LECTURE XII: THE VISIBLE UNIVERSE II: THE GALAXIES WE SEE

AST 204 March 12 2008

These notes contain material from previous classes and more detail than we can cover in class, for the sake of completeness

I. Collapse and Violent Relaxation: The Formation of Dark Matter Halos

We have seen that if we have a tophat perturbation, the body of the perturbation collapses all at the same time, and the surrounding material out to some r_{last} and corresponding m_{last} rain in on top of it it in some slowly tailing-off infall. A more realistic perturbation will have some central volume of roughly constant contrast which will collapse at roughly the same time, will in general not be so violent as the tophat case, but will not differ much in qualitative terms. We saw that the power spectrum of perturbations on the small scales we are considering is roughly k^{-3}, which corresponds to roughly equal 'roughness' on all scales, so the tophat in the real universe will itself be pretty lumpy. These lumps will pull other lumps away from their purely radial orbits a little, so the mass will in fact *not* all bang into itself at the collapse time τ_c. Thus there will be some minimum size at about τ_c, but there will be lots of kinetic energy and the system will be very lumpy (because all the little lumps have grown during the collapse as well.) It will expand again, bounce around, and finally settle down to some equilibrium configuration. Computer simulations show that it is pretty settled in a short time, about $\tau_c/2$ after the initial collapse, so after a total elapsed time of about $3\tau_c/2$.

During these early phases, the system undergoes what is called *violent relaxation*—energy is exchanged between the particles in lumps on all scales, but it is the interaction of individual particles with lumps which is important during this phase. In a steady, equilibrium system made of stars and gas, the total energy of a particle, kinetic plus potential, is conserved, but in this violent phase, though the total energy of the *system* is conserved, the energy of individual particles is not. Instead, energy is exchanged rapidly. The energy of individual particles is not conserved whenever the gravitational *potential* is changing, and in fact it can be shown that for an individual particle,

$$\frac{dE}{dt} = \frac{\partial \Phi}{\partial t}. \tag{1}$$

We will see in numerous instances that the natural timescale in any gravitational problem is of order $1/\sqrt{G\rho}$; we have, in fact, already seen this. The Hubble constant is the inverse of the timescale for expansion of the universe, and is of order $\sqrt{8\pi G\rho/3}$; the collapse time of our perturbation is $\pi\sqrt{r_{max}^3/2Gm} = \sqrt{3\pi/2G\rho}$. When any self-gravitating system does something in response to its own gravity, this is always the timescale. (Just for fun, think about the orbital time of the earth, and show that it is of order $1/\sqrt{G\rho_m}$, where ρ_m is the density the sun would have if it were smeared out over a sphere whose radius is that of the earth's orbit). If the potential is changing on this timescale, the timescale of a particle orbit in the system, then particles essentially forget their initial energy and initial velocities; they are scattered effectively both in velocity and energy.

The result of this is statistically rather interesting. During this phase a particle feels numerous little kicks in each component of velocity which are essentially statistically independent as it interacts with various lumps in the rapidly changing environment, and the total change in velocity is large enough that it forgets what its initial velocity was. There is a very powerful theorem in statistics, the Central Limit Theorem, which says that if some variable (say the x component of the velocity) is a sum of a very large number of statistically independent variables (the little changes) then the sum (the final velocity) has a probability distribution which is *normal*:

$$P(v_x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp(-v_x^2/2\sigma^2), \tag{2}$$

the classical 'bell curve'. Here the *variance* σ^2 of the distribution is the square root of the sum of the mean squares of the kicks in velocity it got along the way. In our example, the mean kick is assumed to be zero, which it must be under the assumption of spherical symmetry; if it is not, the quantity in the numerator in the exponent is $-(v_x - \bar{v}_x)^2$, and $\bar{v}_x$ is the *mean x-velocity*, which is the sum of the *mean* of the kicks. The quantity σ is called the *dispersion* or, more correctly the *standard deviation*, and is often referred to as if 'sigma' was its name–a value which lies three standard deviations from the mean is referred to as a '3-sigma' point. The multiplier, $1/\sqrt{2\pi\sigma^2}$ is just to make the integral (probablilty that v_x has *some* value) unity. Physicists and astronomers do not use the term 'normal', though everybody else in the world does; they instead call this a 'gaussian' distribution, but they mean the same thing. So we can expect the velocity distribution of our final object to be roughly gaussian or normal.

II. Hydrostatic Equilibrium

What is its density distribution like once it settles down? To think about this, we need to consider the general equation which describes a system which has established an equilibrium with its own self-gravity and internal motions.

