Lecture XXIX: The FRW metric

Christopher M. Hirata
*Caltech M/C 350-17, Pasadena CA 91125, USA**
(Dated: April 4, 2012)

I. OVERVIEW

In this lecture, we will consider the general model of a homogeneous, isotropic universe. Only the metric structure is considered here, while the dynamics are described in the next lecture.
Reading:

- MTW Ch. 27.

II. HOMOGENEITY & ISOTROPY

The basic assumption is that the Universe on large scales is homogeneous and isotropic. This is the *cosmological principle*.
Why make this assumption? Initally this was a philosophical principle ("we are not special"), combined with the desire to restrict the form of possible solutions to something we could analyze. But the homogeneous, isotropic models are in good agreement with modern data. For example:

- The cosmic microwave background is isotropic to a few parts in 10^5 (aside from the dipole due to our motion).
- Cosmologically distant objects (quasars, GRBs) appear isotropic to a few parts in 10^2.

Testing homogeneity is harder than testing for the isotropy around a particular observer (us). Of course, if every observer sees an isotropic universe, then the universe is homogeneous, but how can we be sure that we don't live at the center of a spherically symmetric universe? It turns out that there actually are tests of this, involving Thomson scattering of the CMB off of distant electrons. If the Universe were inhomogeneous, one would expect that a distant electron would see a different CMB temperature in different directions. Then after scattering, the CMB sky we see would contain a mixture of different temperature blackbodies superposed by scattering, e.g. we would see

$$I_\nu = \sum_i w_i \frac{2h\nu^3}{c^2(e^{h\nu/kT_i} - 1)}, \tag{1}$$

where $\sum_i w_i = 1$. This mixture of blackbodies is **not** a blackbody spectrum. Since the COBE/FIRAS experiment observes a blackbody to one part in $\sim 10^4$, gross deviations from cosmic homogeneity are not acceptable [1]. (A possible exception would be to imagine a universe in which the original CMB spectrum was narrower than a blackbody and was smeared into a blackbody by electron scattering; in this case the blackbody spectrum must be a striking coincidence. Such a suggestion causes other problems with early Universe physics, and besides we won't consider anything this contrived.)
How shall we formalize our notions of homogeneity and isotropy in GR? Really what we are saying is the following:

- The spacetime $\mathcal{M}$ can be sliced into spacelike surfaces Σ_t of constant cosmic time t.
- The surfaces are homogeneous and isotropic: given any two points $\mathcal{O}, \mathcal{P} \in \Sigma_t$ (at the same t), and given any spacelike unit vectors $\boldsymbol{a} \in T_{\mathcal{O}}\Sigma_t$ and $\boldsymbol{b} \in T_{\mathcal{P}}\Sigma_t$ tangent to Σ_t, there is an isometry (symmetry operation of the spacetime) that takes $\mathcal{O}$ and $\boldsymbol{a}$ to $\mathcal{P}$ and $\boldsymbol{b}$, and fixes the $\Sigma_{t'}$ (i.e. maps $\Sigma_{t'} \to \Sigma_{t'}\ \forall t'$).

In most cases, the "slicing into surfaces of constant time" is unique, but we will encounter a few examples where this is not the case.

*Electronic address: `chirata@tapir.caltech.edu`

III. THE METRIC STRUCTURE – OVERALL FORM

We now consider the implications of homogeneity and isotropy for the metric structure. We begin by defining the forward-pointing unit normal $\boldsymbol{n}$ to the surfaces Σ_t, which is a timelike vector: $\boldsymbol{n}\cdot\boldsymbol{n} = -1$. We define a set of *comoving observers* whose 4-velocity is $\boldsymbol{n}$ and whose trajectories are the integral curves of $\boldsymbol{n}$. We may then define a coordinate system using $x^0 = t$ as the time coordinate, and (x^1, x^2, x^3) labeling which comoving observer passes through a particular event.

