Chapter 18: **Galaxies**

- Hubble type & characteristics of different types of nearby galaxies; introduction to galaxy clusters
- introduction to the expansion of the universe; Cepheids; the Hubble relation
- galaxy evolution; active galaxies; interacting galaxies
- large-scale structure of the universe
- sample problems

Hubble type

Not all galaxies are spirals like the Milky Way. And just to be slightly complicated, not all spirals even look like the Milky Way. We have only been looking at galaxies, and recognizing them *as galaxies*, for about a hundred years. Just as the first stellar spectra were classified based on what they looked like, it makes sense that the first images of galaxies were classified based on what they look like.

The majority of galaxies in the universe today (i.e., nearby) fall into two broad classes (or "morphologies"): spirals and ellipticals. Other galaxies, usually the less-symmetric ones, are called irregular. Some galaxies, of any type, that look a bit odd, get the term peculiar appended to their type; often this describes galaxies that are actively in the process of merging. The main classes of spirals and ellipticals are subdivided. Ellipticals are given a notation that indicates how much they appear to deviate from round: E0 look round, while E6 or E7 are distinctly squashed. For type En, $n = 10(1 - b/a)$, where a and b are the major and minor axes of the ellipse of stars. Spirals are divided into two main subtypes, depending on whether the nucleus of the galaxy is round or barred (elongated). These subtypes are further divided depending on the relative dominance of the bulge vs the disk. Edwin Hubble, for whom the Hubble Space Telescope is named, was one of the first to classify galaxies based on their appearance. He organized these basic types in an order that, graphically, is for fairly obvious reasons called "Hubble's tuning fork". Galaxies do change over time, e.g., as they collide with other galaxies, but Hubble's tuning fork is *not* an evolutionary sequence! Here, with a few additions, is a schematic of Hubble types:

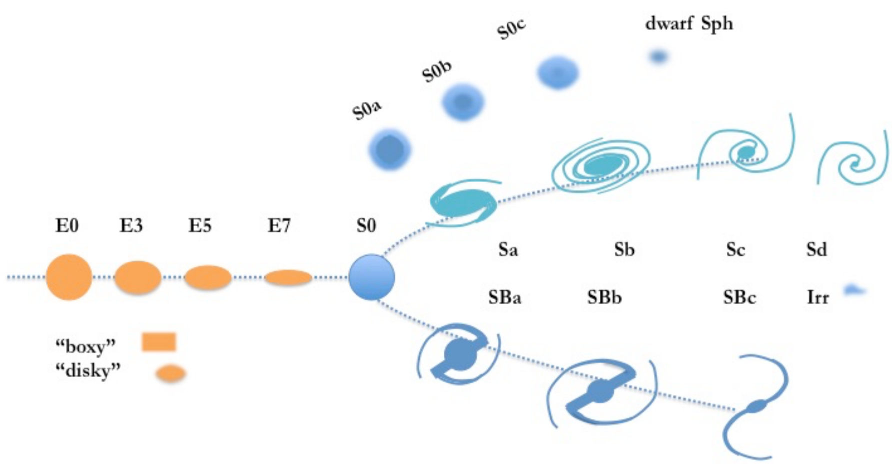


Figure 18.1: Hubble types

Of course, real galaxies don't look like stick figures and they don't all present themselves face-on to our line of sight. That complicates matters. . .a galaxy that's really shaped like an E7 might look like an E1 if seen nearly end on, or we might see a spiral edge on and have no firm idea of whether it has a bar or not. Ellipticals, seen in more detail today, may be classified as "boxy"-looking or "disky"-looking. Some spirals have rings around the whole galaxy, which Hubble's system doesn't account for. Some spirals have even less prominent nuclei than Hubble's type Sc, so we've added an Sd in this figure (and some disk galaxies have no bulge at all). Lenticular, or

S0 galaxies, are a bit like a spiral with no arms; Hubble had one, generic, S0 type where the spirals meet the ellipticals, but S0s vary as well, with differing degrees of central concentration. Dwarf spheroidals, small dim galaxies only well identified within our Local Group (us, M31, and our companions), are not the same as dwarf ellipticals and deserve separate representation. Irregulars exist and may show vague spiral or bar structure.

These types are based on what galaxies look like to us, making them a good first step in understanding galaxy types, but by no mean the last word. Some pictures are appropriate here. We looked at these first two in the section on the Milky Way. NGC 891 has a relatively balanced ratio between the nuclear bulge and the disk. It doesn't seem to have a bar, so it is probably an Sb. NGC 1672 has a bar and its disk dominates over the nuclear bulge, although not as much as is the case for some spiral galaxies. It's probably an SBbc, where the "bc" means that it's intermediate between type SBb and type SBc. Following these two are a selection of Hubble Space Telescope images of spirals, barred spirals, ellipticals, and then a few more odd-looking galaxies. The last is an image of UGC 1382, an example of a low-surface-brightness (LSB) galaxy, not much brighter than the sky background; it is ~220 kpc across, approximately seven times as wide as the Milky Way, but similar in mass. We'll look at their properties in more detail after the pictures. Keep in mind that all of these galaxies are relatively nearby and not actively engaged in merging with other galaxies.

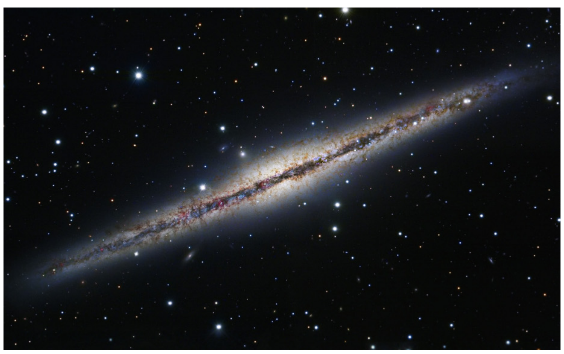


Figure 18.2: NGC 891

http://apod.nasa.gov/apod/ap120526.html

Credit: *Composite Image Data* - Subaru Telescope (NAOJ), Hubble Legacy Archive, Michael Joner, David Laney (West Mountain Observatory, BYU); *Processing* - Robert Gendler

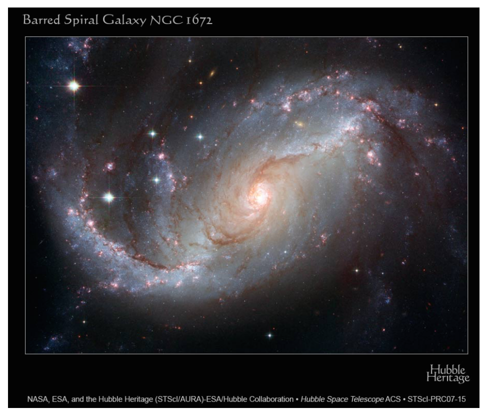


Figure 18.3: NGC 1672

http://heritage.stsci.edu/2007/15/index.html

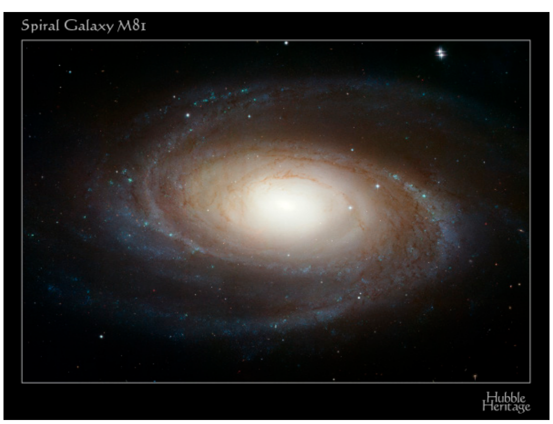


Figure 18.4: M81 is an Sab

http://heritage.stsci.edu/2007/19/index.html

Figure 18.5: NGC 1376 is an Scd

http://heritage.stsci.edu/2010/00/index.html

Figure 18.6: NGC 4319 is an SBab (with a ring)
http://heritage.stsci.edu/2002/23/index.html

Figure 18.7: NGC 1300 is an SBbc
http://heritage.stsci.edu/2005/01/index.html

Figure 18.8: M104 is an Sa
http://heritage.stsci.edu/2003/28/index.html

Figure 18.9: NGC 1132 is an elliptical, and a bit odd. More below.
http://heritage.stsci.edu/2008/07/index.html

Figure 18.10:

M87 is an E0-1, with an interesting jet. More below.
http://heritage.stsci.edu/2008/30/index.html

Figure 18.11:

S0 or lenticular galaxies are like spirals with no arms.
http://heritage.stsci.edu/2006/24/index.html

Figure 18.12: This is a dwarf irregular, similar to the SMC.
http://heritage.stsci.edu/2008/02/index.html

UGC 1382 in optical, UV, and hydrogen emission (shown in green).

Figure 18.13: A *very* large, low-surface-brightness galaxy; NASA/JPL/SDSS/NRAO/ L. Hagen and M. Siebert.
https://carnegiescience.edu/node/2061

The light falls off from center to outskirts a bit differently for spirals and ellipticals. There is no preferred plane for star orbits in an elliptical, but the stars do still concentrate toward the center, a bit like the stars in a globular cluster are somewhat concentrated toward the center of the cluster. If we measure the intensity of the light from an elliptical galaxy as a function of the distance from the center of the galaxy outward, it falls off approximately as $\log I(r) \propto r^{-1/4}$. The intensity curves for the bulges of spiral galaxies often follow a similar relation. In contrast, the intensity of the light in the disk of a spiral galaxy falls off more slowly, as $\log I(r) \propto r^{-1}$.

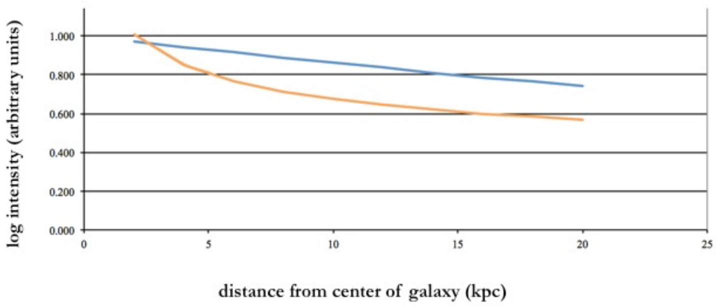


Figure 18.14: Galaxy light curves (disk on top, bulge below)

Lenticular galaxies don't have spiral arms, but their light curves resemble the disk of spirals more than an elliptical.

A somewhat more technical description of the surface intensity profiles is given by the following relation, published by Sérsic in 1963:

$$I(R) = I_0 \cdot e^{\{-b(n)[(R/R_e)^{1/n}-1]\}},$$

where b is a scaling factor, R_e is the effective (half-light) radius of the galaxy, and n is called the Sérsic index. For exponential disks, $n = 1$; for the elliptical profiles, $n = 4$.

Galaxies are extended objects, meaning that we are often going to want to talk not simply about magnitudes, but rather about magnitudes / arcsec^2. It is not particularly difficult to translate from observed mag / arcsec^2 to the number of solar luminosities per parsec^2 at the galaxy that would be needed to generate the observed magnitude. You don't even have to know the distance to the galaxy to do this. The reason is because the flux from the galaxy falls off as the square of the distance, *and*, the area on a galaxy subtended by a given solid angle is going to go up as the square of the distance. Those two effects cancel. Any given star will look fainter if it's farther away, but, if the stars are fairly uniformly distributed on the visible surface of the galaxy, then we've got more stars included in a square arcsecond and that compensates for the fact that those stars look fainter.

Example: Suppose that we observe the disk of a spiral galaxy and find that it has a surface brightness in *B* of 22 mag / arcsec^2. How many $L_\odot$/pc^2 is this at the galaxy?

We need to assume a distance. *Any* distance. Take 10 pc, for instance. Yes, that's *silly*; our galaxy is not at 10 pc. But we know the *B* magnitude of the Sun at 10 pc (5.4), and distance does not matter. Honest. Proceed in two steps.

To convert the numerator, recall that a luminosity ratio is equal to $10^{0.4\Delta m}$. Our numerator in this case will be $10^{0.4(5.4-22)} = 2.29 \cdot 10^{-7} L_\odot$.

To convert the denominator, recall that at 1 pc, 1 AU subtends an angle of 1 arcsecond and that there are 206,265 AU per parsec.

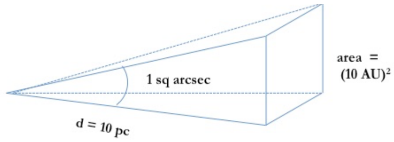


Figure 18.15: The geometry for one square arcsec at a distance of 10 parsecs

Convert that (10 AU)^2 to pc: [10 AU · (1 pc / 206,265 AU)]^2 = $2.35 \cdot 10^{-9}$ pc^2.

Next, put the two pieces together:

$$22\frac{\text{mag}}{\text{arcsec}^2} = \frac{2.29\cdot 10^{-7}\,L_\odot}{2.35\cdot 10^{-9}\,\text{pc}^2} = 97\ L_\odot\,/\,\text{pc}^2.$$

Try this again assuming a distance of 206,265 pc. In this case, the area on the face of the galaxy is just 1 pc^2 (because 206,265 AU = 1 pc). We do need to calculate the magnitude of the Sun at that distance, though:

$$m_{B,\odot} = 5\log d - 5 + M_{B,\odot} = 5\log(206{,}265) - 5 + 5.4 = 26.97.$$

The conversion this time looks like:

$$22\frac{\text{mag}}{\text{arcsec}^2} = \frac{10^{0.4(26.97-22)}\,L_\odot}{(1\ \text{pc})^2} = 97\ L_\odot\,/\,\text{pc}^2.$$

Try this at another distance and see that you get the same result.

At some point the edges of the galaxy will fade into the brightness of the night sky. The sky brightness, in mag/arcsec^2, varies with wavelength and with the presence of other sources of brightness, such as moonlight, and with the altitude of your observatory. For example, from the ground a dark sky might be ~22½ mag/arcsec^2 in the *B* filter, but only ~23½ from space. In the infrared, where our atmosphere is quite bright, we could gain 5+ magnitudes by observing from space. One way to define the size of a galaxy is to estimate it size at a distance from the center where the brightness of the galaxy has fallen to a fixed mag/arcsec^2 limit. Depending on how concentrated the core of a galaxy is, the center might be ~18 mag/arcsec^2 in the *B* filter. One standard size estimate is the Holmberg radius, which is the distance from the center at which the intensity of the galaxy reaches 26.5 mag/arcsec^2 in *B*. That's less than the sky brightness, but it's possible to measure a galaxy that's only adding about 1% to the background sky level.

Something you may have noticed from the images, above, is that there would seem to be more dust in spirals than in ellipticals. There's also usually more gas, as well, and more regions of active star formation. That means that the overall color of spiral galaxies is a bit bluer, on average, than the overall color of ellipticals. Spirals, still forming stars, are likely to have some massive hot blue stars among their star population, whereas ellipticals are less likely to have blue stars.

The dust in galaxies, including the Milky Way, is, no shock, going to affect our ability to adequately determine their colors and luminosities. First we need to account for the dust that is in the plane of our galaxy. How much extinction that's going to cause depends on what direction we are looking. Directly out of the plane toward the galactic poles we find that A_B, the extinction in the *B* filter, is ~0.20 magnitudes. We are a bit out of the plane of the galaxy, and that means we have a little less extinction to the north than to the south (~0.19 vs. 0.21 magnitudes). At lower galactic latitudes, *b*, as a first approximation we could estimate the path length along our line of sight that is within the plane of our galaxy and use an average value for the extinction per kiloparsec. Because dust isn't evenly distributed, though, the resulting extinction turns out to be a somewhat complicated function of both *b* and *l*. On average A_B will be about 0.2 / sin (*b*), but with quite a bit of variation depending on galactic longitude. (As described in Chapter 17, in the *V*-band, A_V will be about 0.18 / sin (*b*), for latitudes greater than ~ 20 degrees, and ~ 1.8 mag/kpc closer to the plane of the Milky Way.)

We might also want to consider how dust in another galaxy is going to affect our observations of it. For ellipticals, not much. Unless we know that we have a peculiar elliptical, one that for some reason actually has dust, we can pretty much ignore extinction in ellipticals. In spirals, we have to consider the inclination of the galaxy to our line of sight. A face-on spiral isn't going to have a lot of extinction, maybe only a few hundredths of a magnitude. A face-on spiral looks pretty much circular; at a moderate inclination, such that the projected lengths of the axes of the galaxy are ~1:2 (rather than the circular 1:1), the extinction might be a few tenths of a magnitude.

In between our galaxy and others there is very little dust. Hot gas, yes, but very little dust, and extinction in the Intergalactic Medium (IGM) can usually be ignored.

