AST 204 February 4 2008

LECTURE I: THE EXPANDING UNIVERSE, THE CRITICAL DENSITY, AND THE AGE OF THE UNIVERSE

I. The Observations, Simplified

We live in an expanding universe, a fact which we have known for about seventy years. What does this mean? We live in a galaxy which is much like billions of others, whose physics we understand fairly well. The earth is in orbit about the sun, an ordinary star much like billions of others in this galaxy. Though we have been observing for far too short a time to have followed the sun very far in *its* orbit about the center of the galaxy, we know that its velocity about the center is roughly what one would expect from what we understand about the mass of the galaxy and simple Newtonian dynamics. We know with some confidence that the galaxy is *not* expanding. Neither is the solar system, or the sun, or the earth. But the universe *is*.

And what that means is that the other galaxies we observe are receding from our own, in a remarkably uniform pattern. We infer that they are receding from their *redshift*, which is normally (and indeed, very persuasively, as we shall see later) interpreted as an ordinary Doppler shift. The observed phenomenon is that features in the spectra of galaxies which belong to well-known atomic transitions in common chemical elements and molecules are observed to be at longer wavelengths than they are observed in those species in a laboratory. All features in the spectrum of a given galaxy are longer in wavelength by a constant factor, as one would expect from a Doppler shift,

$$\lambda_{obs}/\lambda_{lab} \equiv 1 + z \simeq 1 + v/c. \tag{1}$$

The Doppler formula, $\lambda_{obs}/\lambda_{lab} \simeq 1 + v/c$ is easily derived from a consideration of the spacing of spherical wavefronts emanating from a moving source.

The remarkable pattern in the recession velocities of galaxies is embodied in the *Hubble Law*, which is named after its discoverer, Edwin Hubble (though there is some controversy, as there often is in science, over who *actually* has priority), which relates the recession velocity very simply to the distance of a galaxy:

$$v = Hd \tag{2}$$

Here d is the distance and H a number called the *Hubble constant.* The best determinations of H at the present time (the uncertainties are all in the distances, which are poorly determined at best) yield a number of about 22 km/sec velocity per million light years distance, or in units which are in more common use, about 72 km/sec per megaparsec. (a parsec is 3.08×10^{18} cm, about 3.26 light years, and is the unit of distance most commonly used in astronomy.) It is of some interest that quoting these numbers to this accuracy was not possible until 2003, with the announcements of the results from the

WMAP satellite, which we will discuss later. Since observing the spectra of galaxies and interpreting the redshift as a velocity gives us information about only the component of velocity along the line of sight to the object in question, we know nothing about the components perpendicular to the line of sight, but there is other information from the very distant universe which suggests that those velocity components are small, and we will for now assume that they are indeed negligible.

We can now observe objects which are very distant, and observe redshifts which imply velocities which are more than 95 percent of the velocity of light (this requires a relativistically correct expression for the Doppler shift, which we shall see anon). We find no hint of differences between the Hubble laws observed in different directions in the sky, though notions of "distance" become a little fuzzy when discussing objects like these, which, as we shall see, are almost at the edge of that part of the universe which we can even in principle observe.

II. Some Consequences in the Light of Physics

These data taken at face value would seem to imply something quite remarkable and very anti-Copernican: That we are at rest in the very center of expansion, and the rest of the universe is flinging itself wildly away from us, the faster the farther away.

But before we embrace this egocentric idea, we need to think a bit, and remember some of the principles which have emerged in physics in the last century.

The first is the principle of relativity: *There is no absolute rest or absolute motion; one can measure relative motion, but no experiment can distinguish one supposed state of absolute motion from another. There is no preferred rest frame.* This principle is the cornerstone of relativity theory; one of its consequences is the seemingly nonsensical but rigorously experimentally verified statement that the velocity of light is the same to *all* observers in *any* state of motion. Finally, there is a result first derived by Newton for his gravitation theory and later proved in general relativity by the mathematician Birkhoff, and is called in that context Birkhoff's theorem: *In any spherical physical system, the gravitational accelerations and structure of spacetime at some radius are independent of anything outside that radius.*

Armed with these notions, let us think a little more deeply about the observations. Since we see galaxies receding from us, we naturally assume that we are at rest, and they, moving. The principle of relativity tells us that that is not only wrong, but nonsensical. All we know is that they are receding from *us*. We know and can know nothing about our own state of motion, because the concept means nothing.

