THE MILKY WAY

The Milky Way is a galaxy, which can be defined as a system of $10^7$ to $10^{13}$ stars, interstellar gas and dust, and "dark matter" bound together by the combined gravity of these. In the Milky Way's case, the number of stars is roughly 400 billion.

The main structural components of the Milky Way are the disk, bulge and halo. The disk is where we live, and contains most of the stars of the Milky Way, and the vast majority of the interstellar gas and dust. It is about 30,000 pc (or 30 kpc) in diameter (this is about 100,000 light years, so light takes 100,000 years or so to cross the Milky Way - galaxies are big). The stars in the disk are both young and old - some are forming right now out of interstellar gas clouds. Almost all star formation in the Milky Way is in the disk. The disk also has a spiral structure. This is hard to perceive from within the disk but is discerned because we can measure distances to objects in the disk and see that they are concentrated in arms. (The Sun is not on one of the main arms but on a small spur, called the Orion spur, coming off one of the arms.) A side view of the disk would show dust obscuring the central parts of the stellar layer, because most dust (and gas) is in a thinner layer than the stars.

The halo is a spherical shape component that surrounds the disk. It is at least 30,000 pc across. It contains the globular clusters (about 150 in all), each of which contains up to 1 million stars. There are also stars between the globular clusters in the halo. Halo stars are very old and probably the first ever to form in the Milky Way.

The bulge is a spherical component at the center of the disk, about 1 kpc across. It also contains mostly old stars. There is a "supermassive" black hole at the center of the galaxy of about 3 or 4 million times the Sun's mass, which is a site of unusual activity.

Shapley in 1917 was the first to make a good estimate of the size of the Milky Way. He noticed that there were more globular clusters in one direction in the sky than in the opposite direction. By estimating distances to some Horizontal Branch stars in them, he knew they were far away (many kpc), and he guessed that they formed a system centered on the center of the Milky Way. From all this, he was able to infer a distance of the Sun from the center of about 16 kpc. The modern value is 8 kpc.

The orbits of stars in the halo and bulge are generally very elongated and random in orientation, like bees in a hive.

The disk rotates, but not like a rigid object. The Sun's speed around the center is about 225 km/sec, which means it takes 240 million years to go around once (its orbital period). However, the orbital period of stars and gas clouds increases with distance from the center of the Milky Way. This is called "differential rotation." A "rotation curve" is a plot of rotation speed vs. distance from the center of the galaxy. For the Milky Way, for most of the disk, the rotation speed is constant with that distance, at a value of about 225 km/sec.
Spiral arms are best traced in young stars and clusters, interstellar gas clouds, and older stars to a lesser extent. The space between the arms is not empty, there is just less stuff there. Because of differential rotation, if a spiral arm always consisted of the same stars and gas, it would get wound up very tightly in a short fraction of a galaxy's lifetime, and the spiral pattern would become unrecognizable. So we need another theory to explain spiral structure that can last long. The idea is that the spiral arms mark the location of a "spiral density wave", a spiral shaped wave which circulates around the galaxy, keeping its shape, and bringing stars and gas clouds closer together as it passes them. The resulting region of high star and gas density is a spiral arm. It is analogous to a big traffic blockage moving down the road: stars and gas enter the compressed zone and exit the other side, so that the stars and gas that make up the arms are always changing. The arms also are regions where interstellar gas clouds, including the cold, dense, star-forming molecular clouds are squeezed together. This seems to start many of these clouds collapsing under their own gravity – the beginning of the star formation process. So the spiral arms are also sites of much star formation activity. The bright, hot, blue, massive, short-lived stars are born and die in the spiral arms because their lifetimes are too short for them to exit the arms. So the arms stand out in blue stars and in emission nebulae, where the gas is heated and ionized by such stars.

