Lecture XXII: The homogeneous, isotropic Universe

(Dated: November 20, 2019)

I. OVERVIEW

We now turn our attention to *cosmology* – the study of the origin, structure, and evolution of the Universe. We will focus here on cosmological models that are expanding, homogeneous, and isotropic, as observations suggest these are the most relevant to our Universe. I won't do cosmological perturbations in this class.

The homogeneous, isotropic Universe leads to the *Friedmann-Lemaître-Robertson-Walker* (FLRW) metric, and the Einstein equations therein lead to the *Friedmann equations* that govern the expansion history of the Universe and its relation to the matter content. We'll review those equations here, and consider their phenomenology in the next (and last) lecture.

II. THE POSSIBLE METRICS

A. Expanding Universes

Let's first consider an expanding Universe, threaded by a set of *comoving observers*, each of whom observes the same evolutionary history of the Universe and sees the same thing in all directions. (Not all observers can be comoving: for example, we see the Universe as approximately isotropic on large scales, but someone in a rocketship traveling at $0.8c$ with respect to the Milky Way does not. In what follows, I will assume that the Universe is actually evolving so there is a unique comoving observer at each event. The un-evolving cases – Minkowski, de Sitter, and anti-de Sitter – turn out to be limiting cases of FLRW.) Let us take the spatial coordinates x^i to denote which comoving observer passes through an event $\mathcal{P}$, and the time coordinate t to be that comoving observer's proper time. We take t to be "synchronized" in the sense that two events have the same t if the comoving observers at those events see the Universe in the same evolutionary state.

The metric is then:

$$ds^2 = g_{00}\mathrm{d}t^2 + 2g_{0i}\mathrm{d}t\,\mathrm{d}x^i + g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j, \tag{1}$$

and the comoving observer's 4-velocity is $u^\alpha \to (1,0,0,0)$. We see that the normalization of the 4-velocity (identifying t with the comoving observer's proper time) sets $g_{00} = -1$. Furthermore, since t denotes an evolutionary stage of the Universe and hence is physically measurable, the gradient $\boldsymbol{k} = \boldsymbol{\nabla} t$ is physically measurable. Since we presume that the comoving observer sees an isotropic Universe, $\boldsymbol{k}$ cannot have any spatial component in their rest frame and must be a scalar times $\boldsymbol{u}$. That is,

$$\boldsymbol{k} = \boldsymbol{\nabla} t = b\boldsymbol{u} \quad\to\quad \nabla_\mu t = b g_{\mu\nu} u^\nu \quad\to\quad (1,0,0,0) = (-b, bg_{01}, bg_{02}, bg_{03}). \tag{2}$$

This means $g_{0i} = 0$, and we are left to conclude that the metric has the form

$$ds^2 = -\mathrm{d}t^2 + g_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j. \tag{3}$$

We now want to think about the spatial part of the metric. We begin by considering the velocity gradient $\nabla_\mu u_\nu$, which is a tensor and which must appear isotropic in the observer's frame. Since the observer has no preferred spatial direction, this must be of the form

$$\nabla_\mu u_\nu = H(g_{\mu\nu} + u_\mu u_\nu) + I u_\mu u_\nu, \tag{4}$$

where H and I are scalars. (Think about this in a local orthonormal frame: the first term is the 3×3 identity in the spatial directions, and the second term is the time-time component.) Now since $u_\alpha \to (-1,0,0,0)$, the spatial components of this equation read

$$\nabla_i u_j = H g_{ij}, \tag{5}$$

but then

$$\nabla_i u_j = \partial_i u_j - \Gamma^\alpha{}_{ij} u_\alpha = \Gamma^t{}_{ij}. \tag{6}$$

We therefore must have

$$\Gamma^t{}_{ij} = H g_{ij}. \tag{7}$$

Explicitly writing the Christoffel symbol, and using that $g_{ti} = 0$ and $g^{ti} = 0$:

$$\Gamma^t{}_{ij} = -\frac{1}{2} g^{tt} g_{ij,t} = \frac{1}{2} g_{ij,t}, \tag{8}$$

so the metric obeys the differential equation:

$$\partial_t g_{ij} = 2H g_{ij}. \tag{9}$$

Now H can depend only on t by homogeneity. We define $a = \exp \int H \, \mathrm{d}t$ (the normalization is arbitrary), so that $g_{ij} \propto a^2$, with a coefficient of proportionality that may depend on i, j, and spatial location x^i, but not on t; we may thus write the metric in the form,

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + [a(t)]^2 \gamma_{ij}(x^k) \, \mathrm{d}x^i \, \mathrm{d}x^j. \tag{10}$$

The function $H(t)$ is called the *Hubble constant* or *Hubble parameter*, and $a(t)$ is the *scale factor*. Note that spatial separations between neighboring comoving observers grow $\propto a(t)$. A positive Hubble constant corresponds to increasing scale factor, or equivalently to a positive (expanding) velocity gradient $\nabla_x u_x = \nabla_y u_y = \nabla_z u_z = H$. Since the normalization is arbitrary, standard practice in modern cosmology is to set $a = 1$ at the present epoch (but note that Schutz and many older GR references don't do this). Distances re-scaled to their physical length at $a = 1$ are called *comoving distances*.