You can't tell the sizes and masses of the galaxies from the images. The largest and smallest galaxies tend to be ellipticals. Some ellipticals are dwarfs, about the mass of a globular cluster (although with dark matter, which doesn't seem to dominate the lives of globulars). On the other end of the spectrum, most of the most massive galaxies are also ellipticals, in this case, possibly formed as a result of collisions of medium-sized galaxies; some massive galaxies appear to be "super spirals". The most massive elliptically shaped galaxies are often called supergiant ellipticals, or, a bit more technically, type cD (which stands for supergiant diffuse); these behemoths may exceed a Mpc in diameter and 10^{14} solar masses. The theory that they grow from mergers is supported by the observation that they tend to be found in the centers of clusters of galaxies, as well as the fact that some have multiple nuclei, possible remnants of the original separate galaxies that merged to make the supergiant.

Spirals vary a bit in size as well. Some irregular galaxies that show a bit of structure may have been spirals, but perhaps have been tidally stretched by a larger neighbor. The Large Magellanic Cloud, for instance, has a distinct bar, although at the present time it has only about 1% the mass of the Milky Way and a diameter of less than 5 kpc. The Andromeda galaxy, our nearest large neighbor, is approximately half again as massive as the Milky Way. A stream of stars and the metal-rich nature of M31's halo suggest that it is possible that M31 gained some of its mass by stripping it from a companion galaxy, leaving behind its now-smaller companion M32.

One way to estimate the masses of spiral galaxies (as long as we aren't seeing them face-on) is to estimate the velocities of the stars as far out in the disk of the galaxy as possible. This is assuming that we can resolve one side of the galaxy from the other and measure the Doppler shifts of each side independently. We are assuming that the stars' orbits are approximately circular and that we can estimate their actual velocities from their observed radial velocities and an estimation of the inclination of the galaxy to our line of sight. If the galaxy were edge-on to our line of sight, inclination = 90° to the plane of the sky, then $v_{\text{rad}} = v_{\text{circ}}$, where v_{rad} is the radial velocity we observe and v_{circ} is the actual velocity of the star. (At a lesser inclination, $v_{\text{circ}} = v_{\text{rad}} / \sin(i)$.) If we know the distance to the galaxy and its angular size we can estimate the distance from the center to the edge of the disk. We can then use the velocity equation

$$v = \sqrt{\frac{GM}{r}}$$

to estimate the mass of the galaxy. Or, at least, to estimate the mass of the galaxy interior to the orbit of the farthest stars we can see out in the disk.

If we can't resolve the disk of the galaxy, we can still get an estimate of how fast the fastest stars are going by measuring the width of the lines in the spectrum of the galaxy. The spectrum is made up of the light from many stars; some of those stars are at the edges of the disk, meaning they will either be moving toward us or moving away. The lines in the spectrum of our unresolved galaxy will be broadened by virtue of the fact that they are composed of contributions from many stars, some of which will be absorbing distinctly off-center because of their radial velocities.

A similar technique for estimating the masses of elliptical galaxies is grounded in the virial theorem. Recall that the virial theorem tells us that as material fell together to make a spherical system (star, gas planet, elliptical galaxy. . .) one-half of the released gravitational potential energy was radiated away and the other half went into the kinetic energy of the particles that make up the system. In this case, those "particles" are stars. Their gravitational potential energy is ~ GM/r. Their kinetic energies are $\frac{1}{2}mv^2$, as they are for the stars in the disk galaxy. Here, the problem is that the velocities have components in all three spatial directions, but from the Doppler shifts we can only detect the radial components of those velocities. We also are not so likely to be able to distinguish the velocities of the stars out well away from the center of the galaxy, the way we could with the disk of a spiral, particularly if we cannot well resolve the galaxy, because our line of sight will intersect stars at various distances from the center of the galaxy moving at various speeds. As with the unresolved spiral we can use the width of the spectral lines to estimate the stars' radial velocities. What we measure from the width of the spectral lines, converted to speeds, is called the velocity dispersion (σ). We have to multiply $\frac{1}{2}mv_{\text{rad}}^2$ by three because we have to assume that the kinetic energy in the up-down and cross-ways directions is the same as the kinetic energy in the radial direction. That factor of three will give us a total velocity dispersion, not just an average radial velocity dispersion. We also have to estimate a representative distance for the stars from the center of mass of the galaxy, perhaps a bit less than half the total radius of the galaxy if we expect that the mass of the galaxy is at least a bit concentrated toward the center and the average star is thus less than half way out in the galaxy. Then we can use the velocity equation. Same basic idea as with the edges of the disk of a spiral galaxy, just a bit more complicated with masses that are a bit less certain.

The math on that:

$$P.E. \approx \frac{3}{5}\frac{GMm}{R} = 2\cdot\frac{3}{2}m < v_{\mathrm{rad}} >^2 \rightarrow \ < v_{\mathrm{rad}} >^2 \approx \frac{2}{5}\frac{GM}{R_{1/2}},$$

where $R_{1/2}$ is ~$R/2$ and < *quantity* > means the average of *quantity*.

Galaxies are normally found in groups. The groups vary in size quite a lot. The Local Group only contains about 80 galaxies. It's dominated by us and M31. M33, another spiral, is an order of magnitude less massive. The rest are irregulars, dwarf ellipticals, dwarf spheroidals, many of which are either companions to M31 or the Milky Way, or actually in the process of being tidally disrupted, on the way to being absorbed, by one of the large galaxies. Altogether the Local Group is about 10 Mly (or 3-ish Mpc) across. We are not an impressive cluster.

Many clusters contain more galaxies and, not being dominated by just two, are, usually, much more spherical in shape. Relatively nearby clusters are often referred to by the constellation in which direction they are found: The Virgo cluster, the Coma cluster, the Hercules cluster, etc. The Virgo cluster contains at least 1,300 galaxies, and perhaps more than 2,000. The center of the Virgo cluster is dominated by the giant elliptical M87 (with the jet) and several other massive ellipticals, spirals, and lenticulars. The core of the cluster is only ~16.5 Mpc (~54 Mly) away, which makes the whole cluster angularly huge on our sky, ~7-8 degrees across.

The Coma cluster, also with over 1,000 member galaxies, is a bit farther away, ~99 Mpc (~321 Mly). The core of the cluster is dominated by ellipticals and lenticulars. Coma is fairly dense, as clusters go, and only about 20 Mly in diameter. Here is an HST image, offset about one-third of the way from the center of the Coma cluster; at the average distance of the galaxies in this field, this image spans ~520 kpc (~1.7 Mly).

Figure 18.16: Coma cluster
https://www.nasa.gov/images/content/242543main_hstimg_20080610_HI.jpg

In clusters, between the galaxies, we find the Intergalactic Medium (IGM). As mentioned above, the IGM contains very little dust. The material here is a warm-ish plasma at temperatures on the order of 10^6 K. Because that's "warm – to – hot", you could also call this stuff "WHIM". . .the Warm-Hot Intergalactic Medium. Much of the plasma is hydrogen that's left over from when the galaxies formed; it's enriched by material blown out from the cores of active galaxies or those supernova explosions energetic enough to blow fountains of material out of their galaxies. Some of that material is in clouds that are actively (re)accreting to galaxies, making continued star formation possible.

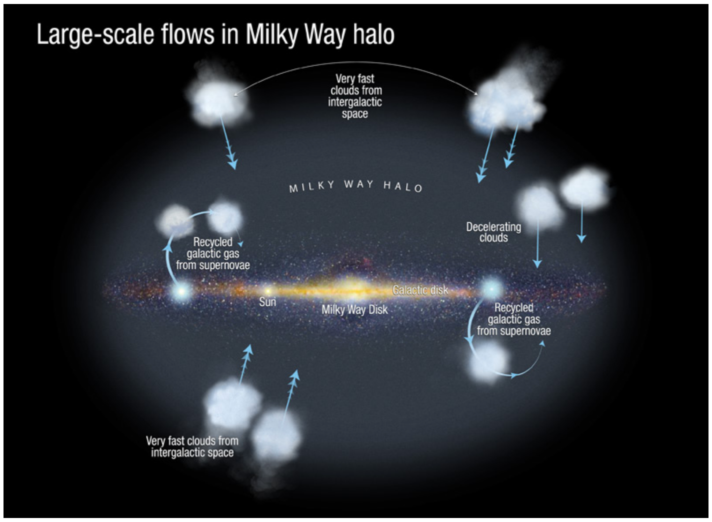


Figure 18.17
Artist's sketch of clouds of gas around the Milky Way

Illustration Credit: NASA, ESA, and A. Feild (STScI)
Science Credit: NASA, ESA, and N. Lehner and C. Howk (University of Notre Dame)

http://hubblesite.org/newscenter/archive/releases/2011/26/image/a/

The plasma is not easy to detect. One particularly successful method makes use of the fact that heavy elements in the WHIM, such as oxygen, that are not totally ionized are capable of absorbing x-rays emitted by more-distant active galaxies. That absorption shows up in spectra taken of the active galaxies with space telescopes such as Chandra and XMM-Newton. Because of its temperature, the WHIM also emits x-rays and contributes to the diffuse x-ray background of the universe.

Observations in the 2020s suggest that roughly three-quarters of the normal ("baryonic") matter in the universe is hiding in the Intergalactic Medium. Fast Radio Bursts (FRBs) are transient high-energy bursts of radio waves with a range of frequencies. The presence of the IGM means that the space across which these radio waves travel is not, quite, a vacuum. Different frequencies will be delayed by different amounts (a quantity called the "dispersion measure") because the IGM effectively has an index of refraction that is not quite equal to one. Measuring the dispersion of radio waves from FRBs from extemely distant galaxies allows us to estimate the density of material along our line of sight and detect the presence of the intergalactic medium.

We will return to this plasma briefly in cosmology when we consider the Cosmic Microwave Background, because the interaction of the CMB with the hot plasma will modify the spectrum of the CMB a little.

Clusters also have dark matter, partly in extended halos around galaxies and partly between the galaxies. The mass-to-light ratios for clusters are often several hundred ($M_\odot / L_\odot$) (i.e., considerably higher than the several tens ($M_\odot / L_\odot$) we find in individual galaxies). One of the best demonstrations of the existence of dark matter comes from observations of what's colloquially called the Bullet Cluster (which is a more vivid name than 1E 0657-56). Two clusters of galaxies, approximately 3.8 billion light years away, have collided. The two clusters' clouds of hot gas experienced drag as they tried to pass through each other. The dark matter didn't. It just barreled right on through. How do we know? The hot gas emits x-rays, so the gas distribution can be mapped by the Chandra x-ray space telescope. The mass distribution overall can be mapped because mass warps spacetime and acts as a gravitational lens, distorting images of more distant objects. The following composite image shows the gas and the mass distribution.

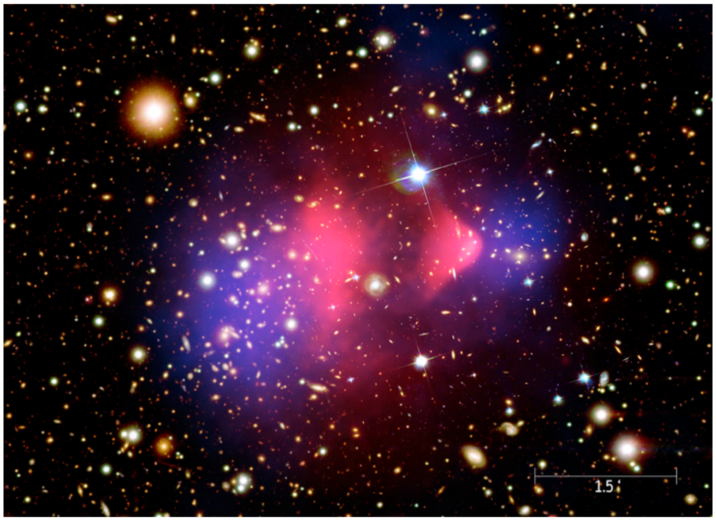


Figure 18.18: Bullet cluster; Credit: X-ray: NASA/CXC/CfA/ M.Markevitch et al.; Optical: NASA/STScI; Magellan/ U.Arizona/D.Clowe et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/ D.Clowe et al.
http://chandra.harvard.edu/photo/2006/1e0657/

In the above image, the pink is the gas. The bullet-shaped blob to the right is the reason for the nickname: it has moved toward the right through the gas, from the more massive cluster of the colliding pair, on the left. The blue is the mass distribution, and the two main chunks have just moved through each other without slowing down. The gas clearly does not provide the majority of the mass in these clusters. These two images are superposed on the background optical image from the Hubble Space Telescope.

Clusters of galaxies are themselves found in superclusters, which are draped around huge voids (bubbles some 10s of Mpc across that are mostly empty) like a film of soap on the surface of soap bubbles. We will return to these as we look at the large-scale structure of the universe, but first we need to consider the expansion of space and a few more methods for determining large distances.

Introduction to the expansion of the universe

As we start to consider galaxies, and, in particular, galaxies that may be billions of light years away, we need to talk about how to determine distances to other galaxies. How, in fact, did we come to know that those fuzzy patches of light are in fact distinct galaxies and not a part of the Milky Way?

In the late 1800s / early 1900s, Edward Pickering served as Director of the Harvard College Observatory. Pickering's tenure as Director coincided with the rise of women's colleges in the U.S. Pickering hired a number of women, initially as "computers", who performed often tedious calculations, examined photographic plates, classified stellar spectra, often, in the process, making quite valuable contributions to research in astronomy. One of these women was Henrietta Swan Leavitt (1868 – 1921). Leavitt studied at Radcliffe College and began working at the Harvard College Observatory in 1893. Leavitt was assigned the task of studying variable stars, particularly variables in the Magellanic Clouds.

A significant number of the stars Leavitt identified belong to a type called Cepheid variables. Cepheids are pulsating variables, stars whose outer layers rise and fall. They have a distinctive light curve, rising in brightness more rapidly than they fall, over periods that range from a few hours to upwards of 100 days. They are named after the star Delta Cephei, the first such star identified. Delta Ceph itself varies by about a magnitude over about 5.4 days.

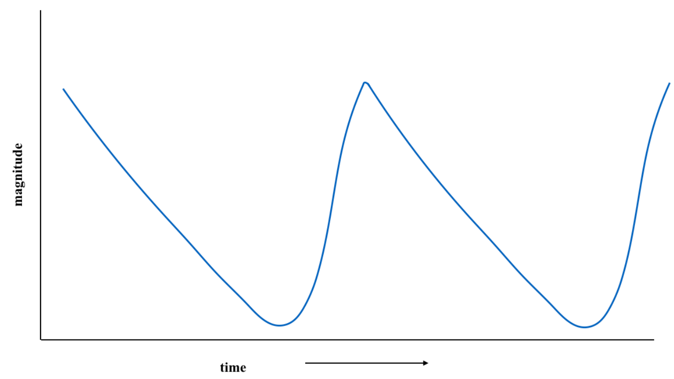


Figure 18.19: A schematic of a Cepheid light curve

The stars in the Magellanic Clouds are not all at exactly the same distance from us. Relative to the distance between us and the Magellanic Clouds, though, the range in distance among the stars themselves is not very large, meaning that for practical purposes, the stars are close enough to the same distance that differences between stars' apparent magnitudes do correspond directly to differences in their absolute magnitudes. In her study of the Cepheids in the Magellanic Clouds, Leavitt made one of the most important discoveries in Astronomy, ever. She found that the Cepheids with the longest periods of variation are the most luminous. This is now called the *Period-Luminosity relation* (P-L relation, or Leavitt's law) and it is wildly important. It took a bit of work to calibrate the magnitudes for a few stars, but once you know a few of them, you can find the absolute magnitudes of any Cepheid simply by observing the period over which it varies and using the P-L relation. That means you can find its distance. The brightest Cepheids are quite luminous, with absolute magnitudes of ~ –5 to –6. That means we can see them in galaxies out to a few million light years, vastly farther than any distance method available before this discovery.

As mentioned, it took a bit of work to calibrate the P-L relation, in part because there are no Cepheids close enough for easy parallax measurements and in part because there turned out to be two distinct types of Cepheids, depending on whether the stars were Population I (solar composition) or Population II (low metals).

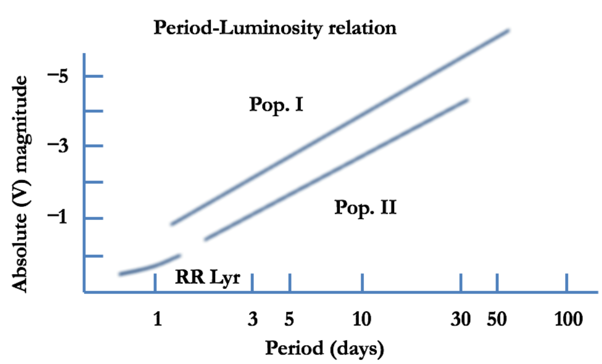


Figure 18.20: A sketch of the P-L relation; RR Lyrae stars are related variables often found in globular clusters

In the early 1920s Edwin Hubble (1889 – 1953) was working at the Mount Wilson Observatory in California, using one of the largest telescopes of the day. On a 1923 photograph of the Andromeda "nebula" Hubble realized, comparing the plate with one taken earlier, that he could identify a Cepheid variable star. When people realized that this star must be roughly a million light years away (today, we'd place M31 at ~ 2½ million light years), the "nebula" became a galaxy and the universe as we knew it suddenly became a lot larger. Here's part of that original photographic (negative) plate. Next to it is a Hubble Space Telescope image of Andromeda; the insets are four recent observations of "V1", that first identified variable in M31.

