neutrinos are thought to somehow re-energize the explosion and allow it to finally break out of the star. However, there are lots of details missing.

C. Nucleosynthesis

The shock propagating outward through the star from the core heats the gas up to $\sim 5 \times 10^9$ K, and this is hot enough to induce nuclear burning in the envelope. This burning changes the chemical composition of the envelope, creating new elements. Much of the material is heated up enough that it burns to the iron peak, converting yet more of the star into iron-like elements.

I say iron-like because the initial product is not in fact iron. The reason is that the most bound element, ^{56}Fe, consists of 26 protons and 28 neutrons, so it has two more neutrons than protons. The fuel, consisting mostly of elements like ^{4}He, ^{12}C, ^{16}O, ^{28}Si, all have equal numbers of protons and neutrons. Thus there are not enough neutrons around to pair up with all the protons to make ^{56}Fe.

Converting protons to neutrons is via β decays is possible, and in fact it is the first step in the *pp*-chain. However, as we learned studying that reaction, β decays are slow, and in the few seconds that it takes for the shock to propagate through the star, there is not enough time for them to occur.

The net result is that the material burns to as close to the iron peak as it can get given the ratio of protons to neutrons available. This turns out to be ^{56}Ni. This is not a stable nucleus, since it is subject to β decay, but the timescale for decay is much longer than the supernova explosion goes on for, so no beta decays occur until long after the nucleosynthetic process is over.

Not all the material in the star is burned to the iron peak. As the shock wave propagates through the star it slows down and heats things up less. The net results is that material farther out in the star gets less burned, so the supernova winds up ejecting a large amounts of other elements as well. Calculating the exact yields from first principles is one of the goals of supernova models.

D. Observations

1. Light Curves

When a supernova goes off, what do we observe from the outside? The first thing, which was only seen for the first time a couple of years ago, is a bright ultraviolet flash from the shock breaking out of the stellar surface. We saw this because Alicia Soderberg, a postdoc at Princeton got very lucky. She was using an x-ray telescope to study an older supernova in a galaxy, when she saw another one go off. The telescope was observing the star as it exploded, and it saw a flash of x-rays as the shock wave from the deep interior of the star reached the surface.

[Slide 7 – Soderberg image]

After the initial flash in x-rays, it takes a little while before the optical emission reaches its peak brightness. That is because the expanding material initially has a small area, and most of that emission is at wavelengths shortward of visible. As the material expands and cools, its optical luminosity increases, and reaches its peak a few weeks after the explosion. After that it decays. The decay can initially take one of two forms, called linear or plateau, but after a while they all converge to the same slope of luminosity versus time.

[Slide 8 – light curve image]

This slope can be understood quite simply from nuclear physics. As we mentioned a moment ago, the supernova synthesizes large amounts of ^{56}Ni. This nickel is unstable, and it undergoes the β decay reaction

$$^{56}_{28}\mathrm{Ni} \;\rightarrow\; ^{56}_{27}\mathrm{Co} + e^+ + \nu_e + \gamma$$

with a half-life of 6.1 days. This is short enough that most of the nickel decays during the initial period of brightening or shortly thereafter.

However, the resulting ^{56}Co is also unstable, and it too undergoes a β decay reaction:

$$^{56}_{27}\mathrm{Co} \;\rightarrow\; ^{56}_{26}\mathrm{Fe} + e^+ + \nu_e + \gamma.$$

This reaction has a half-life of 77.7 days, and it turns out to be the dominant source of energy for the supernova in the period from a few tens to a few hundreds of days after peak. The expanding material is cooling off, and this would cause the luminosity to drop, but the radioactive decays provide a energy source that keeps the material hot and emitting.

By computing the rate of energy release as a function of time via the β decay of cobalt-56, we can figure out how the luminosity of the supernova should change as a function of time. Radioactive decays are a statistical process, in which during a given interval of time there is a fixed probability that each atom will decay. This implies that the number of cobalt-56 decays per unit time that occur in a particular supernova remnant must be proportional to the number of cobalt-56 atoms present:

$$\frac{dN}{dt} = -\lambda N.$$

Here N is the number of cobalt-56 atoms present and λ is a constant. The equation simply asserts that the rate of change of the number of cobalt-56 atoms at any given time is proportional to the number of atoms present at that time.

This equation is easy to integrate by separation of variables:

$$\frac{dN}{N} = -\lambda\, dt \qquad \Longrightarrow \qquad N = N_0 e^{-\lambda t},$$

where N_0 is the number of atoms present at time $t = 0$. The quantity λ is known as the decay rate. To see how it is related to the half-life $\tau_{1/2}$, we can just plug in $t = \tau_{1/2}$:

$$\frac{1}{2} N_0 = N_0 e^{-\lambda \tau_{1/2}} \qquad \Longrightarrow \qquad \lambda = \frac{\ln 2}{\tau_{1/2}}.$$

For ^{56}Co, $\lambda = 0.0089$ / day.

While radioactive decay is the dominant energy source, the luminosity is simply proportional to the rate of energy release by radioactive decay, which in turn is proportional to the number of atoms present at any time, i.e. $L \propto N$. This means that the instantaneous luminosity should follow

$$L \propto e^{-\lambda t} \qquad \Longrightarrow \qquad \log L = -(\log e)\lambda t + \text{constant}.$$

Thus for the cobalt-56-powered part of the decay, a plot of $\log L$ versus time should be a straight line with a slope of

$$-(\log e)\lambda = -0.004 \text{ day}^{-1}.$$

An excellent test for this model was provided by supernova 1987A, which went of in 1987 in the Large Magellanic Cloud, a nearby galaxy. The supernova was observed for more than five years after the explosion, and as a result we got a very good measure of how its luminosity dropped. We can see a clear period when the slope follows exactly what we have just calculated. Once enough of the ^{56}Co decayed, other radioactive decays with longer half-lives took over.

[Slide 9 – light curve of SN1987A]

The effect is even more prominent in type Ia supernovae, which are produced when a white dwarf is pushed over the Chandrasekhar limit. In that case the nuclear reaction burns a much larger fraction of the star to ^{56}Ni, so its decay into cobalt and then iron completely dominates the light curve.

2. Neutrinos

   Supernova 1987A also provided strong evidence for another basic idea in supernova theory: that supernovae involve the neutronization of large amounts of matter, and with it the production of copious neutrino emission. The first detection of supernova 1987A was *not* its light. The shock wave takes some time to propagate through the star after the core collapses. The neutrinos, however, escape promptly, and on February 23, 1987 the Kamiokande II neutrino detector in Japan and the IMB detector in Ohio both measured a burst of neutrinos that arrived more than three hours before the first detection of visible light from the supernova. Burst is perhaps too strong a word, since the total number of neutrinos detected was 20 – neutrinos are hard to measure! Nonetheless, this was vastly above the noise level, and provided the first direct evidence that a supernova explosion involves release of neutrinos.

3. Historical importance

   A brief aside: because of their brightness and the long duration for which they are visible, supernovae played an important part in the early development of astronomy, and in the history of science in general. In November of 1572, a supernova went off that was, at its peak, comparable in brightness to the planet Venus. For about two weeks the supernova was visible even during the day. It remained visible to the naked eye until 1574.

   The 1572 supernova was so bright that no one could have missed it. One of the people to observe it was the Dane Tycho Brahe, who said "On the 11th day of November in the evening after sunset, I was contemplating the stars in a clear sky. I noticed that a new and unusual star, surpassing the other stars in brilliancy, was shining almost directly above my head; and since I had, from boyhood, known all the stars of the heavens perfectly, it was quite evident to me that there had never been any star in that place of the sky, even the smallest, to say nothing of a star so conspicuous and bright as this. I was so astonished of this sight that I was not ashamed to doubt the trustworthyness of my own eyes. But when I observed that others, on having the place pointed out to them, could see that there was really a star there, I had no further doubts. A miracle indeed, one that has never been previously seen before our time, in any age since the beginning of the world."

   [Slide 10 – plate from Tycho's *Stella Nova*]

   Tycho was so impressed by the event that he wrote a book about it and decided to devote his life to astronomy. He went on to make the observations that were the basis of Kepler's Laws. Kepler himself saw another supernova in 1604. The supernovae played a critical role in the history of science because they provided clear falsification of the idea that the stars were eternal and unchanging, which had dominated Western scientific thought since the time of the ancient Greeks. Previous variable events in the sky, such as comets, were taken to be atmospheric phenomena, and there was no easy way to disprove this. With the supernovae, however, they persisted long enough to make parallax observations possible. The failure to detect a parallax for the supernovae provide without a doubt that they were further away than the moon, in the supposedly eternal and unchanging realm outside the terrestrial sphere.

Unfortunately for us, Tycho's supernova was the last one to go off in our galaxy (unless one went off on the far side of the galactic center, where we wouldn't be able to see it due to obscuring dust). A number of astronomers would very much like there to be another one, since astronomical instrumentation has improved a bit since Tycho's day...

