Brief introduction to FRW cosmology lectures for MATH199

Gregory Levine
Hofstra University

February 20, 2003

Contents

1 The FRW metric tensor

1.1 Introduction

Isotropy and homogeneity of matter lead to a universe described by a Friedmann-Robertson-Walker (FRW) metric tensor. Rather than trying to prove this assertion, we proceed more "deductively" by assuming a particular form for the metric tensor, and then developing some conclusions based upon it. These conclusions match our universe very convincingly and lead comsmologists to believe that they are on the right track.

Recalling the form of the (exterior) Schwarschild (S) metric tensor, both spatial and temporal components of $g_{\mu\nu}$ depended explicitly upon the spatial coordinate r. Thus, "observable" time and space intervals were related to coordinate time and space intervals by *space* dependent factors of the form $(1 - r_s/r)^{\pm 1}$. In FRW cosmology, the relationship between observable space intervals and coordinate space intervals will depend upon a factor (the scale factor $a(t)$) that depends explicitly on coordinate *time*. In contrast to the exterior S metric tensor, the FRW metric tensor is not a vacuum solution to the Einstein field equations; rather, it involves a uniform distribution of mass/energy (referred to as "dust"). The density of "dust" in the universe will determine the time evolution of the scale factor. The FRW metric tensor is (in spherical polar coordinates):

$$g_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{a^2(t)}{1-\kappa r^2} & 0 & 0 \\ 0 & 0 & -a^2(t)r^2 & 0 \\ 0 & 0 & 0 & -a^2(t)r^2\sin^2\theta \end{pmatrix} \tag{1}$$

A spacetime interval is then:

$$d\tau^2 = dt^2 - a^2(t)(\frac{1}{1-\kappa r^2}dr^2 + r^2 d\theta^2 + r^2\sin^2\theta d\phi^2) \tag{2}$$

The scale factor $a(t)$ that multiplies the whole spatial sector of the metric tensor leads to the possibility of an expanding (or contracting) universe in that observable spatial intervals change in time. (Note that I have chosen to denote coordinate time by t rather than x^0.)

Also appearing in g_{11} is a space dependent factor involving κ. κ represents the curvature of the spatial part of the metric. If we look at a spatial slice at fixed coordinate time, κ is the scalar curvature ($R = g_{ab}R^{ab}$) of the spatial "sub-manifold" of spacetime ($a, b = 1, 2, 3$). By rescaling $r \to r/\sqrt{|\kappa|}$

and $a \to a\sqrt{|\kappa|}$, κ may be scaled to $\kappa = \pm 1, 0$. It is interesting to note that the conditions of isotropy and homogeneity impose (at a sufficiently "large scale") a universe fulfilling exactly one of the above cases. Although there is now strong evidence that the spatial sector of our universe is geometrically flat ($\kappa = 0$), it is useful (if not mind-bending) to investigate the implications of $\kappa \neq 0$.

1.2 Positive curvature ($\kappa = 1$)

The case of positive curvature lends itself to a simple physical interpretation. First consider the spatial sector of the metric tensor limited to $\theta = \pi/2$. We "compactify" r by making the following substitution (needs to be justified—single coordinate patch?):

$$r = \sin\chi; \quad 1 - r^2 = \cos^2\chi; \quad \kappa = 1$$

The spacetime interval is now:

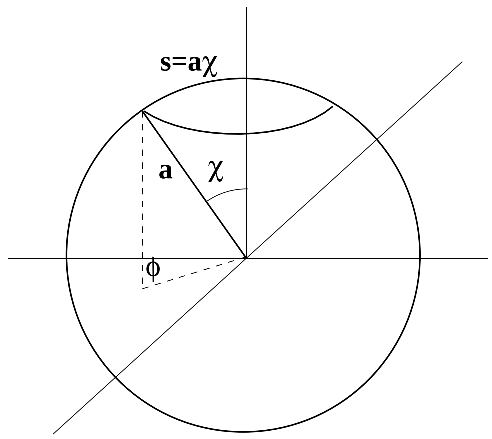
$$d\tau^2 = dt^2 - a^2(d\chi^2 + \sin^2\chi d\phi^2) \tag{3}$$

The spatial sector now looks like S^2 with the former radial coordinate, r playing the role of a *polar angle*, χ. The scale factor, $a(t)$, now serves as the "radius" of S^2, converting between polar angle, χ and observable radial distance, $a(t)\chi$. (To picture the full 3-dimensional spatial sector, you have to picture S^2 as described attached to every "real" polar angle, θ. Alternatively, you can write the equation for S^3 imbedded in R^4 and eliminate one coordiate, etc.) One immediate consequence is that **space is finite**! Observable radial distance, s is related to r by:

$$s = a\chi = a\sin^{-1} r$$

Therefore r is bounded by $r \in [-1, 1]$ and the maximum spatial separation is $s = a\pi$.

The significance of curvature in the spatial sector may be seen by looking at the relationship between radial distance and circumference within FRW space. In a homogeneous, isotropic universe, our radial position is not unique; so we can imagine $\chi = 0$ to be located in this room (call this point O). Consider taking a walk in some direction from O for a (observable) distance of s. Now take a walk visiting the locus of points equidistant from O until you return to where you started (walk in a circle). To calculate



the circumference, c, we use the metric (3) and look at a purely azimuthal coordinate interval ($\phi \in [0, 2\pi]$) for fixed χ. It follows that:

$$c = a \int_0^{2\pi} \sin\chi d\phi = 2\pi a \sin\chi = 2\pi a \sin(s/a),$$

where, in the last equality, we have expressed the coordinate interval χ by the observable radial distance, s, to which it corresponds. To compare positive curvature to the familiar Euclidean case, expand the observable circumference,c, as a series in the observable radius, s:

$$c = 2\pi a \sin(s/a) = 2\pi s - \frac{\pi}{3}\frac{s^3}{a^2} + \dots \quad (< 2\pi s)$$

In a positively curved space, $c < 2\pi s$ with the Euclidean case ($c = 2\pi s$) recovered in the limit $a \to \infty$. It is important to keep in mind that the scale factor $a(t)$ is a parameter not a coordinate and is thus unobservable.