We can write the Hubble law as a vector law if we keep the assumption that the velocities are strictly along the line of sight:

$$\mathbf{v} = H\mathbf{r} \tag{3}$$

or,

$$\begin{aligned} v_x &= Hx \\ v_y &= Hy \\ v_z &= Hz. \end{aligned} \tag{4}$$

Let us ask a very simple question: What would somebody on *another* galaxy observe? Suppose there is an observer on a galaxy with coordinates X, Y, and Z, velocity HX, HY, HZ. Then the velocity of a galaxy at x, y, z with respect to *this* observer's galaxy is just the difference in their velocities,

$$\begin{aligned} v'_x &= Hx - HX = H(x - X) \\ v'_y &= Hy - HY = H(y - Y) \\ v'_z &= Hz - HZ = H(z - Z). \end{aligned} \tag{5}$$

The *distance* from the observer's galaxy to the one at x, y, z has components, of course, $x' = (x - X), y' = (y - Y), z' = (z - Z)$. Thus we have the simple but profound result:

$$\begin{aligned} v'_x &= Hx' \\ v'_y &= Hy' \\ v'_z &= Hy'. \end{aligned} \tag{6}$$

or, equivalently,

$$\mathbf{v}' = H\mathbf{r}'.$$

Thus the observer on another galaxy sees *exactly the same* expansion pattern as we do, and would conclude, if she did not know about relativity and did not think deeply, that *she* was at the center of the expansion.

The clear conclusion of this is that there *is no* center of the expansion, or perhaps more correctly that *everywhere* is the center. There is no preferred place, and clearly no preferred direction.

III. A Suggestion About Origins

Let us do a little thought experiment. Since the universe is expanding, the distances between galaxies tomorrow will be a little larger than they are today; yesterday they were a little smaller. Notice that the Hubble law, written in the following way,

$$d/v = 1/H \tag{7}$$

makes an interesting suggestion. We do not (at the moment, although we'll look at this shortly) know how the velocity changes with time, and so are at liberty to assume for the sake of argument that it has not changed much; let us in fact see what the consequences of assuming that it is strictly constant–that is, the galaxies have had whatever velocities they have now, always. Clearly H has the units of velocity/distance, that is, 1/time, and $1/H$ has the units of time. If H = 72km/s/Mpc, and 1 Mpc $\simeq 3.1 \times 10^{19}$ km, then $1/H$ $= 4.3 \times 10^{17}$ seconds, or about 14 billion years. (one year $\simeq \pi \times 10^7$ seconds, which you should drill into your brains.)

If the velocities do not change, then the distance at a time t previous to the present, which we can call t_0, was, if d_0 is the present distance,

$$
\begin{aligned}
d(t) &= d_0 - v \cdot (t_0 - t) \\
&= d_0 \cdot [1 - H(t_0 - t)]
\end{aligned}
$$

Fourteen billion years ago the distances between *any* two galaxies was — exactly zero ! The density of matter in the universe currently is very small, probably less than about 10^{-6} hydrogen atoms per cubic centimeter; a far, far better vacuum than the best laboratory vacuum (at 10^{-12} Torr, an ultrahigh vacuum, there are about 10^5 molecules per cubic centimeter.) But at very early times the suggestion is strong that the densities are very high indeed, and we will see later that the temperatures were almost certainly likewise...and so the *suggestion* is that about 14 billion years ago the universe may have had its beginning in a fiery cataclysm which has come to be called the Big Bang. We will see shortly that the velocities are *not* likely to be constant, but vary in a way which even strengthens the conclusion that there was a cataclysmic origin.