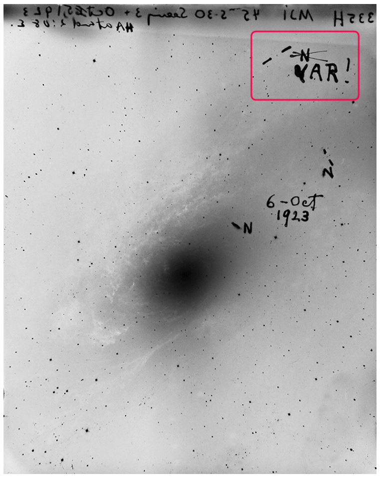


Figure 18.21: Mount Wilson Discovery Plate of M31-V1 (Carnegie Observatories) http://heritage.stsci.edu/2011/15/images/Hubble-M31-100in-w.jpg

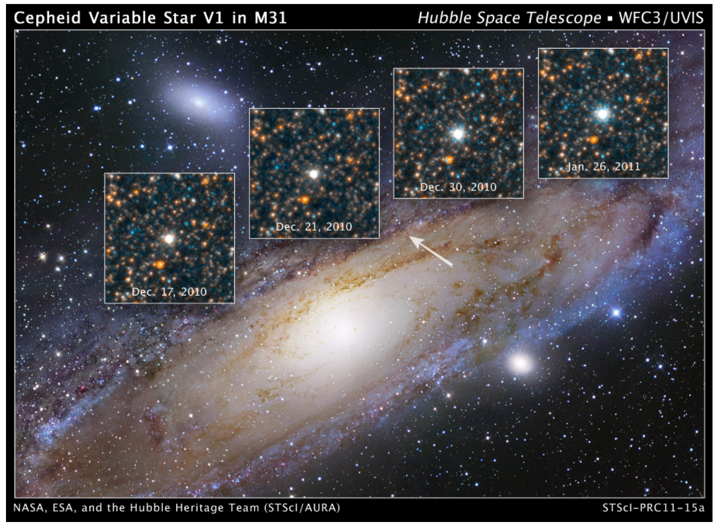


Figure 18.22: Variable V1 in M31 over 5 weeks. http://www.nasa.gov/images/content/551318main_hs-2011-15-a.jpg

In the 1910s, even prior to knowing the distances, it was already possible to observe the spectra of galaxies. Vesto Slipher (1875 – 1969) at Lowell Observatory, in Flagstaff, Arizona, detected the shift in the lines in the spectra of "spiral nebulae" in ~1912. By 1914 he had good enough spectra and angular resolution to tell that these objects were rotating, as well as receding from us.

Hubble, working with Milton Humason (1891 – 1972), was able to combine the redshifts of Slipher and other astronomers with the distances obtained at Mount Wilson to formulate observationally the first "rough draft" of the correlation between distance and redshift now known as Hubble's Law, or the velocity-distance relation. As mentioned above, M31 is headed toward us. Galaxies farther away than our Local Group, though, are almost all receding. And the farther away they are, the faster they are seen to recede from us. For distances out to a few hundreds of millions of light years, the relationship is linear: $v = H_0 \cdot d$, where v is the speed with which the galaxy is receding and d is the distance (more technically, it is called the "proper distance", the distance measured at one time). H_0, the "Hubble constant" is the slope of the relation in the near-ish universe. It's ~ 70 (km/s)/Mpc. Unless it's 67 or 73. . .more on that below. Here's a schematic of what that relationship looks like:

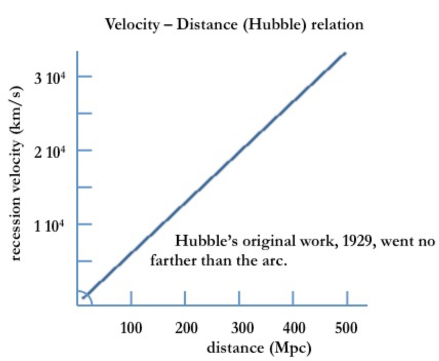


Figure 18.23: velocity - distance relation

In practice, there's quite a bit of scatter because any individual galaxy will have a velocity with respect to the center of mass of whatever cluster it's in. For the galaxies that are less than ~100 Mpc distance, those individual velocities can be a substantial fraction of the galaxy's recession velocity.

Meet z, the "redshift": That recession velocity looks a lot like a radial velocity and you shouldn't be shocked to find the quantity $\Delta\lambda/\lambda$ showing up here. Define $z = \Delta\lambda/\lambda$, meaning that the recession velocity may be identified as $v_{rec} = c\,z$. Note that in this case the stretching of the wavelengths of light is due to the expansion of the universe, meaning that we can have values of z that are substantially > 1 without raising questions about the rules of Special Relativity. If the galaxies themselves were all rushing away from us through space, we could be worried about how such massive objects could be moving so fast. They aren't, so it isn't a problem.

How does this work? We are *not* the center of the universe; galaxies are not receding from us because we are special. Any observer on any galaxy would see the same phenomena, namely the galaxies receding and the more distant galaxies receding more rapidly. Here are some galaxies, spaced one distance unit apart. Let one unit of time pass and let space stretch by a factor of 2:

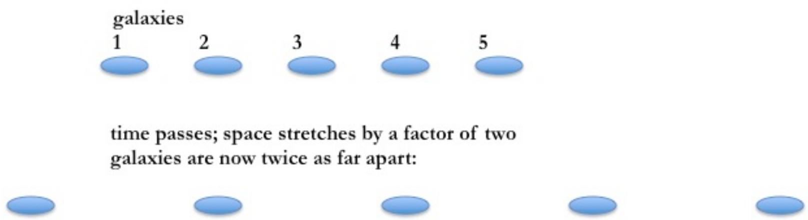


Figure 18.24: Why the velocity distance relation doesn't imply that we are the center of the universe.

Suppose we are riding on galaxy 2. We observe galaxy 3 to have gone from a distance of one unit away to a distance of two units during our one interval of time. We say that galaxy 3's speed is 1 (distance unit / time unit). What did we observe happen to galaxy 4? It went from two units distant to *four* units distant in the same time. We observe its speed to be 2 (d/t). Galaxy 5? It was at 3 units initially, but we find it at 6 units distance after the expansion, a change of three units of distance; we call its speed 3 (d/t). The farther away the galaxy was initially, the farther it will be carried away by the expansion of space and the higher a recession velocity we will observe. But now suppose that you compare notes with someone riding on galaxy 5. They see the speed of galaxy 4 as 1 (d/t), the speed of galaxy 3 as 2 (d/t), the speed of galaxy 2 as 3 (d/t), and the speed of galaxy 1 as 4 (d/t). Observers in galaxy 5 *say the same thing* that we do: more distant galaxies are receding from us faster. Neither galaxy 2 nor galaxy 5 has any claim to call itself the center of the universe.

It would be helpful if we knew H_0, the slope of the distance-velocity relation at the current time. That's not a trivial measurement. There are two basic ways of tackling the question. One is to build up a distance ladder (e.g., nearby stars, Cepheids, more distant objects) that will allow us to determine distances to galaxies out to a few hundred Mpc, calculate recession speeds from redshifts, plot the data, and measure the slope. This method gives a value of ~73.5 ± ~2% (km/s)/Mpc. A second method is to start back in time and estimate forward. The cosmic microwave background (CMB) radiation dates from the time the universe was about 380,000 years old, at which point it had cooled enough for protons and electrons to combine to make hydrogen atoms. Prior to this the universe had been an opaque plasma, as photons kept scattering off of free electrons; because the early universe was opaque, we can't see back before 380,000 years. The CMB is almost, but not entirely, uniform because the universe was pretty much all the same temperature then. But there were density fluctuations which we can today observe as temperature variations in the CMB. High-resolution observations of the CMB in all directions, for example by the European Planck spacecraft, can provide measurements of the angular sizes of the variations in the CMB and of the density of matter at that time. Assuming that the universe evolves smoothly from then until now, it is possible to relate the angular size of the variations and the densities of the various components of the universe (matter, radiation, dark energy) to the expansion rate today. This method of estimating the slope of the Hubble relation gives a value of H_0 ~67.4 ± ~1% (km/s)/Mpc. Not only are these two values of H_0 different, the estimations of the error bars on these two values don't overlap, and we don't yet understand why. (Some argue that the Milky Way is in a local 'void' (see large-scale structure, below) and that objects near us are being gravitationally drawn away by higher densities of matter around the edges of the void, thus preferentially increasing the recession speed of nearby galaxies.) Distances can also be obtained from gravitational wave events, such as the first neutron star merger observed (GW 170817), and compared with the recession velocity for the host galaxy. The neutron star merger emits a fairly well-defined amount of gravitational energy which can be compared with the observed amplitude of the received gravitational wave. One wrinkle is that gravitational waves will be redshifted by the expansion of space just as electromagnetic waves are. The energy released in a neutron star merger isn't emitted isotropically (it's brighter perpendicular to the orbit plane); this should show up in a polarization of the observed gravitational waves. If there's an associated jet of energetic material ejected during the collision and/or possibly also a GRB, we may be able to use radio or gamma-ray telescope observations of that jet to complement the gravitational wave data and determine more precisely the orientation of the orbital plane and thus determine the distance more precisely. With more detections, this technique of using standard sirens is likely in the future to become a third method of determining H_0. The one data point from the 2017 event gave a value of ~70 ± ~5 (km/s)/Mpc. This discrepancy in the determination of the value of H_0 is often called the "Hubble tension".

For galaxies that are not too distant – i.e., for galaxies for which the relationship between distance and velocity can be consider linear – the Hubble relation will permit us to estimate the galaxy's distance from its recession velocity.

Example: suppose we observe a galaxy to have a recession velocity of 14,000 km/s; estimate its distance from the Hubble relation.

$$v_{\text{recession}} = H_0 d \rightarrow d = v_{\text{recession}} / H_0 = (14{,}000 \text{ km/s}) / 70 \text{ (km / s) / Mpc} = 200 \text{ Mpc}.$$

For another approach to determining galaxy distances, let's return to the idea that the more massive a galaxy is the higher the kinetic energy of its components. R.B. Tully and R. Fisher considered the broadening of the 21-cm spectral line of neutral hydrogen and found a correlation between the line width (correcting for the inclination of the galaxy) and the luminosity of spiral galaxies. (Remember that ellipticals don't have much free gas, which is why Tully and Fisher were studying spirals.) It makes sense that a more massive spiral galaxy would be more luminous. Once this relationship was calibrated it became possible to estimate the absolute magnitude of a spiral galaxy from the width of the 21-cm line. This is best done in the infrared, rather than the visible, so that a determination of a galaxy's apparent magnitude won't be adversely affected by dust in that galaxy. The relationship between luminosity and maximum velocity for the galaxy goes approximate as $L \propto v^{\alpha}$, where $\alpha \sim 4$.
There's more than one way to express the Tully-Fisher relation, but in the *I* band we have approximately that

$$M_I = -8.55(\log[\text{width}_{21-cm} - 2.50]) - 21.51 \pm 0.04.$$

This assumes that we have corrected for extinction, especially in the Milky Way, and the inclination of the disk of the galaxy we are observing.

There is a corresponding relation for ellipticals. The Faber-Jackson relation (S.M. Faber and R.E. Jackson) was one of the first to establish a correlation between absolute magnitude and the velocity dispersion of the central regions of elliptical galaxies. This works best for the most luminous ellipticals, for whom the relationship between luminosity and velocity dispersion is similar to what we find with the Tully-Fisher relation for spirals; in this case, $L \propto \sigma^{\gamma}$, where $\gamma \sim 4$. Estimating the absolute magnitude of a less luminous galaxy is tricky, because absolute magnitude implies that we can measure all the light, even from the faint outer reaches of the galaxy. The Faber-Jackson relation was improved by using instead the diameter *D* of the region within which the galaxy's brightness is above a limit of 20 or 21 mag/arcsec^2 in *B* and comparing *D* with the velocity dispersion σ. Yet another improvement comes about if we consider a three-component relationship, using central velocity dispersion (σ), effective radius (R_e, rather than the diameter in the $D - \sigma$ relation), and surface brightness (I_e) at the chosen radius. This three-dimensional relationship is called the *fundamental plane*. It is approximately

$$\sigma^{2.48} \propto I_e^{1.64} \cdot R_e^2.$$

Another method for getting distances to galaxies is to determine the magnitude of the stars at the Tip of the Red Giant Branch (TRGB). A brief recap of stellar evolution for stars that are roughly solar mass: After running out of core hydrogen these stars will leave the main sequence (in a color-magnitude or HR diagram) and, after some amount of time, ascend the Red Giant Branch. The envelopes of RGB stars expand and their core temperature slowly approaches the temperature required for helium fusion. The core is mostly degenerate and isothermal; when the triple-alpha process (three helium fusing to become carbon) kicks in, it does so throughout the core and very rapidly. This is called the 'helium flash', even though, buried down in the star's interior, it's not something we can directly observe. Stars of a broad range of masses and metallicities converge to a narrow range of luminosities and colors at this point and there should be a reasonably well-defined limiting magnitude, the Tip of the RGB, brighter than which red giant stars are not seen. In the near-IR *I* filter, the TRGB is at $M_I \sim -4.0$, with $(V-I) \sim 1.6$. Unlike the Classical Cepheids (Population I stars, brighter than their low-metallicity cousins), which are relatively massive stars and thus likely to be found in the dusty crowded planes of spiral galaxies, RGB stars should be found in a broader range of galactic environments. In particular, TRGB stars should be found in the less-crowded and less-reddened halos of galaxies. So: obtain images in two red or near-IR filters of the halo of a galaxy, create a color-magnitude diagram (CMD), estimate the apparent magnitude of the TRGB, and compare that with the expected absolute magnitude. There is a serious wrinkle, namely the need to avoid confusion with the brighter Asymptotic Giant Branch (AGB) stars, those making their second, post-helium fusion, ascent of the giant branch, so observers need to take care to identify regions in a galaxy where there should be relatively few stars that are already at the AGB stage.

Yet another method for determining distances to galaxies, one that's going to take us to even larger distances, is to observe Type Ia supernovae. The traditional model for a Type Ia supernova (SNe Ia) is that, in the course of expanding as a giant, a companion star dumps mass onto a white dwarf, pushing it over the ~1.4 solar mass Chandrasekhar stability limit and causing it to explode. If this is correct, namely, if all Type Ia supernovae arise from the detonation of a carbon-oxygen white dwarf of known mass, then we have reason to expect that all SNe Ia reach the same absolute magnitude. This provides us with a "standard candle", an object of known, and quite bright, luminosity. Type Ia supernovae reach absolute magnitudes of ~ –19.3, meaning we can see them for quite a ways. SNe Ia have been observed to a $z \sim 1.9$, which means in a galaxy from which the light of the supernova left on its way to us nearly 10 billion years ago. It remains possible that SNe Ia are not, in fact, exactly a standard candle. For instance, observations published in 2018 of 13 Type Ia supernovae detected within the first day or so after exploding, i.e., before reaching maximum brightness, show that in the first few days about half have bluer B−V colors of ~ −0.1 and the other half ~ +0.4. After a few days, this distinction disappeared and all the objects were fairly white. It is possible that Type Ia supernovae might occur when two white dwarfs merge, in which case the total mass involved in the supernova could be larger than 1.4 solar masses and the maximum absolute magnitude brighter. It remains to be seen whether observations such as that of the early color of the event will help distinguish among possible different formation mechanisms. Much of our current picture of the age and expansion rate of the universe rests on an accurate knowledge of the distances to the most distant galaxies, meaning that better understanding the Type Ia supernovae is a critical area of current research.

