The Virial Theorem

In this lecture we will discuss the virial theorem, which relates the kinetic energy (or temperature) of particles to their potential energy. This will give us, in a surprisingly simple way, some deep insight into the average temperature of stars.

As initial motivation for the theorem, suppose that we consider particles that are non-interacting and moving in circular orbits. What relation is there between the kinetic and the potential energy? For a particle of mass m in a circular orbit of radius r around a mass M, the potential energy is $V = -GMm/r$ and the speed is $v = \sqrt{GM/r}$ so that the kinetic energy is $K = \frac{1}{2}mv^2 = \frac{1}{2}GMm/r$. Therefore, $2K + V = 0$, where K is the kinetic energy and V is the potential energy. This means that the total energy is $K + V = V/2 < 0$.

As stated, this may not seem that interesting or relevant. For example, if we consider instead an eccentric orbit then at most places in the orbit it is not true that $2K+V = 0$, and even if it were it's not obvious what that would portend. But if we look a little deeper we find that this already tells us something highly important about gravitation: when things lose energy, they speed up and increase in kinetic energy, which you can think of as an increase in temperature. That means that gravitationally bound systems have, in this sense, negative specific heat! This is a fact of great importance to understanding such systems. This form of the theorem can be shown to apply as a time average for any system with bounded motion (see, e.g., Landau and Lifshitz).

But beyond that, what can we say about the relation between the kinetic and potential energy? As derived in the appendix (from our textbook, but the result is from the 1800s), a more general statement is that for *any* system in which the interaction force scales like $1/r^2$ for distance r (e.g., gravity), we have

$$2K + V = \frac{1}{2}\frac{d^2I}{dt^2} , \tag{1}$$

where I is the moment of inertia of the system, $I = \sum_i m_i r_i^2$, for some collection (indexed by i) of masses m_i at positions r_i. Those masses could be individual mass elements of a star, or individual stars in a cluster, or anything like that.

But it probably still isn't clear why this is a big deal, so let's do the following. Instead of worrying about the theorem at a particular instant, average the terms over a long time. Here "long time" means "long relative to a characteristic dynamical time associated with the system". So, that could be the orbital time, or the free-fall time, or something like that. Also, let's assume that we have a system that is gravitationally bound and not exploding or collapsing; thus, for example, we wouldn't be thinking about a supernova. Then, over a long time, the average of $\frac{1}{2}\frac{d^2I}{dt^2}$ has to be zero. Why? Because if it weren't, then with time the moment of inertia would increase or decrease without bound, and thus (since the m_i are fixed in the definition of I), the size of the system would go to infinity or zero.

This means in practice that the virial theorem implies $\langle 2K\rangle + \langle V\rangle = 0$, where the angle brackets mean an average taken over many dynamical times.

Still not convinced that this is useful? Fine, then, let's use it to calculate the average temperature in the Sun.

We'll do that by realizing that the thermal energy per particle takes the role of the kinetic energy, and that the energy should be about kT per particle. The total gravitational potential energy of the Sun, to a lowest-order approximation, is $GM_\odot^2/R_\odot$; as we noted in the last class there is a correction factor because the Sun is actually centrally concentrated, but that's what we'll take for our rough estimate. The number of particles in the Sun is equal to the mass of the Sun divided by the average mass per particle, or $M_\odot/m$. Again doing a rough-and-ready calculation, let's say that $m = m_p$, the mass of the proton. Then the time-averaged virial theorem becomes

$$GM_\odot^2/R_\odot + 2k\langle T\rangle(M_\odot/m_p) = 0 \ , \tag{2}$$

where the angle bracket reminds us that this relates to the *average* temperature in the Sun. Solving gives $\langle T\rangle = 1.15 \times 10^7$ K, although to our level of precision we should really say $\sim 10^7$ K.

But this is amazing! Detailed models of the Sun do indeed suggest that the interior has a temperature around 10^7 K, and we got that answer by very basic considerations combined with the virial theorem. Sure, our exact value is suspect, but given that the photosphere has a temperature of about 6000 K it is very far from obvious that the interior, on average, has to be more than 1000 times hotter. Similarly (and you can think about this), you can figure out the gravitational potential of a galaxy by seeing how fast its stars move. You can also apply the virial theorem to other systems. For example, here we assign the role of kinetic energy to the thermal energy in the Sun. If the star is instead a white dwarf, then the temperature is much less than what we would get from the theorem. That's because it is actually the degeneracy energy that matters, not the thermal energy, but using the theorem tells us the properties of the interior matter.

However, now we need to ask: have we ignored important things? In our calculation, we considered only gravitational forces. Are we being sloppy by not including other forces? **Ask class:** what other contributions might come in? The other basic forces are the strong and weak nuclear force, and the electromagnetic force. The strong and weak force are important only on very small length scales (on the order of 10^{-13} cm, which is approximately the radius of a proton). Thus these really can be ignored.

