PHY 475/375

Lecture 6

(April 11, 2012)

The Friedmann Equation

The idea of curved space, and therefore the mathematics of describing curved space-time, was developed long before Einstein's general theory of relativity. Einstein's breakthrough with general relativity was to link the curvature of space-time to its mass-energy content. The key equations of general relativity are Einstein's field equations, which can be used to find the linkage between $a(t), \kappa$, and R_0, which describe the curvature of the universe, and the energy density $\varepsilon(t)$ and pressure $P(t)$ of the contents of the universe.

The equation which connects $a(t), \kappa, R_0$, and the energy density $\varepsilon(t)$ of the universe is known as the *Friedmann equation*, after Alexander Alexandrovich Freidmann, who first derived the equation in 1922. It is of interest to note that Friedmann derived his equation in 1922, seven years before Hubble's discovery of the expansion of the Universe in 1929.

While Friedmann derived his equation using the full suite of general relativistic equations, we will first derive a non-relativistic equivalent using Newton's 2^{nd} law of motion and law of gravity, and then state (without proof) the modifications that must be made to obtain the general relativistic form of the Friedmann equation.

Consider a homogenous sphere with a total mass M_s that is constant in time. The sphere is expanding or contracting isotropically, so that its radius $R_s(t)$ is increasing or decreasing with time. Place an infinitesimal test mass m at the surface of the sphere. Newton's law of gravity gives the gravitational force experienced by this test mass m as:

$$F = -\frac{GM_s m}{R_s(t)^2} \tag{4.3}$$

Using Newton's 2^{nd} law of motion ($F = ma$), we can write this force as

$$F = m\frac{d^2R_s}{dt^2} = -\frac{GM_s m}{R_s(t)^2}$$

so that we get the gravitational acceleration at the surface of the sphere as:

$$\frac{d^2R_s}{dt^2} = -\frac{GM_s}{R_s(t)^2} \tag{4.4}$$

Let us now multiply each side of this equation by dR_s/dt

$$\left(\frac{dR_s}{dt}\right)\frac{d^2R_s}{dt^2} = -\frac{GM_s}{R_s(t)^2}\left(\frac{dR_s}{dt}\right) \tag{4.4.a}$$

Note that the left hand side can be written as

$$\frac{1}{2}\,\frac{d}{dt}\left[\left(\frac{dR_s}{dt}\right)^2\right] \tag{4.4.b}$$

We can integrate equation (4.4.a) with the left hand side written as (4.4.b) by writing as

$$\frac{1}{2}\int \frac{d}{dt}\left[\left(\frac{dR_s}{dt}\right)^2\right] dt = -GM_s \int R_S^{-2} dR_s$$

Integrating, we obtain

$$\frac{1}{2}\left(\frac{dR_s}{dt}\right)^2 = -GM_s\left[\frac{R_S^{-2+1}}{-2+1}\right] - U_I$$

where U_I is just a constant of integration; notice that we have chosen a minus sign, opposite to that in your textbook, in order to make it correspond to the general relativistic form that we will eventually write. Cleaning up the right hand side, this becomes

$$\frac{1}{2}\left(\frac{dR_s}{dt}\right)^2 = \frac{GM_s}{R_s(t)} - U_I \tag{4.5}$$

If we rearrange terms as

$$\frac{1}{2}\left(\frac{dR_s}{dt}\right)^2 - \frac{GM_s}{R_s(t)} = -U_I \tag{4.5.a}$$

then the first term on the left can be identified as the *kinetic energy per unit mass*:

$$E_{\text{kin}} = \frac{1}{2}\left(\frac{dR_s}{dt}\right)^2 \tag{4.6}$$

whereas the second term can be identified with the *gravitational potential energy per unit mass*:

$$E_{\text{pot}} = -\frac{GM_s}{R_s(t)} \tag{4.7}$$

Therefore, equation (4.5.a) simply states that the sum of the kinetic energy per unit mass and the gravitational energy per unit mass is constant for an element of matter at the surface of a sphere, as the sphere expands or contracts under its own gravitational influence.

Now, since the mass of the sphere is constant as it expands or contracts, we may write it as

$$M_s = \rho(t)\, V_s(t)$$

where $\rho(t)$ is the density and $V_s(t)$ is the volume of the sphere. They are both time-dependent, so that they can adjust their values to keep their product, the mass (M_s), constant. This gives

$$M_s = \rho(t)\left[\frac{4\pi}{3} R_s(t)^3\right] \tag{4.8}$$

Also, since the expansion is isotropic about the center of the sphere, we may write the radius $R_s(t)$ in the form

$$R_s(t) = a(t)\, r_s \tag{4.9}$$

where $a(t)$ is the scale factor, and r_s is the comoving radius of the sphere.