III. Supernova Remnants

The material ejected by a supernova into space slams into the interstellar medium, the gas between the stars, at a velocity up from a few to ten percent of the speed of light. When this collision happens, it creates a shock in the interstellar medium that heats interstellar gas to temperatures of millions of K. The shocked bubble filled with hot gas is known as a supernova remnant, and such remnants can be visible for many thousands of years after the supernova itself fades from view.

The association of these structures with supernovae can be demonstrated quite clearly by looking with modern telescopes at the locations of historical supernovae. For example, remnants have been identified for both Tycho's and Kepler's supernovae, and another for the Crab supernova (named after the constellation where it is located). The Crab supernova was recorded in 1054 by Chinese astronomers – no one in Europe at the time was paying attention to the sky, or if they were, they didn't bother to write it down.

[Slides 11 - 13 – Tycho's SNR, Kepler's SNR, and the Crab SNR]

We can understand the structure of a supernova remnant using a simple mathematical argument made independently by L. I. Sedov in the USSR and G. I. Taylor in the UK. These authors discovered the solution independently because Taylor discovered it while working in secret on the British atomic bomb project, which was later merged with the American one. It turns out that the problems of a supernova exploding in the interstellar medium and a nuclear bomb exploding in the atmosphere are quite similar physically. Sedov published his solution in 1946, just after the end of World War II, while Taylor's work was still secret.

Consider an idealized version of the supernova problem. An explosion occurs at a point, releasing an energy E. The explosion occurs inside a medium of constant density ρ, and we assume that the energy of the explosion is so large that the pressure it exerts is vastly greater than the pressure in the ambient material, so that the ambient gas can be assumed to be pressureless. This is a very good approximation for both supernovae and nuclear bombs. We would like to solve for the position of the shock front r as a function of time t.

The mathematical argument used to solve this relies on nothing more than fancy dimensional analysis. Consider the units of the given quantities. We have the energy E, density ρ, radius r, and time t, which have units as follows:

$$\begin{aligned}
[r] &= L \\
[t] &= T \\
[\rho] &= ML^{-3} \\
[E] &= ML^2T^{-2}.
\end{aligned}$$

Here L means units of length, T means units of time, and M means units of mass. Thus a density is a mass per unit volume, which is a mass per length cubed. Energy has units of ergs (CGS) or Joules (MKS), which is a mass times an acceleration times a distance, and acceleration is distance per time squared.

We want to have a formula for r in terms t, ρ, and E. It is clear, however, that there is only one way to put together t, ρ, and E such that the final answer has the units of length! The mass must cancel out of the problem, so clearly the solution must involve E/ρ. This has units

$$\left[\frac{E}{\rho}\right] = L^5T^{-2}.$$

We want to obtain something with units of length, so clearly the next step is to cancel out the T^{-2} by multiplying by t^2. This gives

$$\left[\frac{E}{\rho}t^2\right] = L^5.$$

Finally, to get something with units of L and not L^5, we must take the 1/5 power. Thus, the radius of the shock as a function of time must, on dimensional grounds, be given by

$$r = Q\left(\frac{E}{\rho}\right)^{1/5} t^{2/5},$$

where Q is a dimensionless constant. Similarly, the shock velocity as a function of time must follow

$$v = \frac{dr}{dt} = \frac{2}{5}Q\left(\frac{E}{\rho}\right)^{1/5} t^{-3/5}.$$

Actually solving the equations of fluid dynamics shows that

$$Q = \left[\left(\frac{75}{16\pi}\right)\frac{(\gamma_a - 1)(\gamma_a + 1)^2}{3\gamma_a - 1}\right]^{1/5},$$

where γ_a is the adiabatic index of the gas into which the shock propagates.

Taylor used this solution to deduce the energy of the first atomic explosion at Trinity using nothing but photos of the blast wave at different times that had been published in newspapers and magazines. When he published the result in 1950, a number of people were not happy.

For supernovae we generally can't see them expand – the expansion takes too long. However, we can obtain a relationship we can test between the temperature of the shocked material and the radius of the remnant. At the shock the kinetic energy of the expanding gas is converted into heat, so the temperature at the shock, which is a measure of internal energy per unit mass, is simply proportional to the kinetic energy per unit mass, which varies as v^2. Thus we have

$$T_{\mathrm{shock}} \propto v^2 \propto \left(\frac{E}{\rho}\right)^{2/5} t^{-6/5}.$$

Now let us rewrite this in terms of the radius. Solving the first equation for t, we have

$$t \propto \left(\frac{E}{\rho}\right)^{-1/2} r^{5/2},$$

and plugging this into our equation for the shock temperature, we have

$$T_{\mathrm{shock}} \propto \left(\frac{E}{\rho}\right) r^{-3}.$$

Thus the temperature of supernova remnants should decrease as the third power of their size, assuming roughly constant energy and ISM density. Small remnants such as Kepler's, Tycho's, and the Crab are visible in x-rays, but the rapid temperature drop with size ensures that, once they expand significantly, they cool off too much to be visible in x-rays.

Class 18 Notes: Neutron Stars and Black Holes

In the last class we discussed the violent deaths of massive stars via supernovae. We now turn our attention to the compact remnants left by such explosions, with a goal of understanding the structure and properties. We are at the forefront of astrophysical knowledge here, so much of what we say will necessarily be uncertain, and may change. As we discussed last time, the natural object left behind by collapse of the iron core of a massive star is a neutron star, a neutron-dominated object at huge density. We will begin with a discussion of neutron stars, and then discuss under what circumstances we may wind up with a black hole instead. Finally, we will discuss how these objects radiate so that we can detect them.

I. Neutron stars

A. Structure

We begin our discussion by considering the structure of a neutron star. We can make a rough estimate for the characteristic radius and density of such an object by considering that it is held up by neutron degeneracy pressure. The neutrons in the star are (marginally) non-relativistic, so the pressure is given by the general formula we derived for a non-relativistic degenerate gas:

$$P = \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20m} n^{5/3},$$

where m is the mass per particle and n is the number density of particles. If we take the star to be composed of pure neutrons, then $m = m_n = 1.67 \times 10^{-24}$ g, and $n = \rho/m_n$. Plugging this in, we have

$$P = \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20m_n^{8/3}} \rho^{5/3}.$$

This pressure must be sufficient to hold up the star. To see what this implies, we approximate the structure of the star as a polytrope, which is a reasonably good approximation since the pressure is most of the star is dominated by non-relativistic degeneracy pressure, which corresponds to an $n = 3/2$ polytrope. Using the relationship between central pressure and density appropriate to polytropes:

$$P_c = (4\pi)^{1/3} B_n G M^{2/3} \rho_c^{4/3},$$

where B_n is a constant that depends (weakly) on the polytropic index n. Combining this with the pressure-density relation for a degenerate neutron gas, we have

$$\left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20m_n^{8/3}} \rho_c^{5/3} = (4\pi)^{1/3} B_n G M^{2/3} \rho_c^{4/3}$$

$$\rho_c = \frac{4}{9}(20\pi B_n)^3 \frac{G^3 M^2 m_n^8}{h^6}.$$

We can also make use of the relationship between central density and mean density for polytropes:

$$D_n = \frac{\rho_c}{\overline{\rho}} = \rho_c \frac{4\pi R^3}{3M},$$

where $D_n = -[(3/\xi_1)(d\Theta/d\xi)_{\xi_1}]^{-1}$ is another constant that depends on the polytropic index. Plugging this in for ρ_c, we have

$$\begin{aligned} \frac{3M}{4\pi R^3} D_n &= \frac{4}{9}(20\pi B_n)^3 \frac{G^3 M^2 m_n^8}{h^6} \\ R &= \frac{3D_n^{1/3}}{20(2\pi)^{4/3} B_n} \left(\frac{h^2}{G m_n^{8/3}}\right) \frac{1}{M^{1/3}} \end{aligned}$$

Plugging in the values appropriate for an $n = 3/2$ polytrope ($D_n = 5.99$ and $B_n = 0.206$) gives

$$R = 14 \left(\frac{M}{1.4\ M_\odot}\right)^{-1/3} \text{ km}.$$

The choice of 1.4 $M_\odot$ is a typical neutron star mass. This R is only slightly higher than what more sophisticated models get (10 km) for neutron stars that have had a chance to cool off from their initial formation and become fully degenerate.

The slight discrepancy has several causes. First, the fluid isn't pure neutrons; there are some protons too, which do not contribute to the neutron degeneracy pressure. Second, as we'll see shortly, the neutrons are not too far from being relativistic, and this reduces their pressure compared to the fully non-relativsitic pressure we've used. Third, the neutron matter also has a considerably more complex structure than a simple degenerate electron gas, due to nuclear forces between the neutrons. Nonetheless, our calculation establishes that neutron stars have an incredibly high density. The mean density of a star with a mass of 1.4 $M_\odot$ and a radius of 10 km is about 10^{15} g cm^{-3}, comparable to or greater than the density of an atomic nucleus.

Neutron star matter has a number of interesting and bizarre quantum mechanical properties that we only understand in general terms. First, the free neutrons spontaneously pair up with one another. This pairing of two fermions (half-integer spin particles) creates a boson (integer spin). Bosons are not subject to the Pauli exclusion principle, and this allows the pairs of neutrons to settle into the ground quantum state. Since they are already in the ground state, they cannot lose energy, which means that they are completely frictionless. The gas is therefore a superfluid, meaning a fluid with zero viscosity. Similar fluids can be made in laboratories on Earth by cooling bosons, usually helium-4.