What are the forces in such a system? First, there is gravity. The potential energy w of a particle of mass m in a gravitational field is $m\Phi$ (indeed, this is the *definition* of the potential Φ); the potential Φ in a spherical system outside the mass is just, of course, GM/r, where M is the mass of the *system*. The *force* is just the negative of the gradient of the potential energy because the work done in moving a particle from $\mathbf{x}$ to $\mathbf{x} + \mathbf{dx}$ is just the change in potential energy:

$$\text{work} = -dw = \text{force} \cdot dx,$$

or

$$\text{force} = -dw/dx = -md\Phi/dx$$

Where we have written things as if they dependended on only one variable x; in three dimensions, this becomes, clearly

$$\text{force} = \mathbf{f} = -m\nabla\Phi.$$

So suppose the final system is spherical, and consider a little volume element oriented so that it has faces perpendicular to the radius, has thickness dr in the radial direction, and has faces of area A. The density in the volume is ρ, so the mass is $dm = \rho A dr$, and the gravitational force on the mass in the volume is radial and is $-dm \cdot d\Phi/dr = -dm \cdot GM_r/r^2$, where M_r is the mass contained within r. The system is in equilibrium by assumption, so the *mean radial velocity* of particles in the box must be zero. If not, the body would either be expanding or contracting. In general, the velocity of particles at some point in the body can be broken down into two parts:

$$\mathbf{v} = \bar{\mathbf{v}} + \mathbf{u}, \tag{3}$$

where $\bar{\mathbf{v}}$ is the mean velocity in a small region surrounding the point, and $\mathbf{u}$ is the difference from the mean. $\mathbf{u}$ is called the *peculiar, or random, velocity.* In our spherical example, $\bar{\mathbf{v}} = 0$. The peculiar velocities exert *forces* in the system called *pressure forces.* To see this, imagine replacing the imaginary faces of our volume with perfect reflecting surfaces. The particles bouncing off these faces clearly exert a force on the face, but since the mean velocity of the particles is zero at the face, inserting the reflecting face can in no way alter the physics of the problem. What is the force? The particles hit the face at a rate $An(u_r)u_r du_r$, where $n(u_r)du_r$ is the number density of particles at u_r in du_r. The change in momentum of each particle when it hits the face is $2mu_r$. The rate of change of momentum is the force, so for the particles of peculiar velocity u_r the force is $df = 2n(u_r)mu_r^2 du_r$. If we are looking at the lower face, we are only interested in particles with *positive* u_r, because they are the only ones which hit the face. So the total force is

$$f_r = 2A \int_{(u_r>0)} mn(u_r)u_r^2 du_r = A \int mn(u_r)u_r^2 du_r,$$

where in the last expression we recognize that we are integrating only over half the particles (the upward-moving ones), but that just cancels the factor of 2 in front. But the particle mass times the number density is just the mass density ρ. So we can write

$$f_r = A\rho \bar{u_r^2} = A\rho\sigma_r^2$$

The quantity σ_r is called the *radial velocity dispersion*, and the product with rho is called the *pressure*. If particles with different masses, if we have different masses, have different velocity distributions, this is slightly more complex, but you get the idea. Now the box must neither rise nor fall, (more accurately, be accelerated neither up nor down) which means that the gravitational force on the box must just be balanced by the pressure difference on its faces:

$$-\rho A dr d\Phi/dr = -\rho GM_r/r^2 A dr = AdP$$

or

$$\frac{dP}{dr} = \frac{d(\rho\sigma^2)}{dr} = -\rho G M_r / r^2 \tag{4}$$

This is called the *hydrostatic equation* for spherical geometry. If the system is *not* spherical, then the general form of this is more complex and involves the mean velocity as well (rotation, for example, develops centripetal accelerations, and in a rotatating system the mean velocity is not, of course, zero, even though the system is not changing in time) but this will serve us for the moment. This development is the same for a gas, in which the pressure is a familiar concept, and for a system of stars or dark matter particles. The physical difference is profound; in a gas there are lots of collisions, so it does not make much sense to talk about the *orbit* of a gas particle, and the pressure is very intuitive. In a stellar system or a system of dark matter particles, there are essentially *no* collisions, and particles are in independent orbits, but the hydrostatic equation is the same, as is the *definition* of the pressure. A system such as we are thinking about, in which the mean velocity is everywhere either zero or contributes negligibly to the forces, is called *pressure supported* and is referred to as a *hot* system, which has nothing to do with ordinary temperature but does, as in the ordinary thermal case, refer to large random particle velocities. In the opposite limit, in which, all the particles are in regular circular orbits in a disk about the center of mass of the system, all the velocity is mean or bulk velocity and there is no peculiar or random velocity. These systems are called *rotationally supported* and are referred to as *cold* systems.

