Cosmology

Lecture Notes C358

Mathematics University College London 2006

Contents

Chapter 1

Black Holes

1.1 How to fall into a black hole

Remember that the Schwarzschild metric only applies in the empty space surrounding a massive object. For the Earth, this starts (ignoring the thin and light atmosphere) at

$$r = R_{\oplus} = 6371\,\mathrm{km} = 7 \times 10^5 r_{s\oplus}. \tag{1.1}$$

Exercise 1.1 *In the interior of the Earth, we can replace* $r_{s\oplus} = 2GM_{\oplus} = 0.9\,\mathrm{cm}$ *by* $2GM(r)$ *where* $M(r) = 4\pi\rho(r)r^3/3$ *and* $\rho(r)$ *is the mass density inside the Earth at radius* r*. The mean mass density of the Earth is* $5.5 \times 10^3\,\mathrm{kg/\,m^3}$*.*

Suppose that the mass density of the Earth is constant. Show that $r \gg r_s(r)$ *everywhere in the interior of the Earth. At what radius is the ratio* r/r_s *at a minimum? What is this minimum?*

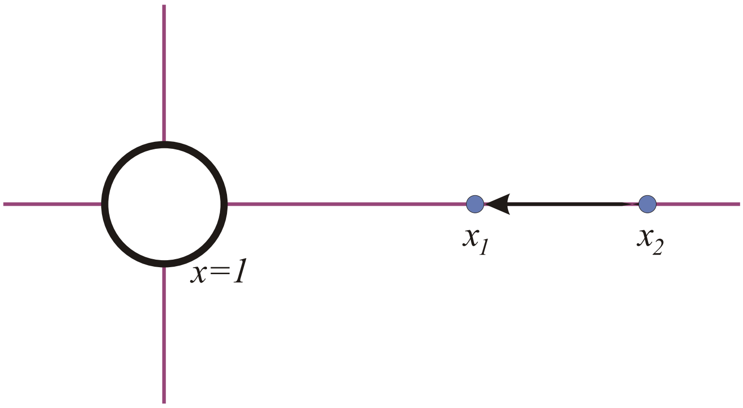
Fortunately, this keeps us well away from the dangerous region near $r = r_s$. But for a Schwarzschild black hole, all the mass is assumed to be crushed into the central singularity at $r = 0$. Thus the Schwarzschild metric is valid throughout all the space $r > 0$. We can explore the curious geometry of space-time near the Schwarzschild radius by imagining a few thought experiments about flying into a black hole (since real experiments would be unpleasant). The Schwarzschild metric is

$$\mathrm{d}\tau^2 = \left(1 - \frac{r_s}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\mathrm{d}r^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2). \tag{1.2}$$

Consider a spaceship falling into a black hole with purely radial motion, so that $\mathrm{d}\theta = \mathrm{d}\phi = 0$. Also convert to the non-dimensional coordinate $x \equiv r/r_s$. Thus $x = 1$ at the Schwarzschild radius, and

$$\boxed{\mathrm{d}\tau^2 = \left(1 - \frac{1}{x}\right)\mathrm{d}t^2 - \left(1 - \frac{1}{x}\right)^{-1} r_s^2\mathrm{d}x^2}. \tag{1.3}$$

First consider a spaceship in free-fall from x_2 to x_1:



A spaceship in free-fall follows a geodesic, so

$$u_0 = k = \text{constant} \tag{1.4}$$

$$\Rightarrow u^0 = \frac{\mathrm{d}t}{\mathrm{d}\tau} = \left(1 - \frac{1}{x}\right)^{-1} k. \tag{1.5}$$

Divide the metric line element by $d\tau^2$:

$$\begin{aligned} 1 &= \left(1 - \frac{1}{x}\right)\left(\frac{\mathrm{d}t}{\mathrm{d}\tau}\right)^2 - r_s^2\left(1 - \frac{1}{x}\right)^{-1}\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2 & (1.6)\\ &= \left(1 - \frac{1}{x}\right)^{-1}\left(k^2 - r_s^2\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2\right) & (1.7)\\ \Rightarrow 1 - \frac{1}{x} &= k^2 - r_s^2\left(\frac{\mathrm{d}x}{\mathrm{d}\tau}\right)^2. & (1.8)\end{aligned}$$

Also,

$$k = \frac{E}{m} \approx \frac{1}{m}(m + \frac{1}{2}mV^2 + m\Phi(x)), \tag{1.9}$$

where

$$\begin{aligned} \Phi(r) &= \frac{-GM}{r} = \frac{-r_s}{2r}, & (1.10)\\ \text{or} \qquad \Phi(x) &= -\frac{1}{2x}. & (1.11)\end{aligned}$$

Suppose the ship *starts* at $t = 0$ with $V \ll 1$ and $x \gg 1$ (i.e. $r \gg r_s$). In this case $k \approx 1$ at $t = 0$. Now, k is a constant, so $k \approx 1$ even as the space-ship approaches the Schwarzschild radius. With this approximation

$$\boxed{\frac{\mathrm{d}x}{\mathrm{d}\tau} \approx -\frac{1}{r_s\sqrt{x}}} \quad \text{(minus because ship is falling inwards).} \tag{1.12}$$

Integrate:

$$\Delta\tau = \int_{\tau_2}^{\tau_1} d\tau = -r_s \int_{x_2}^{x^1} \sqrt{x}\mathrm{d}x. \tag{1.13}$$

So

$$\Delta\tau = \frac{2r_s}{3}\left(x_2^{3/2} - x_1^{3/2}\right). \tag{1.14}$$

If the ship falls all the way to the centre of the black hole (as it must do if it goes past the Schwarzschild radius!) then $x_1 = 0$, and

$$\boxed{\Delta\tau = \frac{2r^s}{3}x_2^{3/2}}. \tag{1.15}$$

This is the proper time to fall to the centre of the black hole.

Now let us watch from a safe distance. Consider a fixed observer at $x_2 \gg 1$. At this distance space-time is almost flat, so proper time for the observer is almost the same as t. What is the coordinate time t when the spaceship crosses $x = 1$?

With τ the proper time on the falling ship, and $k \approx 1$, we have

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= \frac{\mathrm{d}x}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}t} && (1.16)\\ &\approx -\frac{1}{r_s\sqrt{x}}\left(1-\frac{1}{x}\right) && (1.17) \end{aligned}$$

Integrate from x_2 to x_1:

$$\begin{aligned} \Delta t &= t_1 - t_2 = -r_s\int_{x_2}^{x_1}\left(1-\frac{1}{x}\right)^{-1}\sqrt{x}\,\mathrm{d}x && (1.18)\\ &= r_s\int_{x_1}^{x_2}\frac{x^{3/2}}{x-1}\,\mathrm{d}x && (1.19)\\ &= r_s\left(2(\sqrt{x_2}-\sqrt{x_1}) + \frac{2}{3}(x_2^{3/2}-x_1^{3/2}) + \log\frac{\sqrt{x_2}-1}{\sqrt{x_1}-1} - \log\frac{\sqrt{x_2}+1}{\sqrt{x_1}+1}\right). && (1.20) \end{aligned}$$

Note that

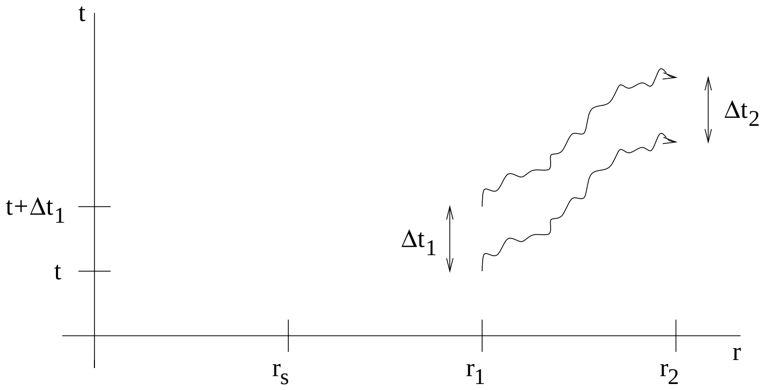
$$\boxed{\Delta t \to \infty \text{ as } x_1 \to 1}. \tag{1.21}$$

Outside observers never see the falling ship passing through the edge of the black hole.