Putting equations (4.8) and (4.9) into the energy conservation equation (4.5), we obtain

$$\frac{1}{2}\left[\frac{d}{dt}\Big\{a(t)\,r_s\Big\}\right]^2 = \frac{G}{a(t)\,r_s}\left[\frac{4\pi}{3}\,\rho(t)\Big\{a(t)\,r_s\Big\}^3\right] - U_I$$

which simplifies to

$$\frac{1}{2}\,r_s^2\,\dot{a}^2 = \frac{4\pi}{3}\,G r_s^2 \rho(t)\,a(t)^2 - U_I \tag{4.10}$$

using the usual convention that $da/dt = \dot{a}$.

Dividing each side of equation (4.10) by $r_s^2 a^2/2$, we obtain

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{2U_I}{r_s^2}\frac{1}{a(t)^2} \tag{4.11}$$

This is a simplified version of the Friedmann equation, derived using Newtonian mechanics. While our derivation based on Newtonian mechanics has given us a form that looks very close to the actual Friedmann equation, the truth is that an isotropically expanding sphere still contains a very special location — the center of the sphere, which violates the principle of homogeneity. Therefore, the correct form of Friedmann's equation must be derived based on general relativity.

Before writing the general relativistic form of the Friedmann equation, though, it is worth considering some insights we can gain from looking at equation (4.11). Notice that the time derivative of the scale factor enters into this equation as $\dot{a}^2$, so a contracting sphere ($\dot{a} < 0$) is simply the time reversal of an expanding sphere ($\dot{a} > 0$). If we consider the case of an expanding sphere, analogous to our situation in the Universe, then the future of the expanding universe falls into one of 3 classes, depending on the sign of U:

- If $U_I < 0$, the right hand side of equation (4.11) is always positive, which makes $\dot{a}^2$ always positive, so that the expansion never stops.

- If $U_I > 0$, the right hand side of equation (4.11) starts out positive (since we are still considering an expanding universe with $\dot{a} > 0$). However, we will eventually reach a maximum scale factor for which the right hand side will be zero. To find it, we can set the right hand side equal to zero, so that

$$\frac{8\pi G}{3}\rho = -\frac{2U_I}{r_s^2}\frac{1}{a_{\max}^2}$$

  from which we get

$$a_{\max}^2 = -\frac{2U_I}{r_s^2}\left(\frac{3}{8\pi G}\right)\frac{1}{\rho} = -\frac{2U_I}{r_s^2}\left(\frac{3}{8\pi G}\right)\frac{4\pi a_{\max}^3 r_s^3}{3M_s}$$

  and finally

$$a_{\max} = -\frac{GM_s}{U_I r_s} \tag{4.12}$$

  So, at the maximum scale factor given by (4.12), the right hand side of equation (4.11) will be equal to zero, and the *expansion will stop*! Since $\ddot{a}$ will still be negative, the sphere (universe) will then contract.

- If $U_I = 0$, we get the boundary case in which $\dot{a} \to 0$ at $t \to \infty$ and $\rho \to 0$.

Now, let us write down (without proof) the correct form of Friedmann's equation including all general relativistic effects. It is

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3c^2}\varepsilon(t) - \frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2} \tag{4.13}$$

Notice that in going from the Newtonian form of the Friedmann equation that we derived in equation (4.11) to the correct general relativistic form in equation (4.13), we have made two changes:

1. The mass density ρ has been replaced by an energy density ε divided by the speed of light squared. This has its basis in Einstein's mass-energy relationship

$$E = \sqrt{p^2c^2 + m^2c^4} \tag{4.14}$$

2. We have replaced the term $\frac{2U_I}{r_s^2}$ with $\frac{\kappa c^2}{R_0^2}$, where κ is the curvature.

   There are then 3 possibilities for κ, corresponding to the 3 cases for U_I already discussed.

   - The case with $U_I < 0$ in which the expansion never stops corresponds to negative curvature ($\kappa = -1$).
   - The case with $U_I > 0$ in which the expansion stops at a maximum scale factor, then begins to contract, corresponds to positive curvature ($\kappa = +1$).
   - The case with $U_I = 0$ corresponds to the special case where the space is perfectly flat ($\kappa = 0$).