But what about the electromagnetic force? If we focus on electrostatic attraction or repulsion, then the force scales with separation r like $1/r^2$, just as gravity does. Moreover, the electrostatic force between an electron and a proton (for example) is overwhelmingly greater than the gravitational force (by almost 40 orders of magnitude!). However, we can have positive and negative electric charges; we cannot have positive and negative masses. This means that all big things in the universe (say, the mass of an asteroid or larger) are effectively electrically neutral in bulk. This means that it is usually okay to assume that gravity dominates on large scales.

Under what conditions will the moment of inertia term in the virial theorem play a significant role? It will only be significant when the bulk velocity is a significant fraction of free-fall (if the velocity is ϵ times free-fall, the correction is of order ϵ). Do we have to worry about this? Again, it depends on the accuracy you want. $I = mr^2$ for single particle, so $\ddot{I} = 2m\dot{r}^2 + 2mr\ddot{r}$. For radial free-fall, $\dot{r} = -\sqrt{2GM/r}$, so $\frac{1}{2}\ddot{I} = 2GM^2/R - GM^2/R = GM^2/R$, which is comparable to V or K. Therefore, the moment of inertia term is only comparable to the other terms if $\dot{r}$ is comparable to free-fall.

For our next question, we can ask whether we have to worry about relativistic corrections, given that so far we have used only Newtonian physics. Let's think specifically about whether we need such corrections for the outer portion of the Sun.

For this, we can recognize that special relativity comes in via the Lorentz factor $\gamma = (1 - v^2/c^2)^{-1/2} = (1 - \beta^2)^{-1/2}$, where v is the speed of the particles, c is the speed of light, and $\beta \equiv v/c$. $\gamma = 1$ is Newtonian, and we see that γ deviates from 1 by of order β^2 to lowest order. From the virial theorem, the characteristic thermal velocity will be about equal to the Keplerian orbital velocity. That would be $\sqrt{GM/r}$, or $v = 4.3 \times 10^7$ cm s^{-1}. Since $c \approx 3 \times 10^{10}$ cm s^{-1}, the ratio is then $\sim 10^{-3}$ and thus the fractional correction is $\sim 10^{-6}$, which is also of the order kT/mc^2 for particle mass m. That can be neglected for the Sun, but if you do the same exercise for more compact stars it can be a larger correction; for example, for neutron stars the equivalent speed ("equivalent" because we actually have degenerate matter and quantum weirdness) is maybe half the speed of light. In the center of the Sun, sharing of energy between electrons and protons means that $T = 1.5 \times 10^7$ K for both, which turns out to mean 0.3% corrections for electrons. This can be even higher for more massive star late in their evolution, so you'd want to keep an eye out for such corrections if you need a precise answer.

Mean Molecular Weights

Before proceeding further, we need to define a quantity called mean molecular weight, which is often denoted by μ. It is simply the average mass per particle (or subset of particles), in units of a proton rest mass. For example, suppose we have only neutral atomic hydrogen. The mean molecular weight is essentially 1, because for each particle (an atom) there is roughly one proton mass (the electron is negligibly massive). Now suppose we have fully ionized atomic hydrogen. The number of particles is now doubled, since where before we had just a single atom, now we have a proton and an electron. The mean molecular weight is then 1/2. The mean molecular weight per electron is $\mu_e = 1$, because for every electron there is one proton. As yet another example, suppose we have pure neutral helium. Here we have four proton masses (really two proton and two neutron masses, but that's close) per atom, so $\mu = 4$. If the helium were completely ionized, we'd now have three times as many particles (one nucleus plus two electrons), so $\mu = 4/3$. The mean molecular weight per electron is $\mu_e = 4/2 = 2$.

The use of this comes in different ways: for example, at a given temperature each particle contributes the same amount of pressure if it's close to an ideal gas, because then $P = nkT$. Thus, for a given pressure the required temperature is greater if there are fewer particles. Consider fusion inside a main sequence star. Everything is fully ionized, and four protons (plus four electrons) are converted into one helium nucleus plus two electrons. The net result is a decrease of pressure, which thus requires that the core heat up.

A Simple Model

Now let's think about a simple stellar model and see what it can tell us about stars. In fact, let's pick something stupidly simple: a star with constant density $\rho = \rho_c$(!) that has a sharp cutoff at the stellar radius R. This is crazy, of course, but it's easy to calculate and gives some insight.