Observing distant galaxies in visible wavelength filter is going to require correcting the observed light for the redshift introduced by the expansion of the universe. Light that we observe in the *B* filter, for instance, could have been emitted by the galaxy in the *U*-filter wavelengths and redshifted into the *B* bandpass. If we're looking for stars, e.g., to identify the galaxy's dominate type, mass, color, etc., of stars, we need to recognize that the galaxy might not have been particularly bright in the near UV and we'd thus be underestimating its brightness if we didn't correct for the fact that the larger amount of blue light (that we are actually looking for) got shifted to the red and we were instead measuring a lesser amount of UV (that got shifted into the *B* bandpass). This is called the *K* correction, and it requires that we estimate the original spectral energy distribution (SED) of our galaxy as well as know its redshift (and what filter we wish to correct). The *K* correction for a spiral in *B* will be larger than for an elliptical. Galaxy spectra are dominated in visible wavelengths by the light from their stars. Ellipticals are less blue than spirals. An elliptical would not have emitted as much *B*-bandpass light as a spiral in the first place, so it will matter less for our estimation of the *B* magnitude of an elliptical that the light in the *B*-bandpass got redshifted off into the *V* (or farther) filter than it would matter for a spiral. If we observe in the infrared, where the spectrum is much flatter and there's less difference in magnitude between, say, the *J* (1.22 μm) and *H* (1.63 μm) bandpasses, there will be less need for a *K* correction for whatever type of galaxy. If we observe a galaxy at a higher redshift, where there is more shift of light from one bandpass to the next, the *K* correction will be larger.

Galaxy evolution, active galaxies, and interacting galaxies

The first galaxies formed when the universe was a few hundred million years old, at $z \sim 14$. Maybe. One problem with answering the question "when did the first galaxies form?" is that we need to have a good idea what it means to be a galaxy. We can't just look for objects at high *z* that look like our neighbors because the first galaxies formed in a very different environment. When the universe was a few million years old it was warm, dark, and transparent. It contained dark matter, hydrogen, helium, and a lot of infrared photons (and a lot of neutrinos, but they are not likely to be significant players in the story of galaxy formation). The infrared photons were left over from an earlier time when the universe was hotter. As the universe expands, all lengths, including light, stretch; the universe is a blackbody and at a few million years the dominant wavelengths of light were in the infrared. Today we see that light as the CMB; then, it was IR. Infrared doesn't have enough energy to ionize hydrogen, so the matter in the universe was neutral.

Matter, both dark and baryonic, must have been a bit clumpy. We know this both because the CMB is not totally smooth and because galaxies exist. As mentioned above, the CMB dates from the time when the universe cooled enough to become neutral. When the universe was as hot as the insides of a star, hydrogen was ionized, there were lots of free electrons, and the universe was opaque because the photons kept scattering off of the free electrons. When it cooled enough to become mostly neutral, at ~3,000 K / 380,000 years old, the photons didn't have anything to scatter off of any more and the universe became transparent. There are degree-scale anisotropies in the CMB that reflect the variations in density of matter at the time when the photons and the matter decoupled, so we know that when the universe was 380,000 years old, matter was clumpy. It's reasonable to assume that the dark matter was also clumpy and that it was all still clumpy a few million years later. Besides, we need clumps: if matter were totally uniform in the early, and expanding, universe, galaxies would not have formed. Gravity must have been strong enough in some regions to overcome the expansion and make matter fall together to make the first objects.

We can't quite see faint enough or far enough back in time to see those very first objects yet. In other words, it is not clear how massive the first objects to collapse were. We are fairly certain that they were not as massive as the galaxies we see around us today nor as small as individual stars; our best estimates are that they were a few $10^{5\text{-}6}$ solar masses, more the size of today's globular clusters.

The earliest stars, Population III, would have been formed almost exclusively of hydrogen and helium. They are likely to have included stars that were more massive than stars today; those stars would have evolved rapidly and done their part to seed the ISM with the first heavy elements which were then incorporated into the second-generation stars.

Pause for a note about why the first stars could be more massive: Stars are hot inside and that produces pressure outward. Today, stars can't form at much more than ~100 solar masses, because above that mass the pressure outward exceeds the force of gravity drawing the material together. This is called the Eddington Limit. Photons hitting atoms of heavy elements is one of the sources of outward pressure. If you don't have heavy elements the opacity of the star's outer layers will be lower, and we can add more mass and make larger stars. The first stars could easily have been 300 (or more) solar masses and lived and died even faster than the most massive stars today.

The earliest stars may have formed in dark matter mini-halos at $z \sim 20\text{-}30$, when the universe was 100-200 million years old. In other words, clumps of dark matter with masses of a few 10^5 solar masses could have formed and attracted enough hydrogen and helium to form the first stars. Clumps merged and those first stars died out. The first objects that deserve to be called galaxies, at $z \sim 12\text{-}14$, had masses on the order of 10^8 solar masses and their stars were Population II stars, already enriched, just a little, by heavy elements formed inside the first generation of stars. The most distant galaxies that we can see show evidence of this enrichment in the form of infrared emission (which we observe redshifted to sub-mm wavelengths) from dust grains — dust is formed from heavy elements formed inside stars. A slightly technical distinction between the dark matter halos and the first galaxies is based on how hot they are as well as their masses: the smaller objects would have cooled by radiation from molecular hydrogen whereas the first galaxies would have been hot enough to have split apart the molecules and be cooling by emission from hydrogen atoms. Recall from the virial theorem that as you form roughly spherical systems of matter, one-half the released gravitational potential energy must get radiated away. The other half goes into the kinetic energy of the system. The first galaxies also have to be massive enough to retain that gas that got heated by the collapse and the stellar activity of the mini-halos.

The James Webb Space Telescope launched in December, 2021, and the first science results were available by July 2022. With a large mirror and infrared capability, one of the objectives for the JWST is to push our observations of early galaxies to higher redshift. Gravitational lenses enhance the capacity of JWST to detect very early galaxies by making those galaxies look brighter than they otherwise would appear. In its first year of operation JWST observations caught galaxies with redshifts of ~13, i.e., meaning that galaxies had definitely already formed by the time the universe was about 330 million years old; subsequent observations have gone even deeper: the record as of May 2025 is $z = 14.44$, corresponding to an age of the universe of roughly 280 million years. Initial results suggest that we are seeing galaxies that had gotten more massive and had more star formation faster than we were expecting. This is exciting and presents a challenge for those studying the formation of structure in the early universe.

Galaxies grew not just by pulling in nearby cold gas but also by merging. Many of the earliest galaxies that we can see are irregular, blob-shaped objects, not the spirals and ellipticals that we see around us today, although some do seem to have recognizable structure. By $z \sim 7$, the most luminous galaxies already have masses of $10^{9\text{-}10}$ solar masses and are well into serious star-formation activity. By $z \sim 3$, many galaxies are clearly what we would today call "peculiar", actively merging and forming stars. During a collision between galaxies the gas and dust collide and for a time a galaxy may be referred to as a "starburst" galaxy, forming well more than the average number of stars per year. More massive galaxies are forming stars more rapidly than we can account for given the amount of gas each galaxy would bring to a merger, so the process of merging needs to include not just accreting additional early galaxies but also becoming massive enough to attract clouds of extra gas from the intergalactic medium. Relatively more of the massive galaxies we see at $z \sim 3$ were disk systems than today, where the distribution of massive galaxies is skewed in favor of ellipticals. Early disk galaxies weren't necessarily spirals, though: it seems to take time to develop arms (and more time to develop bars). By $z \sim 1$, about half of all stars had been formed and galaxies are massive enough that the Hubble sequence is identifiable, i.e., there are now more clear spirals and ellipticals than there are peculiar galaxies. Galaxies don't all grow at the same rate: the Milky Way, for instance, has many small companion galaxies and remnants of small galaxies that have accreted onto the Milky Way can still be identified, especially in the halo of our galaxy. Even today, galaxies have not stopped growing. Mergers still happen, and there are still peculiar colliding starburst galaxies.

If galaxies grow by merging, then there were more galaxies when the universe was younger than there are now. Estimates of how many galaxies we could see in the observable universe (at least in principle, if we could see early faint galaxies in the IR), taking into account how many we see around us now and that there should have been more earlier, suggest that there are should be over a trillion galaxies.

On a slightly esoteric note, galaxies don't all form stars at the same rate, which means that their chemical enrichments, i.e., their abundances of elements heavier than just hydrogen and helium, don't progress in lockstep. As mentioned briefly in the chapter on stellar evolution, the first C, N, and O were introduced into the interstellar medium, and were thus available for star formation, a bit before elements such as iron. In low-mass galaxies without robust star formation rates there are still today carbon-enhanced metal-poor (CEMP) stars. As some of those low-mass galaxies eventually accreted onto the Milky Way, we have some of these CEMP stars in the Milky Way, providing evidence of the universe's conditions at the time of some of the earliest star formation. A few other low-metallicity galaxies can still be found, not that far from home. As of mid-2016, the galaxy AGC 198691, estimated to be between about 25 and 50 million light years away, has the record for the galaxy with the lowest measured metallicity. It's a dwarf galaxy of roughly 10 million solar masses, its light dominated by some recently formed massive blue stars. Looking for remaining dwarf galaxies is also of interest to those trying to understand the nature of dark matter and, by extension, the role it plays in galaxy formation. As of 2016, there are some suggestions that the assumption made in the standard cosmological model that dark matter is predominantly cold, i.e., slow-moving, should result in more small galaxies than are seen today, e.g., as satellites of larger galaxies like the Milky Way.

To complicated matters a bit further, galaxies also don't all have the same proportion of dark matter. Some low-surface brightness galaxies, in particular, seem to be heavily dominated by dark matter. For instance, an ultra diffuse galaxy in the Coma Cluster called Dragonfly 44, roughly the mass of the Milky Way, may be over 99% dark matter. At the other end of the spectrum is a small ultra-diffuse galaxy of approximately $2 \cdot 10^8$ solar masses, NGC 1052-DF2 (also identified in the Dragonfly search program and named because it appears to be associated with the luminous elliptical NGC 1052), has at most a minimal amount of dark matter and possibly none at all. Another small ultra-diffuse galaxy in the same cluster, DF4, also show signs of minimal dark matter. In 2020 a team of astronomers suggested that, because stars in a galaxy are embedded in a more extensive dark matter halo, in tidal interactions between small galaxies and larger neighbors dark matter could be stripped from the small galaxy more readily than the stars could be. In other words, a small galaxy might have formed with a normal amount of dark matter but have lost it in interactions with more massive nearby galaxies. While galaxies such as these may complicate our understanding of galaxy formation, they do have the benefit of bolstering arguments for the reality of dark matter. That fact that galaxies can have different ratios of dark to baryonic matter implies that dark matter is real stuff and not simply an artifact of a misunderstanding of the gravity of baryonic matter.

As the galaxies are growing, so are the massive black holes in their centers ("supermassive black holes", SMBH, or, in smaller objects, "intermediate" mass black holes, just to be sure no one mistakes these for the stellar remnant variety!). Observations of a large number of galaxies show a fairly tight relationship between the mass of the central black hole and the mass of a galaxy's bulge, the latter being determined from the velocities of bulge stars. This is called the $M - \sigma$ relation, where M is the mass of the central black hole and σ is the velocity dispersion of the stars. The average stellar velocity may be zero, indicating equal numbers moving toward and away from us, but the dispersion, or the range of velocities around that average, will be larger for a more massive galactic bulge. (Not all galaxies are neatly symmetric: being a tad contrary, some dwarf galaxies have massive black holes that aren't central, perhaps having gotten kicked off center during interactions with other galaxies.)

The activity at the center of our Milky Way today is quite modest compared to the level of what are termed "Active Galactic Nuclei", or AGNs. An AGN has at least some of the following characteristics: an over-bright nucleus (compared to the rest of the galaxy), high luminosity (over $\sim 10^{37}$ J/s), high levels of emission from sources other than stars and the dust & gas heated by stars, rapid variability, jets, very Doppler-broadened emission lines. The most persuasive model for the high level of activity displayed by AGNs is one in which the activity is driven by quite a lot of material falling onto a massive central black hole; possibly the motion of the infalling material is influenced by the strength and orientation of the magnetic field near the galaxy's core. We see somewhat different properties for the various types of AGN depending on the amount of material available, the size, mass and type of the galaxy, and the angle at which we are observing the central engine.

High energy means that much of the gas near the black hole is likely to be ionized. One effect that will have on the spectrum of the galaxy is that we may see emission lines from ions; e.g., we may have emission not just from atoms of carbon, but from carbon that has been ionized two or three times. In addition to carbon, oxygen, nitrogen, neon, magnesium, sulfur are all abundant enough (produced in early generations of stars) by the time a galaxy becomes an AGN to contribute to the emission spectrum. The atoms don't care where the energy comes from to produce the ionization: it could be UV from hot stars or high-energy synchrotron emission, collisions among the particles themselves, shock waves from supernovae; there are a lot of possibilities. At low gas densities (less than about 10^8 atoms / cm^3), some of the emission lines we see are going to be "forbidden" lines. Recall that the 21-cm emission of neutral hydrogen is a forbidden transition, in that the electron violates a quantum mechanical selection rule in flipping its spin. The notation for a forbidden transition is square brackets: e.g., a particular forbidden green emission line from doubly ionized oxygen would be written [O III] 500.7 nm.

Physics note: permitted emission lines are electric dipole transitions; the forbidden ones are electric quadrupole or magnetic dipole or quadrupole transitions.

Synchrotron emission is the term for a broad range of emission from charged particles (usually electrons) spiraling around magnetic field lines. (See the chapter on the Milky Way for a sketch of the synchrotron spectrum compared to a blackbody spectrum. Properly, at lower electron energies, we'd call this cyclotron emission, but our sources of interest often have high-energy electrons.) Synchrotron emission has a power-law spectrum, meaning the flux goes as

$$F(f) = F_0 f^{-\alpha} \ .$$

If we plotted log flux vs. log frequency, α (called the spectral index) would be slope of the spectrum. Synchrotron emission is polarized. If the electrons are moving relativistically, the radiation is also beamed, meaning it is emitted into a cone in the direction the electron is moving (rather than being emitted isotropically). If we have enough electrons, some of the lower-energy, radio frequency, photons will scatter off the energetic electrons, boosting the photons' energies. This means we won't get as much flux at lower frequencies as the power-law equation would predict. Synchrotron emission is often called non-thermal, which is a bit frustrating because the electrons have kinetic energies and we could easily calculate an associated temperature. What non-thermal means in practice is continuous (not line) emission that isn't blackbody or bremsstrahlung. (The latter being electrons interacting with electric fields of ions, which is very unreasonably called thermal!) Non-thermal also tends to mean that we are not in thermal equilibrium: our particle energies do not follow a Maxwell-Boltzmann distribution. To get back to the emission itself, the fact that we see synchrotron emission from AGNs means we need a model that accounts for the presence of a magnetic field and a means of accelerating electrons to high speeds.

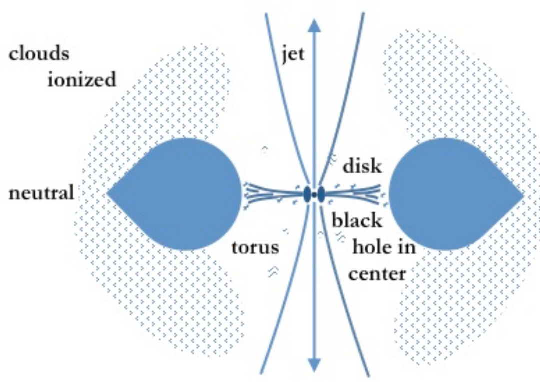


Figure 18.25: A sketch of the general model for an AGN.

The drivers for the activity are gravity and a magnetic field. At the center is a massive, growing, black hole. Material falling toward the black hole will pile up in a rotating accretion disk. Close to the black hole that disk will be quite hot, as material is falling in very fast and hitting the disk with a lot of energy, which can't be radiated away very efficiently. The hot inner regions will puff up and will radiate energy that will fall on the slightly more distant parts of the disk, heat it, and make it flare out slightly. The disk is not very large; for a central black hole of 10^8 solar masses, it might have a radius of ~1 pc. (The Milky Way's central black hole is ~$4 \cdot 10^6$ solar masses.) The disk will get clumpy and break into discrete clouds as we get near the outer torus. These clouds are still close enough to the central black hole that they are orbiting rapidly and they are hot, meaning that the gas is moving rapidly. It's a plasma and it's rotating, so it's no surprise that there are strong magnetic fields near the central black hole. Some of the magnetic field lines point outward, perpendicular to the rotating accretion disk. Plasma flows outward in a jet, expanding into huge lobes well outside the scale of the above sketch. Farther out in the plane of the disk, cooler and moving more slowly, is a larger torus of dust and gas. The hot inner material is emitting x-rays; this outer accretion torus is dense enough to be opaque to all but the highest energy x-rays.

Clouds of hot gas will produce emission lines. If the gas is very rapidly moving, the lines will be Doppler broadened. The regions that produce lines like that are called Broad Line Regions. If the gas is moving moderately rapidly, the lines will be less Doppler broadened and we have Narrow Line Regions. That terminology is a little confusing, because in neither case is the gas moving slowly! In broad lines, the line width implies gas velocities of several thousand km/sec; in narrow regions, it's still moving at several hundred km/sec. The forbidden lines, mentioned above, don't tend to come from the Broad Line Regions, which makes sense if the fastest moving material is in closer to the center where the gas is denser and there are more collisions.