One aspect of superfluidity relevant to neutrons stars is the way they rotate. Rotations of superfluids do not occur as macroscopic rotation like for a normal fluids, but rotate as quantized vortices. If a superfluid is held in a vessel and the vessel is rotated, at first the fluid remains perfectly stationary. Once the rotation speed of the fluid exceeds a critical value, the fluid will rotate in a vortex at that critical speed. If the speed of the vessel is increased further, the fluid will not speed up any more until the next critical speed is reached, at which point the fluid will jump to that rotation speed, and so on. Similar things happen in neutron stars – rotation of the star induces the appearance of quantized vortices in the star, and this may affect its structure.

Another property of neutron star matter is that the residual protons present also form pairs, and these pairs make the fluid a superconductor, with zero electrical or thermal resistance. This makes the star isothermal, and the superconductivity has important implications for the magnetic properties of the star, which we'll discuss in a moment.

B. Maximum mass

Neutron stars are subject to a maximum mass, just like white dwarfs, and for the same reason. The mass radius relation we derived earlier is $R \propto M^{-1/3}$, as it is for all degeneracy pressure-supported stars, so as the mass increases, the radius shrinks and the density rises.

If the star becomes too massive, the density rises to the point where the neutrons become relativistic, and a relativistic gas is a $\gamma = 4/3$ polytrope, which has a maximum mass. We can roughly estimate when the relativistic transition must set in using the same method we did for white dwarfs. The non-relativistic degeneracy pressure is

$$P = \left(\frac{3}{\pi}\right)^{2/3} \frac{h^2}{20 m_n} n^{5/3},$$

and the relativistic equivalent is

$$P = \left(\frac{3}{\pi}\right)^{1/3} \frac{hc}{8} n^{4/3}.$$

Equating these two, we see that the gas transitions to being relativistic at a number density

$$n = \frac{125\pi c^3 m_n^3}{24 h^3} \qquad \Longrightarrow \qquad \rho = \frac{125\pi c^3 m_n^4}{24 h^3} = 1.2 \times 10^{16} \text{ g cm}^{-3},$$

which is only slightly higher than the mean density we have already computed. Thus the gas must be close to relativistic in a typical neutron star. This is not surprising, since the escape velocity from the surface is

$$v_{\text{esc}} = \sqrt{\frac{2GM}{R}} = 0.64c,$$

for $M = 1.4 M_\odot$ and $R = 10$ km, i.e. the surface escape velocity is more than 60% of light speed. Thus the neutrons must be moving around at an appreciable fraction of the speed of light even at the typical neutron star mass.

For the Chandrasekhar mass of a neutron star, note that the relativistic degenerate pressure is

$$P = \left(\frac{3}{\pi}\right)^{1/3} \frac{hc}{8} n^{4/3} = \left(\frac{3}{\pi}\right)^{1/3} \frac{hc}{8m_n^{4/3}} \rho^{4/3}.$$

This is exactly the same formula as for a degenerate electron gas with $\mu_e = 1$, so we can compute the Chandrasekhar mass just by plugging $\mu_e = 1$ into the Chandrasekhar mass formula we derived earlier in the class:

$$M_{\rm Ch} = \frac{5.83}{\mu_e^2} M_\odot = 5.83\ M_\odot.$$

Unfortunately this turns out to a be a pretty serious overestimate of the maximum neutron star mass, for two reasons. First, this estimate is based on Newtonian physics, and we just convinced ourselves that the escape velocity is approaching the speed of light, which means that we must use general relativity. Second, our calculation of the pressure neglects the attractive nuclear forces between neutrons; electrons lack any such attractive force. The existence of an attractive force reduces the pressure compared to the electron case, which in turn means that only a smaller mass can be supported. How small depends on the attractive force, which is not completely understood. Models that do these two steps correctly suggest a maximum mass of a bit over 2 $M_\odot$, albeit with considerable uncertainty because our understanding of the equation of state of neutronized matter at nuclear densities is far from perfect – this is not an area where we can really do laboratory experiments!

C. Magnetic fields

As we discussed a moment ago, one important property of neutron stars is that they are superconductors, i.e., they have nearly infinite electrical conductivity. Therefore, electric currents flow with essentially no resistance and magnetic fields difuse very little in superconductors; fields don't diffuse in or out of them. Therefore the magnetic field within them is said to be "frozen into the fluid"; meaning that any field line that passes through a given fluid element is trapped in that fluid element and moves and deforms with it. A magnetic field that is deformed (stretched or compressed) responds by applying a restoring "Lorentz force" on the fluid.

To see what this implies, suppose that the stellar core out of which a neutron star formed was threaded by an initial magnetic field intensity B_i (also called the magnetic flux density, in units of, for example, gauss, where 1 gauss = 10^{-4} Webers/m^2). The core was a superconductor because of electron degeneracy, so the same magnetic flux that passed through the core must now pass through the neutron star – think of this as the number of field lines passing through the core being the same as the number that now go through the neutron star.

The magnetic flux through a small region of surface area A on the initial core was $A\,B_i$. For the final neutron star, this surface area has shrunk in proportion to the radius squared. Therefore,

$$R_i^2 B_i = R_f^2 B_f \qquad \Longrightarrow \qquad B_f = \left(\frac{R_i}{R_f}\right)^2 B_i.$$

Thus when the core collapses to make a neutron star, the magnetic field that is trapped in the core is enhanced by a factor of $(R_i/R_f)^2$. The initial radius we said last time is about 10^4 km and the final one is 10 km, so the field intensity is boosted by a factor of 10^6 and the magnetic field energy, which is proportional to the square of the field intensity, is increased by a factor of 10^{12}!

We're not sure exactly how strong the magnetic field is before the supernova, but we can take the observed magnetic fields of white dwarfs as a rough guess, since the massive star core is basically an iron white dwarf before it collapses. These cover a very wide range, but typical values are $\sim 10^5$ gauss, which means that we expect neutron stars to have magnetic fields of order $10^{11} - 10^{12}$ gauss, with some going much higher and some much lower. Indeed the highest observed neutron star magnetic fields reach nearly 10^{15} gauss, although 10^{12} gauss is more typical.

To put this in perspective, the Earth's surface magnetic field is around 0.6 gauss, a typical refrigerator magnet is around 100 gauss, the strongest magnets we can make on Earth are well under 10^6 gauss, and the strongest magnetic field ever achieved briefly (using focused explosives) are around 10^7 gauss. Even a 10^6 gauss field cannot be created using conventional materials because the magnetic forces generated exceed the tensile strength of terrestrial materials, i.e. a 10^6 gauss electromagnet would crush itself because steel would not be strong enough to hold it up. A 10^{12} gauss magnetic field is high enough that atoms cannot have a normal structure, and instead the electron orbitals become highly distorted and flattened.

In the stars with the strongest magnetic fields, $\sim 10^{15}$ gauss, known as magnetars, sudden re-arrangements of the magnetic field can generate bursts of gamma rays. One such even on August 27, 1998 was sufficient to ionize large parts of the Earth's outer atmosphere, disrupting radio communications.

[Slide 1 – x-ray light curve of SGR 1900+14]

D. Rotation and Pulsars

The strong magnetic field is particularly important when coupled with another aspect of neutron stars: rapid rotation. Neutron stars are rapid rotators for exactly the same reason they are strongly magnetized: conservation during collapse, in this case conservation of angular momentum. Consider a massive star core rotating with an initial angular velocity ω_i and an initial moment of inertia $I_i = C_i M R_i^2$, where C_i is a constant or order unity that depends on the core's density structure. Its angular momentum is

$$L = I_i \omega_i = C_i M R_i^2 \omega_i.$$

As it collapses it must conserve angular momentum, so its angular momentum after collapse is

$$L = I_f \omega_f = C_f M R_f^2 \omega_f = C_f M R_i^2 \omega_i \qquad \Longrightarrow \qquad \omega_f \approx \omega_i \left(\frac{R_i}{R_f} \right)^2,$$

where we have dropped the constants of order unity.

Thus the angular velocity of the core is enhanced by the same factor of $\sim 10^6$ as the magnetic field. The period, which is simply $P = 2\pi/\omega$, decreases by the same factor. As with the magnetic field, we're not exactly sure what rotation rates should be for massive star cores, but we can guess based on white dwarfs. The fastest rotating of these, which are probably the youngest and closest to their original state, have periods of about an hour, or a few $\times 10^3$ s. The period of a neutron star should be roughly a million times smaller than this, which is a few milliseconds. Thus newborn neutron stars should be extremely rapidly rotating.

The combination of a strong magnetic field and rapid rotation gives rise to an interesting phenomenon: pulsation. The strong magnetic field of the pulsar traps and accelerates charged particles. As these particles move they are confined to move along the strong magnetic field lines, and this causes their trajectories to curve. Any charged particle that accelerates, as the particle must to move in a curved trajectory, emits radiation, and this produces a beam of radiation from the North and South poles of the star. The emitted radiation turns out to be in radio waves.

