N-body problem

There's a lot of interesting physics related to the interactions of a few bodies, but there are also many systems that can be well-approximated by a large number of bodies that interact exclusively by gravity (meaning that, e.g., physical collisions are rare). **Ask class:** can they think of some examples? Stellar systems, e.g., open or globular clusters, or galaxies, are a good example because the distance between stars is much greater than the size of the stars, usually. Dark matter particles in a halo are another example.

It's completely hopeless to attempt to solve such a many-body problem exactly. However, we can get a remarkable amount of understanding of some systems if we go to the extreme of having lots and lots of particles. That's because one can then make statements about the statistical average of various interesting properties. That is, although prediction of the movement of an individual star in a cluster is impossible in the long run, we can say how the cluster as a whole will evolve. Once again we see echoes of thermodynamics: you're not going to predict the motion of an individual molecule, but you can say how the pressure, temperature, etc. of a system will evolve with high accuracy if there are enough molecules.

To start, let's notice something curious about a particle in a circular orbit, where we back up to our one-body motion. This particle has an orbital speed of $\sqrt{GM/r}$, and therefore a kinetic energy of $K = \frac{1}{2}(GMm/r)$, where m is the mass of the particle, M is the mass of the object in the center, and r is the radius of the orbit. The gravitational potential energy is $W = -GMm/r$, so $W = -2K$ or $W + 2K = 0$. This may seem a special statement about circular orbits, but amazingly we can prove a much more general statement, called the *virial theorem*, that applies to self-gravitating systems in general.

To start, note that the equation of motion for some particle i out of a collection of particles is

$$m_i\ddot{\mathbf{r}}_i = -\sum_{j\neq i}\frac{Gm_jm_i(\mathbf{r}_i-\mathbf{r}_j)}{|\mathbf{r}_i-\mathbf{r}_j|^3}\ . \tag{1}$$

Now take the dot product with $\mathbf{r}_i$.

$$m_i\mathbf{r}_i\cdot\ddot{\mathbf{r}}_i = -\sum_{j\neq i}\frac{Gm_jm_i\mathbf{r}_i\cdot(\mathbf{r}_i-\mathbf{r}_j)}{|\mathbf{r}_i-\mathbf{r}_j|^3}\ . \tag{2}$$

Summing over i, we get

$$\sum_i m_i\mathbf{r}_i\cdot\ddot{\mathbf{r}}_i = -\sum_i\sum_{j\neq i}\frac{Gm_jm_i\mathbf{r}_i\cdot(\mathbf{r}_i-\mathbf{r}_j)}{|\mathbf{r}_i-\mathbf{r}_j|^3}\ . \tag{3}$$

Note that this sum is essentially a sum over a square matrix, i by j, that is missing the diagonal. It is therefore the sum of two triangles in this matrix, which may be reexpressed as

$$\sum_i m_i \mathbf{r}_i \cdot \ddot{\mathbf{r}}_i = -\sum_i \sum_{j=1}^{i-1} G m_j m_i \frac{\mathbf{r}_i \cdot (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} - \sum_j \sum_{i=1}^{j-1} G m_j m_i \frac{\mathbf{r}_i \cdot (\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3} \ . \tag{4}$$

Since i and j are dummy indices, we can switch their names in the second term and combine them to get

$$\begin{aligned} \textstyle\sum_i m_i \mathbf{r}_i \cdot \ddot{\mathbf{r}}_i &= -\textstyle\sum_i \sum_{j=1}^{i-1} G m_i m_j (\mathbf{r}_i - \mathbf{r}_j) \cdot (\mathbf{r}_i - \mathbf{r}_j)/|\mathbf{r}_i - \mathbf{r}_j|^3 \\ &= -\textstyle\sum_i \sum_{j=1}^{i-1} G m_i m_j / |\mathbf{r}_i - \mathbf{r}_j| \\ &= W \ . \end{aligned} \tag{5}$$