1.2 Gravitational Red-shift

Suppose a ship hovering at a constant radius r_1 has a transmitter which emits an electromagnetic wave with period P_1. This period equals the proper time (ship time) between successive wave crests, i.e. $P_1 = \delta\tau_1$. Similarly, a receiver on a ship hovering at constant r_2 observes a period $P_2 = \delta\tau_2$. To derive the relation between P_1 and P_2, we must first understand the relation between the coordinate time intervals δt_1 and δt_2. Claim: at fixed x, θ, ϕ,

$$\delta t_1 = \delta t_2. \tag{1.22}$$



Why? The path of the photon should be invariant to translations $t \to t + \delta t$. (If the metric is invariant in time). So the gap between two photon paths should always be the same $\delta t = \delta t_1 = \delta t_2$.

$$\mathrm{d}\tau^2 = \left(1 - \frac{1}{x}\right)\mathrm{d}t^2 - 0 - 0 - 0 \tag{1.23}$$

$$\Rightarrow \mathrm{d}\tau = \sqrt{1 - \frac{1}{x}}\,\mathrm{d}t. \tag{1.24}$$

Thus at x_1 and x_2

$$P_1 = \Delta\tau_1 = \sqrt{1 - \frac{1}{x_1}}\,\Delta t, \tag{1.25}$$

$$P_2 = \Delta\tau_2 = \sqrt{1 - \frac{1}{x_2}}\,\Delta t. \tag{1.26}$$

The ratio of periods P_1/P_2 is equal to the ratio of wavelengths λ_1/λ_2 and is inverse to the ratio of frequencies ν_1/ν_2:

$$\boxed{\frac{\lambda_2}{\lambda_1} = \frac{\nu_1}{\nu_2} = \frac{P_2}{P_1} = \frac{\sqrt{1 - x_2^{-1}}}{\sqrt{1 - x_1^{-1}}}}. \tag{1.27}$$

This expression simplifies for $x_2 \gg 1$:

$$\frac{\lambda_2}{\lambda_1} \approx \sqrt{\frac{x_1}{x_1 - 1}}. \tag{1.28}$$

The redshift is defined as

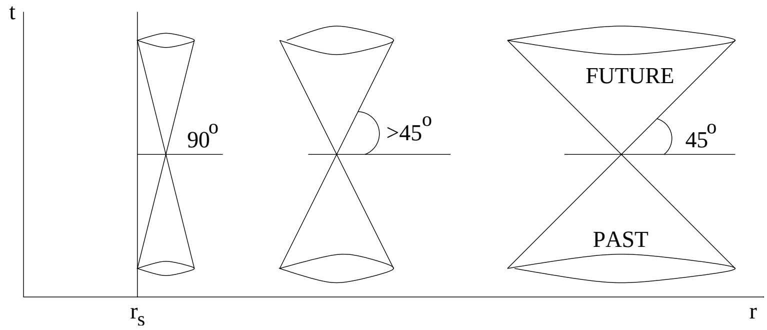
$$z \equiv \frac{\lambda_2}{\lambda_1} - 1. \tag{1.29}$$

As $x_1 \to 1$, the redshift and the time dilation P_2/P_1 both go to infinity.

1.3 Light Cones

1.3.1 Outside

Some of the geometry of space-time can be seen by considering the light-cones in a space-time diagram. We will use a space-time diagram in the t-r plane.



The slope of the light cone is $\mathrm{d}t/\mathrm{d}r$. For light, $\mathrm{d}\tau = 0$.

$$\Rightarrow \mathrm{d}\tau^2 = 0 = \left(1 - \frac{r_s}{r}\right)\mathrm{d}t^2 - \left(1 - \frac{r_s}{r}\right)^{-1}\mathrm{d}r^2. \tag{1.30}$$

$$\Rightarrow \boxed{\frac{\mathrm{d}t}{\mathrm{d}r} = \pm\left(1 - \frac{r_s}{r}\right)^{-1} = \pm\frac{r}{|r - r_s|}} \tag{1.31}$$

Far away from the black hole, spacetime looks like Minkowski space, with 45° light cones:

$$\lim_{r\to\infty}\frac{dt}{dr} = \pm 1. \tag{1.32}$$

However, near the event horizon $r = r_s$ the light cones narrow to zero thickness and 90° slope:

$$\lim_{r\to r_s}\frac{dt}{dr} = \pm\infty. \tag{1.33}$$

Because of the 90° slope, photons (as well as massive particles) cannot escape – they would need a more horizontal slope pointing to the right to get away from the black hole. On the other hand, they seemingly cannot fall in, as they cannot travel to the left. In fact, they do fall in, but outside observers must wait an infinite amount of coordinate time t before this happens.

1.3.2 Inside

For $r < r_s$, $\left(1 - \frac{r_s}{r}\right) < 0$.

$$\Rightarrow \mathrm{d}\tau^2 = \left|1 - \frac{r_s}{r}\right|^{-1}\mathrm{d}r^2 - \left|1 - \frac{r_s}{r}\right|\mathrm{d}t^2 - r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2) \tag{1.34}$$

Note that the $\mathrm{d}r^2$ term has the positive metric coefficient, while the $\mathrm{d}t^2$ term goes negative. This implies that the (negative) r direction becomes the new future time direction: proper time, entropy, and conscious time move toward $r = 0$ rather than along the old t coordinate, which has now become spatial. Thus one can no more escape being crushed into the central singularity than one can escape progressing into the future.

1.4 Kruskal Coordinates

Let us change coordinates to a new system where *all* light cones have 45° angles. We replace t and r by new coordinates u and v. In these coordinates, for radial photon motion

$$\frac{\mathrm{d}u}{\mathrm{d}v} = \pm 1. \tag{1.35}$$

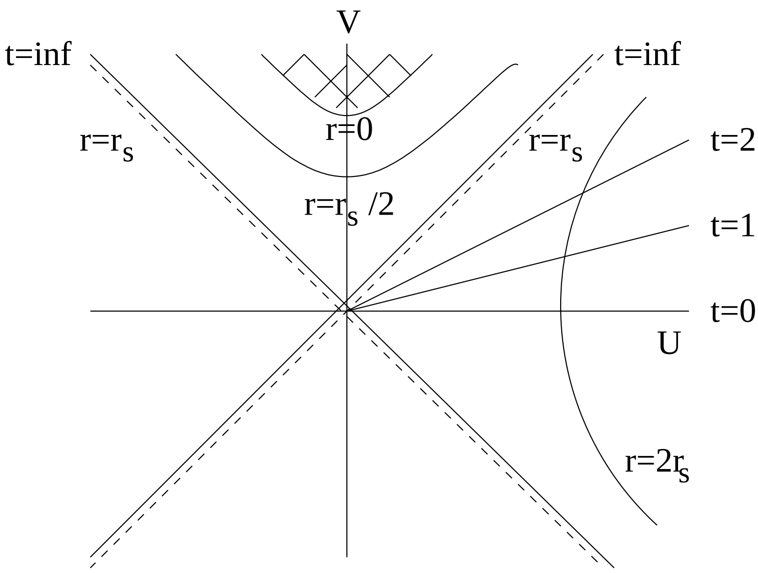
These new coordinates can be shown to be the following functions of t and r:

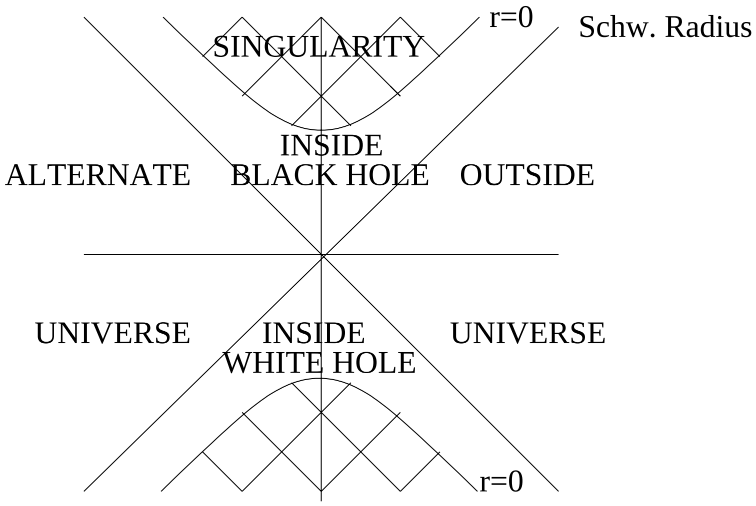
$$u = \left|1 - \frac{r_s}{r}\right|^{1/2} e^{r/2r_s} \begin{cases} \cosh \frac{t}{2r_s}, & r > r_s; \\ \sinh \frac{t}{2r_s}, & r \leq r_s. \end{cases} \tag{1.36}$$

$$v = \left|1 - \frac{r_s}{r}\right|^{1/2} e^{r/2r_s} \begin{cases} \sinh \frac{t}{2r_s}, & r > r_s; \\ \cosh \frac{t}{2r_s}, & r \leq r_s. \end{cases} \tag{1.37}$$

The metric line element in these new coordinates is

$$\boxed{\mathrm{d}\tau^2 = \frac{4r_s^3}{r} e^{-r/r_s} (\mathrm{d}v^2 - \mathrm{d}u^2) - r^2 \mathrm{d}\Omega^2}. \tag{1.38}$$





A white hole is the time-reversal of a black hole.

