6. Gravity, Dark Matter, and Cosmic Expansion

Reading: The reading is sections 4.2 and 4.3, which you should have already read but would be worth revisiting. If you want to get ahead, you could also start reading chapter 6.

Density

Density is mass divided by volume.
It is usually designated with the Greek letter ρ (rho). The metric unit of density is $\mathrm{kg/m^3}$.
The volume of a sphere of radius R is $\frac{4}{3}\pi R^3$.
The mass of a sphere of radius R and constant density ρ is therefore

$$M = \frac{4}{3}\pi R^3 \times \rho.$$

If the density isn't constant (e.g., if the mass is concentrated in the center), what matters is the average density inside the radius R.

Measuring mass inside a galaxy

In the previous section, we equated acceleration in a circular orbit ($a = v^2/r$) to acceleration produced by gravity ($a = GM/r^2$) to obtain the equation

$$M = \frac{v^2 r}{G},$$

from which we can determine the mass M of an object if we know the speed v and radius r of an object that is orbiting it in a circle.
What do we do in a galaxy, where the mass is not all at the center (like it is in the solar system) but spread out?
Newton demonstrated the remarkable fact that *inside* a spherical shell of matter there is no net gravitational force — the pull from different sides of the shell cancels exactly.
If you are at radius r in an extended spherical mass distribution, therefore, your gravitational acceleration is determined entirely by the mass *interior* to radius r; the gravitational effects of all the more distant matter cancels out.
We can therefore generalize our equation to

$$M_{\text{int}}(r) = \frac{v^2 r}{G},$$

v = speed of orbiting object [m/s]
r = radius of (circular) orbit [m]
G = Newton's gravitational constant [$\mathrm{m^3\,kg^{-1}s^{-2}}$]
M = mass *interior* to radius r [kg]
Strictly speaking, this equation only works for a spherical galaxy, but it is not far off even for a flattened (e.g., disk-shaped) galaxy.
We will therefore continue to use it, without worrying about the small inaccuracies for flattened galaxies.

Galaxy rotation curve, simple expectation

It is also useful to consider a rearranged form of this equation:

$$v_{\rm circ}(r) = \sqrt{\frac{GM_{\rm int}(r)}{r}}$$

where $v_{\rm circ}(r)$ stands for the *circular velocity*, i.e., the velocity that an object (such as a star) orbiting in a circle would have at a distance r in a mass distribution $M_{\rm int}(r)$.

By measuring Doppler shifts, we can measure the velocities of stars at different locations inside galaxies.

A plot of the circular velocity in a galaxy against radius (distance from the galaxy center) is called a *rotation curve.*

Suppose we have a spherical galaxy of total mass $M_{\rm tot}$, made up of stars on randomly oriented circular orbits, held together by the gravity of those stars.

For the simplest case, suppose that the average density of stars within the galaxy is constant, with a sharp edge at radius $R_{\rm gal}$.

What do we expect to measure for $v_{\rm circ}(r)$?

Inside radius $R_{\rm gal}$, the mass interior to r is

$$M_{\rm int}(r) = \frac{4}{3}\pi\rho r^3 = \left(\frac{4}{3}\pi\rho R^3\right)/R^3 \times r^3 = (M_{\rm tot}/R^3) \times r^3.$$

i.e., the mass is increasing in proportion to the enclosed volume, hence to r^3.

The circular speed is therefore

$$v_{\rm circ}(r) = \sqrt{\frac{G \times (M_{\rm tot}/R^3) \times r^3}{r}} = \sqrt{\frac{GM_{\rm tot}}{R} \times \frac{r^2}{R^2}} = \sqrt{\frac{GM_{\rm tot}}{R}} \times \frac{r}{R}.$$

Thus, within the galaxy, we expect the circular velocity to rise as we go further out, $v_{\rm circ}(r) \propto r$.

Suppose we find a few stars or gas clouds at $r > R_{\rm gal}$ to measure the circular velocity.

Beyond $R_{\rm gal}$, the interior mass is just $M_{\rm tot}$, so we expect

$$v_{\rm circ}(r) = \sqrt{\frac{GM_{\rm tot}}{r}} \propto \frac{1}{\sqrt{r}}.$$

Outside the edge of the galaxy, the circular velocity should drop in proportion to $1/\sqrt{r}$, just like in the solar system.

Galaxy rotation curve, more realistic expectation

In real galaxies, the density of stars is highest in the middle, and drops off as you go further from the center.

Just as in the simplest case, we expect the rotation curve to rise in the central regions where the density is high, then drop in the outer regions once we are outside most of the mass.

The turnover from rising to falling should be gradual rather than sharp, since the density drops of smoothly instead of instantly.

A realistic rotation curve should therefore look like a smoothed out version of the one we figured out for the constant density, spherical galaxy.

Galaxy rotation curve, observations

In disk galaxies, the stars and gas clouds in the disk are moving on nearly circular orbits (like planets in the solar system).

We can measure Doppler shifts at different locations in the disk to determine $v_{\rm circ}(r)$.

Gas clouds can often be detected far out in the disk, enabling measurements at large radii where there are very few stars left.

As expected, observed rotation curves rise in the middle, then begin to turn over.

But instead of falling at large r, observed galaxy rotation curves stay flat, with constant $v_{\rm circ}(r)$.

The discovery of *flat rotation curves* was made by several astronomers using different methods, with particularly important contributions by the American astronomer Vera Rubin.

Dark Matter Halos

A flat rotation curve implies

$$M_{\rm int}(r) = \frac{v_{\rm circ}^2 r}{G} \propto r.$$

Thus, at large radii where the light from the stars has faded out, the mass of the galaxy is continuing to grow, with $M_{\rm int}(r) \propto r$.