Therefore, $\sum_i m_i \mathbf{r}_i \cdot \ddot{\mathbf{r}}_i$ is the potential energy. Having established that, now consider

$$C \equiv \sum_i \mathbf{p}_i \cdot \mathbf{r}_i = \sum_i m_i \dot{\mathbf{r}}_i \cdot \mathbf{r}_i = \frac{1}{2}\frac{d}{dt}\sum_i m_i |\mathbf{r}_i|^2 = \frac{1}{2}\frac{d}{dt} I \ , \tag{6}$$

where $I \equiv \sum_i m_i |\mathbf{r}_i|^2$ is the moment of inertia. Then

$$\begin{aligned} dC/dt = \tfrac{1}{2} d^2 I/dt^2 &= \textstyle\sum_i m_i \ddot{\mathbf{r}}_i \cdot \mathbf{r}_i + \sum_i m_i \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i \\ &= \textstyle\sum_i m_i \mathbf{r}_i \cdot \ddot{\mathbf{r}}_i + 2 \sum_i m_i |\dot{\mathbf{r}}_i|^2/2 \\ &= W + 2K \ , \end{aligned} \tag{7}$$

where $K \equiv \sum_i m_i |\dot{\mathbf{r}}_i|^2/2$ is the kinetic energy. Therefore, we have derived the virial theorem:

$$W + 2K = \frac{1}{2}\frac{d^2 I}{dt^2} \ . \tag{8}$$

Whew! But what does this mean, physically? First of all, **Ask class:** what is the statement of this theorem for a time-independent system? Time independence means that all time derivatives are zero, so we would simply have $W + 2K = 0$. In fact, that is true as long as the moment of inertia $I = \sum_i m_i |\mathbf{r}_i|^2$ is constant or varying at a constant rate (so that the second derivative is zero). For example, for a particle in a circular orbit, the distance from the center of force is a constant, so I is constant and therefore $W + 2K = 0$, as we found.

But what if the system is varying? At first glance we're hosed, because we'd need to determine $d^2 I/dt^2$ in an extremely complicated situation. However, the key here is to take a time average. If a system is roughly in statistical equilibrium, then the time average $\langle I \rangle$ is constant, so that over time the second derivative is small and can be neglected. In fact, we can go farther than that. Suppose that the "motion" of a system is *bounded* in physical space and momentum space. Then, over a very long time compared to oscillation periods, I goes through maxima and minima, but its time average is some constant. Therefore, for any bounded system, $d^2\langle I \rangle/dt^2$ tends to zero over a sufficiently long period. Then again we have $\langle W \rangle + 2\langle K \rangle = 0$, time-averaged. **Ask class:** can they think of a counterexample that doesn't follow this? An explosion is an example. If all the particles are flying off to infinity, then the system is not bounded and the time-average of $\langle W \rangle + 2\langle K \rangle$ is not zero. For most systems we care about though, it is. For example, a particle in an elliptical orbit has W and K that vary over the orbit, but the time-average follows $\langle W \rangle + 2\langle K \rangle = 0$.

This is a profoundly useful theorem. Again, it doesn't say anything about a particular particle. However, it does relate the physical size of a system to its velocity dispersion and mass, or its physical size to its mass and total energy. Let's look at some examples.

First, consider a star cluster. Suppose it has mass M and radius R. **Ask class:** not bothering with factors of order unity, what is the approximate potential energy of the cluster? It is roughly $-fGM^2/R$, where the factor f depends on the exact mass distribution. Therefore, **Ask class:** what is the approximate typical speed? From the virial theorem, the speed must be $v \approx f^{1/2}\sqrt{GM/R}$. We'd like to be able to use this to estimate the mass of the cluster. **Ask class:** suppose we are observing a globular cluster (nice and spherical!). What quantities can we measure with bearing on the mass, using the virial theorem? We can estimate the radius of the cluster, based on its angular radius and some measure of the distance to the cluster. We can also, star by star or as a whole, measure the radial component of the speed because of Doppler shifts. However, the transverse components are a mystery because usually a cluster is too far away to see proper motions. **Ask class:** for a spherical cluster, what is a reasonable approximation that may allow us to move from the line of sight speed to the total speed? We can assume that there has been enough scrambling of velocity directions that, on average, the speeds are distributed isotropically. Suppose we measure the line of sight speed for a number of stars, and come up with a velocity dispersion. Since we measure one component (call it σ_v), we can assume that the total squared three-dimensional velocity dispersion is $3\sigma_v^2$ for an isotropic distribution.