III. The Virial Theorem

For systems in hydrostatic equilibrium, such as stars or galaxies or planets or clusters, there is an important and simple relation among the various kinds of energy in the system which enables us to do simple calculations of their properties. To see where this comes from, we consider again a spherical system, though the result is quite general. For each particle in the system, we can write Newton's second law as

$$m\ddot{\mathbf{x}} = -m\nabla\Phi = m\mathbf{g}. \tag{5}$$

Here $\mathbf{g}$ is the gravitational acceleration, which is radial, inward, and has magnitude GM_r/r^2. If we dot this relation with the position vector $\mathbf{x}$, we get

$$m\mathbf{x}\cdot\ddot{\mathbf{x}} = -m\mathbf{x}\cdot\nabla\Phi = -mGM_r/r.$$

The last term on the right comes from multiplying the radial component of $\mathbf{x}$, which is r, by the expression for the gravitional force. Now you can easily verify that

$$d^2\mathbf{x}^2/dt^2 = 2\mathbf{x}\cdot\ddot{\mathbf{x}} + 2(\dot{\mathbf{x}})^2,$$

so we can write

$$\frac{d^2}{dt^2}\left(\frac{m\mathbf{x}^2}{2}\right) - m(\dot{\mathbf{x}})^2 = -\frac{mGM_r}{r}. \tag{6}$$

Let us now sum over all the particles in the system. Then Equation (6) becomes

$$\frac{1}{2}\frac{d^2}{dt^2}\int \rho \mathbf{x}^2 dV - 2T = W, \tag{7}$$

Where we have replaced in the first term the sum over all particles by the integral of the ρdV, which is clearly equivalent if we have lots of particles. The sum over the second term in Equation (6) is clearly the total kinetic energy T of the *system*, and the RHS the total gravitational energy W of the system. If the system is in equilibrium, its shape and density distribution are not changing, so the first term, which is just the second time derivative of the *moment of intertia*, vanishes. This is strictly only true for systems of very large numbers of particles; for smaller systems one needs both to sum over particles *and* average over time, but the result is the same:

$$2T + W = 0 \tag{8}$$

which holds instantaneously for systems with very large numbers of particles, and in time average for any bound system in equilibrium. This is called the *virial theorem.* We have derived it for spherical systems, but it is, in fact, true for any equilibrium system; the only difficulty in the general derivation is some algebraic messiness in the potential energy term, the biggest problem being convincing oneself that the term obtained actually *is* the potential energy and some bookkeeping worries about double counting, etc. Look at a book if you are interested. In the general case, it is often useful to separate the velocity into the average (bulk) velocity and the peculiar velocity, as in Equation (3); when we do this, the virial theorem becomes

$$2T_{av} + 3P_{tot} + W = 0 \tag{9}$$

for the case in which the velocity distribution is *isotropic*, so $\sigma_x^2 = \sigma_y^2 = \sigma_z^2$, and here T_{av} is the kinetic energy associated with the mean velocity and P_{tot} the integral of the pressure over the volume of the system:

$$T_{av} = \frac{1}{2}\int \rho \bar{v}^2 dV$$
$$P_{tot} = \int P dV$$

This provides a way of calculating the kinetic energy of the system without having to worry about the details of the motion of individual particles; only the mean velocity and the pressure appear. Why is there no cross term when you square the velocity of Equation (3) and sum everything up???

To see how one might use this relation, ask a question. Suppose we have some system of mass M with some velocity dispersion σ. Roughly how big is it when it is in equilibrium? The gravitiational energy of each particle is $-GmM/\bar{r}$, where $\bar{r}$ is some average radius of the particle in its orbit, so it must be that the gravitational energy of the *system* is, just summing over the particle masses, $W = -GM^2/R_g$, where R_g is some average of the radii over the particle orbits of *all* the particles. The quantity R_g is clearly a reasonable measure of the size of the system, and is called the *gravitational radius* for obvious reasons. The kinetic energy is, if the velocity dispersion σ is reasonably constant over the system, $T = 1/2M\sigma^2$, so the virial theorem says

$$M\sigma^2 = \frac{GM^2}{R_g}$$

or

$$R_g = \frac{GM}{\sigma^2}$$

.

Now it is a happy coicidence that there is a very close but essentially coincidental relation between the gravitational radius and the so-called 'half-mass' radius $R_{1/2}$, the radius containing half the mass of the system, which is usually much easier to compute than R_g and is a much more intuitive notion of the size of a system. Over a very large range in density distributions, from uniform spheres, exponential spheres, power laws, *etc*, *etc*, the relation

$$R_g \approx 2R_{1/2}$$

holds to an accuracy of about ten percent, and better for realistic density distributions. Notice that for a system in equilibrium the *total energy*, sometimes called the *binding energy*, is negative and since $E = T + W$, $2T + W = 0$, it must be that

$$\begin{aligned} E &= -T \\ &= W/2 \end{aligned}$$

IV. The Dark Matter: Isothermal Spheres

After the collapse and violent relaxation, the dark matter in a body which forms out of the fluctuations in the early universe does not change much unless it merges with another. The baryonic matter can do interesting and wonderful things like make stars and planets and people, but the dark matter particles can only move in their more-or-less original orbits. We do not see the dark matter directly, but know it is there and can measure its gravitational effect. How do we do this?