1.3 Negative curvature ($\kappa = -1$)

The hyperbolic case, $\kappa = -1$, does not have an interpretation that is easily visualized. However, making the same style cooridate transformations (with hyperbolic function instead) leads to:

$$s = a \sinh^{-1} r$$

$$c = 2\pi a \sinh(s/a) = 2\pi s + \frac{\pi}{3}\frac{s^3}{a^2} + \dots \quad (> 2\pi s)$$

Thus r is unbounded and space (the range of s) infinite.

2 Einstein field equations and implication for cosmology

Now we turn to the problem of working out the dynamics of $a(t)$ in the presence of uniformly distributed matter. The Einstein field equations in the presence of matter are:

$$R^{\mu\nu} - (\frac{1}{2}R - \Lambda)g^{\mu\nu} = 8\pi G T^{\mu\nu} \quad G \simeq 8.3 \times 10^{-45}\text{m/J} \tag{4}$$

The gravitational constant, $G \equiv G_{\text{Newton}}/c^4$ (where G_{Newton} is the "familiar" G constant entering $F = Gm_1m_2/r^2$), has units of curvature/(energy density) or $\text{m}^{-2}/(\text{J/m}^3)$. Water has a density of 1000 kg/m^3 and therefore an energy density of $\sim 10^{20}$ J/m^3. Water, then, produces a tiny *local* spacetime curvature of approximately 10^{-24} m^{-2}. The "radius of curvature" is approximately 10^{12} m—about Pluto's orbital distance from the sun. This law may be thought of in terms of Gauss' law in electrostatics. The local charge density is proportional to the second (spatial) derivative of the electromagnetic vector potential, $\partial^2 A$, which may be thought of as a "curvature" analogous to the left hand side of the Einstein equations. This curvature is not, however, curvature within a metric space (physical spacetime), but is a curvature of an internal space related to quantum mechanical degrees of freedom.

The cosmologic constant Λ is known to be less than, for instance, the curvature corresponding to the mass density of our solar system ($\sim 10^{-10}$ kg/m^3). Effects on this scale would be readily observable. However, as discussed later, there is presently strong evidence for a non-zero Λ.

The left hand side of the Einstein equations (4) must be calculated from the assumed form of the metric tensor (1). This is most easily done by writing the Euler-Lagrange equations for the FRW spacetime interval (2) and comparing coefficients to the geodesic equation

$$\frac{d^2x^\mu}{d\tau^2} + \Gamma^\mu_{\nu\rho}\frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau} = 0 \tag{5}$$

to read off the components of the Christoffel (Christ—Awful!) symbol.

2.0.1 Example: affine connection and Ricci tensors for S^2

We can use the variational formalism for geodesics and Euler-Lagrange equations to simplify finding the Christoffel components and curvature tensors.

The distance along a curve parameterized by $x^\mu(\tau)$ on a manifold characterized by a metric tensor $g_{\mu\nu}(x)$ is given by:

$$s = \int d\tau \sqrt{g_{\mu\nu}\frac{dx^\mu}{d\tau}\frac{dx^\nu}{d\tau}}$$

Minimizing s with respect to $x^\mu(\tau)$ results in the Euler-Lagrange equation:

$$\frac{d}{d\tau}\frac{\partial L}{\partial(dx^\mu/d\tau)} = \frac{\partial L}{\partial x^\mu}$$

where L is chosen to be the integrand squared. The E-L equation then becomes:

$$2g_{\mu\nu}\frac{d^2x^\nu}{d\tau^2} + \frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}(\frac{\partial g_{\mu\nu}}{\partial x^\rho} + \frac{\partial g_{\mu\rho}}{\partial x^\nu}) = \frac{dx^\nu}{d\tau}\frac{dx^\rho}{d\tau}\frac{\partial g_{\rho\nu}}{\partial x^\mu}$$

This equation may be reorganized to give the geodesic equation (5) (derived by other means in class.)

To see how to use the variation of s advantageously, consider the problem of finding the affine connection (Christoffel symbol) and Ricci tensors for the Riemannian manifold S^2. The metric for S^2 and the corresponding lagrangian may be expressed:

$$ds^2 = d\theta^2 + \sin^2\theta d\phi^2; \quad L = \left(\frac{d\theta}{d\tau}\right)^2 + \sin^2\theta\left(\frac{d\phi}{d\tau}\right)^2$$

The E-L equations become, for this example,

$$\frac{d^2\theta}{d\tau^2} = \frac{1}{2}\sin 2\theta\left(\frac{d\phi}{d\tau}\right)^2; \quad \sin^2\theta\frac{d^2\phi}{d\tau^2} = -\sin 2\theta\frac{d\theta}{d\tau}\frac{d\phi}{d\tau}$$

From these equations, one can "read off" the components of the Christoffel symbol:

$$\Gamma^\theta_{\phi\phi} = \frac{1}{2}\sin 2\theta; \quad \Gamma^\phi_{\theta\phi} = \Gamma^\theta_{\phi\theta} = -2\cot\theta$$

Now returning to the FRW metric—from $\Gamma^\mu_{\nu\rho}$, the Riemann tensor and scalar curvature R are calculated. The results are:

$$R^{00} = -3\frac{\ddot{a}}{a}, \quad R^{ij} = -(\frac{\ddot{a}}{a} + 2\frac{\dot{a}^2}{a^2} + 2\frac{\kappa}{a^2})\delta^{ij}$$

and

$$R = g_{\mu\nu}R^{\mu\nu} = -6(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{\kappa}{a^2})$$

where overdots mean coordinate time derivative.

Some assumptions on $T^{\mu\nu}$ will dictate the dynamics of $a(t)$. Furthermore, there will be an interesting connection between $a(t)$ and the scalar curvature of the spatial section, κ. Before looking at the details of $a(t)$ for our universe, we will look at the general properties of light rays (null vectors) in a FRW universe with some $a(t)$.