B. The possible spatial metrics

We now ask about the possible spatial metrics $\gamma_{ij}(x^k)$. The principle of isotropy tells us that we may choose any one observer and define a spherical coordinate system around them, analogous to what we did for the spherically symmetric star:

$$\gamma_{ij}(x^k) \, \mathrm{d}x^i \, \mathrm{d}x^j = f(r) \, \mathrm{d}r^2 + r^2 (\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2) = \mathrm{d}\chi^2 + [r(\chi)]^2 (\mathrm{d}\theta^2 + \sin^2\theta \, \mathrm{d}\phi^2), \tag{11}$$

where in the second version we changed the radial variable to $\chi = \int \sqrt{f(r)} \, \mathrm{d}r$. The choice of function $f(r)$, or equivalently $r(\chi)$, completely determines the spatial metric. In order for the spatial metric to look locally flat, we will take $f(r) = 0$ at $r = 0$, or $r \approx \chi$ for small χ.

We can assess these spatial metrics by considering the curvature of a 3D slice at $a = 1$. [I will use the pre-superscript $^{(3)}$ to denote 3D quantities.] The spatial metric has non-zero Christoffel symbols:

$$^{(3)}\Gamma^\chi{}_{\theta\theta} = -rr', \quad ^{(3)}\Gamma^\chi{}_{\phi\phi} = -rr' \sin^2\theta, \quad ^{(3)}\Gamma^\theta{}_{\chi\theta} = {}^{(3)}\Gamma^\theta{}_{\theta\chi} = \frac{r'}{r}, \quad ^{(3)}\Gamma^\phi{}_{\chi\phi} = {}^{(3)}\Gamma^\phi{}_{\phi\chi} = \frac{r'}{r},$$

$$^{(3)}\Gamma^\theta{}_{\phi\phi} = -\sin\theta\cos\theta, \quad \text{and} \quad ^{(3)}\Gamma^\phi{}_{\theta\phi} = {}^{(3)}\Gamma^\phi{}_{\phi\theta} = \cot\theta, \tag{12}$$

where the prime denotes a derivative with respect to χ. Then we may compute the Ricci tensor:

$$^{(3)}R_{ij} = {}^{(3)}\Gamma^k{}_{ij,k} - {}^{(3)}\Gamma^k{}_{ik,j} + {}^{(3)}\Gamma^k{}_{lk} {}^{(3)}\Gamma^l{}_{ij} - {}^{(3)}\Gamma^k{}_{lj} {}^{(3)}\Gamma^l{}_{ik}, \tag{13}$$

which evaluates to:

$$^{(3)}R_{\chi\chi} = -2\left(\frac{r'}{r}\right)' - 2\left(\frac{r'}{r}\right)^2, \quad ^{(3)}R_{\theta\theta} = -(rr')' + 1, \quad ^{(3)}R_{\phi\phi} = [-(rr')' + 1] \sin^2\theta, \tag{14}$$

and with zero off-diagonal terms. Now for the metric to be isotropic, the Ricci tensor must be a multiple of the metric tensor. That is,

$$^{(3)}R_{ij} = 2K \gamma_{ij}, \tag{15}$$

where K is a constant. (The factor of 2 is a convention.) That is,

$$-2\left(\frac{r'}{r}\right)' - 2\left(\frac{r'}{r}\right)^2 = 2K \quad \text{and} \quad -(rr')' + 1 = 2Kr^2. \tag{16}$$

The left-hand side of the first equation simplifies to $-2r''/r$, so the overall equations simplifies to $r'' = -Kr$. With the "initial condition" $r = 0$ and $r' = 1$ at $\chi = 0$, we find

$$r = \begin{cases} K^{-1/2}\sin(K^{1/2}\chi) & K > 0 \\ \chi & K = 0 \\ (-K)^{-1/2}\sinh[(-K)^{1/2}\chi] & K < 0 \end{cases}. \tag{17}$$

Inspection shows the second equation is automatically satisfied. Note that the second two cases are analytic continuations of the first to zero and negative K.