IV. Another Remarkable Observation–Isotropy and its Implications

To summarize what we learned above: the expansion of the universe is uniform, described by the simple Hubble law which says that the velocity of recession of a galaxy is directly proportional to its distance, and that the constant of proportionality, the Hubble constant, is independent of direction. We inferred from this and the relativity principle that the Hubble law is universal, that is, that there is no preferred center of the expansion. The fact that it is independent of direction also means that there are no preferred directions, so the expansion is *homogeneous* and *isotropic*.

Now on small scales, one notices that the *distribution* of matter in the universe is very clumpy. The mean density of the Earth is several grams per cm^3, of the Sun about 1 gram per cm^3, and the mean density of the universe perhaps 10^{-30} grams per cm^3. Enormous contrasts indeed. However, if one goes out in the universe to distances of over about a hundred million light years (30 Mpc), and carves out spheres of roughly that size, you find that the contents of those spheres is remarkably similar, both in amounts of mass and statistical structure. Another way to say this is that if you observe the distant universe, objects, say, more than a few hundred million light years away, you cannot easily tell what direction you are looking. To be sure, a particular strange galaxy may be sufficiently strange that there is not another one like it among all the galaxies you can see, but statistics like the total amount of light, total number of galaxies, redshift distribution of galaxies of a given apparent brightness in some angular patch on the sky yield the same answers for any direction to within the accuracy allowed by the statistics of the sample. Thus the universe appears to be isotropic on large scales not only in its velocity structure but in the distribution of galaxies as well.

This result is underscored by yet another observation, which carries the conclusion to seemingly almost ridiculous limits. There is observed a ubiquitous microwave radio signal from the sky which *very* accurately acorresponds to the radiation field from a very cold black body at 2.7K. It is almost universally believed to be the echo of a time when the universe was very hot and dense, and is the signal from the earliest time in the history of the universe which we can observe now. We will discuss it in some detail in future lectures, but suffice for now to say that it is remarkably isotropic–the intensity of this radiation is the same to better than one part in 10^4 over the whole sky (but it is not *exactly* isotropic, which fact has a great deal to do with why we are here). This result suggests strongly that however isotropic the universe is now, it was very much more so early in its history.

Now we learned that the expansion is not only isotropic but homogeneous; that is, there is no preferred origin. The observed isotropy of the galaxies and the microwave background *either* implies that we are very accurately at the center of a spherical distribution of matter and radiation in the universe *or* that there is no center; the distribution appears isotropic to *all* observers. The latter idea is clearly preferable, and is the *Cosmological Principle.* Not only the expansion velocities but *all* physical measurements are universally isotropic (again on sufficiently large scales–if you sit just outside the Galaxy, the side of the sky containing the Galaxy is different from the other side, but if you smear the universe over a hundred million light years or so, the individual features wash out.) If the universe is universally isotropic it *must* be homogeneous as well. If you sit on the side of a lump the direction toward the lump is different from the one away from it, so there can be no lumps.

V. A Simple Newtonian Model for the Expansion

The size of the sphere one must average over to smear out the structure in the universe, about 30 Mpc, corresponds to expansion velocities from the Hubble law of about 2000 km/s, fast by ordinary experience but very much slower than the velocity of light. Newton and Birkhoff both tell us that we can understand a spherical system perfectly by throwing away all the matter outside the radius which interests us, so the fact that the distant universe is moving away from us at a substantial fraction of the speed of light has no influence whatever on things closer to us, and we should be able to understand these things perfectly well in terms of ordinary Newtonian physics.