But how does this get us a mass? For that we have to get some estimate of how the stars are distributed. This can come from observations of the light: we see a projection of the total light, from which we make educated guesses about the three-dimensional distribution. For example, if the stars are distributed uniformly then

$$W = -\frac{3}{5}\frac{GM^2}{R} \ . \tag{9}$$

The virial theorem then gives us

$$\begin{aligned} 2K &= -W \\ 2(\tfrac{1}{2}M3\sigma_v^2) &= 3GM^2/(5R) \\ M &= 5R\sigma_v^2/G \ . \end{aligned} \tag{10}$$

The exact coefficient will be different from 5 if the mass is distributed non-uniformly. However, the point is that one can "weigh" a cluster using the virial theorem. If the stars are moving in a disk (as in a spiral galaxy) then it is more useful to use Kepler's laws to figure out how much mass is interior to the orbit of a particular star, but it's the same idea. The net result of all this is that all galaxies with good data show evidence of greater gravitational attraction than is accounted for by the visible stars and gas. In fact, the bigger the collection of mass, the larger the correction factor; clusters of galaxies show this, too. This is known as "dark matter", and is thought to add up to something like six times as much mass as all the baryons in the universe. Oddly, globular clusters show no evidence for dark matter. The nature of dark matter is a mystery. It has to be something other than baryons, based on constraints from big bang nucleosynthesis. The leading candidate is that dark matter is some form of elementary particle that interacts only by gravity (i.e., it doesn't collide or radiate). Still, no particular candidate has been verified yet.

The virial theorem also allows insight into cosmological structure formation. The basic idea behind structure formation is that in the early universe various processes imprinted fluctuations in the density. That is, some regions of the universe were a little denser than others (and some a little less dense). The slightly denser regions contracted gradually under their self-gravity, and eventually settled into equilibrium. But how much did they eventually contract?

Consider a cloud of particles of mass M and initial radius R. Let the particles interact only by gravity. Assume that at the beginning the particles are moving very slowly, so that their kinetic energies can be neglected. **Ask class:** how can we use the virial theorem to estimate the final radius of the cloud in equilibrium? With purely gravitational interactions, the energy of the system is conserved (e.g., no energy escapes to infinity as radiation). The original energy of the system is purely potential energy, of magnitude $-GM^2/R$. The total energy, $E_{\rm tot} = W + K$, is conserved. From the virial theorem, $K = -W/2$, so $E_{\rm tot} = W - W/2 = W/2$. Therefore, $W/2 = -GM^2/R$. If the equilibrium radius is $R_{\rm eq}$ then $W = -GM^2/R_{\rm eq}$, so $R_{\rm eq} \approx R/2$! The exact amount of contraction again depends on how the stars are distributed, but the radius contracts by approximately a factor of two, so that the density goes up by about a factor of 8.

In cosmological structure formation theory, now beautifully confirmed by results from the Wilkinson Microwave Anisotropy Probe (among other experiments), the initial fluctuations in density are largest at small scales. That means that, all else being equal, structure will form at small scales first. Since the universe is expanding, the average density of the universe is going down with time. Therefore, consider a density enhancement at small spatial scales. It collapses out and forms a system in virial equilibrium at, say, 8 times the surrounding density. Since this happens at an early time, the surrounding density was relatively large and thus the system has high density itself. Now consider a density enhancement at large spatial scales. It also collapses out and forms a system with 8 times the surrounding density, but by this later epoch the surrounding density is less, so the system itself has a lower density in virial equilibrium. What we therefore expect (and see!) is that lower-mass systems have higher average density than higher-mass systems. For example, a globular cluster might have $10^5\,M_\odot$ in a radius of 10 pc, for an average density of $\bar{\rho} \approx 20\,M_\odot$ pc^{-3}. The Milky Way has a mass of about $10^{12}\,M_\odot$ (including its dark matter halo) in a radius of about 10^5 pc, for an average density of about $2 \times 10^{-4}\,M_\odot$ pc^{-3}. Galaxy clusters are even less dense.

It is thought that larger systems were formed by the hierarchical assembly of smaller systems. That is, the little stuff formed clumps, but then those clumps often assembled into bigger super-clumps, and so on. We now see structure on size scales out to ~ 100 Mpc, the scale of superclusters.

There's one more comment to make, which will lead us into the next class. When we start with a low kinetic energy cloud and let it collapse into virial equilibrium, it's a violent process. The particles essentially free fall to the center, where the rapidly changing potential scrambles their velocities. This is a process called violent relaxation (I love that term!). Effectively, it means that a cluster can settle into something close to virial equilibrium in the short time that it takes to fall. After a couple of bounces, it is virialized (meaning that its moment of inertia is changing slowly if at all). You might think that would be the end of it, and that a cluster could remain in this "equilibrium" indefinitely. But it's not true, so in the next class we need to think about how something in virial equilibrium evolves.