There is evidence that about 90% of the matter in the Milky Way is "dark" i.e. gives off no detectable radiation. The evidence comes from the rotation curve. The speed at which a star (or a planet in the Solar System) rotates depends on the amount of matter interior to its orbit. The luminous matter in the Milky Way starts to run out at about 15 kpc from the center. But rotation speeds for whatever stars and gas clouds we can see beyond that radius indicate that the total matter does not run out at that radius. The rotation speeds out there should start falling for stars that are beyond the bulk of the matter (in the same way that rotation speeds of planets in the Solar System fall with distance from the Sun, which dominates the mass in the Solar System). But these rotation speeds stay pretty constant out to 40 kpc from the center. This implies a lot of mass out there that is giving off no radiation. We don't know what this mass is made of. Most of it cannot be "ordinary" matter (i.e. stuff made of protons, neutrons, and electrons). This is a constraint that comes from our understanding of the Big Bang. So while ordinary things like brown dwarfs (small, failed stars), dead white dwarfs, neutron stars and black holes may make up some of the dark matter, most of it must be in the form of some exotic particle that has yet to be understood. Whatever it is, it does not interact with radiation much at all, neither absorbing nor emitting it. It is believed to be in a spherical halo surrounding the disk, something like the stellar halo but obviously much more massive.

The mass of the Milky Way is about 600 billion solar masses.

GALAXIES

Fuzzy "nebulae" in the sky were known about for a long time, but nobody knew they were galaxies like the Milky Way. In the 1920s, Hubble observed individual stars in them and eventually found that their distances placed them well outside our Galaxy.
Hubble also classified galaxies, creating a famous "tuning fork" diagram. There are three main types of galaxy morphology -- elliptical, spiral and irregular. Ellipticals can further be divided, using a number from 0 to 7, depending on how flat they appear (at least on the sky which is all we can observe). So EO's appear round, while E7's are the flattest class. The star orbits in E's are like halo and bulge orbits in the Milky Way: randomly oriented and highly elongated, like bees in a hive.

Spirals are further divided into barred (SB) and unbarred (S). The bar is a linear feature of stars that runs through the center of the galaxy, generally connecting the beginnings of the spiral arms. Bars are wave patterns, like spirals. Within each of these spiral classes are subclasses a, b, and c. The two things that change in going from Sa to Sc is the size of the bulge and how tightly wrapped the spiral arms are, with Sa's having the biggest bulges and the mostly tightly wrapped spiral arms. These are general trends only. The Milky Way is an SBbc (a barred spiral in between b and c). A class seemingly intermediate between E and S is S0, which has an enormous bulge, and a disk, but with no spiral structure. The irregulars have all kinds of shapes.

A further distinction is between "giant" ($10^8$-$10^{13}$ stars) and "dwarf" ($10^6$-$10^8$ stars) galaxies. Dwarfs are the most common kind of galaxies. There are dwarf E's and dwarf irregulars, but not many, if any, dwarf spirals. Among giant galaxies, spirals make up the large majority, then E's, while irregulars are more rare.

E's made all their stars from their interstellar gas early on in their history, and there is generally no star formation now. Spirals still form stars. Irregulars have a variety of histories. These are general rules only.

Distances to nearby galaxies can be determined in a variety of ways, but one important one is by identifying a class of star within them called a Cepheid variable. Variable stars are evolved stars whose brightness varies periodically (an example is an Asymptotic Giant Branch stars undergoing pulsations before ejecting a Planetary Nebula, as we studied earlier). Cepheids are bright (1000's of solar luminosities), massive pulsating variable stars that we can see and isolate even in other galaxies, at least out to about 25 million parsecs (Mpc) - with the Hubble - which is still pretty local as far as galaxy distances go. From Cepheids in the Milky Way whose distance we know from other means, it was found that their variation periods depend on their luminosities (averaged over the period). So for Cepheids in other galaxies, we measure their variation period and that tells us their luminosity. With their apparent brightness, we get the distance to them, and thus to the galaxy.