Emission from near the center of the AGN, i.e., broad emission lines & x-rays, may show rapid variability. For emission to vary on a time scale of weeks implies that the emitting region must be less than lightweeks across. If it were larger, then light from the near and far sides would reach us at sufficiently different times that the signal "bright now" would be washed out. Here's a sketch illustrating this:

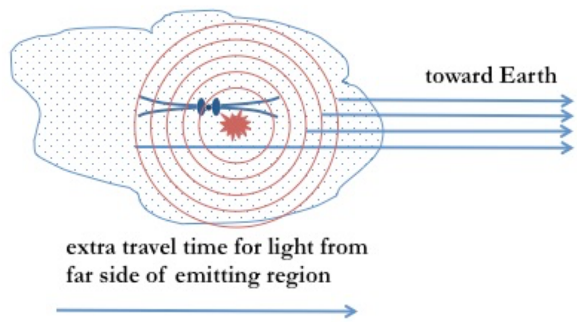


Figure 18.26: Why rapid time variability of an AGN implies small spatial extent.

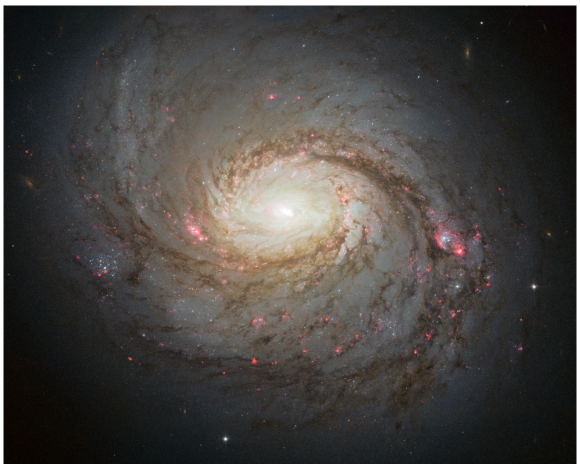
Suppose that a flare occurs near the center of this region, perhaps as a chunk of matter hits the accretion disk; material in the hatched region will brighten as light propagates outward from the flare. Light will propagate toward us not just from the flare itself but from the brightening material around it. Light coming toward us from the near side of the flare has a head start over light coming toward us from the brightening regions on the far side of the flare. If we see the near side back to its normal dim state before we ever get the message that the far side is bright, the signal that a flare happened may pretty much totally cancel out. Only if the region is small enough that it effectively all brightens at once ("at once" meaning over, say, a week or so) will we know that there was a flare. The fact that we observe variability on time scales of weeks says that the regions doing the varying can't be larger than lightweeks across. Note that some active nuclei that are very distant and thus nearly point sources on the sky (e.g., quasars, about which more below) will vary, especially at radio frequencies, over much shorter timescales, perhaps a few hours, for reasons that have nothing to do with the AGN itself. Just as stars low on the horizon will scintillate or twinkle as their light passes through regions in our atmosphere that have different densities, temperatures, motions, etc., radio waves from a point-source-sized AGN may pass through warm and variable streams of gas in the ISM, leading to a similar scintillation effect.

As mentioned above, what sort of activity we are going to see depends in part on the angle with which we are viewing the active region. Let's consider some of the possibilities and look at some of the types of AGNs.

Suppose for instance that we were looking end-on at the jet coming out perpendicular to the inner accretion disk. High-energy electrons are going to be spiraling around the magnetic field lines and emitting synchrotron radiation, which is polarized, continuous over a broad range of frequencies, and beamed in our direction. The amount of material in the jet may vary rapidly, because it derives from the material, possibly clumpy, streaming inward from the sides toward the inner accretion disk, and thus the brightness we observe will vary. We won't see much in the way of emission lines because any signal from those background Broad Line Region clouds is drowned out by the bright jet. This is the consensus model for the majority of the type of AGN called BL Lacertae objects. The prototype, BL Lac, was first thought to be a variable star (variable stars are named for the constellation in which they are located and designated by letters indicating the order of their discovery) because initially telescopes were not powerful enough to detect the surrounding galaxy.

Next, consider the class of galaxies called Seyferts, named for astronomer Carl Seyfert. These are, mostly, spiral galaxies with very bright nuclei, emission lines that arise from either or both the Broad Line Regions and the Narrow Line Regions, and some non-thermal continuous (synchrotron) emission. If we are observing the galaxy nearly perpendicular to the inner disk we might see emission from several Broad Line Regions; if we are observing from enough of an angle that our line of sight to the Broad Line clouds is blocked by the outer torus we might only get the Narrow Line Regions. There are quite a few Seyfert galaxies (~90 nearby; estimates are that between 5 – 15% of galaxies are Seyferts), enough that they have subtypes based on the widths of their emission lines. Those dominated by broad lines are Type 1, those dominated by narrow lines are Type 2, and there are intermediate types 1.2, 1.5, 1.8, and 1.9 for galaxies with a mix of narrow and broad emission lines.

The following is a Hubble Space Telescope image of M77 (NGC 1068), which is a Type 2 Seyfert, and in the next panel, the visible wavelength region of its spectrum.



http://www.spacetelescope.org/news/heic1305/

Figure 18.27: M77; image credit: NASA, ESA & A. van der Hoeven

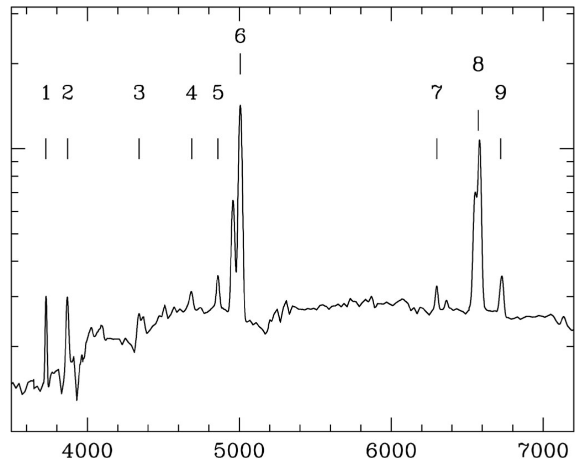


Figure 18.28: M77 (NGC 1068) spectrum

http://ned.ipac.caltech.edu/level5/March04/Risaliti/Figures/figure10.jpg

The emission lines marked in the spectrum are due to: 1) [O II] λ3727 Å; 2) [Ne III] λ3869 Å; 3) Hγ; 4) He II λ4687 Å; 5) Hβ; 6) [O III] λ5007 Å; 7) [O I] λ6300 Å; 8) Hα blended with [N II] λ6586 Å; 9) [S II] λ6732 Å. (Recall that 10 Å = 1 nm.) The widths of the emission lines correspond to gas motions of ~300 – 800 km/s and include both forbidden and permitted transitions. The continuum spectrum underlying the emission lines is dominated by the light from the stars.

NGC 4151 is a Type 1 Seyfert, with several very broad emission lines. The image is a composite of x-rays (from Chandra) shown in blue, optical / Hα (from the Kapteyn Telescope on La Palma) in yellow, and radio / neutral hydrogen (from the VLA) shown in red. The ultraviolet spectrum on the right was obtained with the Hopkins Ultraviolet Telescope flown on the Space Shuttle in 1990.

Figure 18.29: NGC 4151

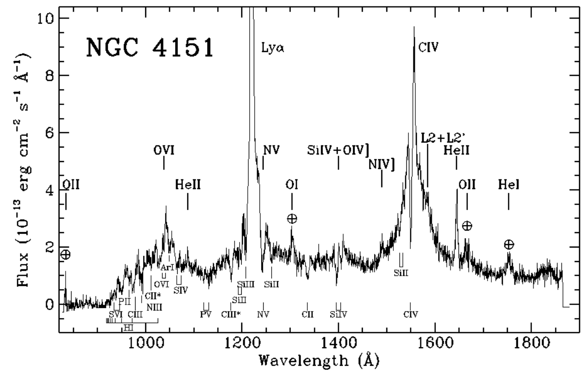


Figure 18.30: NGC 4151 spectrum

X-ray: NASA/CXC/CfA/J.Wang et al.; Optical: Isaac Newton Group of Telescopes, La Palma/Jacobus Kapteyn Telescope, Radio: NSF/NRAO/VLA
http://www.chandra.harvard.edu/photo/2011/n4151/; http://archive.stsci.edu/hut/images/NGC4151_fig1.gif

NGC 1275 (Perseus A) is a cD galaxy about 68 Mpc away, at the center of the Perseus cluster of galaxies. It is a Type 1.5 Seyfert and also a radio galaxy (the designation "A" means it was the first radio source identified in Perseus) and also possibly caught in active collision with another galaxy, something dusty and blue and definitely moving toward NGC 1275. Filaments of gas tens of kpc long, emitting in Hα, appear to stretch out of the core of the galaxy into the hot surrounding intergalactic medium, moving rapidly but constrained by magnetic fields so that the filament material doesn't dissipate into the IGM. NGC 1275 emits at the high-energy end of the spectrum, in x-rays and gamma rays, although at high energies, the galaxy's emission is a bit hard to separate from the emission from the hot intracluster medium. Radio observations show distinct jets coming out from the bright core. The jet on the near side isn't too far off our line of sight, making Perseus A almost a BL Lac object. In the following panels, the image on the left is an HST image in the optical (emphasizing Hα); the one on the right is a composite of optical plus x-rays (Chandra, violet) and radio (VLA, pink). Even for an AGN, this is a pretty wild galaxy, definitely an object of on-going research.

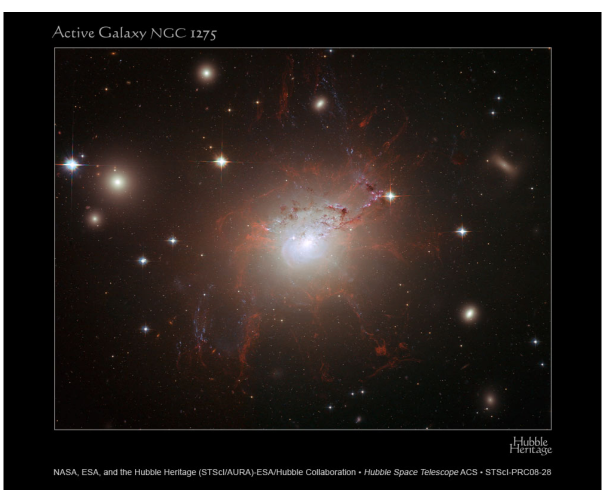


Figure 18.31a
http://hubblesite.org/newscenter/archive/releases/2008/28/image/a/

Figure 18.31b; Image credit: NASA, ESA, L.Frattare (STScI); http://hubblesite.org/newscenter/archive/releases/2008/28/image/c/

Radio galaxies are most often massive elliptical AGNs that are extra luminous at radio frequencies, usually due to synchrotron emission from jets. To the extent that more massive galaxies are more likely to have more massive central black holes and that massive galaxies are more likely to be ellipticals it's not unreasonable that huge jets and radio lobes are more often associated with ellipticals than with spirals. Jets themselves sometimes extend far out into space: observations reported in 2024 showed a pair of jets reaching more than 10 million light years from their host AGN.

Centaurus A (NGC 5128) is a fairly close example of a radio galaxy, only ~4 Mpc away. It is classified as an E0 (or possibly an S0) peculiar galaxy and looks like an elliptical with the dusty disk of a spiral cutting through it. The black hole at the center has a mass of ~55 million solar masses.

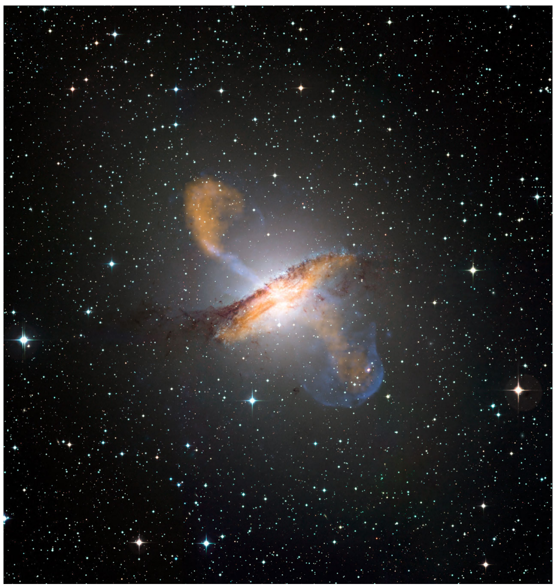


Figure 18.32: Centaurus A; credit: ESO/WFI (Optical); MPIfR/ESO/APEX/A.Weiss et al. (Submillimetre); NASA/CXC/CfA/R.Kraft et al. (X-ray) http://www.eso.org/public/images/eso0903a/

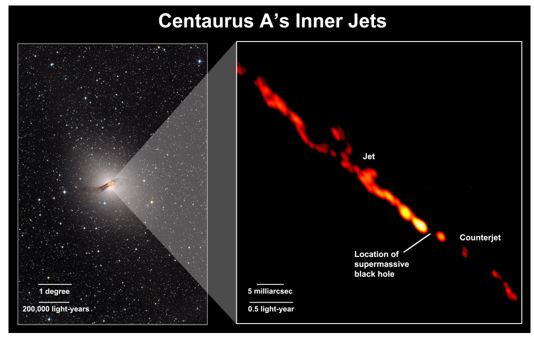


Figure 18.33; credit: NASA/TANAMI/Müller et al. http://pulsar.sternwarte.uni-erlangen.de/tanami/public/gallery/Cen_A_inner_jet.jpg

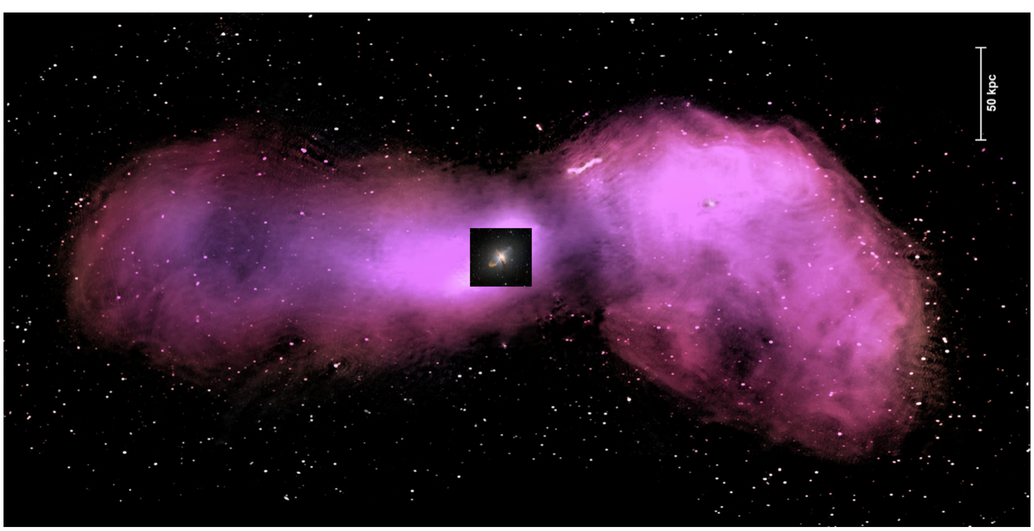


Figure 18.34: Centaurus A; credit – Ilana Feain, Tim Cornwell & Ron Ekers (CSIRO/ATNF). (Rotated 90 degrees to fit on the page.)

Above is a composite picture combining an optical image from the European Southern Observatory, x-rays from Chandra (shown in blue), and 870 micron observations from APEX (in orange), a high-altitude radio dish in the Atacama in Chile. Particles in the jets, moving at nearly half the speed of light, collide with gas surrounding the galaxy and produce x-rays. The inner part of the jet shows knots and kinks; it's not a steady flow of particles. A radio image of the inner part of the jet is shown in the second panel. The jets feed outer radio-bright lobes that are simply immense: the visible galaxy is about 1/3 of a degree across; the full extent of the radio lobes covers 10 degrees of declination on the sky. The radio lobes are shown in the third panel, in 1.4 GHz observations from Australia's CSIRO (where the background white dots are distant radio sources, not nearby stars). The visible galaxy is superposed over the radio.

M87 (NGC 4486 = Virgo A) is another nearby radio galaxy, a giant elliptical at the heart of the Virgo cluster ~16.4 Mpc away and ~37,000 pc across. It is also an E0 or cD peculiar galaxy, although it hasn't got the dust that Centaurus A has. It does, however, have the jets. What look like stars in the outer regions of the galaxy in the HST image are actually globular clusters. On the right is a montage of radio observations, particularly focused on the jet.