[Slide 2 – pulsar schematic]

If the pulsar rotates and the magnetic and rotation axes are not perfectly aligned, this beam will sweep through space, and, if it happens to pass over the Earth, we will see a pulse of radio waves once per rotation period. These objects are therefore called pulsars.

The first pulsar was discovered by Jocelyn Bell as a graduate student at Cambridge in 1967, quite by accident. She was building a radio telescope as part of her thesis, and discovered an extremely regular signal coming from a spot on the sky. Due to its regularity, she at first thought it might be a beacon from an alien civilization, and she actually labelled the signal on the paper record "LGM", with the LGM standing for Little Green Men.

[Slide 3 – pulsar discovery record]

In 1974, the Nobel Prize for physics was awarded for the discovery of pulsars, but it was *not* given to Jocelyn Bell. Instead, it went to her (male) PhD advisor Anthony Hewish. She went on to become a very successful astronomer and a university president. She is now the president of the Institute of Physics in the UK.

Pulsars are extremely regular because their "clock" is the rotation of the neutron star, which has a huge amount of inertia to keep it spinning steadily. However, pulsars do slow down. A rotating magnetic dipole such as a pulsar emits radiation (which is not the same as the radio beam we see), and this radiation reduces the kinetic energy of the pulsar, causing it to slow. The slowdown is very slow: a typical pulsar requires of order ten million years to slow down significantly.

The magnetic dipole radiation is deposited in the material around the pulsar, leading to formation of a structure called a pulsar wind nebula, and these are often observed at the center of supernova remnants. A famous example is the Crab nebular supernova remnant, which hosts a pulsar, and which the x-ray telescope Chandra showed to host a pulsar wind nebula as well. In other cases we also see supernova remnants with pulsar wind nebulae at their centers.

[Slides 4, 5 – crab and G292.0+1.8 pulsar wind nebulae]

II. Black holes

A. The Schwarzschild Radius

We have seen that there is a maximum possible mass for neutron stars. Usually the stellar core of a star that explodes as a supernova is smaller than this limit, and the result is a neutron star. However, it is possible for the core to be pushed above the maximum neutron star mass in rare cases. One way this may happen is if not all of the stellar envelope is ejected, and some of it falls back onto the proto-neutron star. In this case the neutron star may accrete the material and exceed its maximum mass. Another possibility is that a very massive star may encounter the pair instability region in the $(\log \rho, \log T)$ plane before it gets to the iron photodisintegration instability region. In this case the collapsing core may be more massive than $2-3\ M_\odot$, and the result will again be a core that exceeds the maximum possible neutron star mass.

If such a core is created, there is, as far as we know, nothing that can stop it from collapsing indefinitely. A full description of what happens in such a collapsing star requires general relativity, which we will not cover in this class. However, we can make some rough estimates of what must happen using general arguments.

As the star collapses, the escape velocity from its surface rises:

$$v_{\mathrm{esc}} = \sqrt{\frac{2GM}{R}}.$$

Once the radius is small enough, this velocity exceeds the speed of light. The critical velocity at which this happens is called the Schwarzchild radius:

$$R_{\mathrm{Sch}} = \frac{2GM}{c^2} \approx 3\frac{M}{M_\odot}\ \mathrm{km}.$$

Thus a neutron star is roughly $2-3$ Schwarzschild radii in size, and it doesn't take much additional compression to push it over the edge.

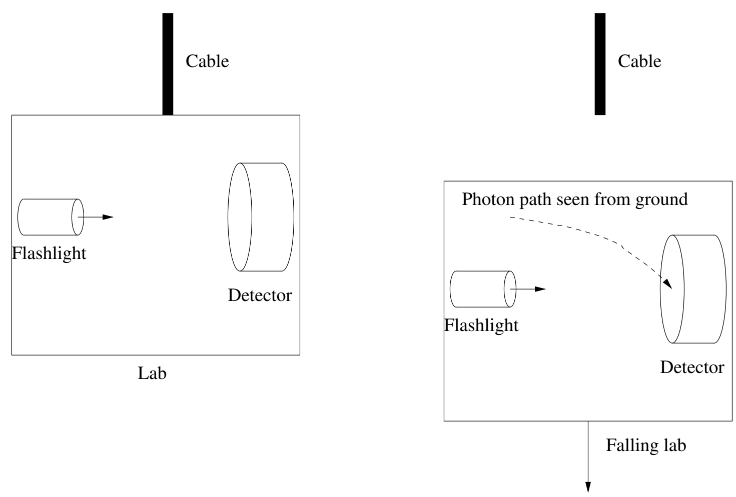
The Schwarzschild radius is the effective size of the black hole. Nothing that approaches within that distance of the mass can escape, since nothing can move faster than light. Because nothing that happens inside the Schwarzschild radius can ever influence events outside it, the Schwarzschild radius is called an event horizon.

B. Light deflection, gravitational redshift, and time dilation

A full description of what happens near the event horizon of a black hole is a subject for a GR class, but we can sketch some basic phenomenology here. Our main tool to do so will be Einstein's equivalent principle, which states that any physical experiment must give identical results in all local, freely-falling, non-rotating laboratories. In other words, if I take a laboratory and put it in deep space anywhere in the universe, or allow it to freely orbit or fall in a gravitational field, I have to get the same results. This seems like a simple statement, but it has profound implications for the effects of gravity.

Our basic tool to understand this will be a simple thought experiment: consider a laboratory inside a sealed elevator, which is suspended from a cable in a gravitational field. The laboratory contains flashlights capable of emitting one photon, and detectors capable of detecting them.

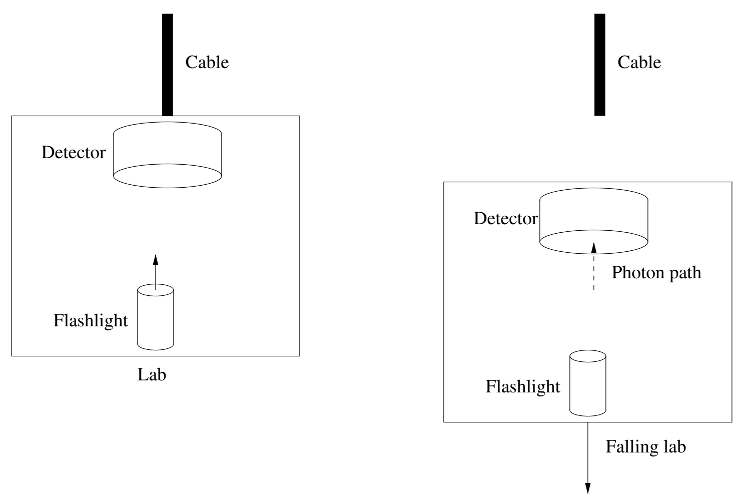
In our first experiment, the flashlight is attached to one of the vertical walls of the lab, and the detector is attached to the opposite horizontal wall. The lab is rigged so that, at the moment the flashlight emits its photon, the cable detaches and the lab is allowed to fall freely.



Since the lab is in free-fall, the physicist in the lab must get the same result he would get if the lab were in free-fall in deep space: the photon travels in a straight horizontal line across the lab to the detector. However, consider what an observer on the ground would see. The light still has to hit the detector – when they compare the reading on the detector after the experiment is over, the physicist on the ground and the one in the lab must read off the same result for whether a photon was detected or not. Since the lab is falling, however, this requires that the photon follow a curved rather than straight path. Otherwise the photon would miss the detector. From the perspective of the physicist on the ground, the gravitational field that accelerated the lab must have also caused the photon's trajectory to bend. The conclusion: gravity bends light.

Of course the amount by which the light bends depends on how strong the gravitational field is. Near a black hole, at 1.5 Schwarzschild radii the gravity is strong enough that photons can go in circular orbits. Light can still escape from this radius if travels radially outward from the black hole, but if it is emitted tangentially, it will instead orbit forever.

Now consider another experiment. This time the flashlight is affixed to the bottom of the experimental chamber, and the detector is affixed to the top. As before, the chamber is rigged so that, at the same moment the flashlight emits its photon, the cable is released and the apparatus begins to fall freely.



Again, we invoke the principle that the physicist in the falling laboratory must obtain the same result as if the lab were in deep space. Thus, the frequency of the light detected by the detector must match the frequency of light emitted by the flashlight, since that is what would happen in deep space.

For the physicist on the ground, this creates a problem. The detector is falling toward the photon, so there should be a Doppler shift that makes the photon appear bluer. However, when that physicist examines the reading on the detector later, it will not show a shift in frequency. The conclusion is that the photon, in climbing out of the gravity well, must have undergone a redshift that counters the blue Doppler shift. Thus, gravity not only bends light, it shifts the frequency of light. Photons that climb out of gravity wells become redder. Reversing the thought experiment so that the flashlight is mounted on the ceiling and the detector is on the floor shows that the converse is also true: photons falling into gravity wells become bluer.