With ways of getting distances, we can now map out the universe in galaxies. We find that they are not randomly distributed, but form structures. The Milky Way is in a structure called the Local Group, which also contains the spirals Andromeda and Triangulum, and a few dozen dwarf galaxies, many discovered recently by the Sloan Survey in New Mexico. A group is bound together by the combined gravity of its galaxies and dark matter. About 15 Mpc away is the Virgo cluster. Clusters contain 100's or 1000's of galaxies. Galaxies in groups and clusters orbit within them, just like stars in a star cluster. The Local Group, Virgo, and several other groups and clusters form a unit called the Local Supercluster, about 40 Mpc across.
Cepheids can only be isolated in galaxies closer than 25 Mpc or so, so we need another distance method to map out larger structures. One is the Tully-Fisher relation, which is a relation between how fast a spiral galaxy rotates and its luminosity. Again, the relation is calibrated using nearby galaxies whose distances are known from Cepheid variables. So you measure a galaxy's rotation speed from the Doppler Shifts of stars or gas clouds in its disk - that tells you its luminosity. Then, you measure its apparent brightness. The distance you get from the luminosity and the apparent brightness using the inverse square law for radiation. The Tully-Fisher relation can be used for more distant galaxies than the Cepheid technique because the rotation speed and apparent brightness are easy to measure for more distant galaxies.

But there is another method that extends our reach even further. In 1912, Slipher observed that most galaxies were receding from us (he noticed an apparent Doppler shift in their spectra). Hubble determined distances to many galaxies using Cepheids, and found that the receding velocity was proportional to the distance. The constant of proportionality is called Hubble's constant, which is somewhere in the range 65-75 km/sec/Mpc, and this law is known as Hubble's Law. It means that a galaxy 1 Mpc away, following Hubble's Law, should be receding at 65-75 km/sec, while a galaxy 10 Mpc away should be receding at 650-750 km/sec. This is actually evidence for the Big Bang, but the immediate application is to galaxy distances: measure the receding speed (the Doppler shift to the red end of the spectrum or the "redshift") and you have the distance. Using this, structures larger than superclusters were found in the 70's and 80's. The superclusters seem to be strung out on one-dimensional "filaments" or two-dimensional "walls" surrounding relatively empty "voids", like a sort of froth or soap-bubble structure. These voids are typically 100 Mpc in size! This the largest scale on which we have seen structure in the universe.

Galaxies sometimes interact, collide and merge, creating distorted shapes due to tidal forces. If two spirals merge, they may create an elliptical. It is not clear whether all E's were made this way or not. Prodigious star formation can happen in mergers. The Milky Way itself is in the process of accreting a small dwarf galaxy right now. About 3 to 4 billion years from now, the Milky Way and the Andromeda spiral will merge.

It is thought that the first stage in galaxy formation is the formation of sub-galactic fragments, wherein the first stars are born. Many of these then merge to form a large galaxy. Dwarf galaxies, such as the several surrounding the Milky Way, are presumably the leftovers of this process. Observations now seem to support this picture.

ACTIVE GALAXIES

Active galaxies are much more luminous than normal galaxies. Their spectra look very different from starlight. Their brightness comes from the very central region, called the Active Galactic Nucleus, where material is falling into a supermassive black hole, which may have a mass of up to a billion solar masses or so. Not all the matter goes into the black hole, and we observe very long jets of material emanating from the galaxy's central region, often ending in two lobes that are bright in radio waves.
How the supermassive black holes form is still a mystery, but they could form from the merger of many smaller black holes, made in an early burst of star formation, in the dense environment at the centers of galaxies.

The most luminous active galaxies are known as Quasars or Quasi Stellar Objects. Quasars and active galaxies in general are usually seen at high redshift, indicating that they are very far away (according to Hubble's Law), and therefore that the light has taken generally billions of years to reach us. We do not see so many in the nearby universe. So this means they were much more common billions of years ago, when the universe was young. They are probably mostly associated with the early stages of galaxy evolution when they were young, still growing, and had not settled into regular shapes. This doesn't mean the black holes have disappeared. They are probably still at the centers of galaxies, but in general much less matter is flowing in toward them as galaxies became more orderly, so they have nothing to feed on.

COSMOLOGY

It seems that the universe is pretty much the same everywhere if you look at large enough scales (shells and voids), so let's begin by adopting the Cosmological Principle (CP): "The universe is roughly homogeneous (same everywhere) and isotropic (same in all directions)." This sounds reasonable enough but leads to strange conclusions.