2.1 Light rays in an FRW universe

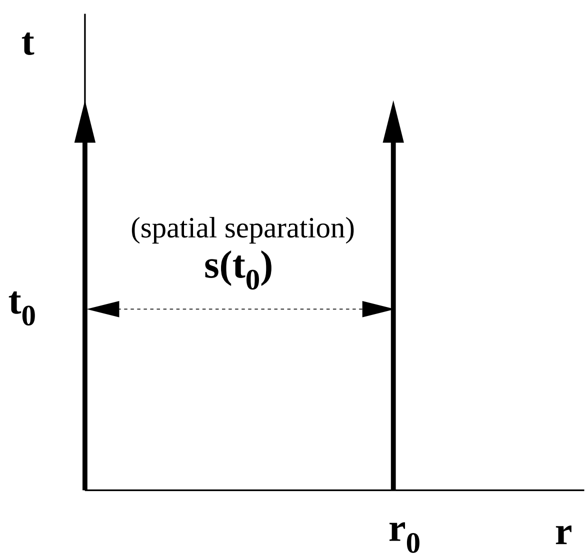
Consider two observers with constant spatial coordinate locations (at $r = 0$ and $r = r_0$). (Their worldlines are actually geodesics; they may be thought of as being in freefall.) At some coordinate time t_0 they have a spatial separation, $s(t_0) = a(t_0)r_0$ in a $\kappa = 0$ (flat) FRW universe. Since $\dot{s}(t_0) = \dot{a}(t_0)r_0$,

$$\dot{s}(t_0) = \frac{\dot{a}(t_0)}{a(t_0)} s(t_0) \tag{6}$$

Note that this equation remains valid when $\kappa \neq 0$. The spatial separation is

$$s(t_0) = a(t_0) \int_0^{r_0} \frac{dr}{\sqrt{1 - \kappa r^2}} = a(t_0) d_0$$

and d_0 plays the role of r_0 above.



Equation (6) describes the apparent relative motion of two observers in freefall (and it is independent of curvature!) In an expanding universe, where $a(t)$ is a monotonically increasing function, the ratio $\dot{a}/a$ is referred to as the Hubble (notso) constant:

$$\dot{s}(t) = H(t)s(t) \tag{7}$$

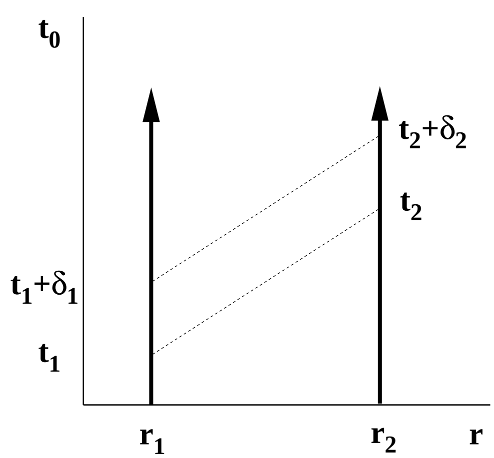
Early measurements of the velocity (relative to us) of stars in nearby galaxies revealed that the velocity was always *recessional* and independent of the direction of observation. The velocity was also seen to be proportional to the distance to the star. This striking observation of the cosmologic relative motion can only be understood within the framework of General Relativity and, furthermore, confirms one of the most straightforward solutions to the Einstein equations—the Big Bang. (see attached fig. 1).

$H(t)$ is measured indirectly by exploiting the doppler shift in the wavelength of a light wave emitted from a source in relative motion to the observer. To see how $a(t)$ is related to redshift, consider again the worldlines of two freefalling observers separated by coordinate interval r_0. The null vector condition $d\tau = 0$ for the FRW metric yields the following equation for the null vector:

$$dt^2 = a^2(t)\frac{dr^2}{1-\kappa r^2}$$

The equation for a null worldline starting at $r = r_1, t = t_1$ and arriving at $r = r_2$ is given implicitly in terms of t_2 by the first equality in:

$$\int_{t_1}^{t_2} \frac{dt}{a(t)} = \int_{r_1}^{r_2} \frac{dr}{\sqrt{1-\kappa r^2}} = \int_{t_1+\delta_1}^{t_2+\delta_2} \frac{dt}{a(t)}$$



However, the equation for a null worldline also between r_1 and r_2, but starting at a later coordinate time $t_1 + \delta_1$ is given by the second equality above. Thus, regarding δ as the period of a light wave traveling between r_1 and r_2, the ratio of the periods is equal to the ratio of the scale factors at emission and observation corrdinate times.

$$\frac{\delta_2}{\delta_1} = \frac{a(t_2)}{a(t_1)} = \frac{\lambda_2}{\lambda_1} \tag{8}$$

The last equality is known as the "cosmologic redshift". The wavelengths of known atomic spectra are compared with astrophysical (extragalactic ones) to determine the ratio of the scale factors at emission and observation. With an $a(t)$ increasing in time, a characteristic photon from a hydrogen atom emitted in a distant galaxy and observed on earth has a longer wavelength than the same photon emitted from a hydrogen atom in the laboratory! (A measurable fraction of the inter-station static on your TV is due to radio frequency photons that are "redshifted" 13 eV UV photons emitted $\sim 200,000$ years after the big bang.)

Because of equation (8), the Hubble parameter $H(t)$ given by equation (7) is a *perfectly observable quantity*, relating $a(t)$ presently to $a(t)$ when the observed light was emitted. In 1928-9, Edwin Hubble made a survey of Cepheid variable stars in nearby galaxies using the 120 in telescope on Mount Wilson, California. The result was the simple plot shown above— the universe is expanding! A more careful treatment of the distance/redshift relationship, taking into account the effect of changing $a(t)$ on luminosity of a source yields (this is a complicated affair.):

$$d_l = H^{-1}(t_0)[z + \frac{1}{2}(1 - q(t_0))z^2 + O(z^3) + \ldots] \tag{9}$$

where d_l is the distance inferred from the absolute luminosity without the effects of cosmology taken into account and the redshift parameter, $z \equiv \lambda_{\text{observed}}/\lambda_{\text{emitted}} - 1$. t_0 is taken to be the "present" time in the evolution of the universe, at which observations are made. The (misnamed) deacceleration parameter, $q(t)$ is defined to be

$$q(t) \equiv -\frac{a(t)\ddot{a}(t)}{\dot{a}^2(t)}$$

Notice that the first term in the expansion for d_l reproduces equation (7) where velocity is identified with redshift (they are in fact proportional in the non-relativistic Doppler effect). The physical significance of $q(t)$ is best understood after introducing the explicit dynamics of $a(t)$.