Consider, therefore, a shell of matter of thickness dr at radius r. It is expanding uniformly away from the origin at a velocity $v = Hr$. If the mean density of the universe is $\bar{\rho}$, the total mass contained within the shell is $m = 4\pi\bar{\rho}r^3/3$. The gravitational acceleration felt by the shell is $-Gm/r^2$, which is the second derivative of the radius with respect to time, or the first derivative of the velocity with respect to time. Thus

$$\dot{v} = -Gm/r^2$$

If we multiply this equation on both sides by v, but call it $\dot{r}$ on the right, we have

$$v\dot{v} = -Gm\dot{r}/r^2$$

or

$$\begin{aligned} \frac{d}{dt}(\frac{v^2}{2}) &= (\frac{d}{dt})(Gm/r), \\ \frac{d}{dt}\left(\frac{v^2}{2} - \frac{Gm}{r}\right) &= 0 \end{aligned} \tag{1}$$

This is just the energy equation for the shell; the first term is the kinetic energy per unit mass, the second, as you may know and in any case will show in the homework, is the gravitational energy per unit mass of the shell; the total energy of the shell is just the mass of the shell,

$$m_{shell} = 4\pi r^2 dr,$$

times the differentiated expression above. The mass is constant, and the derivative of the energy per unit mass is zero, so is constant, and hence both the energy per unit mass and the energy are constant, as we expect from simple dynamics.

Notice, please, that the energy is constant for a given *shell*, but is certainly *not* the same for all shells. In fact, since the velocity (now, say) is proportional to the radius, it is clear that the kinetic energy per unit mass is proportional to r^2. The gravitational energy per unit mass is proportional to the *contained mass*, which is proportional to r^3, divided by the radius, so it, too, is proportional to r^2. So their sum, the *total* energy, is also proportional to r^2 at any given time, but since r is changing, we must be very careful in the way we express this. Let the total energy per unit mass of a shell be εr_0^2, where r_0 is the radius at the present epoch. ε is constant in time *and is the same for all shells.* Thus the equation of motion becomes

$$\frac{v^2}{2} - Gm/r = \varepsilon r_0^2.$$

This is not, actually, very useful, because of the reference to a particular epoch. How do we write this in a more transparent, useful manner?

The radii of the shells are all changing with time, but it is clearly possible to follow a given shell just by *sitting* on it. But we cannot *identify* that shell with a radius unless we also give the time at which that shell *has* that radius. We would like some kind of mathematical label, a new coordinate, which stays with the shell. To this end, let us make the definition of a *comoving coordinate* u:

$$r(t) = R(t)u$$

One can take u as the radius of the shell at some epoch, or any quantity *proportional* to the radius of the shell at some epoch. With this definition u is constant for a given shell; it is called the *comoving or Lagrangian radius for the shell.* Then the total energy is εu^2, the velocity is

$$v(t) = \dot{R}u,$$

and the equation of motion becomes

$$\frac{\dot{R}^2 u^2}{2} - \frac{4\pi G\bar{\rho}R^3u^3}{3Ru} = \varepsilon u^2$$
$$\dot{R}^2 - 8\pi G\bar{\rho}R^2 = 2\varepsilon$$

Since $\bar{\rho}$ is constant in space, this equation is the same for ANY shell, and thus $R(t)$ is a universal function of time, the *expansion factor* of the universe, and we conclude the remarkable result that the expansion of the universe and all of its gross dynamics are described by *ONE* function of the time, $R(t)$. Thus the Hubble constant is

$$\begin{aligned} H &= v/r \\ &= \frac{\dot{R}u}{Ru} \\ &= \frac{\dot{R}}{R} \end{aligned}$$

And is variable in time, in general, but is constant in space.