Hubble's Law seems to indicate that we are special because all galaxies recede from us. But it is easy to imagine a situation in which all galaxies recede from each other. So we are not necessarily at the center of the universe. If the CP is correct, there can be no center, because a center would be a special place, and the universe would not be homogeneous. So Hubble's Law suggests that in the past, everything was separated by much smaller distances, but the CP says there can't be a center of expansion in our 3-d universe. How long ago everything was together depends on the rate of expansion, which is Hubble's constant. In fact, the inverse of Hubble's constant gives an estimate for the age of the universe: something like 13-16 billion years. The Big Bang is the event that started the universe: then, the universe was confined to a single point, and suddenly expanded in all directions. We don't know why it happened. However, the CP demands that it is not that galaxies are expanding through an empty space (again, this is because then you could trace back the expansion to a point, which is a special location and thus violates homogeneity), but that space itself is expanding in all directions, with no center and no edge. The Big Bang was an explosion of space itself, which involved the whole universe.

Note that the expansion of the universe doesn't apply to things held together by gravity or other forces. The Earth, Solar System, Milky Way, and even the Local Group are not observed to be expanding. The expansion is only noticeable on scales larger than individual groups or clusters of galaxies.

An analogy for space expanding is an expanding raisin cake, where the galaxies are the raisins and the dough is space. The raisins do not move through the dough, but the dough expands carry the raisins along. But there you can see a center of expansion, so the analogy is limited. Instead, imagine you are a two-dimensional creature with no perception of the third dimension, but you live on a large, expanding balloon whose curvature is too big to perceive. There is also no
``edge'' to your two-dimensional universe. If you walk around, you come back to where you started. You see every point on the surface expanding from every other point, but with no apparent center because it's in the third dimension you don't know about. Now take this analogy ``up a dimension'' and imagine our 3-d universe having a center of expansion in a fictitious fourth dimension you can't point to. Each point in 3-d space expands from every other point. From the balloon analogy, you might still be tempted to think that the universe is expanding ``into'' something, like empty space. But it isn't, at least in our 3-dimensional universe.

This interpretation of our universe leads to a different interpretation of the redshift. Photons get stretched too as they travel through the expanding universe. So a photon from a distant galaxy, taking perhaps billions of years to get to us, has a much longer wavelength than when it started. So it is ``redshifted'' but not because of the Doppler Shift, but because space expanded during its journey.

A strong piece of evidence in favor of this picture is the Cosmic Microwave Background (CMB). This is radiation that was predicted to be left over from the Big Bang fireball. It was predicted to have a perfect blackbody spectrum and the same intensity in every direction. Since blackbody temperature is related to the wavelength of the brightest radiation, the stretching of wavelengths as the universe expands means that the CMB should always be getting cooler. The CMB was discovered by Penzias and Wilson in 1964. The spectrum was indeed a perfect blackbody, and the best measured temperature is 2.725 K. The radiation has the same intensity in every direction to better than one part in \(10^4\). Therefore, it must fill the universe homogeneously and isotropically, indicating that the Big Bang happened everywhere in the universe. If the Big Bang occurred in one place -- not here -- then the radiation from it could have streamed past us by now, so we wouldn't see it, or it may not have reached us yet, or, if it was just reaching us now, it would be coming from a special direction. If the Big Bang happened here, then the radiation would have left long ago, and we wouldn't see it. The isotropy of the CMB suggests it must have happened everywhere.

The mass and energy in the universe causes the whole thing to have curvature, just as we talked about curved space around black holes. In general, the universe could be curved in three different ways. It could curve back on itself like a closed spherical surface, except in three dimensions. In this case you could travel in any direction and eventually get back to where you started. This is called a closed geometry (it is also said to have ``positive'' curvature. Alternatively, the geometry may be more like a saddle shape, with different directions curving in the opposite sense (i.e. one curves ``up'' while another curves ``down''). This is called an ``open'' geometry or ``negative'' curvature. Finally, between these cases is a ``flat'' universe, which has zero curvature. The best evidence we have favors a flat universe, or at least one with a very tiny degree of curvature.

Whether the expansion will continue forever or eventually halt depends on the contents of the universe. The gravity due to all the mass in the universe should cause the expansion to slow down. The big news in recent years is that the expansion is accelerating. This has been determined using very bright carbon-detonation supernovae in distant galaxies: these are known from nearby galaxies to have a very predictable luminosity, hence the distance to their host galaxies can be determined from their apparent brightness. Then with a redshift from the
supernova's spectral lines, the redshift-distance relation can be studied at large distances, which means in the past. The result is that the expansion was slower in the past than now.