It is readily apparent that the $K = 0$ case is just flat 3D space in spherical coordinates. A universe with $K = 0$ is thus said to be *spatially flat*. (Remember the full 4D manifold may be curved, even though the individual 3D slices are flat!)

For the $K > 0$ case (*closed* or *spherical* universe), the 3D slice is a hypersphere of radius $K^{-1/2}$, with $K^{1/2}\chi$ representing the hyper-colatitude (angle from the North Pole: each dimension we add requires a new angle). The coordinate system has a singularity at $\chi = K^{-1/2}\pi$, which is the South Pole. The closed universe thus has a finite comoving (i.e., at $a = 1$) volume: this is

$$V_c = \int_0^{K^{-1/2}\pi}\int_0^{\pi}\int_0^{2\pi}[K^{-1/2}\sin(K^{1/2}\chi)]^2\sin\theta\,\mathrm{d}\chi\,\mathrm{d}\theta\,\mathrm{d}\phi = 2\pi^2K^{-3/2}. \tag{18}$$

Alternatively, if $K > 0$, one may have a spatial geometry called the *projective sphere* where opposite points on the sphere are identified (i.e., treated as equivalent). Then the maximum value of χ is $K^{-1/2}\pi/2$, i.e., the farthest one can get from the North Pole is the Equator. In the projective sphere, if you walk south past the Equator, you appear in the Northern Hemisphere again on the opposite side. Both the spherical and projective sphere universe models are legal, and are both homogeneous and isotropic, and locally look the same; but they have different global topologies. The comoving volume of the projective sphere is $\pi^2K^{-3/2}$.

If $K < 0$ (*open* or *hyperbolic* universe), then the universe again has infinite comoving volume. Indeed, the volume integral increases exponentially with χ.

This leads us to the overall 4D FLRW metric:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + [a(t)]^2\left\{\mathrm{d}\chi^2 + [r(\chi)]^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right\}. \tag{19}$$

III. THE FRIEDMANN EQUATIONS

We can now consider the dynamical equations that govern the evolution of an FLRW universe. The first thing we want to do is get the Einstein tensor, so that we can write down the Einstein equations. We will use primes to denote ∂_χ and dots to denote ∂_t. We perform our calculations in the usual way: we start with the Christoffel symbols:

$$\Gamma^{\chi}{}_{\theta\theta} = -rr', \quad \Gamma^{\chi}{}_{\phi\phi} = -rr'\sin^2\theta, \quad \Gamma^{\theta}{}_{\chi\theta} = \Gamma^{\theta}{}_{\theta\chi} = \frac{r'}{r}, \quad \Gamma^{\phi}{}_{\chi\phi} = \Gamma^{\phi}{}_{\phi\chi} = \frac{r'}{r},$$

$$\Gamma^{\theta}{}_{\phi\phi} = -\sin\theta\cos\theta, \quad \Gamma^{\phi}{}_{\theta\phi} = \Gamma^{\phi}{}_{\phi\theta} = \cot\theta, \quad \Gamma^{t}{}_{\chi\chi} = a\dot{a}, \quad \Gamma^{t}{}_{\theta\theta} = a\dot{a}r^2, \quad \Gamma^{t}{}_{\phi\phi} = a\dot{a}r^2\sin^2\theta,$$

$$\Gamma^{\chi}{}_{t\chi} = \Gamma^{\chi}{}_{\chi t} = \frac{\dot{a}}{a}, \quad \Gamma^{\theta}{}_{t\theta} = \Gamma^{\theta}{}_{\theta t} = \frac{\dot{a}}{a}, \quad \text{and} \quad \Gamma^{\phi}{}_{t\phi} = \Gamma^{\phi}{}_{\phi t} = \frac{\dot{a}}{a}. \tag{20}$$

The Ricci tensor components are then:

$$R_{tt} = -3\partial_t\frac{\dot{a}}{a} - 3\left(\frac{\dot{a}}{a}\right)^2, \quad R_{\chi\chi} = \partial_t(a\dot{a}) + 2K + \dot{a}^2, \quad R_{\theta\theta} = r^2R_{\chi\chi}, \quad \text{and} \quad R_{\phi\phi} = r^2\sin^2\theta\,R_{\chi\chi}. \tag{21}$$

(The off-diagonal components are zero, and we simplified the radial derivatives by substituting for K.) These components simplify to:

$$R_{tt} = -3\frac{\ddot{a}}{a} \quad \text{and} \quad R_{\chi\chi} = a\ddot{a} + 2\dot{a}^2 + 2K. \tag{22}$$

The Ricci scalar is:

$$R = -R_{tt} + \frac{3}{a^2}R_{\chi\chi} = 6\frac{\ddot{a}}{a} + 6\frac{\dot{a}^2}{a^2} + \frac{6K}{a^2}. \tag{23}$$

This results in the Einstein tensor components:

$$G_{tt} = 3\frac{\dot{a}^2}{a^2} + \frac{3K}{a^2}, \quad G_{\chi\chi} = \frac{G_{\theta\theta}}{r^2} = \frac{G_{\phi\phi}}{r^2\sin^2\theta} = -2a\ddot{a} - \dot{a}^2 - K, \tag{24}$$

and the others are zero.