First of all, what forms do we expect the dark matter to take–that is what *shapes* of the density distributions? Remember that we expected the velocity distribution to be roughly normal and for all particles to have roughly the same velocity dispersion. If we look at the equation of hydrostatic equiblibrium, we thus have, approximately,

$$\begin{aligned} \frac{dP}{dr} &= \frac{d(\rho\sigma^2)}{dr} \\ &= \sigma^2\frac{d\rho}{dr} = \frac{-G\rho M_r}{r^2} \end{aligned} \tag{10}$$

Can we find a simple solution for this equation? Let us look for a power law. Since we know that the hydrostatic equation requires a *decreasing* density with radius, look at solutions of the form $\rho = \rho_0(r/r_0)^{-\alpha}$. If this the the density, the contained mass M_r is

$$\begin{aligned} M_r &= \int_0^r dm \\ &= \rho_0 \int_0^r \left(\frac{r'}{r_0}\right)^{-\alpha} 4\pi {r'}^2 dr' \\ &= 4\pi\rho_0 r_0^3 \frac{1}{3-\alpha}\left(\frac{r}{r_0}\right)^{3-\alpha}. \end{aligned} \tag{11}$$

Note that this is not correct if $\alpha \geq= 3$, in which case the mass is divergent as one approaches $r = 0$, and *no* power law is both well behaved at the origin and contains finite mass as $r \to \infty$. We will see whether there is a satisfactory solution for Equation (10). Substituting our expression for ρ and the expression (11) for M_r into the hydrostatic equation, we get

$$-\alpha\sigma^2\rho_0 r_0^{-1}\left(\frac{r}{r_0}\right)^{-\alpha-1} = \frac{-4\pi G\rho_0^2 r_0}{3-\alpha}\left(\frac{r}{r_0}\right)^{1-2\alpha}$$

Since the powers of r have to be the same on both sides,

$$1 - 2\alpha = -\alpha - 1,$$

or $\alpha = 2$, and

$$\begin{aligned} \sigma^2 &= 2\pi G\rho_0 r_0^2 \\ M_r &= 4\pi\rho_0 r_0^2 r \\ &= 2\sigma^2 r/G \\ \rho &= \sigma^2/2\pi G r^2 \end{aligned} \tag{12}$$

Notice that the *density* goes to infinity at the origin, but it is an integrable singularity and the contained mass is proportional to the radius all the way to the origin. This solution is called the *singular isothermal sphere*. If we demand that the density be finite at the origin, we discover that it falls slowly at first, approximately like $\rho = \rho_c\left[1 + (r/r_c)^2\right]^{-1}$, reaching about half its central value ρ_c at the *core radius* $r_c = 3\sigma^2/(2\pi G\rho_c)$. Thereafter it wiggles a bit and settles down to the behavior of the singular model.

Before we leave the isothermal sphere, let us ask a simple question. Suppose we have a test particle in a circular orbit of radius r about the center of the structure. Then the mass contained within r is $2\sigma^2 r/G$. In a circular orbit, the centripetal acceleration just balances the gravitational force, so the circular *velocity* v_c is

$$\begin{aligned} \frac{v_c^2}{r} &= GM_r/r^2 \\ v_c^2 &= GM_r/r \\ &= 2\sigma^2 \\ v_c &= \sqrt{2}\sigma \end{aligned} \tag{13}$$

and is *constant* with radius. Thus if we can measure the circular velocity of some test particles in a dark matter isothermal halo, we can determine the velocity dispersion in the halo and, if the circular velocity is constant, demonstrate that it is approximately isothermal. This is a very powerful result.

We have seen that the continued infall keeps adding material to the object we have formed; in the problem set you will show that this material forms a halo with a density distribution which is approximately $\rho \propto r^{-9/4}$, only trivially different from the expected r^{-2} in the central parts. This will cut off quite sharply as the last bound shell is approached.

Let us now think about the dynamics of the final object. How can we connect the present quantities (velocity dispersion, size) of the object with its properties before it formed and indeed with its initial conditions?

V. Dynamics Before and After Collapse

Consider the perturbation near the epoch of maximum expansion. It is approximately a uniform sphere, and you will show in the problem set that its gravitational energy (which is its total energy, because it has no kinetic energy at that epoch) is

$$E = -\frac{3GM^2}{5R_{max}}.$$

After it has collapsed and stabilized (we use the term *virialized*, which means that it has settled enough to satisfy the virial theorem), we have $E = -T = W/2$

$$E = \frac{GM^2}{2R_g} \approx \frac{GM^2}{4R_{1/2}}.$$

Thus the half-mass radius of the final configuration is, roughly,

$$R_{1/2} \approx \frac{5}{12}R_{max} \approx 0.4R_{max}$$

So the final dark matter configuration is roughly *half* the size at maximum expansion. If we can determine any two of the half-mass radius, the mass, or the velocity dispersion, we can determine a great deal about the perturbation that made the system, since these three quantities are related by

$$T = \frac{3M\bar{\sigma}^2}{2} = |E| \approx \frac{GM^2}{4R_{1/2}},$$

or

$$\bar{\sigma}^2 = \frac{GM}{6R_{1/2}}$$

so we can, having $R_{1/2}$, get $R_{max} \approx 2.5R_{1/2}$, and thus determine the collapse time

$$\begin{aligned}\tau_c &= \pi \left(\frac{R_{max}^3}{2GM}\right)^{1/2} \\ &\approx 8.2 \left(\frac{R_{1/2}^3}{GM}\right)^{1/2} \\ &\approx 3.3 \frac{R_{1/2}}{\bar{\sigma}}\end{aligned}$$