The visible, stellar portion of a galaxy sits in a much larger, and more massive, “halo” of dark matter, which provides gravity but does not produce light.

A variety of observations (e.g., motions of satellite galaxies) indicate that the extent of a galaxy’s dark matter halos is typically $10-20\times$ the visible extent of the galaxy, and that the mass of the halo is typically 10 times the mass of the stars.

For example, the outer radius of Milky Way’s stellar disk is about 20 kpc, and its dark matter halo extends to about 200 kpc.

The mass of the Milky Way’s halo is about $10^{12}M_\odot$, compared to about $10^{11}M_\odot$ of stars.

Dark Matter

Dark matter is non-luminous matter that is detected via its gravitational effect on visible matter.

(There are attempts to search for it in other ways, but they have so far proven unsuccessful.)

Dark matter was first discovered by the Swiss astronomer Fritz Zwicky in 1933, based on the motions of galaxies in the Coma galaxy cluster.

He argued (correctly) that they were moving too fast to be held together by the gravity of stars alone. Without extra, unseen matter, the cluster would fly apart.

Evidence for dark matter became stronger in the 1970s, when technological advances enabled measurements of galaxy rotation curves at large radius.

These technological improvements included more sensitive detectors for visible light and radio telescopes that could measure motions of hydrogen gas clouds.

In principle, the observations of flat rotation curves and rapid galaxy motions in clusters could be explained by changing the theory of gravity itself instead of invoking a new form of matter — e.g., by saying that gravitational forces fall slower than $1/r^2$ at large distances.

But it is very hard to construct any theory of gravity that explains all of the observational evidence for dark matter.

Escape Speed

Throw a ball up in the air. It falls back down.

Fire it up with a cannon. It goes higher. Then falls back down.

Fire it up faster than 11 km/s (and ignore air resistance). It goes forever, and doesn’t fall back.

11 km/s is the *escape speed* from the surface of the earth.

The general formula for the escape speed at a distance r from a mass M is

$$v_{\rm esc(r)} = \sqrt{\frac{2GM}{r}} = \sqrt{2} v_{\rm circ}(r).$$

It is not surprising that it is larger than but similar to $v_{\rm circ}(r)$.

Gravity and expansion

Consider a spherical region of the expanding universe, with radius r, large enough to contain an "average" amount of matter (both luminous and dark).

The shell at the edge of this sphere is expanding outward with speed $v = Hr$.

Note that I have written H instead of H_0, because the argument we are going to make could also be applied at other times when the Hubble constant is different from today's value (which is what the subscript-0 denotes).

Gravity is pulling back on the shell, causing it to decelerate at a rate

$$a = \frac{GM_{\rm int}(r)}{r^2}.$$

(Our spherical symmetry argument allows us to ignore the effects of all of the matter outside the shell. This is not totally obvious, but it turns out to be true.)

Thus, the expansion of the shell should slow down over time.

The rate of deceleration will depend on $M_{\rm int}(r)$, and thus on the average density of matter in the universe.

Critical Density

It is natural to ask whether the shell is moving above or below the escape speed.

Let's calculate what would be required for it to move exactly at the escape speed:

$$v = Hr = v_{\rm esc}(r) = \sqrt{\frac{2GM_{\rm int}(r)}{r}}$$

implying

$$H^2 r^2 = \frac{2GM_{\rm int}(r)}{r} = \frac{2G}{r} \times \frac{4}{3}\pi r^3 \bar{\rho} = \frac{8\pi G}{3}\bar{\rho} r^2,$$

where $\bar{\rho}$ is the average density of matter in the universe.

We can divide out r^2 and rearrange this into an equation for the *critical density*:

$$\rho_{\rm crit} = \frac{3}{8\pi G} \times H^2.$$

$\rho_{\rm crit}$ = critical density

H = Hubble constant

G = Newton's gravitational constant

If the average density is $\bar{\rho} > \rho_{\rm crit}$, then the shell is moving slower than $v_{\rm esc}(r)$, so it will eventually stop and recollapse.

If the average density is $\bar{\rho} < \rho_{\rm crit}$, then the shell will expand forever.

If the average density is $\bar{\rho} = \rho_{\rm crit}$, then the shell will always slow down, but it will never come to a stop.

Note that r dropped out of our equation. If one shell is above the escape speed, then they all are, and vice versa.

Thus, the average density of the universe determines its fate: expand forever, or recollapse to a "big crunch."

As the universe expands, ρ drops, but H also drops. If the density is above the critical density at one time, it stays above it at all later times, and vice versa.

After plugging in $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and doing lots of unit conversions one gets $\rho_{\rm crit} = 9.2 \times 10^{-27}\,\mathrm{kg/m^3}$.

This is equivalent to about 1 hydrogen atom per cubic meter, roughly a trillion-trillion (10^{24}) times less dense than air.

Warning!

The arguments above are correct given the assumptions that went into them.

However, we will learn later in the course about an important and surprising complication.

Rewinding the Cosmic Movie, Again

Previously: Running Hubble flow backward, assuming galaxies go at constant speed, implies a "big bang" at a time $t_0 = 1/H_0$ in the past.

Gravity slows expansion $\Longrightarrow$ galaxies were moving faster in the past, big bang occurred more recently.

For a universe filled with matter at the critical density, a full calculation gives $t_0 = \frac{2}{3} \times \frac{1}{H_0}$, 33% younger.

For a nearly empty universe, with $\rho \ll \rho_{\rm crit}$, gravity has little effect on expansion, so $t_0 \approx 1/H_0$.

The warning above will also apply to these conclusions.

An important general conclusion: if gravity is always attractive, and the expanding universe is approximately homogeneous on large scales, there must have been a time in the past when everything was tightly packed.

Hubble flow plus gravity makes a strong case for a "big bang," but we still want direct empirical evidence that it happened.