3 The Friedmann-Robertson-Walker metric

3.1 Three dimensions

The most general isotropic and homogeneous metric in three dimensions is similar to the two dimensional result of eq. (43):

$$ds^2 = a^2 \left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2 \right) , \qquad d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2 , \qquad k = 0, \pm 1 . \tag{46}$$

The angles ϕ and θ are the usual azimuthal and polar angles of spherical coordinates, with $\theta \in [0,\pi]$, $\phi \in [0, 2\pi)$. As before, the parameter k can take on three different values: for $k = 0$, the above line element describes ordinary flat space in spherical coordinates; $k = 1$ yields the metric for S_3, with constant positive curvature, while $k = -1$ is AdS_3 and has constant negative curvature. As in the two dimensional case, the change of variables $r = \sin\chi$ ($k = 1$) or $r = \sinh\chi$ ($k = -1$) makes the global nature of these manifolds more apparent. For example, for the $k = 1$ case, after defining $r = \sin\chi$, the line element becomes

$$ds^2 = a^2 \left(d\chi^2 + \sin^2\chi d\theta^2 + \sin^2\chi \sin^2\theta d\phi^2 \right) . \tag{47}$$

This is equivalent to writing

$$ds^2 = dX^2 + dY^2 + dZ^2 + dW^2 , \tag{48}$$

where

$$\begin{aligned} X &= a \sin\chi \sin\theta \cos\phi , \\ Y &= a \sin\chi \sin\theta \sin\phi , \\ Z &= a \sin\chi \cos\theta , \\ W &= a \cos\chi , \end{aligned} \tag{49}$$

which satisfy $X^2 + Y^2 + Z^2 + W^2 = a^2$. So we see that the $k = 1$ metric corresponds to a 3-sphere of radius a embedded in 4-dimensional Euclidean space. One also sees a problem with the $r = \sin\chi$ coordinate: it does not cover the whole sphere. In going from the north pole ($\chi = 0$) to the equator ($\chi = \pi/2$), the variable r ranges from $r = 0$ to $r = 1$; however, in proceeding from the equator to the south pole of the sphere ($\chi = \pi$), $r \sin\chi$ runs back from $r = 1$ to $r = 0$. So the full space has coordinates which are not single valued in r.

The proper distance $d$between the origin and an object at radial coordinate r is given by

$$d = \int_0^r a \frac{dr'}{\sqrt{1-kr'^2}} dr' = a \times \begin{cases} \sin^{-1} r & k = 1 \\ r & k = 0 \\ \sinh^{-1} & k = -1 \end{cases} \tag{50}$$

For the case $k = 1$ we see that d is not defined for $r > 1$, and that for $r = 1$, $d = a\pi/2$. This is the distance along the sphere from the north pole to the equator; to compute the distance from the north pole to the south pole in these coordinates, you need to double the result, to get $d = a\pi$.

3.2 Four dimensions: The Friedmann-Robertson-Walker metric

It is simple to go to the case of interest: four dimensional spacetime. Because of the homogeneity, we can choose the same time coordinate for each point in space, and at each time slice, we must have the isotropic and homogeneous three dimensional metric eq. (46). However, there is no constraint relating the scale factor a at different time slices, which can therefore be a function of time. Aside from isotropy and homogeneity, general relativity requires that locally (eg, near the origin) the line element be invariant under Lorentz transformations:

$$ds^2 = dt^2 - d\vec{x}^2 \ . \tag{51}$$

Thus we arrive at the Friedmann-Robertson-Walker (FRW) metric, which is the most general metric (up to coordinate transformations) fulfilling the cosmological principle:

$$ds^2 = dt^2 - a(t)^2 \left(\frac{dr^2}{\sqrt{1-kr^2}} + r^2 d\Omega^2 \right) \ . \tag{52}$$

These coordinates $\{t, r, \theta, \phi\}$ are called **co-moving coordinates**. The reason is because two objects at different spatial coordinates can remain at those coordinates at all times, while the proper distance between them changes with time according to how the scale factor $a(t)$ changes with time. Picture to dots on a balloon whose coordinates are fixed, while the radius of the balloon changes with time. That the proper distance will change with time is evident — one needs merely to replace a in the three dimensional example eq. (50) by $a(t)$. However, one needs to show that in fact an object can remain at rest at fixed spatial coordinate. The world line of such an object, parametrized by $\lambda = t$, satisfies

$$\frac{dt}{d\lambda} = 1 \ , \quad \frac{dx^i}{d\lambda} = 0 \ , \tag{53}$$

and we have to show that this is indeed a geodesic for our metric. Since the geodesic equation eq. (23) is

$$0 = \frac{d^2 x^\alpha}{d\lambda^2} + \Gamma^\alpha_{\mu\nu} \frac{dx^\mu}{d\lambda} \frac{dx^\nu}{d\lambda} \ , \tag{54}$$

it follows that the world line eq. (53) satisfies this equation provided that Γ^α_{tt} vanishes. However, that is evident if you think of how the geodesic equation is derived from the metric by the Euler-Lagrange equations: since the coefficient of dt^2 in eq. (52) is coordinate independent, there can be no terms proportional to $(dt/d\lambda)^2$ in the geodesic equation.