Notice that the last line of this says that the collapse time is of order the average travel time across the body. Given τ_c, the amplitude of the perturbation can be calculated from the relation we developed last time,

$$\tau_c = \frac{\pi}{H_i} \left(\delta^+\right)^{-3/2}$$

For the Galaxy, the central value of σ is about 160 km/sec, as deduced from the rotation velocity of about 220 km/sec. This value of σ corresponds to a mass of about $10^{10} M_\odot$ per kpc of radius (Equation (3)). We believe that the total mass of the Galaxy is about $10^{12} M_\odot$, so the radius extends to something like 100 kpc, but we do not know exactly what the form of the cutoff is like. The cutoff can *happen* only if the dark matter is colder in the outer parts, so the mean σ will be less than the central value. If we do not worry about this too much, take a mean σ of, say 100 km/sec and a half-mass radius of 50 kpc, we get a collapse time of about 1.7×10^9 years. This is about 7 rotation periods of the Galaxy. The velocities in the collapse near the end must be near the circular velocity, about 200 km/sec.

VI. Angular Momentum

We have seen that gravity causes perturbations to grow, but in the absence of gravity perturbations in *velocity* die away very quickly. You can use the reflecting box trick to show that peculiar velocities are proportional to $1/R$ as the universe expands; all momenta vary thus, and this is another way to get the redshift relation, since the energy of photons is proportional to their momentum. Now angular momentum is a radius times a velocity, so since radii go like R and velocities like 1/R, angular momentum is conserved in the expansion. Cool. Spiral galaxies are typically rotating at 200-300 km/sec today, and a galaxy like the Milky Way has most of its baryonic mass, about $10^{11} M_\odot$ contained within about 10 kpc radius, which corresponds to a baryon density of about $(10^{11} * 2 \times 10^{33}) g/[(4\pi/3) * (3 \times 10^{22} cm)^3] =\approx 2 \times 10^{-23} g/cm^3$. at recombination, the total density was about $3 \times 10^{-21} g/cm^3$ and the baryon density about a tenth of that, say 3×10^{-22}, about a factor of ten higher than the mean spherical density in a spiral galaxy at the present epoch. Thus the mass was contained in a radius smaller by the cube root of 10, about a factor of two, and the velocities thus a factor of two higher, say 500 km/sec. Velocities this high, 1.5×10^{-3} of the velocity of light, would result in fluctuations in the microwave background of this same size, and these are not observed. If one goes back even farther into the past, one soon runs into relativistic velocities, which are clearly at variance with inflation (as are *any* appreciable peculiar velocities), and almost certainly at odds with primordial nucleosynthesis, the exact thermal nature of the background, and many other things. Conserved angular momentum, also, inflates away just as conserved baryon number does in the inflationary scenario, but the observations rule out enough angular momentum early in the universe to explain galaxies even if one does not believe inflation. It seems clear that the rotation does *not* arise from conserved angular momentum from the early universe.

Where else can rotation come from?

Like energy, angular momentum is conserved in the large, but can be exchanged between structures. Consider two nearby tophat perturbations in the universe, neither one completely spherical and aligned randomly with respect to one another. For the sake of argument assume that they are ellipsoids. Then since gravity is a $1/r^2$ force, the *tidal force* exerts a torque on both—the nearer bulge is pulled harder than the farther one, and harder than the center. So relative to the center, the nearer bulge is attracted and the farther is repelled—the mechanism is precisely the same as the tides on the earth, but here we have an intrinsically nonspherical shape. The angular momentum is the torque times the time, and since the shape of the perturbations do not change very much while they are in the linear regime, it is changing the angular momentum in a progressive, steady manner. We could easily estimate the effect knowing what we know, but there is not really time. What we *can* do is figure out how we might characterize it.

How might we characterize the importance of rotation?