The mass interior to some radius r is then $M_r = \frac{4}{3}\pi r^3 \rho_c$. This is true all the way up to the surface $r = R$, where $M_r = M$. Then

$$M_r = \frac{r^3}{R^3} M \ . \tag{3}$$

We now insert this into the Lagrangian form of the equation of hydrostatic equilibrium:

$$\frac{dP}{dM_r} = -\frac{GM_r}{4\pi r^4} = -\frac{GM}{4\pi R^4}\left(\frac{M_r}{M}\right)^{-1/3} \quad . \tag{4}$$

The pressure at the surface is zero, so with this as a boundary condition we can integrate to get

$$P = P_c\left[1 - \left(\frac{r}{R}\right)^2\right] \tag{5}$$

where $P_c = (3/8\pi)(GM^2/R^4)$ is the central pressure. This is actually a lower limit to what the pressure can be in a hydrostatic star. One can also convince oneself that GM^2/R^4 is the only combination of G, M, and R that gives units of pressure, so the central pressure always has to be proportional to that. It is straightforward to verify the virial theorem for this case. If the pressure is supplied by an ideal gas, $P = nkT$ (where, remember, $n \propto 1/\mu$), this implies a central temperature of

$$T_c = 1.15 \times 10^7 \mu (M/M_\odot)(R/R_\odot)^{-1}\,\mathrm{K}\ . \tag{6}$$

This is actually more reliable than what we got using the simple virial estimate, because we're thinking now about the central temperature instead of some "average" temperature.

Appendix: derivation of the virial theorem

So that our notes are self-contained, here we reproduce the derivation of the virial theorem that is in our textbook.

We start by considering $\sum_i \mathbf{p}_i \cdot \mathbf{r}_i$, where there are i particles in the system and each has a radial vector $\mathbf{r}_i$ and a momentum vector $\mathbf{p}_i$ relative to some frame of reference. Assuming nonrelativistic mechanics (which is good for ordinary stars, as we saw above), the time derivative of this quantity is

$$\begin{aligned} \frac{d}{dt}\sum_i \mathbf{p}_i \cdot \mathbf{r}_i &= \frac{d}{dt}\sum_i m_i \dot{\mathbf{r}}_i \cdot \mathbf{r}_i \\ &= \frac{1}{2}\frac{d}{dt}\sum_i \frac{d}{dt}(m_i r_i^2) \\ &= \frac{1}{2}\frac{d^2}{dt^2}\sum_i m_i r_i^2 \\ &= \frac{1}{2}\frac{d^2 I}{dt^2}\ , \end{aligned} \tag{7}$$

because $I = \sum_i m_i r_i^2$.

Now we consider the same time derivative as before but manipulate it differently. We start with

$$\frac{d}{dt}\sum_i \mathbf{p}_i \cdot \mathbf{r}_i = \sum_i \frac{d\mathbf{p}_i}{dt} \cdot \mathbf{r}_i + \sum_i \mathbf{p}_i \cdot \frac{d\mathbf{r}_i}{dt}\ . \tag{8}$$

From Newton's second law we know that $d\mathbf{p}_i/dt = \mathbf{F}_i$, where $\mathbf{F}_i$ is the force on particle i. We also know that since (again in Newtonian mechanics) $\mathbf{p}_i = m_i \mathbf{v}_i$, and generally $d\mathbf{r}_i/dt = \mathbf{v}_i$, $\sum_i \mathbf{p}_i \cdot (d\mathbf{r}_i/dt) = \sum_i m_i v_i^2 = 2K$, where K is the total kinetic energy of the system. Putting this together gives an intermediate result:

$$\frac{1}{2}\frac{d^2 I}{dt^2} = 2K + \sum_i \mathbf{F}_i \cdot \mathbf{r}_i\ . \tag{9}$$

We therefore need to figure out what the last term is. As a first step, we realize that the force on a given particle equals the sum of the forces from all of the other particles. Thus we can write

$$\sum_i \mathbf{F}_i \cdot \mathbf{r}_i = \sum_{i,j;i<j} \left(\mathbf{F}_{ij} \cdot \mathbf{r}_i + \mathbf{F}_{ji} \cdot \mathbf{r}_i\right)\ , \tag{10}$$

where F_{ij} is the force on particle i due to particle j. Newton's third law says that $\mathbf{F}_{ij} = -\mathbf{F}_{ji}$, so we can write this as

$$\sum_{i,j;i<j} \mathbf{F}_{ij} \cdot (\mathbf{r_i} - \mathbf{r_j}) \ . \tag{11}$$

If the force is given by Newtonian gravity, then

$$\mathbf{F}_{ij} = -\frac{Gm_i m_j}{r_{ij}^3} (\mathbf{r}_i - \mathbf{r}_j) \ , \tag{12}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, which means that

$$\sum_i \mathbf{F}_i \cdot \mathbf{r}_i = \sum_{i,j;i<j} \mathbf{F}_{ij} \cdot (\mathbf{r_i} - \mathbf{r_j}) = -\sum_{i,j;i<j} \frac{Gm_i m_j}{r_{ij}} \ . \tag{13}$$

This is simply the total gravitational potential energy V, so we get finally

$$\frac{1}{2}\frac{d^2 I}{dt^2} = 2K + V \tag{14}$$

as promised.