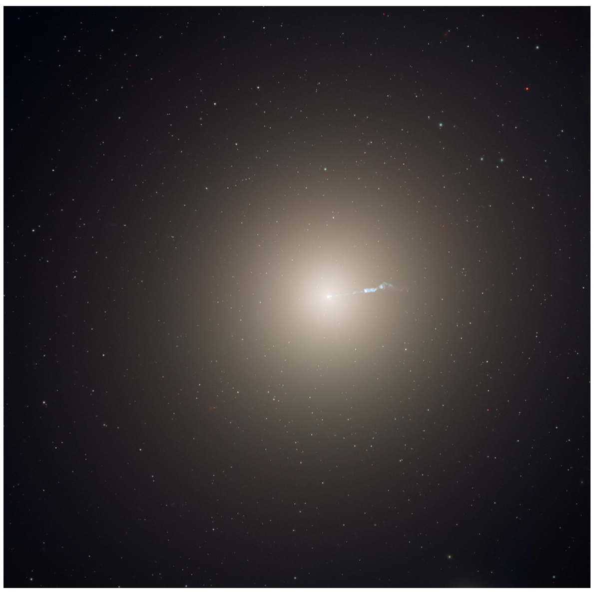


Figure 18.35

http://hubblesite.org/newscenter/archive/releases/2008/30/image/f/

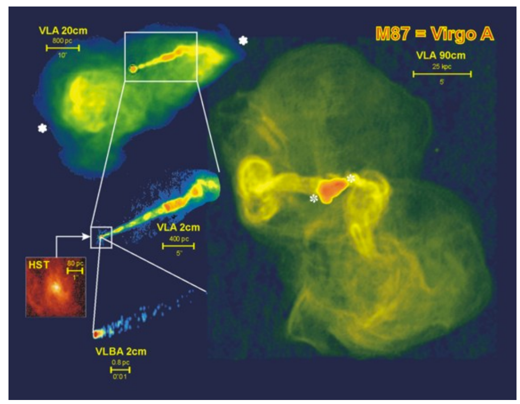


Figure 18.36; courtesy of NRAO/AUI
http://images.nrao.edu/270

In April 2019 the Event Horizon Telescope (EHT) team released an image of the supermassive black hole in the center of M87. EHT is a planet-wide collaboration involving eight high-altitude radio telescopes, a team of over 200 researchers, and several decades of work. The event horizon of the black hole is smaller than the shadow that it casts (see figure 18.37).

In 2021 the Event Horizons team published additional data adding polarization to the intensity map of the light near M87's central black hole. The polarization indicates that the emitting region is hot and magnetized. This will help improve our understanding of the relationship between magnetic fields and the synchrotron-emitting jets from the centers of active galaxies.

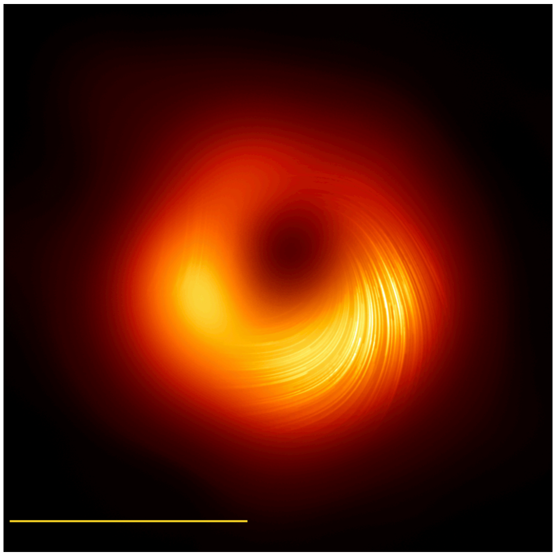


Figure 18.37: Shadow of the 6.5 billion solar mass black hole in the center of M87. The overlay lines indicate the polarization of the light. The scale bar is 50 microarcseconds.
Credit: Event Horizon Telescope Collaboration.
https://apod.nasa.gov/apod/ap210331.html

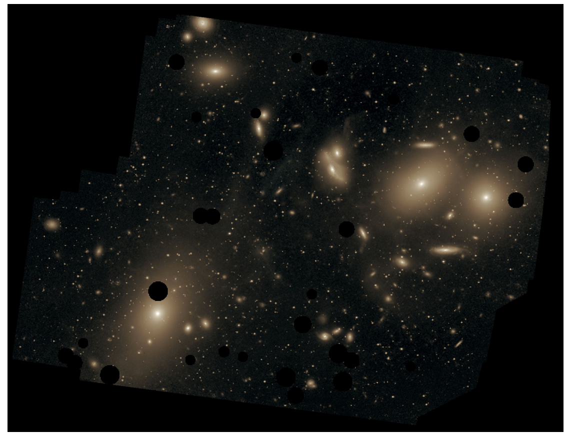
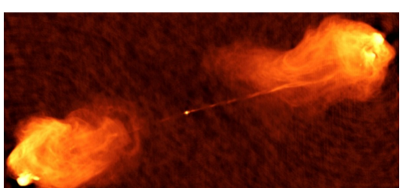


Figure 18.38: The center of the Virgo cluster; M87 is the large galaxy on the lower left.
Credit: Chris Mihos (Case Western Reserve University) / ESO)

http://www.eso.org/public/images/eso0919a/

The right-hand image, above, shows the local environment for M87, i.e., the heart of the Virgo cluster. The black dots are where bright foreground stars have been removed. M87 has a mass of ~$6 \cdot 10^{12}$ solar masses; its central black hole, at ~$6.5 \cdot 10^{9}$ solar masses, is among the most massive known.

The extended lobes of radio galaxies are interacting with the intracluster medium. Some show hot spots and x-ray emission at the ends of the lobes where the plasma in the lobe is slamming into the surrounding gas. Cygnus A is a radio galaxy, ~200 Mpc away, oriented nearly edge-on to our line of sight, so we can clearly see both jets and both radio lobes. Below is a 5GHz map of Cygnus A from observations made at the VLA; the visible galaxy is near the bright central point:



Cygnus A

Figure 18.39 a:
Radio image courtesy of NRAO/AUI
http://images.nrao.edu/110

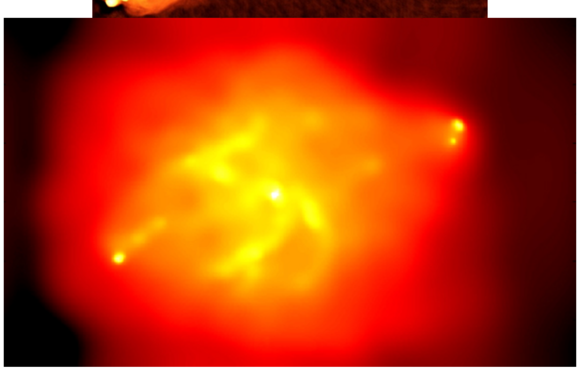


Figure 18.39 b:
Chandra X-ray; credit: NASA/UMD/A.Wilson et al.
http://chandra.harvard.edu/photo/2000/0216/

To approximately the same scale, the lower panel is an x-ray image from Chandra. Notice that the bright spots in the radio image at the ends of the radio lobes correspond to bright spots in the x-ray image, as well. These bright spots are ~90 kpc from the center of the galaxy.

Galaxies move within their clusters. Sometimes this results in jets that are blown backward, deflected by the surrounding gas. NGC 1265 (3C 83.1), another Perseus cluster galaxy, is a classic case, shown in the following image. The galaxy is moving at ~2,000 km/sec with respect to the IGM.

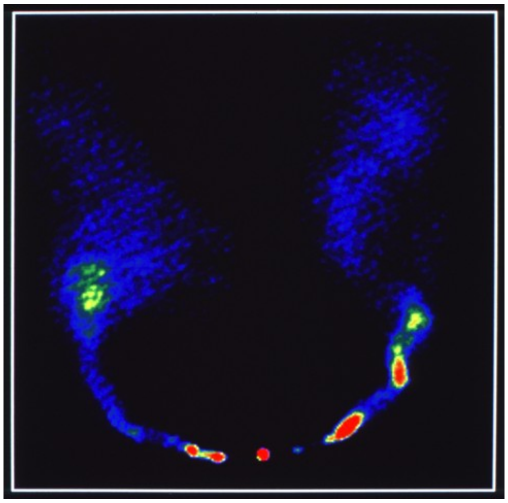


Figure 18.40:
Material blown backward from NGC 1265 by the IGM; image courtesy of NRAO/AUI and C. O'Dea & F. Owen

http://images.nrao.edu/19

The most well known class of AGNs are the Quasars. Quasar stands for Quasi-stellar radio source. The first of these objects were discovered as bright radio sources, eventually identified with optical counterparts, but at least initially telescopes were not capable of resolving the optical counterparts into a bright nucleus and surrounding galaxy. They just looked like stars, albeit radio-bright "stars" with astonishingly large redshifts. Turns out that most aren't radio loud, so today we also use Quasi-stellar object (QSO) to refer to this class of extremely energetic AGNs.

Radio astronomy can claim its origins in the 1930s, when Karl Jansky of Bell Labs first noticed that a particularly annoying source of shortwave radio static interference was none other than the center of the Milky Way. World War II interrupted the nascent discipline, but did provide an impetus for acquiring skills in ionospheric research and the development of radar. After the war radio astronomy made rapid strides, aided by the development of computers. Angular resolution depends inversely on wavelength, meaning that it is harder to discern small angular differences on the sky when observing at longer wavelengths. Radio telescopes are larger for this reason. But to rival optical resolutions an individual radio dish would have to be impractically large. The radio interferometer, digitally combining the signals from several individual antennas, makes decent resolution possible but is computationally intensive. The "3C" prefix to many famous radio sources refers to the Third Cambridge Catalog of Radio Sources, produced in the 1950s by one of the earliest (and still significant) radio astronomy research groups, using an interferometer near Cambridge, England.

Two of the first quasars to be identified are 3C 48 and 3C 273, objects identified as radio sources in the Cambridge surveys of the late 1950s. It took some time before the optical counterparts were identified for these radio sources; determining the position for 3C 273 was aided by lunar occultations. The major breakthrough in our understanding came when Maarten Schmidt, using the 200-inch Hale Telescope on Mount Palomar, obtained a fuzzy spectrum for 3C 273 and realized that the absorption lines he was seeing were normal hydrogen lines, redshifted by a whopping 15.8 %. That's 47,000 km/sec, meaning this was no regular "star". Astronomers Jesse Greenstein and Thomas Matthews soon determined that 3C 48 has a redshift of 0.367. The huge redshifts were a puzzle. On the one hand, applying the Hubble relation would imply that these objects should be at cosmological distances. On the other hand, given the apparent magnitudes, accepting the distances meant having to explain the prodigious energy source needed to power these objects.

Example: 3C 273 has an apparent visual magnitude of ~12.9 and a luminosity distance of ~749 Mpc. How bright is it?

Ignoring extinction, $m_V - M_V = 5 \log(d) - 5$ becomes

$$12.9 - M_V = 5 \log(749 \cdot 10^6) - 5$$

$$M_V = -26.7$$

$$f_{273} / f_{sun} = 10^{0.4\,(4.8 + 26.7)} = 4 \cdot 10^{12}$$

3C 273 is more than 10^{12} times as luminous as the Sun.

It was approximately 20 years, into the early 1980s, before telescopes were large enough and detectors sensitive enough to detect the galaxies surrounding these incredibly bright nuclei. Using today's best values for the expansion and age of the universe, we get that the light from 3C 48 has been traveling toward us for 4 billion years, nearly a third the age of the universe. It's not surprising that it was initially a challenge to accept those numbers for something that looked simply like an odd sort of star.

The image below on the left shows the original spectrum of 3C 273. Below the quasar spectrum is the comparison spectrum used to establish the wavelength scale. Three of the hydrogen Balmer series lines are marked on the quasar spectrum and their unshifted locations are marked on the comparison spectrum. On the right is an HST image of 3C 273, in which you can tell that the quasar has a jet.

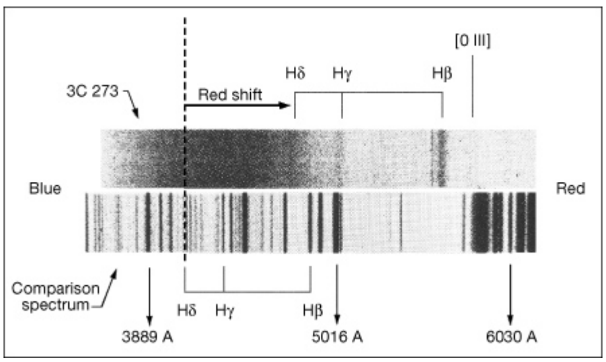


Figure 18.41a: Palomar Observatory optical spectrum of 3C 273; based on J.L. Greenstein and M. Schmidt 1964, *Ap.J.* **140**, "The Quasi-Stellar Radio Sources 3C 48 and 3C 273".

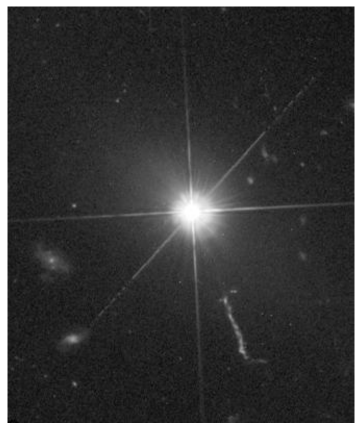


Figure 18.41b: 3C 273; credit: NASA and J. Bahcall http://hubblesite.org/newscenter/archive/releases/2003/03/image/c/

Quasars have been identified as early as $z = 6 - 7$. The most distant known in 2021 is J0313-1806, at $z = 7.64$, meaning we are seeing this object when the universe was only ~ 670 million years old. A decade earlier the most distant quasar record holder was ULAS J1120+0641, at $z = 7.085$, the first quasar identified with a redshift > 7.

Figure 18.42: a *very* distant quasar

http://www.eso.org/public/images/eso1122b/

Credit: ESO/UKIDSS/SDSS

Caption: This image of ULAS J1120+0641, a very distant quasar powered by a black hole with a mass two billion times that of the Sun, was created from images taken from surveys made by both the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. The quasar appears as a faint red dot close to the centre. This quasar is [among] the most distant yet found and is seen as it was just 770 million years after the Big

These objects developed pretty early. Remember that by z of 7-ish we are, on average, just starting to see galaxies with masses of 10^{10} solar masses. It's a bit puzzling to see such energetic galaxies this early, having already developed billion-solar-mass supermassive black holes. One possibility is that some of those early dark matter halos could have been over-dense enough to collapse directly into intermediate-mass black holes, jump-starting the process of forming supermassive black holes.

Over 200,000 quasars have been identified, with the peak of the distribution being between z = 1 to 3. The majority of known quasars have been identified by the Sloan Digital Sky Survey; the majority (perhaps over 90%) are not strong radio sources.

Note that the most distant galaxies are not necessarily quasars; in 2016 astronomers using the Hubble Space Telescope reported observations of an infant galaxy, smaller than the Milky Way but forming stars much more rapidly. GN-z11, as its name suggests, has a z = 11.1. Such an object is about at the limit for the distance to which the HST can detect objects. The light from GN-z11 left on its way to us when the universe was only about 400 million years old! Observations of very young galaxies made JWST and ALMA have identified several potential galaxies even farther away, with redshifts over 13. Understanding how stars and galaxies formed when the universe was so young is a challenge.

Some energetic quasars have jets pointing nearly toward us, akin to the BL Lac objects. The term blazar is used to refer to these objects. Recall that the jets are often irregular, as the plasma is ejected in blobs, rather than a steady stream, and often at relativistic speeds, creating bright shock fronts where it slams into the nearby IGM. High speeds and geometry can collude to make blobs in jets appear to be moving faster than light. This apparent superluminal motion works like this example in the following image:

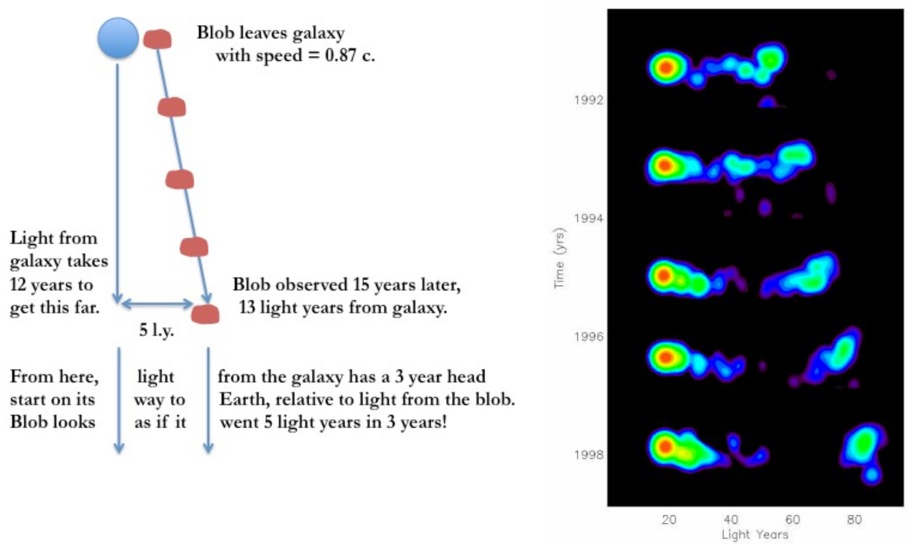


Figure 18.43 a: example of apparent superluminal motion; b: radio image of 3C 279; image courtesy of NRAO/AUI; http://images.nrao.edu/387

In the second panel are VLA observations of 3C 279. In this case, the quasar jet is only 2 degrees off our line of sight and the blob is moving at 0.997 c. That's fast, but it's only when coupled with the geometry that the blob appears to be moving faster than the speed of light.