2.2 The Friedmann equations

We now return to the question of the explicit dynamics of $a(t)$ given some reasonable model of the mass/energy distribution in our universe. Conventional FRW cosmology assumes the existence of a homogeneous and isotropic distribution of mass/energy giving rise to a diagonal stress energy tensor of the form:

$$T^{00} = \rho(t); \qquad T^{ij} = p(t)$$

where ρ is mass density and p is pressure. The Einstein field equations then become:

$$3\left(\frac{\dot{a}^2}{a^2} + \frac{\kappa}{a^2}\right) = G\rho + \Lambda$$

$$2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{\kappa}{a^2} = -Gp + \Lambda$$

G has now been rescaled: $G \to 8\pi G$. These two equations are conventionally written:

$$\dot{a}^2 + \kappa = \frac{1}{3}(G\rho + \Lambda)a^2 \tag{10}$$

$$6\frac{\ddot{a}}{a} = -G(\rho + 3p) + 2\Lambda \tag{11}$$

Let us first look at $\Lambda = 0$. It is convenient to combine equations (10) and (11) to form:

$$\frac{d(\rho a^3)}{dt} = -p\frac{da^3}{dt} \tag{12}$$

This equation may be thought of as a local conservation of energy condition: Work done by the system plus the change in the energy is zero. Notice that a non-zero cosmologic constant Λ in equations (10) and (11) does not change this result.

3 The expanding/contracting universe

3.1 Solution of the Friedmann equations

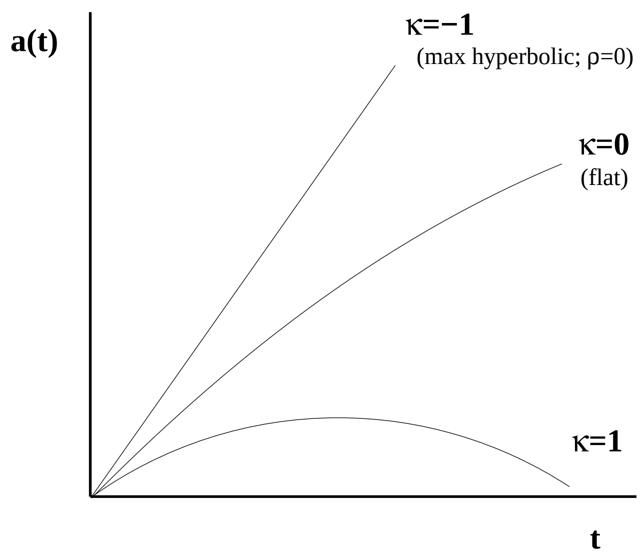
To solve the equations (10) and (11), we need a third condition (a thermodynamic equation of state) relating ρ to p. Presently the energy density in the universe, ρ, reflects the mass energy density of particles rather than their kinetic energy. Similarly, the pressure p resulting from their kinetic energy is also comparitively small. We therefore utilize a simple equation of state, $p = 0$. The conservation equation (12) may be integrated to yield:

$$\rho a^3 \equiv 3M = \text{constant}$$

The constant $3M$ may be loosely regarded as the total mass of a comoving region of size a. (comments: scale a not observable, only ratios; can it be related to total mass within present horizon?) Now equation (10) may be written in a suggestive form:

$$\dot{a}^2 - \frac{GM}{a} = -\kappa \tag{13}$$

Equation (13) looks like the statement of conservation of energy for Keplerian orbit: $K + U = E$, where K is the kinetic energy, P is gravitational potential energy and E is the total energy. In this analog, the curvature, κ, plays the role of $-E$. Just as in the Kepler orbital problem, there are three possibilities corresponding to $\kappa = 1$ ($E < 0$), $\kappa = -1$ ($E > 0$) and $\kappa = 0$ ($E = 0$).



It is a profound implication of General Relativity in a homogeneous setting that the Friedmann equation (10) must have a singularity at some time. Careful work along these lines by Hawking and Penrose have established the existence of an initial singularity (assuming the weak energy condition

$\rho + 3p > 0$). This mathematical feature has come to be named the "Big Bang."

For $\kappa = -1$, negative Gaussian curvature, the universe has a perpetual expansion. The case plotted below corresponds to a $\rho = 0$ universe in which equation (10) is easily integrated to yield a linearly increasing $a(t)$. (A pun: Nothing prevents a massless universe from expanding.) In the Keplerian analog, this case corresponds to an object vastly exceding escape velocity ($E >> 0$). For $\kappa = 1$, positive curvature, the universe expands and then contracts to a "Big Crunch." If $\kappa = 0$, the expansion slows asymptotically to $H = 0$. This is a necessary condition for living in a geometrically flat universe. When we measure the Hubble parameter $H(t_0) = \dot{a}(t_0)/a(t_0)$ at our present time, we are measuring the slope of one the curves depicted below at a particular time. But this is not enough information to know what curve we're on.

Notice that in all cases, the "acceleration" of $a(t)$ is negative. The effect of matter in the universe is always to slow the expansion. This is technically a property of the weak energy condition, as seen from equation (11). However, observations of high redshift galaxies in the last few years has radically changed this picture! I will try to relate my understanding of the present state of affairs in cosmology.

The curvature of the universe is sometimes recast in the form of a critical mass density of the universe, ρ_c, in the following way: If $\kappa = 0$ (flat universe), the present day Hubble parameter $H(t_0)$—which is *observable*—may be related to the present mass density.

$$H^2(t_0) = \frac{\dot{a}^2}{a^2} = \frac{1}{3} G\rho(t_0)$$

Thus

$$\rho_c(t_0) \equiv \frac{3H^2(t_0)}{G}$$

$\rho_c(t)$ is thus the critical density *at a particular time*, t, needed to produce a flat universe. From plots like the original one of Hubble, $H(t_0) \simeq 3 \times 10^{-18}$ (m/s)/m = s^{-1}. (This is about 30 km/sec per million light years.) Putting in this value of the Hubble parameter, one gets a critical density of $3H_0^2/(8\pi G c^4) \sim 1.6 \times 10^{-26}$ kg/m^3. That is approximately 20 Hydrogen atoms per cubic meter.