The general relativistic result is that the frequency of a photon of frequency ν_0 emitted at a radius r_0 around a black hole and received at infinity will have frequency

$$\nu_\infty = \nu_0 \left(1 - \frac{2GM}{r_0 c^2}\right)^{1/2} = \nu_0 \left(1 - \frac{R_{\rm Sch}}{r_0}\right)^{1/2}.$$

Note that this formula has the property that, as $r_0 \rightarrow R_{\rm Sch}$, the observed frequency $\nu_\infty \rightarrow 0$. Thus, light emitted near the event horizon becomes more and more redshifted, until finally at the event horizon it becomes infinitely redshifted and can no longer be observed by the outside world.

An important corollary of this is gravitational time dilation. Consider constructing a clock based around a monochromatic light source. For every crest of a light wave that passes, the clock records one tick. Now consider constructing two such clocks and lowering one near the surface of a black hole. The light coming out of the clock near the black hole will be redshifted, so its frequency will diminish as seen from an observer at infinity. This means that fewer wave crests pass the detector on that clock than pass the detector at infinity in the same amount of time. When the clock near the black hole is pulled back up, it will have recorded fewer ticks than the clock at infinity. The conclusion is that time must slow down near the black hole.

Time dilation follows the same formula as frequency shifting, just in reverse. If a clock at infinity records the passive of a time Δt_∞, then one near the black hole will record a time

$$\Delta t_0 = \Delta t_\infty \left(1 - \frac{R_{\rm Sch}}{r_0}\right)^{1/2}.$$

Thus objects near $R_{\rm Sch}$ appear to outside observers to slow down, until at $R_{\rm Sch}$ they become entirely frozen in time.

III. Accretion power

A. Luminosity

How do we observe neutron stars and black holes? The answer is that, when they're all alone, for the most part we don't. A bare black hole is, by definition, completely free of any kind of emission. A bare neutron star does radiate, but only very weakly. It luminosity is

$$L = 4\pi R^2 \sigma T^4.$$

Neutron stars are born very hot, $T > 10^{10}$ K, but after ~ 1 Myr the star cools and the temperature drops to $\sim 10^6$ K. Plugging in $R = 10$ km with that temperature gives $L = 0.2\ L_\odot$. This is dim enough to make it quite hard to detect any but the nearest neutron stars by thermal emission, particularly since, at this temperature, the emission peaks in the x-ray, and must therefore be studied from space. We have indeed identified some of the nearest and youngest neutron stars, such as the Crab pulsar, by their thermal x-ray emission. However, this is not an option for most neutron stars.

Instead, we tend to find neutron stars and black holes only when they emit non-thermally (e.g. pulsars) or if they are powered by accretion of material from another body. The energetics of this work exactly as they do for protostars, as you worked out on your last problem set. Consider a neutron star of radius R that accretes an amount of mass dM in a time dt. The material falls from rest at infinity, so it has zero energy initially. Just before it arrives at the surface, its potential and kinetic energies must add up to zero, so

$$\frac{1}{2}v^2\,dM - \frac{GM\,dM}{R} = 0 \qquad \Longrightarrow \qquad \frac{1}{2}v^2 = \frac{GM}{R}.$$

When the material hits the surface and stops, its kinetic energy is converted into heat, and then it is radiated away. In steady state all the extra energy must be radiated, so the amount of energy released is

$$dE = \frac{1}{2}v^2\,dM = \frac{GM\,dM}{R}.$$

The resulting luminosity is just the energy per unit time emitted via this process:

$$L_{\rm acc} = \frac{dE}{dt} = \frac{GM}{R}\dot{M}.$$

Note that this is slightly different than the protostar case in that not all of the thermal energy of the material that falls onto the surface of a protostar has to be radiated – about half of it is retained and is used to heat up the star instead. Here, if the temperature of the star is fixed, it will all be radiated. That explains the factor of 2 difference from the protostar case.

Accretion luminosity increases as the radius of the star decreases, which means that it can be a much more potent energy source for compact things like neutron stars than it is for protostars. For example, suppose a star accretes at a rate of $10^{-10}\ M_\odot\ \mathrm{yr}^{-1}$, so that it gains roughly 1 $M_\odot$ of mass over the age of the universe. This is $4-5$ orders of magnitude slower accretion than in the protostar case. For the Sun, $L_{\rm acc} \approx 10^{-3}\ L_\odot$, unnoticeably small. For a white dwarf, $R = 0.01\ R_\odot$, it would be $L \approx 0.1\ L_\odot$, high enough to be brighter than just an isolated white dwarf normally is. For a neutron star, $R = 10$ km, it would be 100 $L_\odot$, and for a black hole, $R \approx 3$ km, it approaches 1000 $L_\odot$! (Black holes don't have surfaces for matter to crash into, but we'll see that they still emit nearly as much as if they did.)

Of course this process cannot produce arbitrarily high luminosities, for the same reason that stars cannot have arbitrarily high luminosities: the Eddington limit. If the luminosity is too high, radiation forces are stronger than gravity, so material will be pushed away from the accreting object rather than attracted to it. The Eddington limit is

$$L_{\rm Edd} = \frac{4\pi c G M}{\kappa},$$

and if we require that $L_{\rm acc} < L_{\rm Edd}$, then we have

$$\frac{GM}{R}\dot{M} < \frac{4\pi c GM}{\kappa} \qquad \Longrightarrow \qquad \dot{M} < \frac{4\pi c R}{\kappa}.$$

Thus there is a maximum accretion rate onto compact objects. The value of κ that is relevant is usually $\kappa_{\rm es}$, since usually the accreting material is hot and fairly low density.

To figure out the wavelength of the emission, we need to estimate the surface temperature of the accreting object. This is given by the normal result

$$L = L_{\rm acc} = 4\pi R^2 \sigma T^4 \qquad \Longrightarrow \qquad T = \left(\frac{GM\dot{M}}{4\pi R^3 \sigma}\right)^{1/4}.$$

If we plug in the maximum possible accretion rate, we get the maximum temperature, which will apply to the brightest objects:

$$T = \left(\frac{GMc}{R^2\sigma\kappa}\right)^{1/4}$$

Thus smaller radii also lead to higher temperatures. Plugging in $R = 0.01R_\odot$ for a white dwarf, $R = 10$ km for a neutron star, or $R = 3$ km for a black hole gives about 10^6 K for a white dwarf, 2×10^7 K for a neutron star, and 4×10^7 K for a black hole. Thus compact objects accreting near the maximum rate should emit primarily in the ultraviolet or x-ray.

B. Binaries, Roche Lobe Overflow, and Disks

We've worked out the energetics, but how does material actually get onto a compact object like a neutron star or black hole? The answer is generally that it must be donated by a companion. Fortunately for us, most stars massive enough to produce neutron stars or black holes are born as members of binary systems. In such a system, the more massive member will evolve off the main sequence first, while the other star is still on the main sequence.

Many binaries are disrupted by the supernova that creates a neutron star or white dwarf, but some remain bound – it's a matter of how much mass is ejected and how asymmetric the explosion is. If the binary remains bound, the result is a main sequence star with a neutron star or black hole companion.

Some time later the companion will begin to evolve off the main sequence. When it does, it will swell into a giant star. However, this may bring the outer parts of its envelope very close to its companion – close enough that they can be gravitationally captured by the companion and accrete onto it. The region around each star where mass is safely bound to that star is known as the star's Roche lobe. (Lobe because it has a teardrop-like shape.) As a star swells into a giant, its outer layers may overflow its Roche lobe.

[Slides 6, 7 – Roche lobe overflow diagram and animation]

The overflowing material falls onto the compact companion. However, it cannot fall directly onto the star because it has too much angular momentum. Instead, it goes into a rotating disk around the compact object, where it is in Keplerian rotation – just like a planet going around a star. The resulting object is called an accretion disk.

If the material in the disk were free of viscosity, it would simply orbit happily forever, just like planets around stars. However, the gas in a disk has some viscosity, due to mechanisms that we won't discuss in this class. The viscosity acts like a frictional drag: blobs of gas somewhat closer to the compact object rub against those somewhat further out, and this friction slows down the inner blobs so that they lose angular momentum and spiral ever close to the central object.

As material in the disk rubs against other material and moves inward, it must heat up to conserve energy: gravitational potential energy is being lost, so it must be converted to heat. In turn, this heating causes the material to radiate. As a result, half the gravitational potential energy of infalling material is radiated away in the accretion disk even before the gas gets to the surface of the compact object. This is why we can see accreting black holes even though accreting material that actually gets to the event horizon simply plunges on through without radiating further – half the energy has already come out in the accretion disk.

Energy release in an accretion disk around a black hole is probably the most efficient form of energy release in the universe. Let's put this in perspective by comparing it to nuclear burning. Consider a single proton. If we burn it to helium in a star, the energy release is $\epsilon = 0.007$ of its total rest energy, where

$$\Delta E = \epsilon m_p c^2$$

is the total energy released, and $m_p c^2$ is the proton's total energy content. If we burn it all the way to iron, the efficiency increases to $\epsilon = 0.009$.