We now begin placing restrictions on the form of the metric. A vector $\boldsymbol{a}$ is tangent to Σ_t if and only if its contravariant t-component vanishes: $a^0 = 0$. Also since the comoving observers' 3 spatial coordinates are fixed, their 4-velocities have $n^i = 0$, and n^0 is the only nonzero component. But since $\boldsymbol{n}$ is the surface normal, we have $n^i a_i = 0$ if $\boldsymbol{a}$ is tangent to Σ_t. It thus follows that vectors tangent to Σ_t have $a_0 = 0$, and hence

$$0 = a_0 = g_{0i}a^i + g_{00}a^0 = g_{0i}a^i \tag{2}$$

for any a^i. This implies that in our coordinate system:

$$g_{0i} = 0. \tag{3}$$

We next consider two points $\mathcal{O}$ and $\mathcal{P}$ on the same spacelike slice Σ_t. Then there is an isometry mapping $\mathcal{O}$ to $\mathcal{P}$ that preserves the slices, so it follows that $\nabla t\cdot\nabla t$ is the same at both $\mathcal{O}$ and $\mathcal{P}$. But in the above coordinate system we have $\nabla_\mu t = (1, 0, 0, 0)$ so

$$\nabla t\cdot\nabla t = g^{00} = \frac{1}{g_{00}}. \tag{4}$$

Since this is the same at $\mathcal{O}$ and $\mathcal{P}$, we conclude that g_{00} can depend only on t and not on (x^1, x^2, x^3). We also note that since $n^\mu = (n^0, 0, 0, 0)$ is timelike, $g_{00} < 0$. We may then re-label the spacelike surfaces with a new time coordinate t' such that

$$t' = \int \sqrt{-g_{00}}\, dt. \tag{5}$$

This new time coordinate is defined up to an additive constant. With the new relabeling, we have $g_{00} = -1$, and we will take this choice from here on. The metric is now reduced to the form:

$$ds^2 = -dt^2 + g_{ij}\, dx^i\, dx^j. \tag{6}$$

We are next interested in how g_{ij} depends on t, x^1, x^2, x^3. We do this by taking any spacelike unit vector $\boldsymbol{a} \in T_{\mathcal{O}}\Sigma_t$ (with $\boldsymbol{a}\cdot\boldsymbol{a} = 1$) and defining the local velocity gradient:

$$H(t, x^i, \boldsymbol{a}) \equiv a^\mu a_\nu \nabla_\mu n^\nu. \tag{7}$$

So far H is just a name but we will discover its importance soon. Note that it is allowed to depend only on t, and not on x^i or a^i. Noting that $a_0 = a^0 = 0$, and that $n^\mu = (1, 0, 0, 0)$, we see that

$$H(t) = a^i a_j \Gamma^j_{i0}. \tag{8}$$

If we use the expansion of the Christoffel symbol and the form of the metric ($g_{0i} = 0$), we find that

$$\Gamma^j_{i0} = \frac{1}{2} g^{jk}\frac{\partial g_{ik}}{\partial t}, \tag{9}$$

so that

$$H(t) = \frac{1}{2} a^i a_j g^{jk}\frac{\partial g_{ik}}{\partial t} = \frac{1}{2} a^i a^k \frac{\partial g_{ik}}{\partial t}. \tag{10}$$

Now since this is independent of the direction of $\boldsymbol{a}$, it follows that $\partial g_{ik}/\partial t$ must be some scalar multiple of g_{ik}:

$$\frac{\partial g_{ik}}{\partial t} = 2H g_{ik}. \tag{11}$$

We then define the *scale factor* $a(t)$ via the relation:

$$a(t) = \exp \int H(t)\,dt, \tag{12}$$

which is defined up to an overall multiplicative constant. We may then write

$$g_{ik}(t, x^j) = \gamma_{ik}(x^j) e^{2\int H(t)\,dt} = \gamma_{ik}(x^j)[a(t)]^2. \tag{13}$$

The metric then takes the form

$$ds^2 = -dt^2 + [a(t)]^2 \gamma_{ij}(x^k)\,dx^i\,dx^j, \tag{14}$$

where γ_{ij} is the spatial metric of some 3-dimensional manifold (the slice Σ_t corresponding to the t where the scale factor is 1). This 3-dimensional spatial manifold must be homogeneous and isotropic. This metric is the *Friedmann-Robertson-Walker metric*, and $H(t)$ is called the *Hubble rate*.