This number, calculated entirely from astronomical observations, is surprisingly close (in order of magnitude) to estimates of the mass density (also from astronomical observations). Commonly, $\rho(t)$ at any time t is expressed as a ratio to $\rho_c(t)$ at the same time t:

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_c(t)}$$

A simple survey of galactic luminous matter gives $\Omega(t_0) \sim 0.01 - 0.03$. However, measurements of the mass of galaxy and galaxy cluster haloes (mass inferred inertially by measurements of rotational velocities) greatly increase this number to $\Omega \sim 0.4$. Thus, *almost all* of the matter in the universe is unseen. Most of this so-called "dark matter" must be non-baryonic because nucleosynthesis limits the baryonic contribution to approximately $\Omega \sim 0.03$. Some of the matter might be massive neutrinos; but most, it seems, must be "exotic." "dark matter," then, refers to matter that can be detected inertially but is non-luminous and probably "exotic."

Until five years ago, the evidence for $\Omega = 1$ was purely circumstantial: the observed $\Omega \sim 0.4$ was simply too close to $\Omega = 1$ to be a coincidence (remember the orders of magnitude.) Because Ω remained stubbornly below $\Omega = 1$ despite many attempts to close the gap by observations, some astrophysicists began to think that maybe we live in a slightly hyperbolic ($\kappa = -1$) universe after all. $\Omega > 1$ (a closed, finite universe) seems very unlikely based on the present accounting of matter.

Two observations in the last five years radically changed this problem and its solution: 1) Compelling independent evidence that we live in a flat $\kappa = 0$ universe was obtained. (note: I won't talk about this—recent CMB data). 2) Surveys of high redshift galaxies revealed that the deacceleration parameter, $q(t)$, was *negative*; that is, the expansion universe is accelerating ($\ddot{a} > 0$)! Before getting to these recent developments, we have to examine cosmology before the cosmologic constant.

3.2 The age of the universe

At approximately $200,000$ years after $t = 0$, the universe entered its matter-dominated (as opposed to radiation-dominated) phase. The simple equation of state $p = 0$ becomes valid because the pressure from photons after decoupling becomes negligible. Therefore the Friedmann equation may be easily integrated to find the explicit dependence of the scale factor $a(t)$ on time during this epoch (for $\kappa = 0$). This calculation also serves to illustrate that the only length scale in a homogeneous universe is the spacetime curvature induced by the constant density.

Suppose at the present time, t_0, the density of matter is assessed locally and since, by assumption, the universe is homogeneous, $\rho(t_0)$ is the density everywhere in the universe at coordinate time t_0. When the local energy conservation equation (12) is integrated, the constant of integration is

$$M = \rho(t_0)a^3(t_0)$$

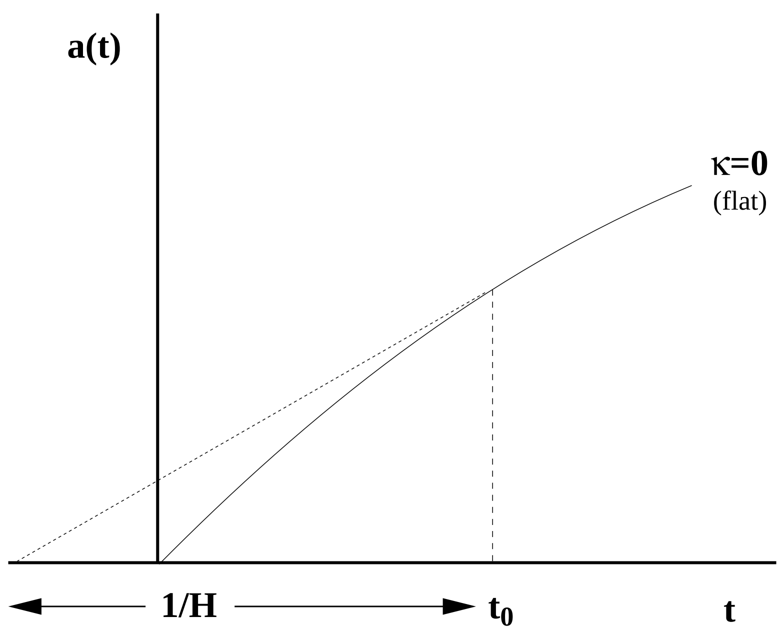
Now the Friedmann equation (13) may be integrated yielding

$$a(t) = \frac{3}{2}\left(\frac{GM}{3}\right)^{1/3} t^{2/3} = \frac{3}{2}\left(\frac{G\rho(t_0)}{3}\right)^{1/3} a(t_0)t^{2/3}$$

But $G\rho(t_0)$ is a curvature—call it r_0^{-2}—with physical dimensions of m^{-2}. Thus the scale at one time, t, is related to the scale factor at another time, t_0, by

$$a(t) = a(t_0)\left(\frac{t}{r_0}\right)^{2/3}$$

where r_0 depends on the effective "curvature" induced by the distribution of matter at t_0.



Looking at the $\kappa = 0$ curve depicted in fig. 5, the age of the universe may be estimated by the present day Hubble parameter $H(t_0)$. A simple estimate is from the slope of the scale factor presently, $t_0 = H^{-1}$ (shown by the dotted line in the figure). The estimate may be improved by the calculation above for the dynamics of the scale factor in a matter dominated universe. This will give

$$t_0 = \frac{2}{3} H^{-1}(t_0)$$

3.3 The flatness problem

Although the universe (according to CMB) appears to be flat, flatness ($\Omega = 1$) is difficult to achieve in a dynamical setting. $\Omega(t)$ is a time dependent quantity and it is useful to consider the dynamics of Ω itself. Writing $\rho = \rho_c \Omega = (\dot{a}^2/a^2)\Omega$, the first Friedmann equation becomes:

$$\dot{a}^2(t)(\Omega(t) - 1) = \kappa \tag{14}$$

This equation gives the sensible results that: $\Omega > 1$ corresponds to a $\kappa = 1$ (positively curved) universe; $\Omega < 1$ corresponds to a $\kappa = -1$ (negatively curved) universe; $\Omega = 1$ corresponds to a $\kappa = 0$ (flat) universe.