In contrast, consider the same proton being accreted onto a black hole. The energy released is

$$\Delta E = \frac{GMm_p}{2R_{\text{Sch}}},$$

where the factor of 2 assumes that we only get the half of the energy that comes out in the accretion disk, while the rest is swallowed by the black hole. Plugging in $R_{\text{Sch}} = 2GM/c^2$, this is

$$\Delta E = \frac{m_p c^2}{4} \qquad \Longrightarrow \qquad \epsilon = \frac{\Delta E}{m_p c^2} = \frac{1}{4}.$$

This calculation is quite approximate, since we have neglected a number of important general relativistic effects that occur near the Schwarzschild radius. A more sophisticated treatment gives $\epsilon \approx 0.1$.

Nonetheless, this means that accreting a proton onto a black hole releases roughly 10 times as much energy as burning the same proton to iron. It releases 10%

as much energy as the maximum possible amount, which would be released by annihilating the proton with an anti-proton. For this reason, accreting black holes are some of the brightest objects in the universe – indeed, early in the history of the universe, they dominated the total light output of the cosmos. These objects are called quasars. They are powered by black holes much larger than that created by any star, with masses up to $10^9\ M_\odot$.

[Slide 8 – radio / optical image of NGC 4261]

The image shows jets of material being ejected from the galaxy by the quasar. Notice the scale, and how the jet compares in size to the galaxy: each jet is about 100,000 light years (30 kpc) long, larger than the entire galaxy. That's what accretion power can do.

Class 19 Notes: The Stellar Life Cycle

In this final class we'll begin to put stars in the larger astrophysical context. Stars are central players in what might be termed "galactic ecology": the constant cycle of matter and energy that occurs in a galaxy, or in the universe. They are the main repositories of matter in galaxies (though not in the universe as a whole), and because they are the main sources of energy in the universe (at least today). For this reason, our understanding of stars is at the center of our understanding of all astrophysical processes.

I. Stellar Populations

Our first step toward putting stars in a larger context will be to examine populations of stars, and examine their collective behavior.

A. Mass Functions

We have seen that stars' masses are the most important factor in determining their evolution, so the first thing we would like to know about a stellar population is the masses of the stars that comprise it. Such a description is generally written in the form of a number of stars per unit mass. A function of this sort is called a mass function. Formally, we define the mass function $\Phi(M)$ such that $\Phi(M)\,dM$ is the number of stars with masses between M and $M + dM$.

With this definition, the total number of stars with masses between M_1 and M_2 is

$$N(M_1, M_2) = \int_{M_1}^{M_2} \Phi(M)\,dM.$$

Equivalently, we can take the derivative of both sides:

$$\frac{dN}{dM} = \Phi$$

Thus the function Φ is the derivative of the number of stars with respect to mass, i.e. the number of stars dN within some mass interval dM.

Often instead of the number of stars in some mass interval, we want to know the *mass* of the stars. In other words, we might be interested in knowing the total mass of stars between M_1 and M_2, rather than the number of such stars. To determine this, we simply integrate Φ times the mass per star. Thus the total mass of stars with masses between M_1 and M_2 is

$$M_*(M_1, M_2) = \int_{M_1}^{M_2} M\Phi(M)\,dM$$

or equivalently

$$\frac{dM_*}{dM} = M\Phi(M) \equiv \xi(M).$$

Unfortunately the terminology is somewhat confusing, because $\xi(M)$ is also often called the mass function, even though it differs by a factor of M from $\Phi(M)$. You will also often see $\xi(M)$ written using a change of variables:

$$\xi(M) = M\Phi(M) = M\frac{dN}{dM} = M\frac{d\ln M}{dM}\frac{dN}{d\ln M} = \frac{dN}{d\ln M}.$$

Thus ξ gives the number of star per logarithm in mass, rather than per number in mass. This has an easy physical interpretation. Suppose that $\Phi(M)$ were constant. This would mean that there are as many stars from $1-2\ M_\odot$ as there are from $2-3\ M_\odot$ as there are from $3-4\ M_\odot$, etc. Instead suppose that $\xi(M)$ were constant. This would mean that there are equal numbers of stars in intervals that cover an equal range in logarithm, so there would be the same number from $0.1-1\ M_\odot$, from $1-10\ M_\odot$, from $10-100\ M_\odot$, etc.

Often we're more concerned with the distribution of stellar masses than we are with the total number or mass of stars. That is because the distribution tends to be invariant. If we examine two clusters of different sizes, then dN/dM will be different for them simply because they have different numbers of stars. However, they may have the same fraction of their stars in a given mass range. For this reason, it is generally common to normalize Φ or ξ so that the integral is unity, i.e. to choose a pre-factor for Φ or ξ such that

$$\int_0^\infty \Phi(M)\, dM = 1$$

and similarly for $\xi(M)$. If a mass is normalized in this way, then $\Phi(M)\, dM$ and $\xi(M)\, dM$ give the fraction of stars (fraction by number for Φ and fraction by mass for ξ) with masses between M and $M + dM$.

B. The IMF

One can construct mass functions for any stellar population. However, the most useful sort of mass function is the one for stars that have just formed, since that determines the subsequent evolution of the population. This is known as the initial mass function, or IMF for short. The IMF is not the same as the mass function at later times, because stars lose mass over their lives, and some go supernova and disappear completely. Thus the IMF is distinct from the present-day mass function, or PDMF.

Observationally, one can attempt to determine the IMF in two ways. The most straightforward way is to look at star clusters so young that no stars have yet lost a significant amount of mass, and none have yet gone supernova. Since the lifetime of a massive star is only $3-4$ Myr, such clusters must be younger than this. Young clusters of this sort are rare, so we don't have a lot of examples where we can do this. To make matters worse, such young clusters also tend to be still partially enshrouded by the dust and gas out of which they formed, making it difficult to determine stars' masses accurately.

A less simple method is to survey field stars that are not part of clusters and measure the PDMF, and then try to extrapolate back to an IMF based on an understanding of mass loss and stellar lifetimes as a function of mass. This is tricky because we only understand those things at a rough level. The great advantage of the method is that it gives us an absolutely immense number of stars to use, and thus provides great statistical power. This is important in determining the IMF for stars that are rare, and thus are unlikely to be present in the few clusters where we can use the first method.

[Slides 1 and 2 – IMFs from clusters and field stars]

Regardless of which method is used, observations tend to converge on the same result for the IMF of stars larger than $\sim 1\ M_\odot$. For these stars, $\Phi(M) \propto M^{-2.35}$, or equivalently $\xi(M) \propto M^{-1.35}$. This value of $-2.35/-1.35$ for the exponent is known as the Salpeter slope, after Edwin Salpeter, who first obtained the result. This result means that massive stars are rare both by number and by mass, since Φ and ξ are strongly declining functions of M.

At lower masses the IMF appears to flatten out, reaching a peak somewhere between $0.1\ M_\odot$ and $1\ M_\odot$ before declining again below $0.1\ M_\odot$. Some people claim there is a rise again at lower masses, but such claims are still highly controversial – very low mass objects are extremely difficult to find due to their low luminosities, and it is not easy to infer their masses. These two factors make these observations quite uncertain.

I should point out that all of these results are empirical. We don't have a good theory for why the IMF looks like it does. It seems to be very constant in the local universe, but we can't rule out the possibility that it might have been different in the distant past, or that it might depend on the environment where the star formation takes place, and change in environments that are simply not found in our own galaxy. A number of people are working on the problem.

For convenience we sometimes just ignore the flattening below $1\ M_\odot$, and assume that the Salpeter slope holds over a range from $M_{\rm min} = 0.1 M_\odot$ to $M_{\rm max} = 120 M_\odot$. A mass function of this sort is known as a Salpeter IMF. The normalization constant in this case, obtained by requiring that the integral be 1, is

$$\begin{aligned} 1 &= \int_{M_{\rm min}}^{M_{\rm max}} \Phi(M)\, dM = \int_{M_{\rm min}}^{M_{\rm max}} A M^{-2.35}\, dM = \frac{A}{-1.35}\left(M_{\rm max}^{-1.35} - M_{\rm min}^{-1.35}\right) \\ A &= \frac{1.35}{M_{\rm min}^{-1.35} - M_{\rm max}^{-1.35}} = 0.060, \end{aligned}$$

where we are working in units of solar masses. Similarly, for the IMF in terms of mass,

$$\begin{aligned} 1 &= \int_{M_{\rm min}}^{M_{\rm max}} \xi(M)\, dM = \int_{M_{\rm min}}^{M_{\rm max}} B M^{-1.35}\, dM = \frac{B}{-0.35}\left(M_{\rm max}^{-0.35} - M_{\rm min}^{-0.35}\right) \\ B &= \frac{0.35}{M_{\rm min}^{-0.35} - M_{\rm max}^{-0.35}} = 0.17. \end{aligned}$$

Thus if we take

$$\Phi(M) = 0.060 M^{-2.35} \text{ and } \xi(M) = 0.17 M^{-1.35},$$

for $M = 0.1 - 120\ M_\odot$, these functions give us a reasonable estimate of the fraction of stars by number and by mass in a given mass range.