The energy that's powering the AGN ultimately derives from material falling into the black hole. How much mass does it take? If we could convert *all* of the rest mass of the infalling material into energy, we'd get out mc^2. Theoretical work on black holes says that it is not possible to convert all the mass into radiation. The maximum possible is about 0.42 mc^2, and in practice it's likely to be less than that, perhaps only ~0.1 mc^2 would get converted to radiation and the other 90% of the mass would simply add to the mass of the black hole.

Example: How much mass would be required to produce an energy output of $10^{12}\ L_\odot$?

The Sun's luminosity is $3.85 \cdot 10^{26}$ J/s. Multiply that by 10^{12} to get the energy needed and set that equal to 0.1 mc^2:

$$3.85 \cdot 10^{38} \text{ J/s} = 0.1\ m\ (3 \cdot 10^8 \text{ m/s})^2$$

$$m = 4.2 \cdot 10^{22} \text{ kg/s}$$

Convert that to mass / year: $(4.2 \cdot 10^{22} \text{ kg/s}) \cdot (3.16 \cdot 10^7 \text{ s/yr}) = 1.4 \cdot 10^{30} \text{ kg / yr} = 0.7\ M_\odot \text{ / yr}$.

That's nearly a solar mass per year, 90% of which is adding to the mass of the black hole, and 10% being radiated away.

Suppose we asked this question using the virial theorem? In other words, mass falling onto an accretion disk around a black hole will lead to a loss of gravitational potential energy, one-half of which could get radiated away. The change in the gravitational potential energy per unit time for an object of mass M and radius R accreting mass at a rate dm per unit time is given by

$$\frac{dE}{dt} \approx \frac{GM}{R}\frac{dm}{dt}.$$

Solve this for the accretion rate, using the luminosity above and assuming that 1/2 the energy can be radiated away; we can approximate the radius to be that of a black hole's event horizon, namely ~ 3 km per solar mass:

$$\frac{dm}{dt} \approx \frac{2 \cdot L \cdot R}{GM} \approx \frac{2 \cdot 3.85 \cdot 10^{38} \text{ J/s} \cdot (3{,}000 \text{ m}/M_\odot) \cdot (10^8 M_\odot)}{(6.67 \cdot 10^{-11} \text{kg m}^3/\text{s}^2) \cdot 10^8 M_\odot \cdot (2 \cdot 10^{30} \text{ kg})} \approx 1.7 \cdot 10^{22} \text{kg/s} \rightarrow$$

$$1.7 \cdot 10^{22} \text{kg/s} \cdot \frac{3.16 \cdot 10^7 \text{s/yr}}{2 \cdot 10^{30} \text{ kg}/M_\odot} \approx \frac{1}{4} M_\odot \text{/ yr}.$$

Given the approximations we're making, that's basically the same answer as we got from using $E = mc^2$.

Sounds like a semi-infinite power source, if we could just tap the black hole's energy. Technological issues aside, there is some evidence that suggests that supermassive black holes won't just keep growing indefinitely by swallowing more and more gas from their surroundings. There's that accretion disk around the black hole; at some point, enough gas has fallen onto the accretion disk to form new stars, which are much less likely to lose angular momentum and fall into the black hole than is the cluttered, interacting, energy-losing material in an accretion disk. Observationally, we don't tend to see SMBHs larger than a few 10^{10} solar masses.

As mentioned above, galaxies grow in large part by merging with other galaxies. Let's look at some examples from the relatively nearby universe, where we can see the interactions in more detail. Off the handle of the Big Dipper, in the constellation Canes Venatici, the grand design spiral M51 (NGC 5194), also called the Whirlpool, has been warped by the close passage of a smaller companion. The smaller irregular galaxy, NGC 5195, is farther away from us than its larger companion. Bright pink-red emission nebulae and blue stars indicate numerous regions of active star formation, possibly aided by ripples set up in the larger galaxy as the smaller one plowed past. M51 is about 7 Mpc away, and about $1.6 \cdot 10^{11}$ solar masses.

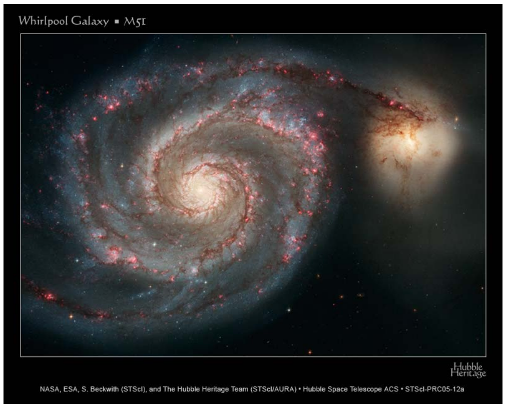


Figure 18.44
http://hubblesite.org/newscenter/archive/releases/2005/12/image/a/format/web_print/

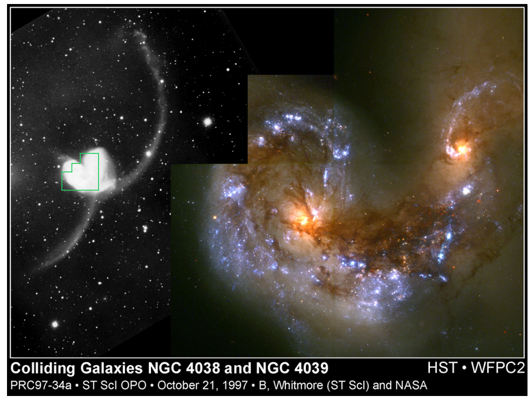


Figure 18.45
http://www.spacetelescope.org/images/opo9734a/

Sometimes spiral arms get more than just bent by the tidal interaction with a passing galaxy. In the right-hand panel, above, are NGC 4038 and 4039, often called the Antennae because of the long trails of stars that have been flung out into intergalactic space by the collision of the two galaxies. When galaxies collide, the stars are likely to miss each other. Clouds of gas and dust, on the other hand, are going to collide. The central region of the Antennae is lit up like a fireworks display as several hundred new clusters of young stars are forming. The Antennae are about 15 Mpc away; they may have started colliding nearly a billion years ago, slowly approaching and passing through each other. In another half a billion years their nuclei may have fallen into each other, merging into one, possibly elliptical, galaxy.

M82 (NGC 3034), sometimes called the cigar galaxy, is a rather odd-looking starburst galaxy located across the Big Dipper from M51, about 3.7 Mpc from us. Its activity may have been triggered by the close passage of M81, an innocent-looking nearby spiral. The following HST image, on the left, emphasizes its blue stars and hydrogen Hα emission. The right-hand image adds x-rays from Chandra (in blue) and infrared from Spitzer (in red). There are a lot of explosive events happening in M82.

Figure 18.46a
http://hubblesite.org/newscenter/archive/releases/2006/14/image/a/

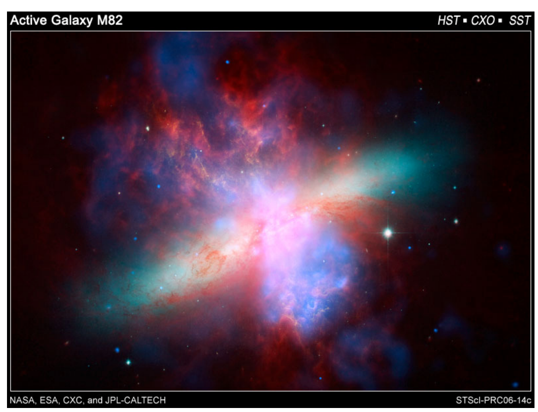


Figure 18.46b
http://hubblesite.org/newscenter/archive/releases/2006/14/image/f/

When we looked at Centaurus A, above, it was principally in the context of its radio emission. But that dusty lane across the center is not normal for an elliptical galaxy. The Spitzer space telescope took images of Centaurus A in the infrared and it appears from the Spitzer data that the dust may be the remnants of a spiral galaxy, twisted a bit during the merger to resemble a parallelogram (below left). In addition, CO gas, which lies in the plane of a spiral galaxy along with the neutral hydrogen, is present in the dusty disk of Centaurus A. The right-hand panel is CO emission at 1.3 mm wavelength, Doppler shifted because of the motion of the gas within the galaxy; the violet color, on the left side, indicates gas that is moving toward us while the light blue on the right side indicates gas moving away. The CO observations were made with the Atacama Large Millimeter Array (ALMA) in the high desert in Chile.

Figure 18.47a: IR Centaurus A; credit: NASA / JPL-Caltech / J.Keene
http://www.spitzer.caltech.edu/images/1192-ssc2004-09a1-Dusty-Elliptical-Galaxy-Centaurus-A

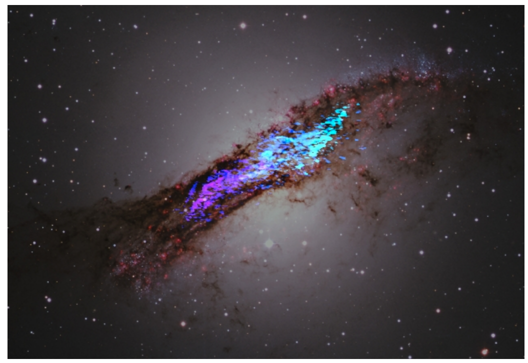


Figure 18.47b: Centaurus A CO emission; credit: ALMA (ESO/NAOJ/NRAO), T.A. Rector (University of Alaska Anchorage). Visible-light image: ESO
http://www.nrao.edu/pr/2012/almacena/

Massive elliptical galaxies definitely seem to form from the mergers of many smaller galaxies. NGC 1132 is a massive elliptical galaxy with thousands of globular clusters and dwarf companion galaxies about 90 Mpc away, in the direction of the constellation Eridanus. In this HST image, on the left, you can also see quite a few more distant (and generally more red) galaxies. What doesn't show, in visible wavelengths, is the x-ray glow around the foreground galaxy, bright enough to indicate an amount of gas that would usually be associated with a cluster of galaxies. The image on the right overlays Chandra x-ray observations in blue and purple. Mapping the dark matter halo in and around NGC 1132 also indicates an amount of mass normally found in a cluster of many tens of galaxies. Some astronomers call groups such as this, i.e., NGC 1132 and its dwarf companions, a fossil group, meaning that it is likely to be the remnant of a group of galaxies that has almost entirely merged to form this one giant elliptical system.

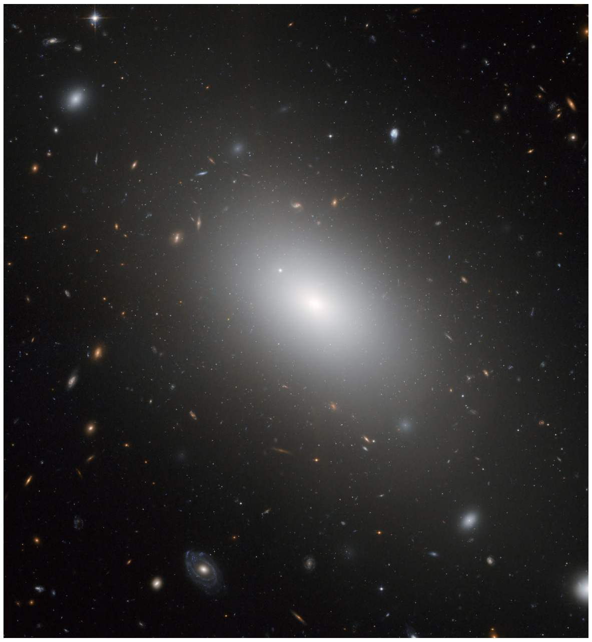


Figure 18.48 a: NGC 1132; credit: NASA, ESA, and the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration. Acknowledgment: M. West (ESO, Chile)
http://www.spacetelescope.org/images/heic0804a/

b: with x-rays; Credit: NASA, ESA, M. West (ESO, Chile), and CXC/Penn State University/G. Garmire, et al.

http://www.spacetelescope.org/images/heic0804b/

And while we are on the topic of collisions, yes, we and M31 (Andromeda) are headed toward each other at ~300 km/sec and will begin to collide in a bit less than 4 billion years.

Large-scale structure

Galaxies today mostly live in clusters and those clusters are found in superclusters, strung together around huge voids like the film around soap bubbles. Clusters don't seem to form quite as early in the evolution of the universe as galaxies themselves. The earliest cluster we know (as of 2016) is called CL J1001 +0220. Multi-wavelength observations of this cluster show evidence of seventeen galaxies, several of which are massive starburst galaxies, and a large amount of hot, x-ray emitting gas, at a $z \approx 2.5$, or roughly 11 billion years ago. ALMA, the Atacama Large Millimeter/submillimeter Array, has detected signs of possible cluster formation, "protoclusters" at least as early as 12.4 billion years ago. Superclusters and the large-scale structure of voids surrounded by filamentary strings of galaxy clusters, may take longer to form, or at least longer to become well-enough defined for us to detect them today. The Lynx supercluster, at a $z \approx 1.3$, or roughly 8.8 billion years ago, is a contender for most-distant-known supercluster. A distant and somewhat amorphous collection of young galaxies, nicknamed Hyperion, was identified in 2018 as a possible proto-supercluster. It has a $z \approx 2.4$ (so we are seeing it roughly 11 billion years ago), a total mass on the order of $5 \cdot 10^{15}$ solar masses, and a size comparable to current superclusters, but its galaxies are more loosely connected, presumably because gravity has had less time to impart the denser structures we associate with a supercluster.

Our Local Group of galaxies is part of the Virgo Supercluster, named for the largest and most central of the 100 or so clusters of galaxies that make up our Local Supercluster. The Virgo Supercluster is over 30 Mpc across and contains on the order of 10^{15} solar masses of matter. The luminosities of the galaxies only add up to about 3

$\cdot 10^{12}$ solar luminosities, meaning we have a mass / light (M / L) ratio of about 300, in solar units, well more than the M / L ratio of ~64 for the Milky Way galaxy.

The motions of individual galaxies will include a component that is due to the expansion of the universe and another, a peculiar velocity, that is indicative of the way each individual galaxy is responding to its local gravitational field. Just as we can describe the motions of stars within a galaxy as a response to the gravitational force they feel, so, too, do galaxies respond to the gravitational pull of other galaxies. This is easier to measure for galaxies that aren't too far away, because, while the peculiar velocities might be similar anywhere, more distant galaxies have a larger expansion velocity which, as we move farther away, will soon be overwhelmingly larger than the peculiar velocities. It's also easier to measure distances more accurately for nearby galaxies, and we need the distance to know what portion of a galaxy's velocity is due to the expansion.

In 2014 astronomers R. Brent Tully and colleagues published a map based on the peculiar velocities of nearby galaxies and determined that the Virgo Supercluster is one lobe of a larger supercluster that they are calling Laniakea. As they state in the abstract to their 2014 paper:

> Galaxies congregate in clusters and along filaments, and are missing from large regions referred to as voids. These structures are seen in maps derived from spectroscopic surveys that reveal networks of structure that are interconnected with no clear boundaries. Extended regions with a high concentration of galaxies are called 'superclusters', although this term is not precise. There is, however, another way to analyse the structure. If the distance to each galaxy from Earth is directly measured, then the peculiar velocity can be derived from the subtraction of the mean cosmic expansion, the product of distance times the Hubble constant, from observed velocity. The peculiar velocity is the line-of-sight departure from the cosmic expansion and arises from gravitational perturbations; a map of peculiar velocities can be translated into a map of the distribution of matter. Here we report a map of structure made using a catalogue of peculiar velocities. We find locations where peculiar velocity flows diverge, as water does at watershed divides, and we trace the surface of divergent points that surrounds us. Within the volume enclosed by this surface, the motions of galaxies are inward after removal of the mean cosmic expansion and long range flows. We define a supercluster to be the volume within such a surface, and so we are defining the extent of our home supercluster, which we call Laniakea. *Nature*, 2014, **513**, pp 71-73.