1.5 Other Properties

a. A static BH is completely specified by 3 quantities:

- Mass : M
- Angular momentum : J
- Charge : Q

b. Hawking: The area of a BH always increases (ignoring the Hawking effect below). For $J = Q = 0$, the area of a black hole is simply the area of the sphere at the horizon $r = r_s$,

$$A(M) = 4\pi r_s^2 = 16\pi G^2 M^2. \tag{1.39}$$

As a consequence, a BH cannot split into 2 smaller pieces:

$$2A(\frac{M}{2}) = 2\left(16\pi G^2 \left(\frac{M}{2}\right)^2\right) \tag{1.40}$$

$$= 8\pi G^2 M < A(M). \tag{1.41}$$

c. The Hawking effect : Hawking's area theorem is like the second law of thermodynamics for the increase of entropy S:

$$\frac{\mathrm{d}A}{\mathrm{d}t} \geq 0 \Leftrightarrow \frac{\mathrm{d}S}{\mathrm{d}t} \geq 0. \tag{1.42}$$

In thermodynamics the change in energy of a gas at constant volume and temperature T is

$$\mathrm{d}E = T\mathrm{d}S. \tag{1.43}$$

Consider a $J = Q = 0$ BH. By the equivalence of mass and energy, $E = M$, so by equation (1.39),

$$\begin{aligned} A &= 16\pi G^2 E^2 && (1.44) \\ \mathrm{d}A &= 32\pi G^2 E \mathrm{d}E && (1.45) \\ \Rightarrow \mathrm{d}E &= \left(\frac{1}{32\pi G^2 M}\right) \mathrm{d}A. && (1.46) \end{aligned}$$

Next,let[1]

$$\begin{aligned} T &= \frac{\hbar}{8\pi k G M} && (1.47) \\ S &= \frac{k}{4\hbar G} A && (1.48) \end{aligned}$$

then $\mathrm{d}E = T\mathrm{d}S$ as before. This implies black holes radiate Black body radiation at temperature T. The black body luminosity is

$$\begin{aligned} L &= \left|\frac{\mathrm{d}E}{\mathrm{d}t}\right| = \sigma A T^4, && (1.49) \\ \sigma &= 5.67 \times 10^{-8} \mathrm{Watts\,m^{-2}\,sec^{-1}\,K^{-4}}. && (1.50) \end{aligned}$$

Exercise 1.2 *What is the Schwarzschild metric line element inside r_s with arbitrary displacements $\mathrm{d}t$, $\mathrm{d}r$, $\mathrm{d}\phi$, and $\mathrm{d}\theta$ (use modulus signs to show which coefficients are positive or negative).*

Show that for any motion of a spaceship inside a black hole, with arbitrary $\mathrm{d}t/\mathrm{d}\tau$, $\mathrm{d}r/\mathrm{d}\tau$, $\mathrm{d}\phi/\mathrm{d}\tau$, and $\mathrm{d}\theta/\mathrm{d}\tau$

$$\left|1 - \frac{r_s}{r}\right|^{-1} \left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 > 1. \tag{1.51}$$

A spaceship enters $r = r_s$ at $\tau = 0$. Show that the spaceship will arrive at $r = 0$ within a proper time τ_{max} where $\tau_{max} = \pi GM$ no matter how it fires its engines. You may use the integral

$$\int x^{1/2}(1-x)^{-1/2}\,\mathrm{d}x = -x^{1/2}(1-x)^{1/2} + \sin^{-1}\left(x^{1/2}\right) \tag{1.52}$$

without proof.

Exercise 1.3 *A rotating black hole with angular momentum J and mass M has the Kerr metric*

$$\begin{aligned} \mathrm{d}\tau^2 = \; & \tfrac{\Delta - a^2\sin^2\theta}{\rho^2}\mathrm{d}t^2 + 2a\tfrac{2GMr\sin^2\theta}{\rho^2}\,\mathrm{d}t\mathrm{d}\phi - \tfrac{(r^2+a^2)^2 - a^2\Delta\sin^2\theta}{\rho^2}\sin^2\theta\mathrm{d}\phi^2 \\ & - \tfrac{\rho^2}{\Delta}\mathrm{d}r^2 - \rho^2\mathrm{d}\theta^2; \end{aligned}$$

[1] We need quantum mechanics to prove these - at the moment this is just a fiddle with constants to get the correct units.

where $a = J/M$, $\Delta = r^2 - 2GMr + a^2$, *and* $\rho^2 = r^2 + a^2\cos^2\theta$.

Consider the conserved quantities along a geodesic near a rotating black hole. Letting $k = u_0$ *and* $h = -u_3$, *express these conserved quantities in terms of* $\mathrm{d}t/\mathrm{d}\tau$ *and* $\mathrm{d}\phi/\mathrm{d}\tau$.

Exercise 1.4

The luminosity $(-dE/dt)$ *of a black-body of area* A *and temperature* T *is* $L = \sigma A T^4$ *where* $\sigma = 5.67 \times 10^{-8}$ *Watts* $\mathrm{m}^{-2}\,sec^{-1}\,K^{-4}$. *The Hawking temperature of a black hole is* $T = \hbar c^3/8\pi kGM$. $(\hbar = 1.055 \times 10^{-34}\,Jsec$, $G/c^2 = 7.425 \times 10^{-28}\,\mathrm{m}\,Kg^{-1}$, $k = 1.4 \times 10^{-23} JK^{-1})$. *Recall that the area of a static black hole is* $4\pi r_s^2$, *where* $r_s = 2GM/c^2$. *Derive the equation for* dM/dt *due to the Hawking effect, and calculate the lifetime of a solar mass* $(M = 2 \times 10^{30} Kg)$ *black hole. What is the initial mass of a black hole formed at the big bang which would be just disappearing now (after* 15×10^9 *years)?*

Chapter 2

The Einstein Field Equations

2.1 Curvature

2.1.1 Intrinsic and Extrinsic Geometry

If a two-dimensional manifold M is imbedded in Euclidean three-space $\mathbb{E}^3$, then we can often deduce its shape and curvature by inspection. For example, the radius R of a two-sphere S^2 can be found by measuring the distance to its centre in $\mathbb{E}^3$. But the manifold S^2 only contains the surface of the sphere; points closer to the centre in $\mathbb{E}^3$ are not on S^2. Thus we have to go outside of S^2 (even if three-dimensional people call this "inside") in order to measure the distance to the centre. Using geometrical information coming from outside the manifold is called *extrinsic* geometry.

Intrinsic geometry, on the other hand, only uses information available on the manifold itself. Suppose we wished to measure the radius R of a sphere intrinsically. One method would be to draw a circle on the sphere. Also on the sphere (i.e. on the surface: with intrinsic geometry we never go inside or outside!) there will be a point which can be regarded as the centre of the circle (see figure **??**). The distance on geodesics drawn from the point to the circle is constant. Let this distance be s. (Note that the distance s to a circle from its centre is generally called its 'radius', at least for circles on a Euclidean plane.) If the sphere has coordinates (θ, ϕ), and the circle is a curve of constant θ (i.e. a latitude line) then the centre of the circle is at the North pole, and

$$s = R\theta. \tag{2.1}$$

However, again staying on the surface, we can also measure the circumference C of the circle. The circumference provides a second possible definition of radius, i.e.

$$r \equiv \frac{C}{2\pi}. \tag{2.2}$$

Note that $r \neq s$. In fact, from figure 2.1.3,

$$r = R\sin\theta = R\sin\frac{s}{R}. \tag{2.3}$$

We can determine R without ever leaving the surface by measuring and then comparing r and s. For example, in the limit $s \to 0$,

$$\lim_{s\to 0} \frac{1}{s}\frac{\mathrm{d}^2 r}{\mathrm{d}s^2} = -\lim_{s\to 0} \frac{1}{sR}\left(\frac{s}{R} - \frac{1}{6}\left(\frac{s}{R}\right)^3 + \cdots\right) \tag{2.4}$$

$$= R^{-2}. \tag{2.5}$$

A second method of determining R also involves drawing a circle. There are actually two points on S^2 equidistant from any circle (as measured on geodesics); for a $\theta =$ constant circle these would be the two poles. The sum of the distances to these two points is πR.

2.1.2 Definitions of Radius

Suppose a circle on some arbitrary manifold M with centre $P \in M$ is defined to be a set of points of equal geodesic distance from P. Then we have two possible definitions of radius: the geodesic distance, and the circumference (length of the circle) divided by 2π. Both of these definitions will be useful in our study of spherically symmetric spacetimes, in particular the Schwarzschild and Robertson–Walker metrics.

2.1.3 Geodesic deviation

We can also measure the radius of curvature of a sphere by examining how neighboring geodesics move away from each other. For simplicity, consider two geodesics passing through the North Pole. The distance ξ between two points separated by longitude angle ψ at co-latitude θ is

$$\xi = (R\sin\theta)\psi. \tag{2.6}$$

The arc length travelled from the North Pole is $s = R\theta$, so

$$\xi = R\psi \sin\left(\frac{s}{R}\right). \tag{2.7}$$

Expanding,

$$\xi = R\psi\left(\frac{s}{R} - \frac{1}{6}\frac{s^3}{R^3} + \cdots\right). \tag{2.8}$$

$$\implies \frac{\mathrm{d}^2\xi}{\mathrm{d}s^2} = -\frac{\psi s}{R^2} + O(s^3). \tag{2.9}$$

Thus we can determine R from the the second derivative of the distance *between* geodesics.

$$\frac{1}{R^2} = -\lim_{s\to 0}\left(\frac{1}{\psi s}\frac{\mathrm{d}^2\xi}{\mathrm{d}s^2}\right) \tag{2.10}$$

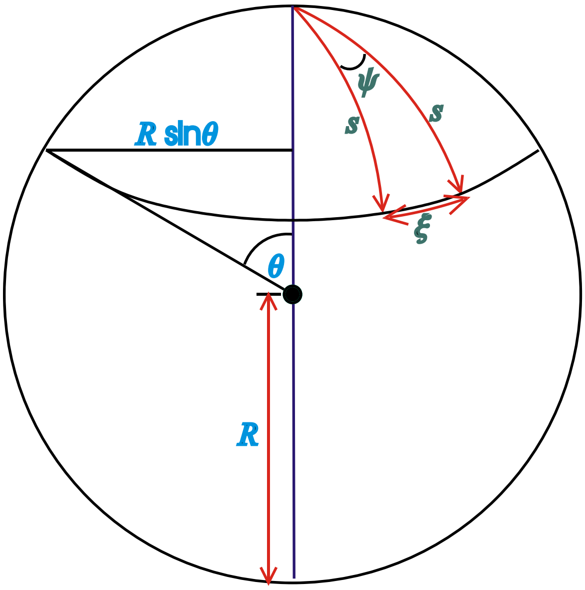


Figure 2.1: two neighboring geodesics on a sphere.