As an example, suppose we wanted to compute what fraction of stars (by number) are more massive than the Sun in a newborn population. The answer is

$$f_N(> M_\odot) = \int_1^{120} 0.060 M^{-2.35} = \frac{0.060}{1.35}\left(1^{-1.35} - 120^{-1.35}\right) = 0.045.$$

Similarly, the fraction of the mass in stars above $M_\odot$ in mass is

$$f_M(> M_\odot) = \int_1^{120} 0.17 M^{-1.35} = \frac{0.17}{0.35}\left(1^{-0.35} - 120^{-0.35}\right) = 0.40.$$

Thus only 4.5% of stars are more massive than the Sun, but these stars contain roughly 40% of all the mass in newborn stars.

C. Star Clusters

As we've discussed several times, stars a born in clusters that are relatively coeval, i.e. all the stars in them are born at the same time plus or minus a few Myr. This means that for many purposes we can approximate the stars in a star cluster as all having been born in a single burst. Everything that happens subsequently is due simply to aging of the stellar population. Star clusters therefore constitute the simplest example of what happens as stellar populations age.

We have already seen that lifetimes of stars decrease monotonically with mass, so it is convenient to introduce for a given cluster the turn-off mass, M_t, defined as the mass of star that is just now leaving the main sequence. In a given cluster no stars with masses above M_t remain on the main sequence, while those with lower masses are all on the main sequence. As clusters age, M_t decreases, since lower and lower mass stars evolve off the main sequence.

We can use the turn-off mass plus the IMF to figure out how the population of stars in the cluster changes with time. As a simple example, we can estimate what fraction of the stars in a cluster by number will still be on the main sequence. The total fraction of stars by number is simply

$$f_N = \int_{M_{\min}}^{M_t} \Phi(M)\, dM = \frac{0.060}{1.35}\left(M_{\min}^{-1.35} - M_t^{-1.35}\right).$$

This expression does not fall below 0.5 unless $M_t < 0.168$, which never happens, since the universe is not old enough for stars with masses so low to have left the main sequence. Thus the majority of the stars in a cluster are always on the main sequence, even for the oldest clusters. This is because the most common stars are those with low masses, which have not yet had time to leave the main sequence.

As a somewhat more complex problem, we can try to estimate what fraction of the stellar mass in the cluster will remain when the turnoff mass is M_t. Stars that go supernova will eject most of their mass at speeds well above the escape speed from a cluster, so the supernova ejecta will simply escape, reducing the mass of the cluster. Most neutron stars probably escape as well, because the supernovae are not perfectly symmetric, and tend to give the neutron stars kicks that are well above the escape velocity. Similarly, the mass ejected from stars that are turning into white dwarfs will also escape the cluster due to its high temperature, although the white dwarfs themselves will remain.

Putting all this together, the mass in the cluster comes in two parts: main sequence stars and remnant white dwarfs. The fraction of the original mass contributed by main sequence stars is given by a calculation just like the one we just did for number, except using $\xi(M)$ instead of $\Phi(M)$:

$$f_{M,MS} = \int_{M_{\min}}^{M_t} \xi(M)\, dM = \frac{0.17}{0.35}\left(M_{\min}^{-0.35} - M_t^{-0.35}\right).$$

For white dwarfs, we will make the simple approximation that they all have masses of 0.6 $M_\odot$ regardless of their initial mass, so that the fraction of the original star's mass left in the white dwarf is $0.6/M$, where M is in $M_\odot$. We also approximate that all stars with initial masses below $M_{NS} = 8M_\odot$ form white dwarfs, while more massive ones go supernova and leave nothing behind in the cluster. Thus the fraction of the original mass in the form of leftover white dwarfs is

$$\begin{aligned} f_{M,WD} &= \int_{M_t}^{M_{NS}} \frac{0.6}{M}\xi(M)\, dM = 0.6 \times 0.17 \int_{M_t}^{M_{NS}} M^{-2.35}\, dM \\ &= \frac{0.6 \times 0.17}{1.35}\left(M_t^{-1.35} - M_{NS}^{-1.35}\right). \end{aligned}$$

Adding these two up, we obtain an expression for the fraction of the original cluster mass that remains:

$$\begin{aligned} f_M &= \frac{0.17}{0.35}\left(M_{\min}^{-0.35} - M_t^{-0.35}\right) + \frac{0.6 \times 0.17}{1.35}\left(M_t^{-1.35} - M_{NS}^{-1.35}\right) \\ &= 1.09 - 0.49 M_t^{-0.35} + 0.076 M_t^{-1.35} \end{aligned}$$

Obviously this expression is valid only when $0.6M_\odot < M_t < M_{NS}$. If $M_t > M_{NS}$, then only main sequence stars are left, and we only get their contribution:

$$f_M = 1.09 - 0.49 M_t^{-0.35}.$$

The slide shows this function.

[Slide 3 – fraction of mass remaining in a cluster]

Thus clusters lose about 35% of their original mass once the turnoff mass declines to 0.6 $M_\odot$. They lose more than 20% from supernovae during their first $\sim$ 10 Myr of life, when the turnoff mass declines to around 10 $M_\odot$. This mass loss process can be important in disrupting star clusters. They are held together by gravity, and as mass is ejected they become less tightly bound. Some of them dissolve entirely as a result of mass loss.

A final game we can play with the IMF and cluster is to ask how their luminosities evolve over time. The initial luminosity of the cluster is simply given by

$$L_0 = N_* \int_{M_{\min}}^{M_{\max}} L(M)\Phi(M)\, dM,$$

where N_* is the total number of stars in the cluster and $L(M)$ is the luminosity of a star of mass M. In other words, we find the total luminosity simply by integrating the luminosity per star as a function of mass against the fraction of stars per unit mass, all multiplied by the total number of stars.

Later on, when some stars have turned off the main sequence, the luminosity is given by a similar expression, but with $M_{\max}$ replaced by M_t:

$$L_1 = N_* \int_{M_{\min}}^{M_t} L(M)\Phi(M)\, dM.$$

This implicitly assumes that white dwarfs contribute negligible luminosity, which is a pretty good approximation.

The fraction of the original luminosity that remains is therefore given by

$$f_L = \frac{L_1}{L_0} = \frac{\int_{M_{\min}}^{M_t} L(M)\Phi(M)\, dM}{\int_{M_{\min}}^{M_{\max}} L(M)\Phi(M)\, dM}$$

To evaluate this, we will again make a very simple approximation: we will take $L(M) = L_\odot(M/M_\odot)^3$ for all stars. Obviously this breaks down at both the low and high mass ends, but it is good enough to give us a rough picture of what happens to clusters' luminosities as they age. Inserting this value for $L(M)$ into our expression for f_L and canceling constants that appear in both the numerator and denominator, we have

$$f_L = \frac{\int_{M_{\min}}^{M_t} M^{0.65}\, dM}{\int_{M_{\min}}^{M_{\max}} M^{0.65}\, dM} = \frac{M_t^{1.65} - M_{\min}^{1.65}}{M_{\max}^{1.65} - M_{\min}^{1.65}}.$$

Since $M_t \gg M_{\min}$ even in the oldest clusters (which have $M_t \sim 0.7 M_\odot$), we can drop the $M_{\min}$ terms, and we have

$$f_L \approx \left(\frac{M_t}{M_{\max}}\right)^{1.65}.$$

Thus the luminosity decreases as something like the turnoff mass to the 1.7 power. Recall that the lifetime of a 12 $M_\odot$ star is only a few tens of Myr, so this means that $M_t/M_{\max} \sim 0.1$ in a cluster that is about 20-30 Myr old, which in turn means that its luminosity has decreased by a factor of about $10^{1.7} = 50$! Thus clusters fade out very quickly. The young ones are extremely bright, but they lose much of their brightness in their first few tens of Myr. Thereafter they dim greatly.

It should be noted that this calculation does ignore one significant effect, which is important for very old clusters: the luminosity of red giants. Although there aren't many such stars present in any given cluster at a given time due to the short times that stars spend as red giants, they are so bright that they can dominate the total luminosity once the massive stars have faded from view. In very old stellar populations, with $M_t \lesssim M_\odot$, the total luminosity is generally dominated by the red giants.

Although we have done this for the total luminosity, we could just as easily have done it for luminosity in some specific color, say blue light, but taking into account how the surface temperature varies with mass as well. Since the surface temperatures of massive stars are higher than those of low mass stars, they emit more in the blue, and thus the blue luminosity fades even more quickly than the total luminosity. The red luminosity falls more slowly. Of course if one really wants to get precise answers, the way to do it is with numerical models of the stars' mass-dependent luminosity and surface temperature, not with rough analytic fits.

D. Stellar Population Synthesis

Figuring out how clusters fade in time, and how the fraction of the mass in white dwarfs and main sequence stars varies with time, is just the simplest example of a more general idea called stellar population synthesis. For star clusters, we assumed all the stars formed in a single burst with a given IMF, and then we computed how the stellar population would look at later times. Obviously we could easily generalize this to the case of a cluster that, for some reason, had two distinct bursts of star formation at different times. The total stellar population would just have properties given by the sum of the two bursts.