Normally in cosmology we will define the scale factor to be $a = 1$ at the present, but other conventions are in use.

IV. THE POSSIBLE SPATIAL GEOMETRIES

Our next task is to solve for the possible spatial geometries $\gamma_{ij}(x^k)$. These must be isotropic around any choice of origin $\mathcal{O}$; therefore we may write them in spherical polar coordinates (the proof of this is similar to our study for spherical stars):

$$ds_3^2 = d\chi^2 + f(\chi)(d\theta^2 + \sin^2\theta\,d\phi^2). \tag{15}$$

We must then find which functions $f(\chi)$ lead to a homogeneous isotropic 3-manifold. A simple way to begin this is to consider the 3-dimensional Ricci tensor associated with Eq. (15). The Christoffel symbols are:

$$\begin{aligned}
\Gamma^\chi_{\theta\theta} &= -\frac{1}{2}f', \\
\Gamma^\chi_{\phi\phi} &= -\frac{1}{2}f'\sin^2\theta, \\
\Gamma^\theta_{\phi\phi} &= -\sin\theta\cos\theta, \\
\Gamma^\theta_{\chi\theta} = \Gamma^\phi_{\chi\phi} &= \frac{f'}{2f}, \quad \text{and} \\
\Gamma^\phi_{\theta\phi} &= \cot\theta,
\end{aligned} \tag{16}$$

where the prime $'$ denotes $d/d\chi$. The sums are $\Gamma^j_{j\chi} = f'/f$, $\Gamma^j_{j\theta} = \cot\theta$, and $\Gamma^j_{j\phi} = 0$. The nontrivial parts of the Ricci tensor are then

$$\begin{aligned}
R_{\chi\chi} &= \left(\frac{f'}{f}\right)' - \frac{f'^2}{2f^2}, \\
R_{\theta\theta} &= -\frac{1}{2}f'' + \csc^2\theta - \frac{f'^2}{2f} + \frac{f'^2}{2f} - \cot^2\theta = -\frac{1}{2}f'' + 1, \quad \text{and} \\
R_{\phi\phi} &= -\frac{1}{2}f''\sin^2\theta + \sin^2\theta - \cos^2\theta - \frac{f'^2}{2f}\sin^2\theta - \cos^2\theta + \frac{f'^2}{2f}\sin^2\theta + 2\cos^2\theta = \left(-\frac{1}{2}f'' + 1\right)\sin^2\theta.
\end{aligned} \tag{17}$$

In order for the 3-manifold to be spatially homogeneous and isotropic, it is necessary (though may not be sufficient) for $R_{ij} = 2Kg_{ij}$, where K is a constant (independent of χ). That is, we need

$$2K = \left(\frac{f'}{f}\right)' - \frac{f'^2}{2f^2} = \frac{-\frac{1}{2}f'' + 1}{f}. \tag{18}$$

[The factor of 2 is arbitrary and introduced for consistency later.] In our search for spatially homogeneous solutions, we use the first and last expressions in Eq. (18) to find

$$f'' = -4Kf + 2, \tag{19}$$

which is a simple linear ODE. If $K \neq 0$, it admits solutions

$$f = \frac{1}{2K} + A\cos(2K^{1/2}\chi) + B\sin(2K^{1/2}\chi), \tag{20}$$

where A and B are constants of integration. Regularity at $\mathcal{O}$, however, requires that $f = f' = 0$ at $\chi = 0$. This forces on us the choice $A = -1/(2K)$ and $B = 0$. Then we may simplify to

$$f = \frac{1 - \cos(2K^{1/2}\chi)}{2K} = \frac{1}{K}\sin^2(K^{1/2}\chi). \tag{21}$$