However, consider the hypothetical plausibility of an $\Omega(t_0) = 1/2$ universe if it were observed presently. Integrating equation (14) back in time assuming some simple power law for $a(t) = c(t/t_0)^\alpha$ yields

$$\Omega(t) - 1 = \frac{\kappa t_0^2}{c^2 \alpha^2} (\frac{t_0}{t})^{2\alpha - 2}$$

When the universe was, say, 10^{-6} the age it is today ($t = 10^{-6} t_0$), the factor of c must have been exquisitely fine-tuned to a precision of $O((t/t_0)^{\alpha - 1})$ to achieve $\Omega(t_0) = 1/2$. However, if for some reason $\Omega(t) = 1$ at any time t, $\Omega = 1$ at all times.

3.4 The horizon problem (for $\kappa = 0$)

In the radiation dominated epoch, the universe was composed of photons and a highly ionized hydrogen plasma. At a time known as the recombination time, $t_r \sim 2 \times 10^5$ years, the universe cooled enough for hydrogen to form a bound state and the universe became largely transparent to photons.

Consider the maximum spatial coordinate displacement, r_r a photon might have had from $t = 0$ to recombination at $t = t_r$. This is given by the equation for the null vector, where $a(t) = \alpha t^{1/2}$ in radiation dominated epoch:

$$dt^2 = a^2(t) dr^2$$

Integrating,

$$r_r = \int_0^{t_r} \frac{dt}{a(t)} = \frac{2}{\alpha} t^{1/2}$$

The *observable* distance that the photon travels, $h_r = a(t_r) r_r$, may be thought of as defining a region of the universe that is causally disconnected from every other such region. h_r is called the *horizon distance* at the time of recombination. Regions of space larger than this at the time of recombination may not be entirely homogeneous because photons have not been able to traverse the entire size of the region to equilibrate it. The horizon distance, in terms of t_r is given by:

$$h_r = a(t_r) r_r = \alpha t_r^{1/2} \left(\frac{2}{\alpha} t_r^{1/2} \right) = 2t_r$$

Now suppose we want to know the *present* size of the largest possible causally connected region at the time of recombination. We must evolve the horizon distance at recombination, h_r, forward in time to the present. The present observable distance, d_p, corresponding to the horizon distance at recombination is:

$$d_p = h_r \frac{a(t_0)}{a(t_r)} = h_r \frac{t_0^{2/3}}{t_r^{2/3}} = 2 t_r^{1/3} t_0^{2/3}$$

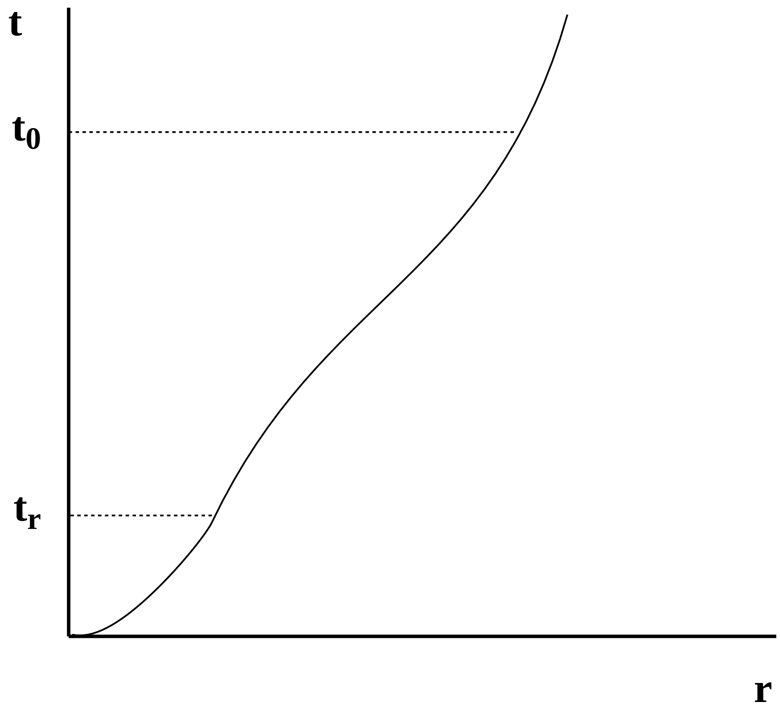
where the scale factor after recombination is taken to increase according to the matter dominated power law.

Now compare the size of these regions to horizon distance at present. To estimate the size of the present horizon we multiply the scale factor (presently) by the coordinate displacement cooresponding to the horizon, which is obtained by integrating the null vector condition as above:

$$h_p = a(t_0) \int_0^{t_0} \frac{dt}{a(t)} = 3t_0$$

In the last step the scale factor for a matter dominated universe ($a(t) = \beta t^{2/3}$) is used. The ratio h_p/d_p determines the number of regions that were causally disconnected at the time of recombination which exist within our observable horizon at present. This is an interesting number in that its inverse,

$$\frac{d_p}{h_p} = \frac{t_r^{1/3} t_0^{2/3}}{3t_0} = \frac{2}{3} \left(\frac{t_r}{t_0} \right)^{1/3} \sim 0.013 = 0.74^\circ,$$



is the angular separation between them in the observable sky. Photons from one such region, which have been dramatically redshifted to microwave frequency, reflect the characteristics—in particular the temperature—of that region. Each region would be expected to exhibit variations in temperature because there is no mechanism by which they might have communicated with one another before recombination. (Their size is defined to be a horizon distance.)