Dynamically, that is. One measure is clearly how much *kinetic energy* is associated with rotation versus the total; we can make a dimensionless measure of this by setting a parameter λ

$$\lambda^2 = C T_{rot}/T_{tot} \approx \frac{C}{3}\frac{v_{rot}^2}{\sigma^2} \tag{5}$$

Where C is some number of order unity we can choose later for our convenience, and v_{rot} is some mean rotation velocity. Now from the virial theorem, T_{tot} is the binding energy of the structure, $|E|$, and T_{rot} is $M v_{rot}^2/2$, where v_{rot} is some mean rotation velocity. The angular momentum $J = M R v_{rot}$, where R is some measure of the size of the object. Thus

$$\lambda^2 = \frac{C}{2}\frac{J^2}{MR^2|E|}$$

But $E \approx -GM^2/4R$, $R \approx GM^2/4|E|$, so

$$\lambda = 2\sqrt{2C}\frac{J\sqrt{|E|}}{GM^{5/2}}. \tag{6}$$

If we take $2\sqrt{2C} = 1$, $C \approx 1/8$, this is actually the usual *definition* of λ. The quantity λ, which is called the *angular momentum parameter*, is thus approximately 0.4 for a cold disk ($T_{rot} = T_{tot}$ and is zero for a nonrotating spherical, pressure supported object. How big can we get λ from tidal torques during the formation of structure? It is not hard to see that the effect of the torquing is pretty small on average–the structures as they grow are probably not long, thin bars or dumbbells, which you need to transfer angular momentum efficiently; the near bulge is not *that* much closer to the other perturbation than the far one, it is unlikely that the major axes of the shapes are at 45 degrees to the separation, which one needs for most efficient transfer, *etc., etc.* Both reasonable rough estimates and numerical experiments indicate that, on average, λ is about 0.05, with a large statistical spread, independent statistically of the details of the power spectrum or anything else that has been thought about. It is certainly not big enough to explain disks; with the value of C indicated, the value of lambda for a cold disk is about 0.4; we are short by something like a factor of 8 in rotation velocity and 60 in energy. But spiral disk galaxies are very common, and in fact *most* galaxies are disks.

So the rotation of galaxies is not primordial and apparently cannot be produced by gravitational torques during formation. But wait— we see the fast rotation and thin disks in the *baryons*, *not* in the dark matter. Might this be something we should think about?

VII. The Baryons in the Collapse

What is happening to the baryons during this time? This depends on many factors, not all of which are very well understood in detail, but the general picture is clear. Baryonic material, which is mostly hydrogen gas with small amounts of helium at this time, differs from the dark matter in one very important respect. Gas clouds can collide with one another, but the dark matter particles cannot. We can assume that the gas and dark matter are well mixed at the beginning; actually the gas will be smoother than the dark matter for a while after τ_{eq} because the dark matter perturbations can grow immediately, while the baryonic perturbations are kept from growing by the radiation pressure, but it has been shown that as soon as the hydrogen recombines, lumps in it catch up with the dark matter lumps very quickly. During the expansion of the perturbation, the baryons and dark matter move together; there is no reason for them to separate, because they feel the same gravitational field, and different parts of the perturbation have not tried to collide with each other yet.

When they do, what happens?

The gas, which started all of this out at about 3000K, kT=0.3 eV, at recombination, has been cooling off as the perturbation and the universe expanded. It has been doing so *adiabatically*, that is, neither radiating energy nor absorbing it, because it is too cold to excite even the first level in hydrogen (at about 9 eV) by collisions and the radiation is too cold to excite it radiatively (which is why it recombined in the first place). We can extend our argument about the reflective boxes in an expanding universe to consider the perturbation–the momentum of a particle is inversely proportional to the size of the box, or to its volume to the -1/3 power. But the temperature is proportional to the *square* of the momentum (kinetic energy $\approx$ 3kT/2), so the temperature goes as the -2/3 power of the volume, or as the inverse square of the size. So by the time maximum expansion has been reached, a factor of several hundred larger than the size at recombination, the gas should be very cold, only a few degrees Kelvin. *colder*, in fact, than the radiation, since the latter cools only *inversely* as the size of the universe, and though the perturbation is not expanding as fast as the universe, the inverse square dependence wins.

Now the speed of sound in a gas is of order the velocity of particles in the gas. As a rough rule to remember, a gas of mostly hydrogen at 10,000K has random motions and soundspeed of about 10km/sec, so the particle motions and sound speed are small compared to the velocities in the perturbation as it collapses. These velocities go as the square root of the temperature, of course, since the kinetic energy of the motion is of order kT.

When two of these very cold gas clouds collide with some velocity v_{coll}, the kinetic energy of the collision is immediately randomized into heat in the resulting *shock wave* — the motion is supersonic with respect to the gas, and the disturbance cannot propagate into the gas as fast as the gas itself is moving. Remember that we are talking about velocities in the case of forming galaxies of order 100 km/sec or more. A proton with that velocity has an energy of about 50 eV, more than enough to ionize a hydrogen atom, so the gas after the collision is expected to be ionized. The collision kinetic energy is all in one direction, and the energy gets shared with motions in all three directions after the shock. A collision with a relative velocity of 200 km/s (so, say, two clouds falling in, each moving at 100 km/s) results in random velocities after the collision of about 60 km/sec and temperatures of about 250,000K; most of the mass falls in much faster, and collisions with relative velocities of 400 km/sec result in million-degree gas. This is actually closer to the *characteristic temperature* of the gas — we can ask ourselves what temperature the gas must be in order to have a density distribution like the dark matter, and the answer, if one looks at the hydrostatic equation, is that the ratio of pressure to density must be the same.

