Class 3 Notes: Hydrostatic Balance and the Virial Theorem

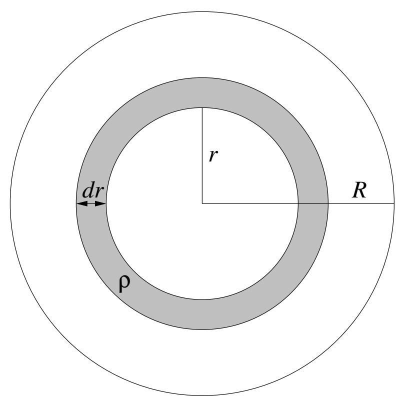
Thus far we have discussed what observations of the stars tell us about them. Now we will begin the project that will consume the next 5 weeks of the course: building a physical model for how stars work that will let us begin to make sense of those observations. This week we'll try to write down some 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: $v_{\rm ff} = \sqrt{2GM/R}$. 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

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

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 $t_{\rm dyn} = 3000$ 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 $\delta Q_{\mathrm{nuc}} = q\, dm\, \delta t$.

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

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

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.