In order to apply the Friedmann equation to the study of the real universe, we must connect it to observable properties. For instance, the Friedmann equation can be tied to the Hubble constant H_0. Recall that we wrote $H(t) \equiv \dot{a}/a$. This allows us to rewrite the Friedmann equation in the form:

$$H(t)^2 = \frac{8\pi G}{3c^2}\varepsilon(t) - \frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2} \tag{4.19}$$

At the present moment

$$H_0 = H(t_0) = \left(\frac{\dot{a}}{a}\right)_{t=t_0} = 70 \pm 7 \text{ km s}^{-1} \text{ Mpc}^{-1} \tag{4.20}$$

Note that purists tend to call $H(t)$ the *Hubble parameter*, and H_0 the *Hubble constant*, where H_0 is the value of $H(t)$ in the present day.

Using the convention that a subscript of "0" indicates the value of a time-varying quantity evaluated at the present, the Friedmann equation evaluated at the present moment is

$$H_0^2 = \frac{8\pi G}{3c^2}\varepsilon_0 - \frac{\kappa c^2}{R_0^2} \tag{4.21}$$

since $a(t_0) = 1$.

Equation (4.21) for the Friedmann equation at the current time gives a relation among H_0, which tells us the current rate of expansion, ε_0, which tells us the current energy density, and κ/R_0^2, which tells us the current curvature. Now, while it is simple to measure the curvature in principle, it turns out to be difficult to do in practice.

In principle, we could determine the curvature of the Universe simply by drawing a very large triangle and measuring the angles α, β, and γ at the vertices. Since the sum of these angles can be found by generalizing equations (3.5), (3.8) and (3.10) to be

$$\alpha + \beta + \gamma = \pi + \kappa A/R_0^2$$

then if $\alpha + \beta + \gamma > \pi$, the Universe would be positively curved, and if $\alpha + \beta + \gamma < \pi$, the Universe would be negatively curved. Moreover, by measuring the area of this triangle, we could figure out the radius of curvature R_0.

In practice, however, the area of the largest triangle we could draw would be much too small (compared to R_0), so the deviation of $(\alpha + \beta + \gamma)$ from π would be too small to measure.

About all we can conclude from geometric arguments is that if the Universe is positively curved, it can't have a radius of curvature that is significantly smaller than the current Hubble distance, $c/H_0 \approx 4300$ Mpc, otherwise we should be able to see more than one image of the same galaxy. This is because if our Universe is positively curved, it has finite size, with a circumference currently equal to $C_0 = 2\pi\, R_0$. If $C_0 \ll ct_0 \sim c/H_0$, then photons will have circumnavigated the Universe. As an extreme example, suppose the Universe is positively curved with a circumference of 10 million LY. Looking toward the galaxy M31 (Andromeda), which is 2 million LY from us, we would see an image of Andromeda as it was 2 million years ago. But, we should also be able to see the light ray that had circumnavigated the 10 million LY circumference, so we would also see Andromeda as it was $10 + 2 = 12$ million years ago. Finally, looking in the direction opposite Andromeda, we should be able to see an image of Andromeda from photons which had traveled $(10 - 2) = 8$ million LY, hence an image of Andromeda as it was 8 million years ago, and so on. Since we don't see periodicities of this kind, we conclude that if the Universe is positively curved, its radius of curvature must be very large, comparable to, or larger than, the current Hubble distance c/H_0.

So, since we cannot measure the curvature by geometric means, we must turn to indirect methods of determining κ and R_0. If we measured H_0 and ε_0, we could use equation (4.21) to determine the curvature. In fact, even without knowing the current density ε_0, we can use equation (4.21) to place a lower limit on R_0 in a negatively curved universe. If we assume ε_0 is non-negative, then for a given value of H_0, the product κ/R_0^2 is minimized in the limit $\varepsilon_0 \to 0$. So in the limit of a totally empty universe with no energy content, the curvature is negative (taking $\kappa = -1$), with a radius of curvature given by

$$\Big[R_0\Big]_{\min} = \frac{c}{H_0} \tag{4.22}$$

This is the minimum radius of curvature that a negatively curved universe can have, assuming of course that curvature is correctly described by general relativity. Since we know that the Universe contains matter and radiation (so $\varepsilon > 0$), the radius of curvature must be larger than the current Hubble distance c/H_0 if the Universe is negatively curved.

In summary, we have shown that the radius of curvature of the Universe must be larger than the current Hubble distance c/H_0, regardless of whether it is positively curved or negatively curved.

The following was done in class on Monday (Apr 16), but is included here for continuity.