For the dark matter, $P/\rho = \sigma^2$. For an ideal gas, $P = nkT = \rho kT/\mu$, where μ is the mean molecular weight. The gas is mostly hydrogen, about a quarter by mass helium, and there is one free electron per hydrogen atom and 2 per helium atom. If one takes a gram of material, therefore, there are $.75/m_p$ protons, $.25/(4m_p)$ helium nuclei, and $.75/m_p + 2 * .25/(4m_p)$ electrons, a total of about $1.7/m_p$ particles. So the mean molecular weight μ is $m_p/1.7 \approx 0.6m_p$, and the temperature is $\mu\sigma^2/k$. For the Galaxy, with $\sigma = 160km/sec$, the equivalent temperature is 1.8 million degrees.

How does such a gas cool? At the lowest temperatures considered here, primarily through recombination. A proton captures an electron and the energy to bind the electron is carried away by a photon. Subsequently more energy is released as the atom cascades to the ground state. The atom will be ionized again soon, but the energy required to do so comes out of the ionizing electron. Basically every recombination/ionization loses the ionization energy plus the mean kinetic energy of an electron, so this energy times the rate of recombinations is the energy loss rate. The recombination rate clearly is proportional to the number density of electrons times the number density of protons, or, since they are proportional (and roughly equal) to each other, to the square of the number density. The rate also depends on the electron velocity; it is much easier for a proton to capture a slow electron than a fast one, but there are more collisions per unit time if the electrons are moving rapidly. These do not cancel, and there is a slowly decreasing velocity (temperature) dependence left.

At higher temperatures, the primary cooling process is *bremsstrahlung*, or *free-free* radiation; this is radiation which occurs as an electron which is not bound to a proton is accelerated by its electric field as it whips past. This dominates for temperatures higher than a few hundred thousand degrees. It must also clearly be proportional to the square of the number density, and here there is a slowly *increasing* temperature dependence, which is why the switch to this process at high temperatures. A simple approximate expression for this cooling (for a gas with the primordial composition) is

$$j = 2.4 \times 10^{-27} T^{1/2} n_e^2 \text{ erg/sec cm}^3,$$

which is the rate of energy loss per cubic centimeter per second from the gas.

A useful thing to look at is the cooling *time* t_{cool}, which is just the characteristic time it would take for the gas to radiate all of its energy by this process. The internal energy of the gas is $(3/2)nkT \approx 3n_e kT$, so

$$\begin{aligned} t_{cool} &\approx \frac{3n_e kT}{j} \\ &\approx 1.7 \times 10^{11} \frac{T^{1/2}}{n_e} \quad \text{sec} \\ &\approx 5 \times 10^6 \frac{T_6^{1/2}}{n_e} \quad \text{yr} \end{aligned}$$

Where T_6 is the temperature in millions of degrees. We saw that the spread-out density of the galaxy was about $3 \times 10^{-22} g/cm^3$, or a number density of protons or electrons about 200 per cubic centimeter. The cooling time for this is incredibly short, only 25,000 years. Even if the baryons are fully distributed like the dark matter, 10 times as large radii, the cooling time is of order 25 million years, short compared to the collapse time.

Suppose there were no cooling. Then, though collisions are important, the energy of the gas would be conserved, and we would expect it to occupy roughly the same volume and be distributed in roughly the same way as the dark matter. Then if one turned on cooling suddenly, at least for a structure like the galaxy, the baryons would radiate away their internal energy on a timescale short compared to the gravitational timescale. But the internal energy density is just proportional to the pressure, so the gas can no longer support itself, and it shrinks with respect to the dark matter.

How far? Well, until something stops it, because as it gets denser and denser it radiates more and more efficiently. What can stop it? Only two things, as far as we know. We know, and there will be much discussion of this later, that gas in galaxies eventually makes stars, by processes which are more-or-less well understood but much too complex to calculate accurately. Once this happens, the 'gas' of stars behaves just like the dark matter; physical collisions between stars are incredibly rare, and the stars do not lose energy any more, but orbit in the combined potential of other stars and the dark matter.

The other thing that can stop it is angular momentum. Let us think about this carefully. The dark matter has an amount of angular momentum from tidal torques which corresponds to some angular momentum parameter, which in the mean is about 0.05. The baryons and dark matter were well mixed while most of this was going on, so the ratio J/M is similar for both. If now the baryons sink in the potential of the dark matter, they rotate faster, of course, because their angular momentum is conserved and the radius is decreasing. Let us see if we can calculate what we expect.

