I. The Past and Future of the Universe

Last time we derived the Friedman equation, which describes the expansion of the universe:

$$\dot{R}^2 - 8\pi G \bar{\rho} R^2 / 3 = 2\varepsilon$$  \hspace{1cm} (2)

and is exact for the whole universe, because it is the same for all shells and is exact for small, nearby shells. 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:

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

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

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

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.

II. 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 gravitational 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.

III. 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. We have done nothing which would demand that $\varepsilon$ 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,

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

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 the mean density is greater than $\rho_c$, the energy $\varepsilon$ 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 $2\varepsilon$, 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 $2\varepsilon$. 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

\[
\dot{R}^2 = \frac{8\pi G \bar{\rho}_0 R_0^3}{3R},
\]

\[
\sqrt{R} \dot{R} = \left( \frac{8\pi G \bar{\rho}_0}{3} \right)^{1/2} R_0^{3/2}
\]

\[
2 \frac{d}{dt} R^{3/2} / 3 = \left( \frac{8\pi G \bar{\rho}_0}{3} \right)^{1/2} R_0^{3/2}
\]

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

\[
R^{3/2} = \text{const} \cdot t,
\]

\[
R = \text{const} \cdot t^{2/3}.
\]

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 it 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 \( \Omega \), 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 \( \Omega \), we can write the Friedman equation as

\[
H^2 (1 - \Omega) = 2\epsilon / R^2.
\]

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,

\[
\frac{1 - \Omega}{\Omega} \propto R,
\]

so that \( \Omega \to 1 \) at early times. You should think about what happens at late times—it is clear that if ever \( \Omega \) is unity it remains so, and if ever is less than 1 remains so and if ever
greater 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 alarms, because it says that our ambitious program to explain the universe on the basis of laboratory physics cannot quite succeed.

IV. The Redshift—a Remarkable Relationship

We have seen that we see a redshift because galaxies are moving away from us. It should be clear from the discussion above and in the last lecture that there is no mysterious physics in the expansion—they are moving locally under quite ordinary Newtonian forces, coasting and decelerating slowly because of their mutual gravitation. The cataclysm which started the expansion is more than a bit mysterious (and though people have begun to think quantitatively about it in the last twenty years or so, it is still accurate to describe it as mysterious; few of our colleagues would take us to task for this.) Can we relate the redshift in any simple, global way to the expansion? The answer is yes, and the result is very beautiful.

Consider a galaxy at some distance $r_s = R(t_s)u_s$. As the light comes to us from this galaxy, it, of course, passes through all radii intermediate. Let $r + dr$ and $r$ be two such, and let $\lambda + d\lambda$ and $\lambda$ be the wavelengths of some spectral feature emitted from the source at $\lambda_s$ and observed by observers on galaxies at $r, r + dr$ at those wavelengths and at times $t, t + dt$. Then

$$\frac{\lambda + d\lambda}{\lambda} = (1 + dv/c) = (1 + H(t)dr/c) = (1 + H(t)dt),$$

or

$$d\lambda/\lambda = H(t)dt = \dot{R}dt/R = dR/R.$$

And so we obtain the simple result that

$$\lambda = \text{const} \cdot R.$$
Thus photons as they travel through the universe have wavelengths which simply scale with the expansion. Recalling that the redshift $z$ is defined as

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}},$$

it is clear that

$$\lambda_{\text{obs}}/\lambda_{\text{em}} = 1 + z = R_0/R_{\text{em}}.$$

Thus the redshift factor $(1 + z) = R_0/R$ is often (usually) used as a parameter to describe the expansion of the universe: The light of the redshift 6.4 quasar left the source at a time when the universe was $1/(1 + 6.4)$ times as large as it is now. How old was it then? We do not know the exact cosmological model, but if $\Omega = 1$ then recall that $R \propto t^{2/3}$ or $t \propto R^{3/2}$, so the universe was $(1/7.4)^{3/2} = 0.050$ of its present age—we are looking back at this object 95% of the way to the Big Bang; if the universe now is 13 billion years old, it was only 650 million years old then.

V. The Light Cone and the Observable World

Let us consider the paths of incoming light rays in the universe. At each place along the incoming ray we must have

$$dr = cdt,$$

$$R(t)du = cdt,$$

or

$$du = cdt/R(t),$$

$$u = c \int_{t_e}^{t_0} \frac{dt}{R(t)}.$$

Now again the exact evaluation of this integral depends on the details of the cosmological model, but as an example again take the $\Omega = 1$ model, and let us normalize the comoving radius to be the physical radius at the present epoch, so $R_0 = 1$ and $R(t) = (t/t_0)^{2/3}$. Then the integral is trivial, and

$$u_{\text{em}} = 3ct_0 \left(1 - \left(\frac{t}{t_0}\right)^{1/3}\right),$$

$$= \left(\frac{2c}{H_0}\right) \left(1 - (1 + z)^{-1/2}\right).$$

The remarkable thing about this integral is that it converges!! As $t \to 0$ or, equivalently, $z \to \infty$, $u \to 2/H_0$. Thus, even though the universe was very small in the past there is a maximum comoving radius to which we can see, and so there is a most distant object we can see. Note that this is not a peculiarity of the $\Omega = 1$ model; we saw that all models converge to the $\Omega = 1$ one near $t = 0$, so even though the exact form of $u_{\text{em}}(z)$ is different,
its qualitative behavior at high redshift, which is just dependent on $R \propto t^{2/3}$ for small $t$, is unchanged.