Therefore, looking at the sky in the microwave band, we should first of all expect to see these primordial photons; second, we should see variations in the spectrum of these photons when comparing regions separated by more than $\sim 1^\circ$. The horizon problem is simply that Cosmic Microwave Background (CMB) shows no such variations and is highly isotropic.

However this calculation of the angular size of causally disconnected regions is useful for another reason. As I mentioned before, there is more than circumstantial evidence that $\kappa = 0$ (a flat universe.) Evidently, there was a “homogenizing” mechanism that produced the featureless microwave sky that we see (a popular theory is discussed in the next section.) However, spatial fluctuations on this smooth background *were not* homogenized and they may be probed for angular correlations. Remarkably, the angular correlations have been measured quite accurately and reveal the 1° features consistent with flat FRW models. (The estimate I presented here was for a flat universe but ignored the details of the different radiation dominated and matter dominated epochs of the universe.) Therefore, independent of any measurements of the quantity of luminous or inertial matter, the flatness of spatial sector of the universe is well established.

3.4.1 Digression on inflation

The story is a little more complicated than this. Here is a very brief version: There is a mechanism by which both the horizon and flatness problems may be solved called "Cosmic Inflation." This is a beautiful theory but it involves quantum field theory (QFT) to state precisely.

Owing to QFT, there might have been a time in the very early universe in which a large cosmologic constant dominated the stress-energy tensor. Looking at the first Friedmann equation (10) for $\kappa = 0$, $\rho = 0$ and $\Lambda \neq 0$, this equation exhibits exponential growth in the scale factor. Solving the equation for $\kappa \neq 0$ shows that curvature gets forced to zero in some sense by the huge inflation of the scale factor $a(t)$. Going back to the "curvature experiment" on the $\kappa = 1$ at the beginning of this chapter, exponentially inflating $a(t)$ leaves the pseudo-S^2 representing the universe locally flat, in that deviations from Euclidean behavior go as $O(s^3/a^3)$.

It is interesting to examine the effect of a cosmologic constant from the Einstein-Hilbert action. A space with $\rho = 0$ and $\Lambda \neq 0$ with an action functional

$$S = \frac{1}{G}\int d^4x\sqrt{-g}(R - 2\Lambda) \tag{15}$$

defines a de-Sitter space. Spacetime tries to "minimize" curvature and, because of Λ, maximize the product of the eigenvalues of the metric tensor. Without any energy density ρ to hold it back, spacetime inflates exponentially.

To see how a non-zero Λ might be generated, consider the matter part of the action:

$$S_{\text{mat}} = \frac{1}{G}\int d^4x g_{\mu\nu}T^{\mu\nu} \tag{16}$$

where the classical (or first quantized) stress-energy tensor is given by:

$$T^{\mu\nu} = \frac{1}{2}\sum m_n \int d\tau_n \delta^4(x - x_n)(\dot{x}_n^\mu \dot{x}_n^\nu - V\{x_n^\mu\})$$

The stress-energy tensor is simply related the Lagrangian of a point particle moving on a manifold described by the $g_{\mu\nu}$.

Comparing equations (15) and (16), the combination $\delta^4(x-x_n)V\{x_n^\mu\}/\sqrt{-g}$ behaves like a cosmologic constant, Λ, if we ignore (the usually dominant) kinetic and mass energy density term in the matter action. Dividing by $\sqrt{-g}$ is necessary so that Λ appears multiplying $\sqrt{-g}$ and is thus coupled to spacetime. ($\delta^4(x)/\sqrt{-g}$ is the invariant delta function.) Therefore, such a potential energy term—or its expectation value in a quantum mechanical setting—would contribute a cosmologic constant if the time derivative terms in (16) were for some reason negligible.

Considering the nonrelativistic case first, the virial theorem keeps $\langle T\rangle - \langle V\rangle = 0$ where T is the nonrelativistic kinetic energy. The rest mass energy of particulate matter then, typically, keeps the potential energy term from dominating the stress-energy tensor. For this reason, "regular" matter cannot produce a cosmologic constant.

To examine how a phase transition in a quantum field theory may generate a cosmologic constant, we restate the Higgs mechanism in the easier setting of a superconductor. The non-relativistic Euclidean action for fermions with an attractive interaction is given by:

$$S = \int_0^\beta d\tau \int dx \left(\bar{\psi}_\alpha(\partial_\tau - \partial_x^2)\psi_\alpha - u\bar{\psi}_\alpha\bar{\psi}_\beta\psi_\beta\psi_\alpha\right) \tag{17}$$

where $\alpha, \beta = \uparrow, \downarrow$. Decoupling the four fermion interaction with a complex order parameter field, $\Delta \leftrightarrow \psi_\alpha\psi_\beta$ and rewriting ψ as a spinor $\psi = (\psi_\uparrow, \psi_\downarrow)$, the action becomes

$$S = \int_0^\beta d\tau \int dx \left(\bar{\psi}(\underbrace{\partial_\tau - \sigma_3\partial_x^2}_{-g_0^{-1}} + \sigma_1\Delta)\psi + \frac{1}{u}\Delta^2\right) \tag{18}$$

and the functional integration is now over Δ also: $Z = \int D\psi D\bar{\psi}D\Delta \exp -S$. g_0 is identified as the free fermi field propagator. The fermions may be integrated out to write the effective action in terms of Δ. By restricting the phase to zero, we will only access the amplitude mode. Fluctuations in the amplitude are what what particle physicists call the Higgs boson. Now the effective action is expanded in the field amplitude to quartic order—which produces the necessary symmetry change for a phase transition:

$$S = -\mathrm{tr}\log(-g_0^{-1} + \sigma_1\Delta) + \mathrm{tr}\frac{1}{u}\Delta^2 \tag{19}$$

$$= \quad -\frac{1}{2}\text{tr}g_0\Delta g_0\Delta + \frac{1}{u}\text{tr}\Delta^2 - \frac{1}{4}\text{tr}(g_0\Delta g_0\Delta)^2 + O(\Delta^6) \tag{20}$$