VIII. How Rotation Stops the Collapse

We saw in the last lecture that tidal torques produce a value of the angular momentum parameter of about 0.05, and that the parameter in terms of simple quantities is about

$$\lambda \approx 0.4\sqrt{T_{rot}/T_{tot}} \approx 0.24 v_{rot}/\sigma \approx 0.05,$$

so

$$\frac{v_{rot}}{\sigma} \approx 0.2$$

Now as the gas cools, it shrinks, and if it shrinks by a factor F, it is clear by angular momentum conservation that $v_{rot} \propto 1/F$; But we have seen that a particle with a velocity of $\sigma\sqrt{2}$ is in a circular orbit in the dark matter isothermal sphere, so this velocity is reached when

$$v_{rot} \approx 0.2\sigma/F = \sigma\sqrt{2},$$

or $F \approx 0.14$. Thus the gas sinks into the dark matter halo by about a factor of 7 before it finds itself spinning fast enough to support itself. Its density rises by about a factor of $7^3 \approx 350$, of course, in the process. It has sunk into the dark matter structure, which has a density *distribution* which is proportional to $1/r^2$, so the baryons find themselves in a place where the *dark matter density* is a factor of $7^2 \approx 50$ higher, so the *ratio* of baryonic to dark matter is up by a factor of about 7, $1/F$. If there is 7 times as much dark matter as baryons, when the baryons have finished settling in, *there are comparable amounts of dark and baryonic matter in the region which the baryons occupy, and that region is of order 7 times smaller than the extent of the dark matter.* By this time and at these densities the cooling time of the gas is very short. There is nothing to support it along the axis of rotation, and it settles into a thin disk until it cools to about 10,000K, at which temperature it recombines; once it is neutral the cooling essentially stops. This 10^4K temperature is magic, as you will learn later. It takes only 1 eV per particle to heat it from dead cold to this temperature, then 13.6eV to ionize it *at this temperature*, but once it is ionized another 10eV takes it to 10^5 degrees, so it is a temperature where a *lot* of gas typically lives. It is also true that cooling by recombination is very efficient for ionized gas at 10,000K, so it is hard to heat it even when it is fully ionized.

Before we go back and take a critical look at this picture, let us review the numbers for the Galaxy. First, the observations say that the circular velocity is about 220 km/sec and that there is about $10^{11} M_{\odot}$ of baryonic matter (gas and stars) within a radius of about 15 kpc, and rather little outside. The evidence is strong for our galaxy and absolutely convincing for others (where the measurement can be made more easily) that the rotation velocity is accurately constant at distances far beyond where the baryonic contribution to the mass is important. This is one of the strong arguments for a dark halo; the density distribution must be $1/r^2$ in order to produce a constant rotation velocity, just as we might expect from the picture we have been describing. There is a body of evidence which suggests that there is *much* more dark matter than baryonic, though the flat rotation curves of galaxies merely demand that it *exist*; the measurements cannot be performed far enough out to really measure how much there is. There are several ways, which we will discuss later, to get at how much there is, and all agree that there is of order 7-10 times as much as baryonic matter. If this is right, the total mass of the Galaxy is about $10^{12} M_{\odot}$ and the dark halo extends to of order 100 kpc, as we discussed earlier. In order that the circular velocity be 220 km/sec, the velocity dispersion of the dark halo must be about $220/\sqrt{2} \approx 160$ km/sec. The disk is a factor of 7-10 smaller than the halo, and everything fits pretty well.

The galaxy is in a small group of galaxies called the *Local Group*, which has of order 15 members, but it is completely dominated by the Galaxy and the great spiral in Andromeda, M31, a galaxy probably a little more massive than our own but not dissimilar to ours in general properties. If one thinks about the tidal torquing hypothesis, it should be that the axis of rotation of the two objects, if there are but two objects, should be roughly perpendicular to the line connecting them. In the Local Group, this is true of both M31 and the Galaxy within about 15 degrees.

It would appear that tidal torques *do* work, and work spectacularly well.

At this point let us think about what this general picture would look like *without* dark matter. If there were only gas, then the Galaxy would have had to contract as a gas cloud only under its own gravity; in that case, we can look at the evolution of the angular momentum parameter directly, since the energetics are simple; we do not worry about the large-mass, fixed dark halo. We must go from λ of roughly 0.05 to about 0.4. Recall that

$$\lambda \propto J|E|^{1/2}.$$

J is conserved, and $E \propto 1/r$, so to make λ increase by a factor of 8, E must increase by a factor of 64, and the system must shrink by a factor of 64, not 7 or 8 as we had in the dark matter halo. Thus the maximum expansion radius must be huge, (about 2.5 Mpc) and since the mass is only the baryon mass, $10^{11} M_{\odot}$, the collapse time is very, very long, several hundred times the age of the universe. So tidal torquing fails spectacularly to work unless the universe is dominated by dark matter.

Galaxies which are not spiral disk galaxies exist, and they exist especially commonly in large *clusters* of galaxies. The picture above suggests that disks should form, but it is probably so that in regions of high density disks form and form some stars and are then disrupted by collisions/ mergers. This picture is not completely clear but qualitatively explains what is seen, that in clusters one finds more massive galaxies than in the field, but primarily ones with very little gas and often no disk at all—that is, *elliptical* galaxies.