2.2 The Equation of Geodesic Deviation

While the methods above work fine for a familiar shape like a sphere, we need more general tools for measuring curvature. In particular, we should be able to calculate curvature directly from the metric and its derivatives. First we obtain a general formula for geodesic deviation.

We first imagine a two-dimensional surface filled with geodesics moving more or less parallel to each other (see figure 2.2). Crossing these geodesics is another family of curves (not necessarily geodesics). Thus the surface resembles a woven cloth, with the warp threads the geodesics, and the weft threads the family of crossing curves. If the geodesics deviate away from each other, then the crossing curves must travel farther to get from one geodesic to the next. This extra travel is what we will calculate in order to measure geodesic deviation.

Suppose we start with a single curve $\gamma(\sigma)$ on a manifold, where σ measures distance along the curve. At each point (labelled by σ) on this curve, we send out a geodesic into the manifold. These geodesics will be parameterized by distance τ away from the original curve $\gamma(\sigma)$. If we do this smoothly, we can cover a surface $\mathcal{S}$ with these geodesics.

Any point on the surface $\mathcal{S}$ can be found by specifying σ, telling us which geodesic it is on, and τ, telling us how far along the geodesic it is. Thus $\mathcal{S} = \mathcal{S}(\sigma, \tau)$. Let $\overline{\mathbf{U}}(\sigma, \tau)$ be the tangent vector to a geodesic passing through the point at (σ, τ):

$$U^a(\sigma, \tau) = \frac{\mathrm{d}\mathsf{X}^a(\sigma, \tau)}{\mathrm{d}\tau}. \tag{2.11}$$

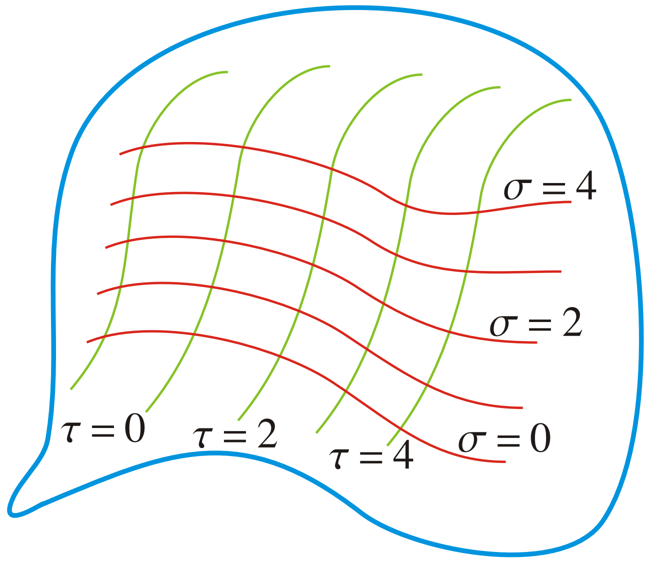


Figure 2.2: Family of geodesics. Each geodesic is labelled by one value of σ, and is parameterized by τ.

The geodesics have constant σ, by definition. We could also draw curves through the points of constant τ (for example, or original curve $\gamma(\sigma)$ has $\tau = 0$). Let $\overline{\xi}(\sigma, \tau)$ be the tangent vector to these curves:

$$\xi^a(\sigma, \tau) = \frac{\mathrm{d}\mathsf{X}^a(\sigma, \tau)}{\mathrm{d}\sigma}. \tag{2.12}$$

Consider two geodesics at, say, $\sigma = 0$ and $\sigma = 1$. The change in coordinates between the two at $\tau = 0$ is

$$\begin{aligned} \delta \mathsf{X}^a &= \mathsf{X}^a(1,0) - \mathsf{X}^a(0,0) && (2.13)\\ &\approx \left.\frac{\mathrm{d}\mathsf{X}^a}{\mathrm{d}\sigma}\right|_{(0,0)} \delta\sigma && (2.14)\\ &= \xi^a(0,0), && (2.15) \end{aligned}$$

as $\delta\sigma = 1$. So if they are very close to each other then $\xi^a(0,0)$ must be small; if they are far apart, then $\xi^a(0,0)$ must be large. Now let us slide along the geodesics to $\tau > 0$. If the geodesics diverge as τ increases, then ξ^a must also be increasing with τ. From the analysis of geodesics on spheres in section 2.1.3, the second derivative gives us the essential information about the curvature of the manifold. Using covariant derivatives leads us to the quantity $D^2\overline{\xi}/D\tau^2$.

Now each $D/D\tau = \overline{\mathbf{U}} \cdot \nabla$, so this quantity involves two copies of $\overline{\mathbf{U}}$. It also involves one copy of ξ. Thus we can reasonably guess that if we calculate this quantity, it will have the form

$$\left(\frac{\mathrm{D}^2\overline{\xi}}{\mathrm{D}\tau^2}\right)^a = R^a{}_{bcd}U^bU^c\xi^d, \tag{2.16}$$

where $R^a{}_{bcd}$ provides a table of coefficients for the $\overline{\xi}$ vector and the two copies of $\overline{\mathbf{U}}$. As $U^bU^c\xi^d$ has three upper indices, $R^a{}_{bcd}$ must have three lower indices, plus one upper to match the left hand side. As this *equation of geodesic deviation* is a tensor equation, and everything else in the equation is a tensor, $R^a{}_{bcd}$ must also be a tensor. It is known as the *Riemann Curvature Tensor* after its discoverer, Bernhard Riemann.

2.3 The Riemann Curvature Tensor

Actually, Riemann first defined his famous tensor by considering the commutator between two covariant derivatives. He showed that, given any vector $\overline{\mathbf{V}}$,

$$(\nabla_c\nabla_d\overline{\mathbf{V}})^a - (\nabla_d\nabla_c\overline{\mathbf{V}})^a = R^a{}_{bcd}V^b, \tag{2.17}$$

where

$$R^a{}_{bcd} = \partial_c\Gamma^a{}_{bd} - \partial_d\Gamma^a{}_{bc} + \Gamma^e{}_{bd}\Gamma^a{}_{ec} - \Gamma^e{}_{bc}\Gamma^a{}_{ed}. \tag{2.18}$$

Comments:

a. This equation can be derived as a straightforward exercise, using the definitions for the covariant derivative. Note that the right hand side has no derivatives of V^a. The absence of ordinary second derivatives makes sense: $\nabla_b\nabla_c$ contains $\partial_b\partial_c$, whereas $\nabla_c\nabla_b$ contains $\partial_c\partial_b$; but ordinary derivatives commute, so these two terms cancel. Less obviously, terms involving the first derivatives of V^a also cancel.

b. In flat space (Euclidean space or Minkowski space-time), $R^a{}_{bcd} = 0$, since $\Gamma^a{}_{bc}$ and its derivatives vanish.

c. In an LIF, $\Gamma^a{}_{bc} = 0$, *but* the derivative terms $\partial_c\Gamma^a{}_{bd} - \partial_d\Gamma^a{}_{bc}$ do not vanish for a curved manifold.

2.3.1 The Riemann Tensor and Geodesic Deviation (Optional)

We must prove that the tensors in equations (2.16) and (2.17) are the same. First, as $\overline{\mathbf{U}}$ is the tangent (or velocity) vector to a geodesic,

$$\boxed{\frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\tau} = \overline{\mathbf{U}}\cdot\nabla\overline{\mathbf{U}} = 0}. \tag{2.19}$$

We will also find the two theorems useful:

Theorem 2.1

$$\boxed{\overline{\mathbf{U}}\cdot\nabla\overline{\xi} = \overline{\xi}\cdot\nabla\overline{\mathbf{U}}}. \tag{2.20}$$

In terms of directional derivatives,

$$\boxed{\frac{\mathrm{D}\overline{\xi}}{\mathrm{D}\tau} = \frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\sigma}}. \tag{2.21}$$

Proof 2.1 This equation follows from the definitions of $\overline{\mathbf{U}}$ and $\overline{\xi}$ (equations 2.11 and 2.12):

$$\begin{aligned}
\frac{\mathrm{D}\overline{\xi}}{\mathrm{D}\tau} = (\overline{\mathbf{U}} \cdot \nabla\overline{\xi})^a &= (\frac{D\overline{\xi}}{D\tau})^a & (2.22)\\
&= \frac{\mathrm{d}\xi^a}{\mathrm{d}\tau} + \Gamma^a{}_{bc}U^b\xi^c & (2.23)\\
&= \frac{\mathrm{d}^2\mathsf{X}^a}{\mathrm{d}\tau\mathrm{d}\sigma} + \Gamma^a{}_{bc}U^b\xi^c; & (2.24)\\
\frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\sigma} = (\overline{\xi} \cdot \nabla\overline{\mathbf{U}})^a &= (\frac{D\overline{\mathbf{U}}}{D\sigma})^a & (2.25)\\
&= \frac{\mathrm{d}U^a}{\mathrm{d}\sigma} + \Gamma^a{}_{bc}\xi^b U^c & (2.26)\\
&= \frac{\mathrm{d}^2\mathsf{X}^a}{\mathrm{d}\sigma\mathrm{d}\tau} + \Gamma^a{}_{bc}\xi^b U^c & (2.27)\\
&= \frac{\mathrm{d}^2\mathsf{X}^a}{\mathrm{d}\tau\mathrm{d}\sigma} + \Gamma^a{}_{bc}U^b\xi^c. & (2.28)
\end{aligned}$$

In the last line we used the commutativity of partial derivatives, and the condition $\Gamma^a{}_{cb} = \Gamma^a{}_{bc}$.