Now we can compare this to the stress-energy tensor:

$$T_{tt} = \rho, \quad T_{\chi\chi} = p, \quad T_{\theta\theta} = r^2 p, \quad \text{and} \quad T_{\phi\phi} = r^2\sin^2\theta\, p \tag{25}$$

(the other components are zero), and use the Einstein equation:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = 8\pi T_{\mu\nu}. \tag{26}$$

(We include the cosmological constant, since it appears to be important to the real Universe.) This leads to two independent equations – one for the tt part of the Einstein equation and one for the spatial part (the $\chi\chi$ component, divided by a^2):

$$3\frac{\dot{a}^2}{a^2} + \frac{3K}{a^2} - \Lambda = 8\pi\rho \quad \text{and} \quad -2\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2} - \frac{K}{a^2} + \Lambda = 8\pi p. \tag{27}$$

These two equations are usually written in a slightly different form. The first equation is usually written in terms of $H = \dot{a}/a$, and is known as the *first-order Friedmann equation*:

$$3H^2 = 8\pi\rho + \Lambda - \frac{3K}{a^2}. \tag{28}$$

We may also add $-\frac{1}{2}$ times the ρ equation plus $-\frac{3}{2}$ times the p equation to get the *second-order Friedmann equation*:

$$3\frac{\ddot{a}}{a} = -4\pi(\rho + 3p) + \Lambda. \tag{29}$$

A. Continuity equation

The first and second-order Friedmann equations imply some relations between the evolution of density and pressure. If we take the time derivative of the first-order Friedmann equation, and divide by 2, we get:

$$3H\dot{H} = 4\pi\dot{\rho} + \frac{3HK}{a^2}. \tag{30}$$

However, the second-order equation written in terms of $\dot{H} = (a\ddot{a} - \dot{a}^2)/a^2$ gives

$$3\dot{H} + 3H^2 = -4\pi(\rho + 3p) + \Lambda \tag{31}$$

or – substituting in the first-order Friedmann equation again –

$$3\dot{H} + 8\pi\rho + \Lambda - \frac{3K}{a^2} = -4\pi(\rho + 3p) + \Lambda \quad \rightarrow \quad 3\dot{H} - \frac{3K}{a^2} = -12\pi(\rho + p). \tag{32}$$

Substituting in Eq. (30) gives

$$4\pi\dot{\rho} = H\left(3\dot{H} - \frac{3K}{a^2}\right) = -12\pi H(\rho + p) \tag{33}$$

or

$$\dot{\rho} = -3H(\rho + p). \tag{34}$$

This is the continuity equation, and is the time component of $\nabla_\mu T^{\mu\nu} = 0$. As always in general relativity, the dynamical equations imply the continuity equation. When numerically solving for cosmic expansion, the first-order Friedmann equation and the continuity equation are the usual starting points; the second-order Friedmann equation, while conceptually important, is not as convenient (and it contains no new information).

Since $3H = \dot{V}/V$, where V is the physical volume of the Universe, the continuity equation can be re-cast as something like the first law of thermodynamics as the Universe expands:

$$\partial_t(\rho V) = -p\dot{V}. \tag{35}$$

In particular, pressureless matter will have its energy density decay as $\rho \propto V^{-1} \propto 1/a^3$. But other forms of matter might have other scaling behaviors as the Universe expands.

B. Deceleration vs. acceleration

One of the most impressive consequences of the second-order Friedmann equation is that $\ddot{a}$ can have either sign. We can see that energy density $\rho > 0$ gives a negative contribution to the right-hand side of Eq. (29). Pressure also appears on the right-hand side, but pressure can be either positive or negative; positive pressure (e.g., radiation) slows down the expansion of the Universe, but negative pressure (e.g., a potential-dominated scalar field) accelerates the expansion. The cosmological constant has a similar effect: $\Lambda > 0$ accelerates the expansion. Whether the observed cosmic acceleration is really due to a cosmological constant or due to negative pressure in some new dynamical field is a major open question in cosmology.