This result makes sense, of course–if the Hubble constant, which is the logarithmic derivative of the scale factor, and the scale factor itself, varied from place to place, the density could not *remain* constant spatially, and the cosmological principle would be an accident of a particular time–not a very satisfactory situation. We said at the outset that the equations of motion we derived would be correct for small distances, because the physics of small velocities is accurately Newtonian physics. Thus the equation we have derived,

$$\dot{R}^2 - 8\pi G\bar{\rho}R^2/3 = 2\varepsilon \tag{2}$$

is exact for the whole universe, because it is the same for all shells and is exact for small, nearby shells. This is called the *Friedman equation.*

It should be noted, not surprisingly, that this result is consistent with (and indeed demands) the cosmological principle. Shells of radius smaller than some radius of interest expand more slowly, and shells larger expand more rapidly, so shells never cross. We may take as origin anywhere in the universe as we have seen, and the expansion is spherically symmetric about that place, so the density is just inversely proportional to the volume of a sphere centered on the (arbitrary) origin:

$$m = 4\pi r^3\bar{\rho}/3 = 4\pi R^3u^3\rho/3 = const.,$$

so, since u is constant (and, of course, is $4\pi/3$) we can write

$$R^3\bar{\rho} = const,$$

and the density is proportional to $1/R^3$, a function of time alone, and is spatially constant for all time if it is at any time. Thus the density, Hubble constant, and anything else one can measure about the expansion are homogeneous for all time.

VI. Looking Back: How old is the universe?

If we look at the Friedman equation, it is clear that the second term is proportional to $1/R$, since we just showed that the density is proportional to $1/R^3$. So as we look back in time to times when R was smaller, the gravitional term becomes larger and larger. Since the *difference* between the gravitational and kinetic terms is constant, the kinetic term also gets larger and larger, and so the time derivative $\dot{R}$ gets larger and larger. Thus the recession velocity of any particle, which is just proportional to $\dot{R}$, decreases with time, which is exactly what one expects with gravitation; the universe is, in this simple model, *decelerating.* The other side of that coin is that the velocities get higher and higher as one goes into the past, and so the scale factor *MUST* go to zero at some finite time in the past, and indeed at a time *less long ago* than $1/H$, since the motion is accelerated. Thus the conclusion is inescapable that there was indeed a Big Bang, a time when the density was infinite and the scale factor zero.

It is also clear that since the kinetic and potential terms both become very large in the distant past, their difference, the total energy, becomes negligible in comparison to either, so the total energy plays essentially no role in the early evolution of the universe. This is certainly *not* the case in the late evolution, as we shall see now.

VII. Looking Forward: The Cosmic Energy–Will the Universe Expand Forever?

It should be clear from its construction that the value and even the units of R can be anything; all we demand is that $R(t)u = r(t)$, the physical distance, measured in any convenient units. When we discuss Einstein's general relativity, we will learn that one can profitably associate R with a particular scale which causes 2ε to take the values $(1/c^2, 0, -1/c^2)$ depending on the sign of ε and whether it is nonzero or not. The comoving radius u in this case is dimensionless.

Let us next think about precisely this issue. We have done nothing which would demand that ε has a particular sign, and indeed it can be positive, negative, or zero. First consider the case in which it is zero. This says that our shell has zero energy—its kinetic energy is just balanced by its potential, and this is true, because of the conservation of energy, for all time. Thus the shell is always expanding at exactly its escape velocity from the mass interior. In this case,

$$\dot{R}^2 = \frac{8\pi G\bar{\rho}R^2}{3}$$

or, equivalently,

$$H^2 = \frac{8\pi G\bar{\rho}}{3}.$$

The corresponding value of the *density*, ρ_c,

$$\bar{\rho} = \rho_c = \frac{3H^2}{8\pi G}$$

is called the *critical density.* Its interpretation is quite simple. At a given epoch, say the present one, in which we measure the Hubble constant H, the density determines the gravitational field, and hence the gravitational energy. If $\bar{\rho} > \rho_c$, the energy ε is *negative*—that is to say, the shells are *bound* to the interior mass. Look at the terms in equation (2). The gravitational term, remember is proportional to $1/R$, since $\bar{\rho}$ is proportional to $1/R^3$, so it decreases monotonically as the universe expands. Eventually it is the same size as 2ε, and at that point $\dot{R}$ vanishes. Since the gravitational force is still inward, the *acceleration* is still negative, so the universe turns around and collapses, just as a ball thrown upward from a gravitating body with less than its escape velocity will turn around and return.