The alternative case is that $K = 0$. In this case, $f'' = 2$ and with the boundary condition at $\mathcal{O}$ we find $f = \chi^2$. This is simply Eq. (21) with the limit $K \to 0$ in accordance with l'Hôpital's rule. It is readily seen that the other equality in Eq. (18) is obeyed.

We are left with the spatial metric

$$ds_3^2 = d\chi^2 + \frac{\sin^2(K^{1/2}\chi)}{K}(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{22}$$

Its homogeneity and isotropy are easily checked: this is the metric for a 3-sphere of radius $K^{-1/2}$. In the limit $K \to 0$, it becomes flat 3-dimensional Euclidean space in spherical coordinates. If the spatial metric of the Universe has $K > 0$, then we say that it is *closed*, and if $K = 0$ we say that it is *spatially flat*. The spatially flat 3-metric is

$$ds_3^2 = d\chi^2 + \chi^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{23}$$

An interesting property of this metric is that it is possible to have $K < 0$. In this case, the Universe is *open* and we take the analytic continuation of the sine function:

$$ds_3^2 = d\chi^2 + \frac{\sinh^2[(-K^{1/2})\chi]}{-K}(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{24}$$

In short, the spatial geometries have the following properties:

- **Spatially flat**, $\mathbb{R}^3$. This is standard Euclidean space, with infinite volume and the usual rules of geometry. Our own universe appears to be spatially flat to a very good approximation.

- **Closed**, S^3. This space has the geometry of a 3-sphere, and the usual coordinate system corresponds to hyperspherical coordinates. The maximum distance from the origin is the antipodal point at $\chi = \pi K^{-1/2}$, where there is a coordinate singularity. The volume of the space is finite: $V_3 = 2\pi^2 K^{-3/2}$. The closed space exhibits non-Euclidean features: the interior angles of a triangle add to $> \pi$, they Pythagorean theorem reads $a^2 + b^2 > c^2$, etc.

- **Open**, $\mathbb{H}^3$. This is a "hyperbolic space," with the same topology as Euclidean space. However it exhibits non-Euclidean features: the interior angles of a triangle add to $< \pi$, they Pythagorean theorem reads $a^2 + b^2 < c^2$, etc. Not only is the volume of an open space infinite, but it is exponentially infinite in the sense that the volume of a sphere of radius r is $V_3 = 2\pi(-K)^{-3/2}[\sinh(\sqrt{K}r) - \sqrt{K}r]$, which increases exponentially with radius (instead of as $\propto r^3$).

- **Projective space**, $\mathbb{R}P^3$. This is an alternative topology of the closed universe, in which antipodal points are identified: $(\chi, \theta, \phi) = (\pi K^{-1/2} - \chi, \pi - \theta, \pi + \phi)$. Locally it looks like a closed universe, but has only half the volume: $V_3 = \pi^2 K^{-3/2}$. The unique region of the projective space has $\chi < \pi/(2\sqrt{K})$, i.e. is the region between the North Pole (origin) and equator; if one passes the equator one reappears on the opposite side of the sphere. This is the only nontrivial topology of any of the spaces we have considered that is globally homogeneous and isotropic. [2]

---

[1] See e.g. Goodman, PRD 52:1821 (1995) for details.

[2] Other possibilities, such as toroidal compactifications of $\mathbb{R}^3$ in which one identifies points $(x^1, x^2, x^3) = (x^1 + m_1 L, x^2 + m_2 L, x^3 + m_3 L)$ with $m_i \in \mathbb{Z}$, are not globally isotropic.