Laniakea is centered on the Hydra-Centaurus Supercluster. Virgo is an offshoot, as is the Pavo-Indus Supercluster. It appears to have a mass on the order of 10^{17} solar masses, in at least 300 clusters, stretching across ~160 Mpc. The *Nature* publication includes a video to help illustrate the motions of the galaxies. The following image is a frame from a preview video posted by the National Radio Astronomy Observatory. The Milky Way is the dark blue dot at (0, 0). Each blue dot is a galaxy. Virgo is highlighted in yellow. The darkest blues indicate voids, i.e., regions with much lower densities of galaxies. The white and dark blue arrows represent the flow lines, showing the directions galaxies are moving. The orange contour represents the approximate extent of the Laniakea Supercluster.

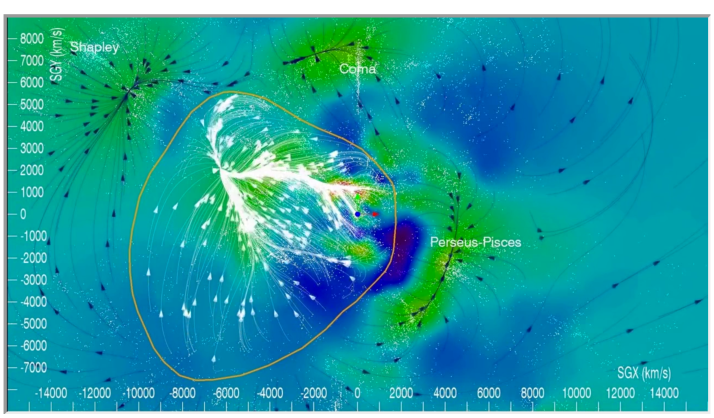


Figure 18.49: Laniakea Supercluster
https://public.nrao.edu/news/pressreleases/superclustergbt

Earlier estimations of galaxy flow had identified a concentration of mass, dubbed the “Great Attractor”, roughly in the direction of the sink of the flow lines of the galaxies in today’s Laniakea. That direction lies in the plane of our galaxy. With modern large telescopes and long-wavelength detectors it is possible to see through the disk of our galaxy to much lower galactic latitudes than it was a few decades ago. The somewhat archaic term “Zone of Avoidance” refers to the band around the sky, roughly 10 degrees on either side of the plane of the Milky Way, that other galaxies appear to “avoid”. Of course they are there, but before we knew about interstellar dust it must have seemed odd not to be able to see other galaxies along this band of the sky. Today the distribution of our relatively “local” voids and superclusters is an area of active research.

Let’s get a bit farther away. Astronomers at the Harvard-Smithsonian Center for Astrophysics made one of the first galaxy redshift surveys in 1985 and were able to demonstrate that the distribution of galaxies is not random but that there are voids and filaments.

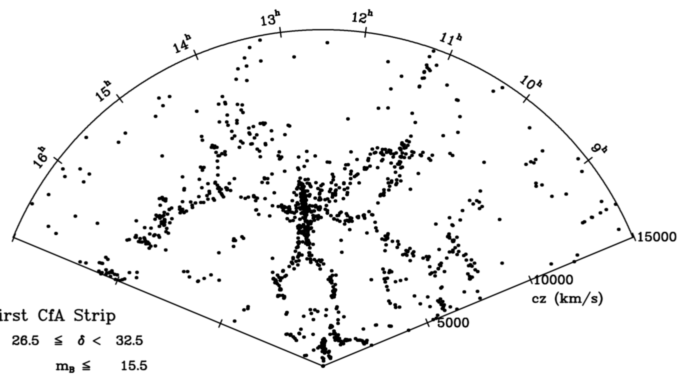


Figure 18.50: Lapparent, Geller, and Huchra, Harvard-Smithsonian Center for Astrophysics Redshift Survey 1985.

We are at the apex of the wedge. A value of cz = 15,000 is z = 0.05, so if all the redshift were due to expansion, the most distant galaxies would be about 200 Mpc away.

https://www.cfa.harvard.edu/~dfabricant/huchra/zcat/

The center of this slice is in the direction of the constellation Coma Bernices and the North Galactic Pole (near 12^h50^m RA, +27° Decl.). These galaxies are not that far away from us, meaning that a non-negligible fraction of their velocities is due to the peculiar motion of the galaxies themselves, particularly in the Coma Cluster. The radial elongation of the clump of galaxies near the center is an elongation in velocity, not a real physical elongation of the cluster in space. On the other hand, the voids are real. Voids, generally, are bubbles ranging from ~10 Mpc to over 100 Mpc diameter with very few galaxies, less than about 10% the average matter density, surrounded by long filaments of clusters and superclusters of galaxies. Some of these filaments stretch for over a billion light years and are referred to as “great walls”.

The motions of the galaxies in the Virgo cluster demonstrate the excess of intracluster velocity over the Hubble flow (expansion) velocity. In addition, some galaxies are still falling toward the cluster. In terms of radial velocities, this results in the fact that some of the galaxies on the near side of the cluster have higher radial velocities than some of the galaxies on the far side of the cluster. The following sketch of the distribution of radial velocities of Virgo cluster galaxies is based on recent studies by astronomer Igor Karachentsev and colleagues. The velocities and distances are relative to the center of the Local Group. The diagonal line represents the Hubble relation; if galaxies’ velocities were solely due to the expansion of the universe, their velocities would fall on or near this line. Most galaxies in the core of the Virgo cluster lie in the region outlined by the large oval. The large range of radial velocities here represents the galaxies’ motions within the cluster and is in accord with what we would expect from the virial theorem. Fewer, but a still-significant number of galaxies occupy the regions outlined by the smaller ovals, indicative of galaxies that are still falling in toward the center of the Virgo cluster, galaxies whose motion is not yet fully governed by the virial theorem.

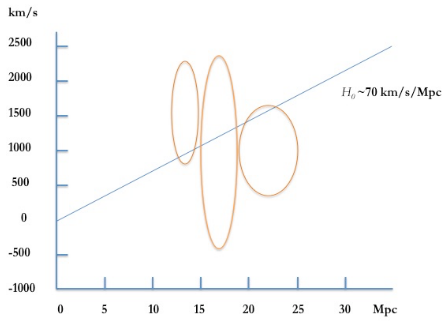


Figure 18.51:
Hubble relation and distribution of galaxies in the Virgo cluster

As mentioned above, more distant clusters will have larger expansion velocities and their intracluster velocities will constitute a much smaller fraction of their observed radial velocities. Infalling galaxies, though, is something to keep in mind. Cluster masses are partly estimated from the internal kinetic energy of the cluster. If a cluster is not finished forming, is not "virialized", but we erroneously assume that it is, then our mass estimate may be in error.

Currently, the most extensive map of the universe comes from work done by the astronomers and data analysts with the Sloan Digital Sky Survey (SDSS). The SDSS observations have been made using a 2.5-m wide-field telescope at Apache Point Observatory in the mountains in southern New Mexico. New observations will also be made using a similar-sized telescope at Las Campanas Observatory in Chile. Through the summer of 2014, the SDSS has mapped about 1/3 of the sky. For our purposes we are most interested here in their observations of distant galaxies, although they have made observations closer to home as well. The wedges in the figure below are centered near the north and south galactic poles, and the dark wedges to the sides represent the "zone of avoidance", i.e., the plane of the Milky Way, where we can see very few galaxies. The colors represent the colors of the galaxies, with red generally meaning older stars.

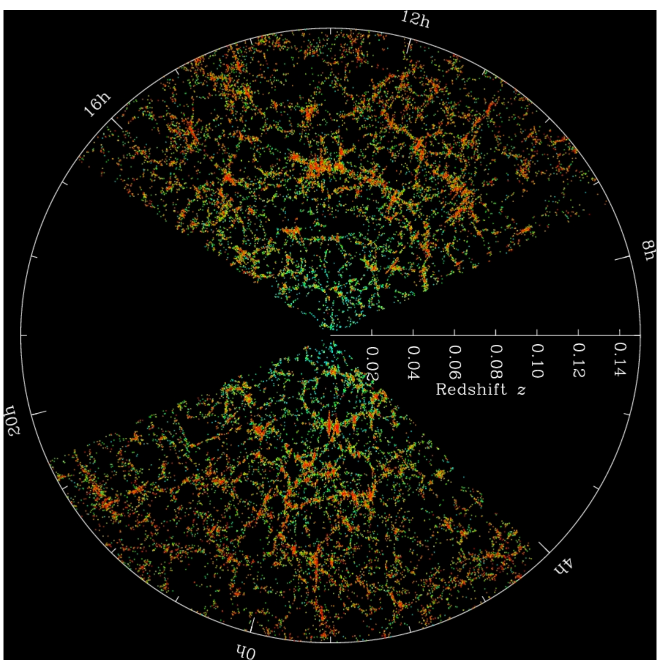
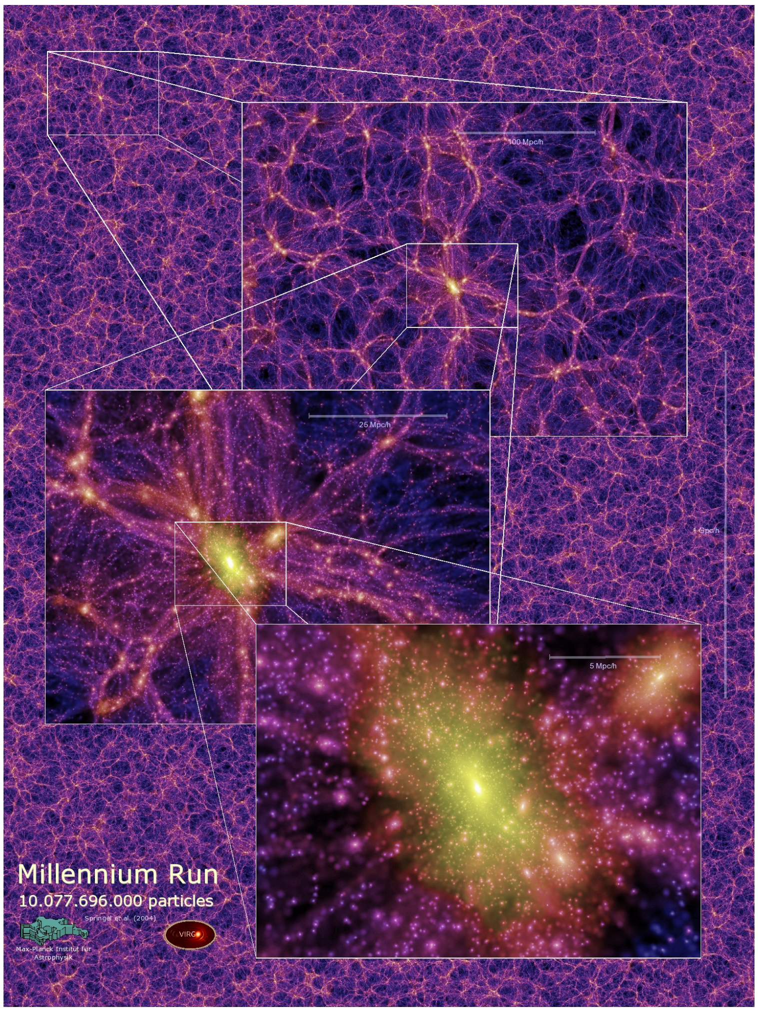


Figure 18.52: Slices through the SDSS 3-dimensional map of the distribution of galaxies. Earth is at the center. The outer circle is about 2 billion light years away. These wedges contain galaxies within ±1.25 degrees declination.

Credit: M. Blanton and the Sloan Digital Sky Survey,
http://www.sdss3.org/science/gallery_sdss_pie2.php

We would like to know how the large-scale structure of the visible matter corresponds to the large-scale distribution of dark matter. We would also like to know how the large-scale structure has changed as the universe has evolved over the past 13.8 billion years. At this point it helps to turn to the work of scientists such as those in the VIRGO Consortium who perform intense supercomputer simulations of the formation of structure in the universe, and compare their results to observational data such as the SDSS map, above. The VIRGO Consortium's

Millennium simulation followed the trajectories of over 10 billion particles (~20 million galaxies) in a virtual box more than 600 Mpc on a side, watching structures evolve from z ~18 to today (z = 0). The left-hand panel, below, is a poster showing the dark matter distribution for z = 0. The scale bars in the figures are in units of $h = H_0 / 100$, meaning that in the last box the scale bar of 5 Mpc/h corresponds to about 7 Mpc. The dense yellow focus of the poster is the locus of a rich cluster of galaxies. In the right hand panel are medium-sized boxes showing the evolution of structure in the dark matter distribution at four different values of z. The earliest corresponds to a universe about 200 million years old, the time of the first dark matter halo protogalaxies. In depth, i.e., into the page, these slices are about 20 Mpc thick. This simulation does an excellent job reproducing the filament / void structure we observe around us today.



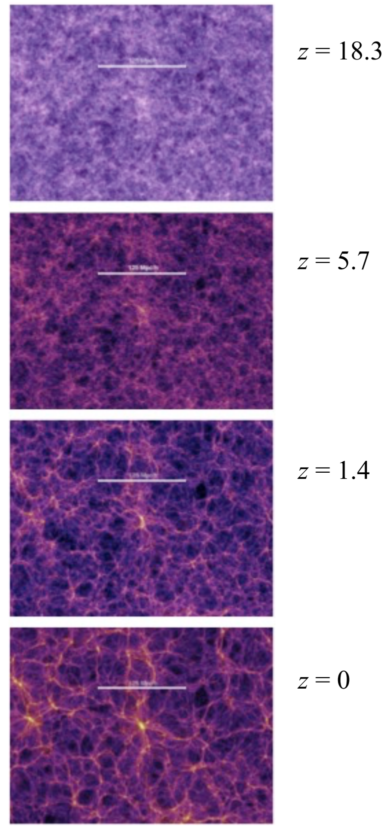


Figure 18.53: structure formation simulation; scale bar ~180 Mpc

http://www.mpa-garching.mpg.de/galform/virgo/millennium/index.shtml

Sample problems

1. Could we readily detect the angular motion of a star in another galaxy? Consider an extremely luminous star (so we can definitely see the star) in a face-on spiral galaxy 20 Mly away. Suppose that the star is 10 kpc from the center of its galaxy, and moving under the gravitational influence of 200 billion solar masses. Observe this star's position (e.g., relative to some wildly distant quasar) over a period of ten years. Assume that we have access to large telescopes, good cameras, and can with moderate confidence resolve angles as small as 0.001 arc seconds. Can we tell that our star has moved? Hint: this is a two-part problem; you need both how far the star has moved and then whether we can resolve that distance.

2. Consider a galaxy receding from us at 35,000 km/sec.

a) At what wavelength will we observe the K line of ionized Ca, rest wavelength of 393.37 nm, in the spectrum of this galaxy?

b) Roughly how distant is the galaxy?

3. Suppose that we are observing the disk of a face-on spiral galaxy where we measure a brightness of 19 mag / arcsec^2. Assume the absolute magnitude of the Sun is ~ 5.

a) Convert this to $L_\odot$ / pc^2.

b) Explain why it doesn't matter what distance you use for the calculation in part a).

c) Convince yourself that you can also work the problem backwards; i.e., take your result from a) and show that you can work backward to get 19 mag / arcsec^2.

d) If you can see stars to a depth of about 1 kpc into the disk, and if the Sun is representative of an average star, *and* assuming you don't have to worry about extinction (a bit of a stretch for both of the latter two), what is the density of stars per cubic kiloparsec in this galaxy's disk? and, despite the approximations, is your answer reasonable?

4. If H_0 ~ 70 km/s / Mpc and *if* it were ok to assume that $H(t) = H_0$, then you could invert H to find the age of the universe. Do it.

5. M31 is ~2.5 Mly away. Estimate the apparent magnitude m_V of a star in M31 with $M_V = -6$. We see M31 at a galactic latitude $b \approx 25°$. Assume an average extinction and that the disk of our galaxy is ~ 1 kpc thick.

6. Describe the role the Cepheid Period-Luminosity relation played in the discovery that M31 is a separate galaxy.

7. Describe various types of AGNs.

8. Reading carefully? Briefly explain / define

a) Hubble types
b) Local Group
c) Laniakeia
d) SMBH
e) Bullet Cluster
f) distance-velocity relation
g) Population III star
h) WHIM

Answers to selected problems are on the next page:

1. No; the star moves ~619 AU in 10 years, which, at a distance of 20 Mly, is an angle on the sky of $\sim 10^{-4}$ ″.

2. a) ~439 nm
b) ~ 500 Mpc

3. a) 1070 $L_\odot$ / pc^2
d) ~ 1 $\cdot 10^9$ stars / kpc^3 or 1 star per cubic parsec, which isn't too bad.

4. About 13.9 billion years.

5. If you ignore extinction within M31 and use ~ 1 mag / 1000 pc within the disk of the Milky Way, $m_V \sim 19.6$.