This maximum $u_{em}$ to which we can see is called the particle horizon. Note that it gets bigger every day; in comoving coordinates it expands at three times the speed of light (!). Its edge defines the observable universe. Cosmologists carry on about whether the universe is finite or infinite, but in fact it would appear that this is a question which is in principle unanswerable; we can see only a finite part of it in any case.

VI. The Cosmic Microwave Background

Thus armed with some understanding of how electromagnetic waves propagate in the universe we are ready to undertake an understanding of one of the most important discoveries in astrophysics of last century, certainly one rivaling in importance the discovery of the expansion itself. Physicists for many years had wondered about conditions in the universe near the epoch of the big bang, and by the middle 1960s several calculations had been done assuming (as seemed intuitively likely, though there was no real evidence) that the universe was very hot as well as very dense at very early times. In 1965 two physicists at Bell Labs at Holmdel NJ, Arno Penzias and Bob Wilson, who were trying to calibrate a new radio telescope very accurately to measure absolute fluxes in the microwave region of the radio spectrum noticed that hard as they tried, there seemed to be a signal coming from the sky which they could not find any source for in their equipment. Finally convincing themselves it was real, they undertook to measure it carefully. They looked at only one wavelength, but it appeared to be a very uniform background which corresponded to radiation from a blackbody at about 3K, 3 degrees above absolute zero. Hundreds of measurements since, culminating in the beautiful measurements by the COBE (COsmic Background Explorer) satellite have confirmed the existence, thermal nature, and incredible isotropy of this radiation. We now know that it is very accurately a radiation field corresponding to a black body at 2.7K, and that it is isotropic to about one part in $10^5$ over the whole sky. Measuring any deviation from isotropy was an enormous challenge for many years, though COBE first made a reliable but crude map, since improved in detail over small regions of the sky by balloon experiments like BOOMERanG and recently done exquisitely by the satellite WMAP.

What does this mean? First of all, even though the radiation is incredibly feeble by terrestrial standards, it completely dominates the thermal energy density of the universe. The spectrum of a black body is given by the Planck function

$$B_\nu(T) = \frac{2h\nu^3}{c^2}(e^{h\nu/kT} - 1)^{-1}$$

where $\nu$ is the frequency, $h$, Planck’s constant, $k$ the Boltzmann constant, and $T$ the absolute temperature in Kelvins. $B_\nu$ is the specific radiant intensity, the flux of energy incident on a unit area per unit frequency in unit solid angle. The energy density associated with this radiation field is obtained by multiplying by a factor of $4\pi/c$— the $4\pi$ factor
simply to account for all solid angles, the $1/c$ simply because the flux of energy is the velocity times the energy density, and the velocity here is clearly the velocity of light. Integrating the Planck energy density over frequency yields the simple relation

$$U_r(T) = aT^4$$

where a difficult integral (which is $\pi^4/15$) and all the physical constants are combined into the radiation constant $a = 8\pi^5k^4/15c^3h^3$, which has the value $7.56 \times 10^{-15}\text{erg cm}^{-3}\text{K}^{-4}$. At 2.7K, the energy density is

$$U_r = 4.0 \times 10^{-13}\text{erg/cm}^3,$$

about 0.25 electron volts per cubic centimeter. For comparison, starlight in the Galaxy has a comparable energy density, but by the time you are a Megaparsec (a typical galaxy-galaxy spacing) from a galaxy, that energy density is down by a factor of ten thousand (though you need to be a little careful–there are lots of galaxies, all contributing). When you do the calculation carefully, the factor is more like 100, but still large. If you do not count the kinetic energy of the expansion itself (which may be infinite), but just look at the random orbital motions, you get an even smaller number; the random motions induced by gravity about ‘pure’ expansion are about $300\text{km/s} = 3 \times 10^7\text{cm/s}$, and the matter density is probably about $1 \times 10^{-30}\text{g/cm}^3$, so the kinetic energy density $U_K$ associated with these motions is

$$U_K = \rho v^2/2 \approx 4.5 \times 10^{-16}\text{erg/cm}^3,$$

about a factor of a thousand smaller than the blackbody radiation density. The only energy density which is not small compared to this is the total rest mass energy density $U_m$ in the universe; for this one multiplies the mass in each cubic centimeter by $c^2$:

$$U_m = \rho c^2 \approx 9 \times 10^{-10}\text{erg/cm}^3,$$

a factor of about 2500 larger than the energy density of the radiation field. Thus if one could magically annihilate all the matter in the universe, one would produce much more energy than is in the radiation field at present.