Theorem 2.2 Consider a two-dimensional submanifold parameterized with coordinates σ and τ, with corresponding tangent vectors $\overline{\mathbf{U}}$ and $\overline{\xi}$. Let $\overline{\mathbf{V}}$ be a vector. Then the commutation of the directional derivatives $\mathrm{D}/\mathrm{D}\sigma$, $\mathrm{D}/\mathrm{D}\tau$ is given by

$$\left(\left[\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}}{\mathrm{D}\sigma} - \frac{\mathrm{D}}{\mathrm{D}\sigma}\frac{\mathrm{D}}{\mathrm{D}\tau}\right] V\right)^a = R^a{}_{bcd}V^b\xi^c U^d. \tag{2.29}$$

Proof 2.2 First,

$$\begin{aligned}
\left(\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}}{\mathrm{D}\sigma}\right)\overline{\mathbf{V}} &= \frac{\mathrm{D}}{\mathrm{D}\tau}\left(\overline{\xi} \cdot \nabla\right)\overline{\mathbf{V}} & (2.30)\\
&= \frac{\mathrm{D}\overline{\xi}}{\mathrm{D}\tau} \cdot \nabla\overline{\mathbf{V}} + \overline{\xi} \cdot \frac{\mathrm{D}}{\mathrm{D}\tau}\left(\nabla\overline{\mathbf{V}}\right) & (2.31)\\
&= \frac{\mathrm{D}\overline{\xi}}{\mathrm{D}\tau} \cdot \nabla\overline{\mathbf{V}} + \xi^d\left(U^c\nabla_c\right)\left(\nabla_d\overline{\mathbf{V}}\right). & (2.32)
\end{aligned}$$

Similarly,

$$\left(\frac{\mathrm{D}}{\mathrm{D}\sigma}\frac{\mathrm{D}}{\mathrm{D}\tau}\right)\overline{\mathbf{V}} = \frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\sigma} \cdot \nabla\overline{\mathbf{V}} + U^c\left(\xi^d\nabla_d\right)\left(\nabla_c\overline{\mathbf{V}}\right). \tag{2.33}$$

Note that the first terms on the right of equation (2.32) and equation (2.33) are equal, by theorem 2.1. The commutator of the two directional derivatives becomes simply

$$\left(\left[\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}}{\mathrm{D}\sigma} - \frac{\mathrm{D}}{\mathrm{D}\sigma}\frac{\mathrm{D}}{\mathrm{D}\tau}\right] V\right)^a = U^c \xi^d \left(\nabla_c \nabla_d - \nabla_d \nabla_c\right) V^a \tag{2.34}$$

$$= U^c \xi^d R^a{}_{bcd} V^b. \tag{2.35}$$

We can now more easily prove the relation between geodesic deviation and the Riemann tensor.

Theorem 2.3

$$\left(\frac{\mathrm{D}^2\overline{\xi}}{\mathrm{D}\tau^2}\right)^a = R^a{}_{bcd} U^b U^c \xi^d. \tag{2.36}$$

Proof 2.3 Using successively theorem 2.1, theorem 2.2, and then the basic property of a geodesic ($D\overline{\mathbf{U}}/D\tau = 0$),

$$\left(\frac{\mathrm{D}^2\overline{\xi}}{\mathrm{D}\tau^2}\right)^a = \left(\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}\overline{\xi}}{\mathrm{D}\tau}\right)^a = \left(\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\sigma}\right)^a \tag{2.37}$$

$$= \left(\frac{\mathrm{D}}{\mathrm{D}\sigma}\frac{\mathrm{D}\overline{\mathbf{U}}}{\mathrm{D}\tau}\right)^a + \left(\left[\frac{\mathrm{D}}{\mathrm{D}\tau}\frac{\mathrm{D}}{\mathrm{D}\sigma} - \frac{\mathrm{D}}{\mathrm{D}\sigma}\frac{\mathrm{D}}{\mathrm{D}\tau}\right] U\right)^a \tag{2.38}$$

$$= 0 + R^a{}_{bcd} U^b U^c \xi^d. \tag{2.39}$$

2.4 Einstein's Field Equations

The geodesic equation tells us how matter responds to geometry (metric). How does geometry respond to matter?

2.4.1 Analogy with Electro-Magnetism

$$\text{E-M:} \quad \frac{dU^a}{d\tau} = \frac{q}{m} U_b F^{ba} \tag{2.40}$$

$$= \left(\frac{q}{m} F^{ba} \eta_{bc}\right) U^c \tag{2.41}$$

$$\text{Gravity:} \quad \frac{dU^a}{d\tau} = \left(-\Gamma^a{}_{bc} U^b\right) U^c \qquad \text{Geodesic Eqn} \tag{2.42}$$

We see that inertial forces are proportional to Γ, and EM forces are proportional to F. Thus, an analogy exists between $F \leftrightarrow \Gamma$.

Now,

$$F_{ab} = \partial_b \phi_a - \partial_a \phi_b$$

and Γ contains 1st derivatives of g_{ab}. So the analogy gives the relation

$$\phi_a \leftrightarrow g_{ab}. \tag{2.43}$$

EM potential $\leftrightarrow$ Metric.

Next look at the Maxwell source equation:

$$\partial_b \partial^b \phi^a - \partial^a \partial_b \phi^b = j^a. \tag{2.44}$$

In this equation, second derivatives of the potential ϕ are proportional to the source j^a.

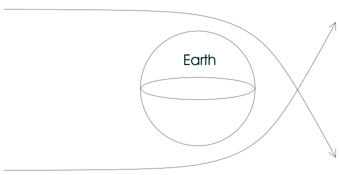
Newtonian gravity is similar. With gravitational potential Φ and source the mass density ρ,

$$\nabla^2 \Phi = 4\pi G \rho. \tag{2.45}$$

Step 1: The analogy with Maxwell's equations suggests that the Einstein Field Equations should have the form

$$\text{expression with 2nd derivatives of } g_{ab} = \text{source} \tag{2.46}$$

Step 2:

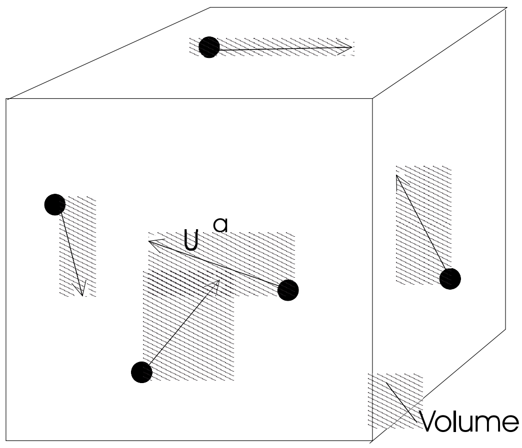


The Riemann Tensor has 2nd derivatives of g and detects curvature, so we'll guess that the LHS is formed from $R^a{}_{bcd}$. Geodesics deviate near matter, implying that matter curves spacetime.

Step 3: The source in EM is current density, i.e.

$$j^a = \begin{pmatrix} \text{charge / volume} \\ \text{electric current / volume} \end{pmatrix} \tag{2.47}$$

e.g. For N charges q_i $i = 1, \dots, N$ travelling on paths with 4-velocities U_i^a



$$j^a = \frac{1}{\text{volume}} \sum_{i=1}^{N} q_i U_i^a \tag{2.48}$$

Can we replace q_i by masses m_i = rest mass of ith particle?
NO. Two reasons:

a. We would like all forms of energy to be included in the source term

b. It gives repulsive forces between two positive masses (all masses are positive)

Instead, replace q_i by 4-momentum p_i^b.
Thus, the source becomes

$$T^{ab} = \frac{1}{\text{volume}} \sum_{i=1}^{N} p_i^b U_i^a \tag{2.49}$$

$$T^{ab} = \frac{1}{\text{volume}} \sum_{i=1}^{N} m_i U_i^a U_i^b \tag{2.50}$$

This tensor now includes kinetic energy as well as rest mass.

Step 4: T^{ab} is a second rank tensor, and so the LHS of the field equation is second rank. There are two ways to form a 2nd rank tensor from the Riemann tensor:

a. Let $R_{bd} \equiv R^a{}_{bad}$ (Ricci Tensor), or

b. Let $\mathfrak{R} \equiv R^b{}_b = g^{bd} R_{bd}$ (Ricci Scalar), in which case $g_{ab}\mathfrak{R}$ is 2nd order.