Einstein actually thought that there could be something that causes the expansion to accelerate (because he preferred a static universe where this acceleration balanced the deceleration due to gravity), but later declared that he had made a big mistake after Hubble discovered that the universe was expanding. But it looks like he might have been right. The acceleration is described by a number called the Cosmological Constant. If this constant were zero, the fate of the universe would just depend on how much mass there is in it, and if there is enough, then gravity can eventually stop the expansion. Our best estimate, even including dark matter, is that there is not enough mass to stop the expansion (only 30% of what is necessary). But the acceleration implies a non-zero Cosmological Constant. The expansion of the universe seems to be accelerating, and gravity can't stop it. Nobody yet understands what is providing the repulsive force that must be responsible for the acceleration. We call it "dark energy" but we have no idea of its nature. However, the acceleration implies that it is more important than dark matter in determining the expansion of the universe.

What is the fate of an accelerating universe? Eventually, you will not be able to see anything except the most local galaxies. If the acceleration grows with time, then all matter will eventually be ripped apart.

A problem with the Big Bang is the "horizon" problem. When we observe the CMBR, we are seeing the universe at an early time when the light was first able to stream freely through the universe (just as we see the layer of the Sun from which light can stream freely towards us). This is only 300,000 years or so after the Big Bang. The CMB photons have been streaming freely through the universe ever since. Considering the CMB observed in two opposite directions on the sky, there has not been enough time for the photons to travel between these locations. So they should not have had any communication with each other (light is the fastest way to communicate). Yet such regions appear identical. Inflation is an idea which suggests that in the earliest times (tiny fractions of a second after the Big Bang), the universe went through a stage of rapid expansion. Before then, all parts could communicate with each other, exchange energy, etc. causing temperatures to even out (if you put a hot object in contact with a cooler one, eventually they will come to the same temperature by exchanging heat). After then, regions expanded so far from each other that they could no longer communicate (the current situation). So the CMB looks the same everywhere because during this very early time, all parts of the universe were able to end up with the same temperature. Inflation also predicts the universe has an essentially flat geometry, and the latest observations of the CMBR are consistent with this.

Very early on, the universe is thought to have been dominated by very energetic radiation. Matter first formed by the process of pair production, which is seen in physics experiments. Photons with sufficient energy can produce so-called particle-antiparticle pairs, such as electrons and positrons, converting pure energy into mass, according to Einstein’s $E=mc^2$. Particles and anti-particles when they run into each other will annihilate, producing pure radiation again. So such pairs were coming in and out of existence all the time in the early universe. But for some reason not yet fully understood, there was a tiny preference for matter over antimatter, and so the
annihilations left an excess of matter, which became the matter that we know in the universe. Electrons, protons and other particles formed.

The early universe was expanding and cooling all the time. During 100-1000 seconds after the Big Bang, the density and temperature were quite similar to the cores of stars, and that means conditions are suitable for fusion. Protons fused together to form helium nuclei. However, to fuse helium into heavier elements requires higher temperatures, and by this time it was too late (only trace amounts of lithium were made). Eventually the universe became too cool for fusion of hydrogen to helium, at which point 25% of the mass of the universe had been turned into helium. This is what the Big Bang theory predicts, and it has been confirmed by measuring the helium in the atmospheres of the oldest stars we can find (whose chemical make-up is unaffected by any fusion occurring in stars, unlike the Sun's atmosphere, for example).

We cannot say what happened before the Big Bang, or even if it makes sense to talk of ``before''. The very ideas of space and time are not clear when we consider the Big Bang, where so much mass and energy is confined to such a tiny volume. We don't have the physics to understand this yet, but it would involve the unification of General Relativity (which describes strong gravity) and quantum mechanics (which describes tiny scales).

The Big Bang theory has been successful in explaining the expansion of the universe. But like any theory worth considering, it not only explained what had been observed, but had testable predictions for what had yet to be observed. It predicted the existence and properties of the CMBR, and the correct helium abundance (and lack of any heavier elements) in the oldest of stars.