3.3 The redshift of light

Now I will demonstrate that light gets red-shifted in the FRW universe due to the time dependent scale factor. Suppose there is a galaxy at rest in the co-moving coordinates, with radial coordinate r_1, and we are at the origin. Light is emitted by this galaxy at time t_1 with frequency ν_1 and received on earth at time t_0 with frequency ν_0. We wish to determine the relation between ν_0 and ν_1.

Light travels on '**null geodesics** — that is, on geodesics with $ds = 0$. Note that in flat space where $ds^2 = dt^2 - d\vec{x}^2$, $ds = 0$ simply means that the speed of light is $c = 1$. However, it is possible to choose locally flat coordinates about any point in curved space, so having light travel on null geodesics means that any observer along its path sees it go by with velocity $c = 1$. Thus along a radial light path in the FRW universe,

$$\frac{dt}{a(t)} = \frac{dr}{\sqrt{1-kr^2}} \ . \tag{55}$$

Now consider the emission of two subsequent crests of a light wave:

crest #1	t	r		crest #2	t	r
emitted:	t_1	r_1		emitted:	$t_1 + \delta t_1$	r_1
received:	t_0	0		received:	$t_0 + \delta t_0$	0

(56)

It follows from eq. (55) that

$$\int_{t_1}^{t_0} \frac{dt}{a(t)} = \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{dt}{a(t)} = \int_{r_1}^{0} \frac{dr}{\sqrt{1-kr^2}} \ . \tag{57}$$

Subtracting the first integral from the second, and assuming $\delta t_{0,1} \ll \frac{a}{\dot{a}}$, we get

$$\frac{\delta t_0}{a(t_0)} = \frac{\delta t_1}{a(t_1)} \tag{58}$$

or, since $\delta t_{0,1} = 1/\nu_{0,1}$, the redshift z is given by

$$1 + z = \frac{\nu_1}{\nu_0} = \frac{a(t_0)}{a(t_1)} \ . \tag{59}$$

Note that the redshift only depends on the ratio of the scale factor at reception to the scale factor at emission. It is *not* simply a function of the relative motion of the source at the time of emission; in fact if the universe were to contract for a while, and then expand, it is possible for a source which is approaching us to emit light which we see as shifted to the red.

It is useful to consider eq. (59) in terms of the wavelength λ instead of the frequency ν:

$$1 + z = \frac{\lambda_0}{\lambda_1} = \frac{a(t_0)}{a(t_1)} \ . \tag{60}$$

We see that the wavelength of light just contracts and stretches with the scale factor, and that fact explains the whole redshift phenomenon. In this view it is not much like the Doppler shift described in §1.2 in the context of special relativity.

3.4 Redshift of a thermal spectrum

Suppose we view a thermal source with temperature T_1 with absolute luminosity L. The power emitted per unit area per unit frequency is given by the Planck formula

$$S_\nu = \frac{2\pi h}{c^2} \frac{\nu_1^3}{e^{h\nu_1/kT_1} - 1}. \tag{61}$$

Integrating S_ν over frequency ν_1 and multiplying by the area A of the source gives

$$L = A \frac{\pi^4}{15h^4} (kT)^4 \ , \tag{62}$$

so the total energy emitted by the object per time dt_1 in a frequency interval $[\nu_1, \nu_1 + d\nu_1]$ is

$$dE_1 = AS_\nu d\nu_1 dt_1 = L\frac{15}{\pi^4}\left(\frac{h\nu_1}{kT_1}\right)^3 \frac{1}{e^{h\nu_1/kT_1}-1}\left(\frac{h}{kT_1}\right) d\nu_1\, dt_1 \ . \tag{63}$$

The energy received on earth per unit area in a time dt_0 in a frequency interval $[\nu_0, \nu_0 + d\nu_0]$ is then obtained from dE_1 above in the following way:

1. One must divide by the area of the spherical shell of light, when received; the relevant part of the metric is $ds^2 = a(t)^2 r_1^2 d\Omega^2$, and so the area of the light pulse is $(4\pi a_0^2 r_1^2)$, where $a_0 = a(t_0)$.
2. One must account for the fact that the photons emitted with frequency ν_1 are received with frequency $\nu_0 = \nu_1/(1+z)$ and have correspondingly lower energy
3. We have also seen that photons emitted over a time dt_1 will be received over a time interval $dt_0 = dt_1(1+z)$.