From there, however, it is clear that we can generalize even further and consider an arbitrary star formation history, i.e. we consider an object within which the star formation rate is a specified function $\dot{M}(t)$, where t is a negative number representing the time before the present. At every time t, the problem of figuring out how that stellar population looks today is exactly the same as the calculation we just went through for star clusters, and one can then simply add up, or integrate, over all times t to figure out the present-day appearance of the stellar population. This is the basic idea of stellar population synthesis.

The power of the technique is that we can run it in reverse. We can take an observed stellar population and try to figure out what formation history would yield something that looks like that. For example, if we see a high luminosity and a blue color for a given mass of stars, we can infer that the stars must have formed quite recently. In contrast a low luminosity and reddish color imply an older stellar population.

Given a bunch of stars in an HR diagram this technique can get quite sophisticated. An example is a recent paper by Williams et al. that used this method to determine the star formation history in different parts of a nearby galaxy. They divided the galaxy into annuli, and in each annulus they made an HR diagram using stars as observed by the Hubble Space Telescope. They then tried to find a star formation history that would reproduce the observed HR diagram.

[Slides 4-6 – star formation history in NGC 2976 inferred by Williams et al. using stellar population synthesis]

This is one example of stellar population synthesis. One can also do this in more distant galaxies where individual stars cannot be resolved by adding up the spectral features for stars at different ages. Whenever you hear a galaxy or a cluster described as consisting of young or old stars, it's a good bet that a technique like this was used to reach that conclusion.

II. Stars and the ISM

Stars are only one component of the baryonic (normal matter) mass in a galaxy. In between them is a sea of gas known as the interstellar medium, which we discussed very briefly a few weeks ago in the context of star formation. The ISM is very diffuse: its mean density is ~ 1 atom per cm^3. However, there are a lot of cubic cm in interstellar space, and, as a result, the mass of the ISM is considerable. In the Milky Way, the ISM has a mass equal to about 10% of the total stellar mass. There is a continuous cycling of matter between stars and the ISM, and it is this cycling that will be the final topic in the class.

A. Star Formation

One side of the cycle is the conversion of interstellar gas into stars, a process that we discussed briefly in the context of star formation a few weeks ago. Star formation is a major problem in astrophysics, and we don't have a full theory for what controls the rate at which gas turns itself into stars. Thus this discussion is mainly going to focus on empirical results and problems, with little hints at possible solutions.

Star formation appears to take place only in gas that is in molecular form – i.e. H_2 rather than atomic hydrogen. This is apparent just from looking at images of galaxies, and quantitative comparison confirms it. This is likely to be for the reason we outlined when we discussed star formation last time. Only molecular gas is cold enough to enable collapse to stars. Other types of gas are warm enough that their pressure prevents collapse.

[Slides 7-8 – star formation plus HI maps, and correlation between SFR, atomic, and molecular gas]

Within molecular gas, however, the star formation rate is significantly lower than one might expect, as can be illustrated using the example of the Milky Way. Our galaxy contains about $10^9\ M_\odot$ in molecular clouds. Recall that we calculated the free-fall time, the time gas requires to collapse when it is not supported, to be

$$t_{\rm ff} = \sqrt{\frac{3\pi}{32G\rho}},$$

where ρ is the gas density. Typical densities in giant molecular clouds are $n \sim 100$ cm^{-3} (much more than the average $n \sim 1$ cm^{-3}), so $\rho \sim n m_{\rm H} \sim 2 \times 10^{-22}$ g cm^{-3}. Plugging this in, we have $t_{\rm ff} \sim 5$ Myr.

Thus if these clouds were collapsing to form stars in free-fall, the star formation rate would be

$$\dot{M} \sim \frac{10^9\ M_\odot}{5\ \text{Myr}} \sim 200\ M_\odot\ \text{yr}^{-1}.$$

The problem is that the observed star formation rate is about 100 times smaller than this. There is a similar discrepancy in other galaxies. Therefore something must be inhibiting the collapse of molecular clouds into stars.

No one is entirely sure what the origin of the discrepancy is, and that's another big problem in astrophysics. One possibility is that the winds and radiation from young, newly-formed massive stars disrupts the molecular clouds out of which they form. This process limits the star formation rate. For example, the ionizing radiation from stars with high surface temperatures can heat up cold molecular gas to $\sim 10^4$ K, causing it to expand rapidly and disrupt the cloud of which it is part. Disruptions like this are observed, and often make for spectacular images.

B. Gas Return and Chemical Enrichment

The cycle of material isn't simply one-way. Star formation converts interstellar medium gas into stars, but stars also return gas to the interstellar medium, via mechanisms we have already described: ejection of the envelopes of red giants, and formation of supernovae. The calculation of what fraction of the mass is returned is just the inverse of the one we already performed to see what fraction of the mass would remain in stellar form in a cluster.

Formally, we define the return fraction ζ for a stellar population as the fraction of mass that is ejected. Making the same assumptions as before, the fraction that is ejected by stars that go supernova is taken to be 1 (which is not a bad approximation, since such stars eject at least 80% of their mass). The mass that remains in white dwarfs is taken to be 0.6 $M_\odot$ independent of the mass of the progenitor. Finally, stars that are still on the main sequence return a negligible amount of their gas.

Thus the return fraction is

$$\zeta = \int_{M_t}^{M_{NS}} \frac{M - 0.6 M_\odot}{M} \xi(M)\, dM + \int_{M_{NS}}^{M_{\rm max}} \xi(M)\, dM,$$

where the first term represents stars that make white dwarfs, which return a fraction $(M - 0.6M_\odot)/M$ of their mass to the ISM, and the second term represents stars that go supernova and return all their gas to the ISM.

Evaluating the integrals for the Salpeter mass function just gives one minus the stellar mass fraction we found earlier:

$$\zeta = 1 - f_M = 0.49M_t^{-0.35} - 0.076M_t^{-1.35} - 0.09,$$

where M_t is in solar masses, as before. For $M_t = 0.7$, roughly the turnoff mass for the oldest stellar populations, this gives $\zeta = 0.34$. Thus old stellar populations eventually return roughly 1/3 of their gas to the ISM. The first 20% or so of this is returned via supernovae in a few tens of Myr. The rest comes out much more slowly via red giants, asymptotic giants, and planetary nebulae produced by lower mass stars that take a very long time to reach the main sequence. A majority of the mass is in stars that stay on the main sequence longer than the age of the universe.

The gas that is returned to the ISM can make new generations of stars. Perhaps more important, it carries with it the products of nuclear burning: metals. The universe was born composed almost entirely of hydrogen and helium, and all the heavier elements were made in stars and then ejected in supernovae or by mass-losing giant stars. This process gradually alters the chemical composition of the gas, enriching it with metals.

The legacy of that process is reflected in the present-day chemical composition of stars and galaxies. Recall that the metal fraction in the Sun is $Z = 0.02$. Other stars have different metal fractions, and there is a correlation between the mass in metals and the age of the star. This is a signature of the gradual enrichment of the ISM by stellar processes over the age of the universe.

[Slide 10 – age-metallicity relation for solar neighborhood stars from Carraro et al. 1998]

We also see a correlation between the mass of a galaxy and its metallicity. This is a signature of two effects. First, more massive galaxies have formed more stars and have turned a larger fraction of their mass into stars, thereby producing more metal processing. Second, massive galaxies have larger escape speeds, which makes it harder for supernova ejecta to escape from the galaxy.

[Slide 11 – mass-metallicity relation from Kewley & Ellison 2008]

Thus the metal content of stars and galaxies provides us with direct evidence for stellar mass return to the ISM.

C. Cosmological Infall

The cycling between gas and stars in a galaxy is part of the story, but it's not all of the story. The reason is that, if you try to balance the books between mass going into stars and mass returned to the ISM, things don't add up. Stars only return 1/3 of the mass that goes into them over the entire age of the universe, which means that most of the mass that goes into stars doesn't come back into gas.

If there were a large enough gas supply in galaxies to keep fueling star formation for the age of the universe, this wouldn't be a problem. Unfortunately, we don't have that much gas. The total mass of interstellar gas in the Milky Way inside the Sun's orbit is a few times 10^9 $M_\odot$, and we have already mentioned that the star formation rate is a few $M_\odot$ per year. This means that it should take roughly 10^9 yr, or 1 Gyr, to use up all the available gas.

The problem is that the Milky Way and the universe are about 10 Gyr old, so there isn't enough gas to keep things going. There might have been more gas in the past, but then we face the uncomfortable proposition that we live at a special time, as do all other star-forming galaxies, when the gas supplies are just about to dry up. This seems to require an unreasonable amount of luck and coordination.

Instead, the preferred explanation is that the Milky Way isn't done growing. Our galaxy continues to acquire new material from intergalactic space. The majority of the baryonic mass in the universe must be out there in the gas between the galaxies – there simply isn't enough mass in the galaxies to account for the amount that we believe is there based on the models of the early universe.

This gas is called the intergalactic medium, and hunting for it, and its infall onto the Milky Way and similar galaxies, is a major project in astronomy right now. So far, no one has found it, although there are tantalizing hints that it may have been observed.