The field equation therefore has the form

$$R^{ab} + C_1 g^{ab}\mathfrak{R} = C_2 T^{ab} \tag{2.51}$$

with C_1, C_2 constants.

Step 5: Mass - Energy Conservation implies that the stress-energy tensor has zero 4-divergence:

$$(\nabla_a T)^{ab} = 0, \tag{2.52}$$

just like charge conservation $\partial_a j^a = 0$.

Thus the left hand side should have zero divergence as well:

$$\nabla_a \left(R + C_1 g \mathfrak{R}\right)^{ab} = 0. \tag{2.53}$$

This condition sets $C_1 = -\frac{1}{2}$.

Step 6: Correspondence with Newtonian Gravity gives $C_2 = 8\pi G$, and so we get the *Einstein Field Equation*

$$\boxed{G^{ab} \equiv R^{ab} - \frac{1}{2} g^{ab}\mathfrak{R} = 8\pi G T^{ab}}. \tag{2.54}$$

Exercise 2.1

a. *The Riemann tensor satisfies $R_{abcd} = R_{cdab}$. Use this result to show that the Ricci tensor*

$$R_{ab} \equiv R^c{}_{acb} = g^{cd} R_{dacb} \tag{2.55}$$

is symmetric, $R_{ab} = R_{ba}$.

b. *Let $\mathfrak{R} \equiv R^a{}_a$ and $\mathfrak{T} \equiv T^a{}_a$. From the Einstein equation*

$$R^{ab} - \frac{1}{2} g^{ab}\mathfrak{R} = 8\pi G T^{ab} \tag{2.56}$$

derive the equation

$$\mathfrak{R} = -8\pi G \mathfrak{T} = -8\pi G T^a{}_a. \tag{2.57}$$

c. *Show that the Ricci tensor vanishes in empty space (where $T^{ab} = 0$).*

2.5 Spherically Symmetric Spacetimes

2.5.1 The General Form of the Metric

Here we derive a general form for spherically symmetric metrics such as the Schwarzschild metric and the Robertson–Walker metric used in cosmology. Suppose space is spherically symmetric. This means that two of the spatial coordinates (call them $\mathsf{X}^2 = \theta$ and $\mathsf{X}^3 = \phi$) behave like the usual coordinates on S^2. Let the other two coordinates be a time coordinate $\mathsf{X}^0 = t$ and a radial coordinate $\mathsf{X}^1 = \rho$ (here ρ may be chosen to be either η or r as described in section 2.1.2).

We assert that the general form for a spherically symmetric metric line element is

$$\mathrm{d}\tau^2 = A(\rho,t)\mathrm{d}t^2 - B(\rho,t)\mathrm{d}\rho^2 - C^2(\rho,t)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2), \tag{2.58}$$

with A, B, and C arbitrary functions. In terms of the metric tensor,

$$g_{ab} = \begin{pmatrix} A(\rho,t) & 0 & 0 & 0 \\ 0 & -B(\rho,t) & 0 & 0 \\ 0 & 0 & -C^2(\rho,t) & 0 \\ 0 & 0 & 0 & -C^2(\rho,t)\sin^2\theta \end{pmatrix}. \tag{2.59}$$

To demonstrate the validity of this form, we must show that this metric satisfies spherical symmetry, and also that it provides the most general form. First, for constant t and ρ, the line element is identical to that of a two-sphere of radius $C(\rho,t)$, i.e. spatial distances are given by

$$\mathrm{d}s^2 = C^2(\rho,t)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2). \tag{2.60}$$

Furthermore, the metric is independent of ϕ, and only depends on θ in the $g_{33} = -C\sin^2\theta$ term. Thus rotations in θ and ϕ preserve the spherical symmetry.

Have we left something out which could modify the metric but still preserve spherical symmetry? The functions A, B, and C can only depend on ρ and t, so they are as general as possible. However, we must still demonstrate that off-diagonal components of the metric all vanish. Consider for example g_{13}. Suppose that $g_{13} \neq 0$ and consider a small step from (t,ρ,θ,ϕ) to $(t,\rho+\delta\rho,\theta,\phi+\delta\phi)$. This step has a net spatial line element given by

$$\mathrm{d}s^2 = B(\rho,t)(\delta\rho)^2 + 2g_{13}\delta\rho\delta\phi + C^2(\rho,t)(\delta\phi)^2. \tag{2.61}$$

Consider two such steps, one where $\delta\phi > 0$ (eastward) and one where $\delta\phi < 0$ (westward). The $\mathrm{d}s^2$ calculated for the two steps will differ by

$$\Delta\mathrm{d}s^2 = 4g_{13}\delta\rho\,\delta\phi. \tag{2.62}$$

But by symmetry eastward and westward steps should give the same result. Thus we must have $g_{13} = 0$. Similar considerations make all the off-diagonal components vanish, except for g_{01}, which does not involve angles so must be handled separately.

A step from (t,ρ,θ,ϕ) to $(t+\delta t,\rho+\delta\rho,\theta,\phi)$ involves the g_{01} coefficient. Here time reversal symmetry will be needed to remove g_{01}. If a step forward in time ($\delta t > 0$) is assumed to have the same proper time squared $\mathrm{d}\tau^2$ as a step backward in time ($\delta t < 0$), then we need $g_{01} = 0$. Thus the vanishing of this component requires more than simply the assumption of spherical symmetry. Nevertheless we shall assume time reversal symmetry in all that follows.

2.5.2 Derivation of the Schwarzschild Metric

The Schwarzschild metric is static, so the functions A, B, and C are independent of t. We will choose the radial coordinate to be $\rho = r$, i.e. the definition of radius which fits the area formula $Area = 4\pi r^2$. Thus $C = r$, and

$$\mathrm{d}\tau^2 = A(r)\mathrm{d}t^2 - B(r)\mathrm{d}r^2 - r^2\mathrm{d}\Omega^2; \qquad \mathrm{d}\Omega^2 \equiv \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2. \tag{2.63}$$

We must now find the two functions $A(r)$ and $B(r)$. The Schwarzschild metric applies to the spacetime *outside* a spherically symmetric mass distribution (such as a planet, a star, or a black hole). The mass density outside is zero, so the stress-energy tensor T^{ab} vanishes. By the Einstein field equation, the Einstein tensor also vanishes:

$$G^{ab} = 0. \tag{2.64}$$

The task, then, is to calculate the Einstein tensor for the metric given by equation (2.63). Setting the components equal to zero then gives differential equations for $A(r)$ and $B(r)$. This can be readily done with a computer algebra program; one finds that the Einstein tensor is diagonal. The 0–0 component is

$$G_{00} = \frac{A}{r^2B^2}\left(B^2 - B + rB'\right) \qquad \left(B' = \frac{\mathrm{d}B}{\mathrm{d}r}\right). \tag{2.65}$$

Setting this to zero gives the differential equation

$$B^2 - B + rB' = 0 \tag{2.66}$$

with solution

$$B = \frac{r}{r - C_1} = \left(1 - \frac{C_1}{r}\right)^{-1}, \tag{2.67}$$

where C_1 is a constant of integration. The 1–1 component gives

$$G_{11} = \frac{1}{r^2A}\left(A - AB + rA'\right) = 0. \tag{2.68}$$

Using the solution for B leads to the solution

$$A = \frac{C_2}{B} = C_2\left(1 - \frac{C_1}{r}\right), \tag{2.69}$$

with C_2 another constant of integration. The G_{22} and G_{33} components add nothing new: they vanish when these two solutions are used.

We need to determine the constants of integration. First, for large r we should recover the Minkowski metric for flat space. In other words,

$$\lim_{r\to\infty} A(r) = 1 \qquad \Rightarrow \qquad C_2 = 1. \tag{2.70}$$

Finally, correspondence with Newtonian physics gives

$$C_1 = r_s = 2GM, \tag{2.71}$$

where M is the net mass of the central object. This can be seen just by deriving the orbit equations as we have already done.

Chapter 3

Cosmological Models

Recall some definitions:

- *Homogeneous* - A physical system is homogeneous if it is invariant to translation $X^a \to X^a + X^c$ where X^a =const. (looks the same everywhere).
- *Isotropic* - Invariant to rotations (looks the same in all directions)

We will assume that, over large enough scales, the universe looks the same everywhere. Thus we assume the universe is homogeneous in space but not in time (If the universe were homogeneous in time then it could not evolve). The scale where the universe looks uniform is somewhat controversial, but is probably at least 100 Mpc = 10^8 pc [1]. In other words, observers find that the number of galaxies in any volume of size (100 Mpc^3) will be roughly the same.

We will also assume the universe is isotropic. This requires a special set of observers. An observer moving rapidly relative to surrounding galaxies will not see isotropy. Light from galaxies observed in the direction of motion will acquire a blueshift and light from the opposite direction a redshift (in addition to the cosmological redshift discussed below). The special set of observers who see an isotropic universe defines a special reference frame:

- *Cosmological reference frame (cosmic frame)* A set of coordinates in which cosmological quantities are homogeneous and isotropic.
- *Comoving observer* An observer at rest in the cosmic frame.
- *Cosmic time* The proper time t measured by a comoving observer, starting with $t = 0$ at the big bang.