To define other useful terms, let us remind ourselves once again that the full general relativistic form of the Friedmann equation with a Robertson-Walker metric is

$$H(t)^2 = \frac{8\pi G}{3c^2}\varepsilon(t) - \frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2} \tag{4.23}$$

In a spatially flat universe, the Friedmann equation then takes on a particularly simple form

$$H(t)^2 = \frac{8\pi G}{3c^2}\varepsilon(t) \tag{4.24}$$

This means that for a given value of the Hubble parameter $H(t)$, we can define a critical energy density

$$\varepsilon_c(t) \equiv \frac{3c^2}{8\pi G}H(t)^2 \tag{4.25}$$

This allows us to define two cases:

- If the energy density $\varepsilon(t)$ is greater than this critical value, i.e., $\varepsilon(t) > \varepsilon_c(t)$, the universe is positively curved ($\kappa = +1$).
- If the energy density $\varepsilon(t)$ is smaller than this critical value, i.e., $\varepsilon(t) < \varepsilon_c(t)$, the universe is negatively curved ($\kappa = -1$).

Since we know the current value of the Hubble parameter to within 10%, we can compute the current value of the critical energy density to within 20%:

$$\varepsilon_{c,0} = \frac{3c^2}{8\pi G}H_0^2 = \frac{3(3\times 10^8\ \mathrm{m\ s^{-1}})^2}{8\pi(6.67\times 10^{-11}\ \mathrm{m^3\ kg^{-1}s^{-2}})}\left(\left\{70\pm 7\right\}\frac{\mathrm{km}}{\mathrm{s}}\ \frac{1}{\mathrm{Mpc}}\ \frac{\mathrm{Mpc}}{3.1\times 10^{19}\ \mathrm{km}}\right)^2$$

from which we obtain

$$\varepsilon_{c,0} = (8.3\pm 1.7)\times 10^{-10}\ \mathrm{J\ m^{-3}} \equiv 5200\pm 1000\ \mathrm{MeV\ m^{-3}} \tag{4.26}$$

The critical energy density is frequency written as the equivalent mass density

$$\rho_{c,0} \equiv \frac{\varepsilon_{c,0}}{c^2} = (9.2\pm 1.8)\times 10^{-27}\ \mathrm{kg\ m^{-3}} = (1.4\pm 0.3)\times 10^{11}\ \mathrm{M_\odot Mpc^{-3}} \tag{4.27}$$

Even though this is not a large density by even interstellar standards (interstellar space has a number density of about 1 atom $\mathrm{cm^{-3}}$ even in its most tenuous spaces, so $1.67\times 10^{-21}\ \mathrm{cm^{-3}}$), when averaged over 100 Mpc scale voids, the Universe has a mean density that is close to the critical density.

In discussing the curvature of the Universe, it is more convenient to use not the absolute energy density ε, but the ratio of the energy density to the critical energy density. So, in talking about the energy density of the universe, cosmologists use the dimensionless density parameter:

$$\Omega(t) \equiv \frac{\varepsilon(t)}{\varepsilon_c(t)} \tag{4.28}$$

The most conservative limits on Ω state that the current value of the density parameter lies in the range $0.1 < \Omega_0 < 2$.

In terms of the density parameter, the Friedmann equation can be written in yet another useful form, which we can obtain by first dividing both sides of equation (4.23) by $H(t)^2$:

$$\frac{H(t)^2}{H(t)^2} = \frac{8\pi G}{3c^2 H(t)^2}\varepsilon(t) - \frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2 H(t)^2}$$

From equation (4.25), we have $8\pi G/3c^2 H(t)^2 = 1/\varepsilon_c(t)$, so the above equation becomes

$$1 = \frac{\varepsilon(t)}{\varepsilon_c(t)} - \frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2 H(t)^2}$$

so that, upon rearranging terms and inserting $\Omega(t)$ from the definition in equation (4.28) we get

$$1 - \Omega(t) = -\frac{\kappa c^2}{R_0^2}\frac{1}{a(t)^2 H(t)^2} \tag{4.29}$$

Note that since the right hand side of equation (4.29) cannot change sign as the universe expands, neither can the left hand side. So, if $\Omega < 1$ at any time, it remains less than 1 for all time; similarly if $\Omega > 1$ at any time, it remains greater than 1 for all times, and if $\Omega = 1$ at any time, it remains equal to 1 for all times.

At the present epoch, the relation among curvature, density, and expansion rate can be written in the form:

$$1 - \Omega_0 = -\frac{\kappa c^2}{R_0^2 H_0^2} \tag{4.30}$$

since $a(t_0) = 1$ at the present time.

Equation (4.30) can be rearranged as

$$\frac{\kappa}{R_0^2} = \frac{H_0^2}{c^2}\left(\Omega_0 - 1\right) \tag{4.31}$$

So, if we know Ω_0, we know the sign of the curvature κ. If we also know the Hubble distance c/H_0, we can compute the radius of curvature R_0.