Therefore the energy we receive on earth per unit area is

$$\begin{aligned} dE_0 &= \left(\frac{1}{1+z}\right)\frac{L}{4\pi(a_0r_1)^2}\frac{15}{\pi^4}\left(\frac{h\nu_1}{kT_1}\right)^3\frac{1}{e^{h\nu_1/kT_1}-1}\left(\frac{h}{kT_1}\right)d\nu_0\,dt_0 \\ &= \left(\frac{1}{1+z}\right)^2\frac{L}{4\pi(a_0r_1)^2}\frac{15}{\pi^4}\left(\frac{h\nu_0}{kT_0}\right)^3\frac{1}{e^{h\nu_1/kT_0}-1}\left(\frac{h}{kT_0}\right)d\nu_0\,dt_0 \\ &= \ell\frac{15}{\pi^4}\left(\frac{h\nu_0}{kT_0}\right)^3\frac{1}{e^{h\nu_1/kT_0}-1}\left(\frac{h}{kT_0}\right)d\nu_0\,dt_0 \end{aligned} \tag{64}$$

where we have defined the **apparent luminosity**

$$\ell = \left(\frac{1}{1+z}\right)^2 \frac{L}{4\pi(a_0r_1)^2} \tag{65}$$

and

$$T_0 \equiv \frac{T_1}{1+z} \ . \tag{66}$$

Note that the expressions for $\frac{dE_0}{dt_0 d\nu_0}$ and $\frac{dE_1}{dt_1 d\nu_1}$ take the same functional form, except for the overall normalization, the substitution of ν_0 for ν_1, and of T_0 for T_1. This means that the spectrum observed still looks like blackbody radiation, but with a temperature red-shifted by a factor of $1/(1+z)$.

3.5 Measures of distance

To observationally determine the parameters of the FRW metric, k and $a(t)$, one important tool will be to compare the redshift of an object with its distance. However, while it is theoretically straightforward to compute the proper distance to an object

$$\int_0^r \frac{dr'}{\sqrt{1-kr'^2}} \ , \tag{67}$$

that is not a quantity readily measured. Instead one usually measures distances in one of the three following ways.

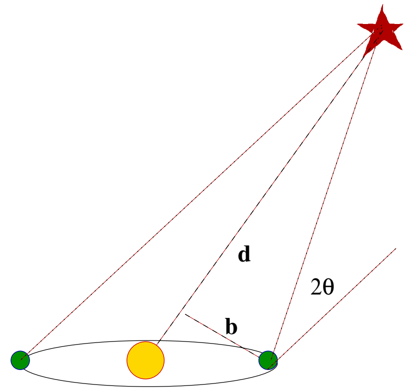

Figure 3: Determining a distance by parallax

3.5.1 Parallax

For relatively nearby objects ($\lesssim 50kpc$) one can use parallax to measure the distance of an object. The idea is to look at the same object from opposite sides of the earth's orbit about the sun, and record the angle 2θ by which the orientation of the telescope must be shifted between the two observations. If the radius of the earth's orbit projected onto the plane perpendicular to the starlight propagation x1is b, then in Euclidean space, the distance to a distant object at $r = r_1$ is given by $d = b/\theta$ (see Fig. 3). Therefore one defines the **parallax distance** $d_p = b/\theta$. In the FRW space one can show that

$$d_p = a(t_0)\frac{r_1}{\sqrt{1 - kr_1^2}} \ . \tag{68}$$

This looks rather strange; for example, with $k = r_1 = 1$, this distance diverges, even though we know that the $k = 1$ space is compact, and that the greatest [proper distance one can get from any other point in that space is $a\pi$! However, the above result makes sense. Recall that for $k = 1$, the spatial slices of our space are 3-spheres. If the star is at the north pole, and we are observing from two point on the equator, then our telescope always points due north, and $\theta = 0$. It follows that $d_p = \infty$ even though the star is a finite proper distance away. I will not derive the above result, but you can find it in Weinberg's gravitation book, and other sources.

3.5.2 Angular distance

Another way to measure distance is to measure the angular extension of a distant object whose size is known, as in Fig. 4. If the object has diameter D, then a suitable definition of **angular distance** for distant objects is $d_A = D/\theta$. In the FRW metric, D is the proper width of the object at the time of the emission of light. Using $ds^2 = \ldots - a(t)^2 r^2 d\theta^2 - \ldots$, this gives us

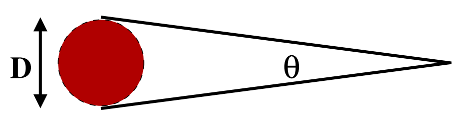

Figure 4: Determining a distance by angular extension

$D = a(t_1)r_1\theta$, for an object of angular extension θ emitting light at time t_1, so that

$$d_A = r_1 a(t_1) \ . \tag{69}$$

3.5.3 Luminosity distance

Suppose we have a source whose **absolute luminosity** L is known, defined as the total energy output of the source per unit time. If we draw an imaginary sphere about the source intersecting our position, then in Minkowski space, the power received on earth is $\ell = L/A$, where $A = 4\pi d^2$, d being the distance to the source. Therefore a sensible definition of **luminosity distance** is

$$d_L = \sqrt{\frac{L}{4\pi\ell}} \ . \tag{70}$$

With the FRW metric, we use the apparent luminosity for ℓ, given in eq. (65). As a result the luminosity distance is given by

$$d_L = r_1 a(t_0)(1+z) \tag{71}$$