The microwave background provides the best method of determining the cosmological reference frame. In this frame the microwave background should be perfectly isotropic, apart from small scale fluctuations associated with primeval galaxies. In fact, we observe

[1] 1pc = 1 parsec = 3.26 light years

a large scale (dipole) fluctuation in the temperature of the microwaves due to the *proper motion* of the Milky Way with respect to the cosmic frame. We are travelling some $600\,\mathrm{km\,s^{-1}}$ in the direction of the Virgo cluster; microwaves observed in this direction are blue-shifted to higher frequencies.

3.1 The Robertson–Walker Metric

The homogeneity hypothesis implies that the universe expands uniformly, i.e. at the same rate everywhere. The universe at one cosmic time defines a three dimensional manifold. The shape of this manifold does not change in time; it just gets larger. In other words, distances grow larger according to some function $a(t)$, which is independent of position. We will call $a(t)$ the *expansion parameter*. We assume $a(t) \geq 0$ to avoid negative spatial distances.

Let us go back to the general form for a spherically symmetric metric, equation (2.59). We need to find the functions A, B, and C.

- A: Recall the condition that the clocks of comoving observers measure coordinate time t, i.e. $\mathrm{d}\tau = \mathrm{d}t$. For these observers the general form gives $\mathrm{d}\tau^2 = A\mathrm{d}t^2$, so $A = 1$.

- B: We will choose radial coordinate $\rho = \eta$ (see §2.1.2). The spatial distance from the coordinate origin to a sphere labelled by η scales with $a(t)$; more precisely, this distance is $s = a(t)\eta$. The choice of η as radial coordinate is convenient for solving the field equations in order to find $a(t)$; later, however, we will transform to the r coordinate when we discuss astronomical distance measures. Now, if $s = a(t)\eta$ then $\mathrm{d}s^2 = a^2(t)\mathrm{d}\eta^2$, which tells us that $B = a^2(t)$.

- C: The area of a sphere at radius η should scale with $a^2(t)$. Thus the effective radius is $C = a(t)r$, where $r = F(\eta)$ for some function $F(\eta)$. To keep the effective radius positive, we require $F(\eta) > 0$.

With these considerations, the metric line element takes the form (the *Robertson–Walker metric*)

$$\boxed{\mathrm{d}\tau^2 = \mathrm{d}t^2 - a^2(t)\left(\mathrm{d}\eta^2 + F^2(\eta)\mathrm{d}\Omega^2\right)}\,. \tag{3.1}$$

We now only need to determine the function $F(\eta)$. By homogeneity, the Ricci scalar curvature should be a function only of time. The scalar curvature of the Robertson–Walker metric is

$$\mathfrak{R}(t) = \frac{2}{F^2 a^2(t)}\left(1 - 2FF'' - F'^2\right). \tag{3.2}$$

By homogeneity, $\mathfrak{R}(t)$ should be independent of η. Also, we see that $\mathfrak{R}(t)\,a^2(t)$ is independent of time. With some foresight, we define the constant

$$k \equiv \frac{\mathfrak{R}(t)\,a^2(t)}{6}\,. \tag{3.3}$$

There are three general solutions of equation (3.2), depending on whether the curvature is positive, negative, or zero:

$$F(\eta) = \begin{cases} |k|^{-1/2} \sin |k|^{1/2}\eta & k > 0 \\ \eta & k = 0 \\ |k|^{-1/2} \sinh |k|^{1/2}\eta & k < 0. \end{cases} \tag{3.4}$$

We can clean up the expressions for $k \neq 0$ by rescaling $a(t)$ and η. Define

$$\widetilde{a}(t) = |k|^{-1/2} a(t), \tag{3.5}$$

$$\widetilde{\eta} = |k|^{1/2}\eta. \tag{3.6}$$

With these rescalings, k becomes ± 1. For $k = +1$, $F(\widetilde{\eta}) = \sin\widetilde{\eta}$, while for $k = -1$, $F(\widetilde{\eta}) = \sinh\widetilde{\eta}$.

Note that both of our geometric measures of radius survive this rescaling. Recall again that $s = a(t)\eta$ gives the geodesic distance from the coordinate origin to the sphere labelled by η. This distance has the same form, i.e. $s = \widetilde{a}(t)\widetilde{\eta}$. Similarly the effective radius corresponding to the area of the sphere becomes $\widetilde{a}(t)F(\widetilde{\eta})$.

We now have the final form of the Robertson–Walker metric. To simplify the notation, let us drop all the tildes (wiggles), and just write a and η instead of $\widetilde{a}$ and $\widetilde{\eta}$.

To summarize, the constant $k = -1, 0, 1$ gives the sign of the Ricci curvature

$$\mathfrak{R}(t) = \frac{6k}{a^2(t)}, \tag{3.7}$$

and we can express the function F as

$$F(\eta) = \begin{cases} \sin\eta & k = +1 \\ \eta & k = 0 \\ \sinh\eta & k = -1 \end{cases}. \tag{3.8}$$

3.2 Examples of Homogeneous Isotropic Spaces

Let us consider a spatial slice of the Robertson–Walker manifold, i.e. the universe at one cosmic time. Spatial distances are governed by the line element

$$\mathrm{d}s^2 = a^2(t)\left(\mathrm{d}\eta^2 + F^2(\eta)\mathrm{d}\Omega^2\right). \tag{3.9}$$

The S^3 Universe

One example of a manifold with the $k = +1$ metric is the 3-sphere S^3.

We can picture the 3-sphere in terms of an imbedding into four dimensional Euclidean space $\mathbb{E}^4$:

$$S^3 = \left\{(x, y, z, w) \in \mathbb{R}^4 | \, x^2 + y^2 + z^2 + w^2 = a^2\right\}. \tag{3.10}$$

Here (x, y, z, w) are Cartesian coordinates for $\mathbb{E}^4$. We will recover the spatial part of the Robertson–Walker metric by converting to *hyperspherical coordinates*. First define an angle η by

$$w = a \cos\eta, \qquad 0 \le \eta \le \pi. \tag{3.11}$$

Then

$$\begin{aligned} x^2 + y^2 + z^2 &= a^2 - w^2 &(3.12)\\ &= a^2(1 - \cos^2\eta) &(3.13)\\ &= a^2 \sin^2\eta. &(3.14) \end{aligned}$$

Thus for $\eta =$ constant we have an ordinary 2-sphere S^2 with radius $a \sin\eta$. For this 2-sphere, we can write down the usual spherical coordinates:

$$\begin{aligned} z &= (a \sin\eta)\cos\theta, &(3.15)\\ x &= (a \sin\eta)\sin\theta\cos\phi, &(3.16)\\ y &= (a \sin\eta)\sin\theta\sin\phi. &(3.17) \end{aligned}$$

If we constrain the Euclidean line element

$$\mathrm{d}s^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 + \mathrm{d}w^2 \tag{3.18}$$

by the condition $x^2 + y^2 + z^2 + w^2 = a^2$, we recover the S^3 line element

$$\mathrm{d}s^2 = a^2(t)\left(\mathrm{d}\eta^2 + \sin^2(\eta)\,\mathrm{d}\Omega^2\right) . \tag{3.19}$$

Exercise 3.1 *The volume of a 3-manifold described by the metric g_{ab} and coordinates x^1, x^2, x^3 is given by*

$$\mathcal{V} = \int\int\int \sqrt{|\mathrm{det}g|}dx^1dx^2dx^3 \tag{3.20}$$

a. *What is the volume of a 3-sphere S^3 with radius a?*

b. *For a $k = -1$ universe the angle η is unbounded: $0 \le \eta < \infty$. Thus the volume of the universe is infinite (in the simply connected topology). However, what is the volume integrated over the range $0 \le \eta \le \pi$? What is the ratio of this volume to that of the $k = +1$ (S^3) universe?*

The $k = 0$ Universe

One possibility for the spatial part of the $k = 0$ universe is simply Euclidean 3-space $\mathbb{E}^3$. Another possibility is the 3-torus $\mathbb{T}^3$.

The $k = -1$ Universe

Here the simplest example, $\mathbb{H}^3$, is called the hyperboloid or pseudo-sphere.

3.3 The Stress Energy Tensor

How do we find $a(t)$, the cosmic expansion parameter? Recall the Einstein Field equations:

$$\underbrace{G^{ab} = R^{ab} - 1/2g^{ab}\mathfrak{R}}_{\text{Metric and its derivatives (1st and second)}} = \underbrace{8\pi G T^{ab}}_{\text{source of energy}} \tag{3.21}$$

where

- G^{ab} is the Einstein tensor;
- R^{ab} is the Ricci tensor;
- $\mathfrak{R} = R^a_a = g_{ab}R^{ab}$ is the Ricci scalar;
- T^{ab} is the stress energy tensor.

The components of the stress energy tensor can be given physical interpretations:

$$T^{ab} = \begin{pmatrix} b=0 & & b=1,2,3 & \\ \text{energy density} & \vdots & \text{energy flux} & a=0 \\ \cdots\cdots\cdots & \cdots & \cdots\cdots\cdots & \\ \text{momentum density} & \vdots & \text{momentum flux} & a=1,2,3 \end{pmatrix}. \tag{3.22}$$

For a set of particles of mass m

First we call T^{00} the energy density ρ. Nothing is truly at rest in space-time; if a lump of matter just sits around with zero three-velocity, it still moves forward into the future. Thus $\rho = T^{00}$ could be called the flux of energy in the t direction. Continuing along the top row, T^{01} is the flux of energy in the x direction, and so on. In the left-most column, T^{10} gives the density of x momentum. From equation (2.50), this is equal to T^{01}, the flux of energy density in the x direction.