On the other hand, if $\bar{\rho} < \rho_c$, the kinetic energy term will always exceed the gravitational energy term because their difference is positive, and as the gravitational term goes to zero as R becomes very large, the kinetic term will tend to a positive constant, just 2ε. Thus the velocity of any particle tends to a constant, and its distance will increase linearly with time. The universe will expand forever, and gravitation will become negligible at very late times. A ball or spacecraft thrown upward with *greater* than its escape velocity will escape the gravitating body and will still have finite velocity when it gets very far from it.

In the case in which the density is just the critical density, the energy is exactly zero (and, of course, remains so). The universe still expands forever, but more and more slowly as time goes on there is no excess of kinetic energy to give a particle constant velocity, and all velocities tend to zero as time goes on; gravitation neither wins nor loses, but always maintains the same relative importance to kinetic energy. In this case, we can easily solve the Friedman equation: If we let $\bar{\rho} = \bar{\rho}_0 R_0^3/R^3$, where the subscript 0 refers to some arbitrary epoch, we have

$$\begin{aligned}
\dot{R}^2 &= \frac{8\pi G \bar{\rho}_0 R_0^3}{3R}, \\
\sqrt{R}\dot{R} &= \left(\frac{8\pi G \bar{\rho}_0}{3}\right)^{1/2} R_0^{3/2} \\
2\frac{d}{dt} R^{3/2}/3 &= \left(\frac{8\pi G \bar{\rho}_0}{3}\right)^{1/2} R_0^{3/2}
\end{aligned}$$

so if we set the zero of time at that point at which $R = 0$, the big bang,

$$\begin{aligned}
R^{3/2} &= const \cdot t, \\
R &= const \cdot t^{2/3}.
\end{aligned}$$

In this simple case, the age is simply related to the Hubble constant;

$$H = \frac{\dot{R}}{R} = \frac{2}{3t}$$

so the universe is 2/3 as old as one would guess from extrapolating back with constant velocity. If the energy is negative, the universe has been decelerating more rapidly and if positive, less rapidly than if the energy were zero, so the universe is younger than $2/(3H)$ in the first case and older in the second, but is always younger than $1/H$.

The relative importance of the kinetic and gravitational terms, or, equivalently, the density with respect to the critical density, is usually summarized in the *density parameter* Ω, defined as

$$\Omega = \frac{\rho}{\rho_c} = \frac{8\pi G\bar{\rho}}{3H^2}$$

If we remember that $H = \dot{R}/R$ and with the above definition for Ω, we can write the Friedman equation as

$$H^2(1-\Omega) = 2\varepsilon/R^2.$$

Recalling that $H^2\Omega$ is proportional to $\bar{\rho}$ which is proportional to $1/R^3$, we reach the same conclusion as we did earlier with the original form of the equation, that the energy term on the right becomes negligible in early times when R is very small. In particular, we have

$$\frac{1-\Omega}{\Omega} \propto R,$$

so that $\Omega \to 1$ at early times. You should think about what happens at late times—it is clear that if ever Ω is unity it remains so, and if ever is less than 1 remains so and if ever greater than 1 remains so. It goes to infinity when the universe is bound and just stops, and goes to zero in the far future if it is unbound, but is always approximately 1 at very early times, when the total energy is negligible.

It is tantalizing, though not strictly accurate, to think that a universe with $\Omega = 1$ is a structure with exactly zero energy, so that if one were to ask the (quite probably nonsensical) question "How much energy did it take to create the universe?" a plausible answer would be "Zero".

We will see in a few lectures that this simple description is, in fact, *too* simple; there is another source of gravitation in space which we have not accounted for, namely the *vacuum* itself, which apparently gravitates very, very weakly. The effect is so small that we would probably never have discovered it were it not for the effect on the universe, but its existence should raise alarm, because it says that our ambitious program to explain the universe on the basis of laboratory physics cannot quite succeed.