The terms may be reassembled into a *relativistic* Minkowski dynamical action for Δ:

$$\int d^2x \left(-g^{\mu\nu}\partial_\mu\Delta\partial_\nu\Delta - a\Delta^2 - b(\Delta^4 + \Delta_0^4) \right) \tag{21}$$

where a changes sign at the transition. Below the transition, the order parameter field acquires an expectation value, $\Delta = \Delta_0$. Particle physicists add the last term to force the potential energy in the ground state of the broken symmetry phase to be zero. As mentioned before, ordinary matter cannot have a cosmologic constant because its rest mass energy dominates the stress-energy tensor. The way a phase transition in QFT solves this problem is that if the transition is passed through quickly, the system will remain temporarily in the symmetric state $\Delta = 0$. Now, because of the counterterm, the expectation value of the potential energy is *positive*—a cosmologic constant, $\Lambda > 0$. Fluctuations in $\Delta(x)$ which contribute to the kinetic energy (the $g_{\mu\nu}$ term in the action) continue to grow small as the universe expands. Eventually, the positive potential energy will dominate and a de Sitter condition is reached.

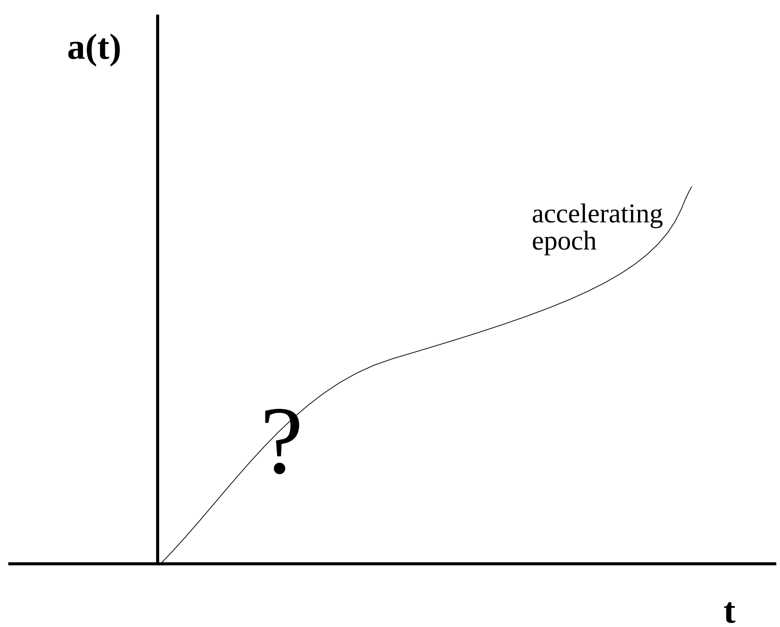
(Side Note: If one studies small fluctuations in the amplitude in the broken symmetry state to quadratic order (i.e the Higgs boson), it is seen that a is proportional to Δ_0. It is interesting to note that the Higgs boson cannot be seen in a SC because its minimum excitation energy lies above the SC gap (i.e. it has a mass proportional to the SC gap). The Higgs excitation, therefore, interacts strongly with quasiparticles which must have an energy larger than the gap energy (these are "normal state" electrons that comprise the broken superconducting Cooper pairs of the condensate) and are damped by them.)

3.5 The accelerating universe

Observations of the Hubble velocity of very distant supervovae have provided measurements of the change in the scale factor, $a(t)$, extending back from our present time, t_0, to much earlier times. The expectation (other than of a few visionaries) had been that one would observe an increasing Hubble parameter $\dot{a}/a$ as one looked at more distant objects. (It may be useful to carefully recall the discussion of null vectors and cosmologic redshift given in section 2.1 to convince yourself that looking at distant objects in fact corresponds to probing a at earlier times. Looking at the relative velocity of distant objects is probing the slope of the curve in figure 5 at earlier times.)

To the great surprise of almost everyone, the Hubble parameter *decreased* rather than increased for these distant supernovae. The universe is not slowing down and expanding, it is *accelerating* and expanding! The data is presnted in the figure I at the end of this document. Unfortunately, it is not a graph of scale factor versus time but rather a graph of observed quantities: redshift, z, and luminosity distance, d_l (this is the distance that is inferred naively from luminosity without taking into account FRW geometry.) To relate these two quantities within an FRW model we have to use the equation given earlier for d_l (equation (9)).

The upturn at large z then corresponds to positive $\ddot{a}$ (or negative $q(t_0)$; I guess everyone expected the acceleration to be negative so they defined q with a minus sign in front.) Roughly speaking, an upturn at large z corresponds in a plot of a to a feature like that depicted in figure 7.



The acceleration has forced cosmologists to include $\Lambda \neq 0$ in the Friedmann equations (10,11). With the CMB strongly indicating a flat universe ($\kappa = 0$), the first Friedmann equation may be written in form like equation

(13):

$$\dot{a}^2 - \frac{GM}{a} = \frac{1}{3}\Lambda a^2 \qquad (22)$$

(noting that the local conservation of "energy" equation (12) still holds.)

At large times, the a^2 term will dominate giving an accelerating, "runaway" solution. The data may be compared to solutions of equation (22) for different M and Λ as in figure I. The values of Λ are typically stated in terms of the $\Omega(t) \equiv \rho(t)/\rho_c(t)$ ratio defined earlier. Supernova data constrain the value of Ω_Λ to about 0.6-0.75 along the "flat universe line" of figure II (this is implicitly using the CMB results to conclude $\kappa = 0$). This leaves the matter portion of Ω constrained to 0.25-0.4. The value of Ω_M 0.4 is remarkably consistent with the careful, inertial surveys of matter which always yielded results substantially below what was needed to make a flat universe.

Thus the cosmologic data is reasonably consistent with FRW cosmology that includes a cosmologic constant. However, the source of $\Lambda \neq 0$ is still deeply mysterious. Several highly imaginative proposals have been forwarded to explain it—some are even testable! Simply imposing $\Lambda \neq 0$ presently seems awkward in that it forces the final (post-inflation) vacuum to be, once again, metastable. Furthermore, $\Lambda \neq 0$ introduces yet another fine tuning problem analogous to the flatness problem which was a primary motivation of cosmic inflation.