But was this always so? Think about how the radiation field changes with time. Each photon has a wavelength which as it travels through space overtakes galaxies which are ever receding from where it was last, and the wavelength grows proportional to $R$; the frequency $\nu$ and hence the energy $h\nu$ decreases by the inverse of this. The shape of the blackbody spectrum does not change, but as the frequencies increase as you go back in time ($R\nu = \text{const}$ for each such photon) the temperature must increase in step. (Think about a photon at the peak of the blackbody spectrum–at a time when $R(t)$ is half its present size, the frequency of that photon is twice its present frequency, and it is still at the peak, because all the other photons have been moving in frequency with it–so the shape of the curve at that time is still like a blackbody, but at twice the temperature.) So it would appear that $T \propto 1/R$. 

7
But what about the normalization? That works, too. Each photon has an energy which scales like $1/R$, but the density of photons clearly scales like $1/R^3$ as the universe expands, so the energy density scales like energy $\times$ number density, so

$$U_r \propto 1/R^4 \propto T^4 \propto (1 + z)^4$$

What have we learned? If the universe is filled with a blackbody radiation field at some temperature $T$ at any epoch and that field does not interact with anything in a way to change the number or energy of the photons, then the universe at any other epoch is also filled with a blackbody field, and the temperatures at the two epochs are related by

$$T_1/T_2 = R_2/R_1.$$ 

But note that the radiation energy density increases as one goes back in time like $R^{-4}$ because both the energy and the number density are changing with $R$, but the matter density and hence also its rest mass energy density increase only as $R^{-3}$. So even if their ratio now is 2500, this ratio decreases as one goes back in time:

$$\frac{U_r}{U_m} \propto \frac{R^{-4}}{R^{-3}} = 1/R.$$ 

Thus at a redshift factor $(1 + z)_{eq} \approx 2500$, the energy density of radiation is equal to that of even the rest energy of matter, and ever earlier it was greater. The temperature $T_{eq}$ at this epoch is about

$$T_{eq} \approx 2.7K(1 + z)_{eq} \approx 6800K.$$ 

The mass density at that time was about $1 \times 10^{-30}(1 + z_{eq})^3 \approx 1.6 \times 10^{-20}$, or about $10^4$ hydrogen atoms per cubic centimeter, still a good vacuum but comparable to the density in fairly high density regions in the interstellar medium in the galaxy. The age of the universe at this time was about $t_0(1 + z_{eq})^{-3/2}$, about 100,000 years. Thus at redshifts greater than about 2500, the universe was radiation dominated, meaning that even the rest energy of the matter becomes insignificant. Clearly our Newtonian ideas do not work in this case, and we will have to reexamine such things as the Friedman equation and our conclusions about the inevitability of the big bang—but note that the energy density, which is, after all, what gravitates, goes up into the past even faster than in the case of matter alone, so we should not expect, and indeed will not receive, a respite here.

It is interesting that something else very significant happens at about this time. If one has a very dilute hydrogen gas at these temperatures and densities, a significant fraction of the gas becomes ionized. Even though the average photon energy is only about $kT \approx 0.6$ eV and it takes 13 eV to ionize a hydrogen atom, there are a few photons in the high-energy thermal tail, and so vastly many more quantum states available in the continuum than in the bound states that the gas becomes essentially completely ionized at a temperature of about 4000K, at $(1 + z_{eq}) \approx 1500$. This is called, rather inappropriately, the recombination
epoch, but it should perhaps more properly be called the combination epoch—prior to this, the universe was completely ionized, and after, at least for a long while, completely neutral, and the transition happens very quickly indeed. The significance of this period to the discussion at hand is that when the universe was ionized there were lots of free electrons around, and free electrons can scatter radiation with the Thompson cross section, \( \sigma_T \approx 6.6 \times 10^{-25} \text{cm}^2 \). The number densities are about \( 3 \times 10^3 \) at this epoch (this is a little later than \( t_{eq} \), and the calculation is exactly as before). A photon will then travel about \( 1/n\sigma \), about \( 5 \times 10^{20} \text{cm} \), about 1.5 kiloparsec, before scattering off an electron. To travel this far takes about 4000 years, much less than the age of the universe then, so the universe is essentially opaque prior to recombination—after, it is almost completely transparent.

But electron scattering does not change the frequency very much, nor does it make or destroy photons, so the blackbody character of the photons remains; at some much earlier time, the photons were actually made; we will consider this epoch later.