Let us now look at the components of the tensor that have two spatial indices. In fluid mechanics and solid mechanics the pressure tensor is defined to be $P^{ij} = T^{ij}$ for $i, j = 1, 2, 3$. For example,

- $P^{33} = z$ momentum transferred in z direction
- $P^{23} = y$ momentum transferred in z direction
- e.t.c.

3.3.1 The Stress Energy Tensor for an Isotropic Medium

In an isotropic medium, all off-diagonal components vanish. e.g. if $P^{23} > 0$ then the y momentum travels in the *positive* z direction, $\Rightarrow$ positive z is different from negative z. Also, $P^{11} = P^{22} = P^{33} \equiv p$ otherwise x direction might be different from y or z.

$$\Rightarrow P^{ij} = \begin{pmatrix} p & 0 & 0 \\ 0 & p & 0 \\ 0 & 0 & p \end{pmatrix} \tag{3.23}$$

Finally, momentum density must vanish, otherwise if x momentum > 0 this distinguishes $+x$ from $-x$. As energy flux = momentum density, it also vanishes. Thus, for an isotropic medium

$$\boxed{T^{ab} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix}}\,. \tag{3.24}$$

3.4 The Evolution Equations

Note that the stress energy tensor derived above is diagonal. Fortunately, the Einstein tensor for the Robertson–Walker metric is also diagonal, so there is no inconsistency in the Einstein Field Equations. The 0–0 component gives

$$\frac{3(k + \dot{a}^2)}{a^2} = 8\pi G\rho. \tag{3.25}$$

The 1–1, 2–2, and 3–3 components all say the same thing:

$$-(k + \dot{a}^2 + 2a\ddot{a}) = 8\pi G p a^2. \tag{3.26}$$

It will be convenient to recast the second equation in a form reminiscent of the first law of thermodynamics. First we state the end result. The expansion of the universe is governed by two evolution equations:

a.

$$\boxed{\dot{a}^2 + k = \frac{8\pi G}{3}\rho a^2}\,; \tag{3.27}$$

b.

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}a}(\rho a^3) = -3pa^2}\,. \tag{3.28}$$

The first follows immediately from the 0–0 equation. The second requires some work. First,

$$\begin{aligned} k + \dot{a}^2 &= \frac{8\pi G}{3}\rho a^2 && (3.29) \\ \Rightarrow \frac{8\pi G}{3}\rho a^2 + 2R\ddot{a} &= -8\pi G p R^2, && (3.30) \end{aligned}$$

which gives us an expression for $\ddot{a}$:

$$\boxed{\ddot{a} = -\frac{4\pi G}{3}(\rho + 3p)\,a}\,. \tag{3.31}$$

Next,

$$\begin{aligned} \ddot{a} &= \frac{\mathrm{d}\dot{a}}{\mathrm{d}t} = \frac{\mathrm{d}\dot{a}}{\mathrm{d}a}\frac{\mathrm{d}a}{\mathrm{d}t} = \dot{a}\frac{\mathrm{d}\dot{a}}{\mathrm{d}a} && (3.32) \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}a}\left(\dot{a}^2\right) && (3.33) \\ &= \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}a}\left(\frac{8\pi G}{3}\rho a^2 - k\right) && (3.34) \\ &= \frac{\mathrm{d}}{\mathrm{d}a}\left(\frac{4\pi G}{3}\rho a^2\right), && (3.35) \end{aligned}$$

since k is a constant. So by equation (3.31),

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}a}\left(\rho a^2\right) &= a^2\frac{\mathrm{d}\rho}{\mathrm{d}a} + 2a\rho = -(\rho + 3p)\,a && (3.36) \\ \Rightarrow a^2\frac{\mathrm{d}\rho}{\mathrm{d}a} + 3a\rho = -3pa. && && (3.37) \end{aligned}$$

Multiply by a and rearrange:

$$a^3\frac{\mathrm{d}\rho}{\mathrm{d}a} + 3a^2\rho = \frac{d}{da}(\rho a^3) = -3pa^2. \tag{3.38}$$

We now have the final form.

3.5 Cosmological Quantities

It will be useful to define:

a. Hubble parameter

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} \tag{3.39}$$

b. Critical density

$$\rho_c(t) \equiv \frac{3}{8\pi G} H^2(t) \tag{3.40}$$

c. Acceleration parameter

$$Q(t) \equiv -q(t) \equiv \frac{-a\ddot{a}}{\dot{a}^2} \tag{3.41}$$

d. Omega

$$\Omega(t) \equiv \frac{\rho(t)}{\rho_c(t)} \tag{3.42}$$

Also, the present time is $t = t_0$. All quantities with the subscript 0 are evaluated at the present time, e.g. $H_0 = H(t_0)$.

3.5.1 Evolution in terms of cosmological quantities

We can recast the first evolution equation in terms of H_0 and Ω_0. The ratio $8\pi G/3$ is constant, so

$$\frac{H^2(t)}{\rho_c(t)} = \frac{H_0^2}{\rho_{c0}}. \tag{3.43}$$

Thus equation (3.27) becomes

$$\dot{a}^2 + k = \left(\frac{H_0^2}{\rho_{c0}}\right) \rho a^2 \tag{3.44}$$

$$= H_0^2 \frac{\rho}{\rho_0} \frac{\rho_0}{\rho_{c0}} a^2, \tag{3.45}$$

i.e.

$$\boxed{\dot{a}^2 + k = H_0^2 \Omega_0 \frac{\rho}{\rho_0} a^2}. \tag{3.46}$$

If we divide by a^2,

$$H^2(t) + \frac{k}{a^2(t)} = H_0^2 \Omega_0 \frac{\rho}{\rho_0}. \tag{3.47}$$

Ω_0 and k

Suppose we evaluate this equation at $t = t_0$:

$$H_0^2 + \frac{k}{a_0^2} = H_0^2 \Omega_0 \frac{\rho}{\rho_0}, \tag{3.48}$$

$$\Rightarrow k = a_0^2 H_0^2 (\Omega_0 - 1). \tag{3.49}$$

Since $a_0^2 H_0^2$ is positive,

$$k > 0 \Leftrightarrow \Omega_0 > 1 \tag{3.50}$$

$$k = 0 \Leftrightarrow \Omega_0 = 1 \tag{3.51}$$

$$k < 0 \Leftrightarrow \Omega_0 < 1. \tag{3.52}$$

This shows ρ_c really is the "critical density"; by the definition of Ω,

$$\Omega_0 = 1 \Leftrightarrow \rho_0 = \rho_{c0}. \tag{3.53}$$

We can also obtain an expression for $\Omega(t)$ from equation (3.49). Rearranging, we obtain

$$\Omega_0 = 1 + \frac{k}{a_0^2 H_0^2} = 1 + \frac{k}{\dot{a}_0^2}. \tag{3.54}$$

Suppose the trilobites studied cosmology 500 million years ago. They would obtain this same equation, but with their t_0 dated 500 million years before ours. In other words, this equation holds true for any time t:

$$\boxed{\Omega(t) = 1 + \frac{k}{\dot{a}(t)^2}}. \tag{3.55}$$

The expansion parameter for $k \neq 0$

If $k \neq 0$ then we can find an expression for a_0. Taking the absolute value of equation (3.49) gives

$$\boxed{a_0 = \frac{1}{H_0 |\Omega_0 - 1|^{1/2}}}. \tag{3.56}$$

3.6 Matter dominated solutions ($p = 0$)

There are many different sources of energy in the universe; ordinary matter is only one of them. For sources such as radiation and vacuum energy, the pressure p is important. But a fluid or cloud of particles has pressure

$$p \approx \rho V^2, \tag{3.57}$$

which is negligible compared to ρ if $V \ll 1$ (i.e. if the molecules move at non-relativistic velocities).

The most relevant example for cosmology is a collection of galaxies. We treat each galaxy as a single molecule in a fluid. The typical velocity of a galaxy is roughly $300\,\mathrm{km\,s^{-1}} = 10^{-3}c$, so $p \approx 10^{-6}\rho$. If the galaxies provide the dominant contribution to the total mass-energy density of the universe, then $p \approx 0$.

The second evolution equation then tells us

$$\frac{\mathrm{d}}{\mathrm{d}a}(\rho a^3) \approx 0, \tag{3.58}$$

implying that ρ decreases as the cube of a,

$$\rho \sim a^{-3}. \tag{3.59}$$

As $\rho = \rho_0$ at $a = a_0$, the precise expression is

$$\boxed{\rho = \rho_0 \left(\frac{a}{a_0}\right)^{-3}}. \tag{3.60}$$

Physically, ρ measures mass per unit volume. For matter, the mass remains constant as the universe expands, but the volume increases as a^3.

The first evolution equation becomes

$$\boxed{\dot{a}^2 + k = (H_0^2 \Omega_0 a_0^3)\, a^{-1}}. \tag{3.61}$$

3.6.1 Qualitative Description

Rewrite equation (3.61) as:

$$\frac{1}{2}\dot{a}^2 - \left(\frac{H_0^2 \Omega_0 a_0^3}{2}\right) a^{-1} = -\frac{k}{2} \tag{3.62}$$

Let

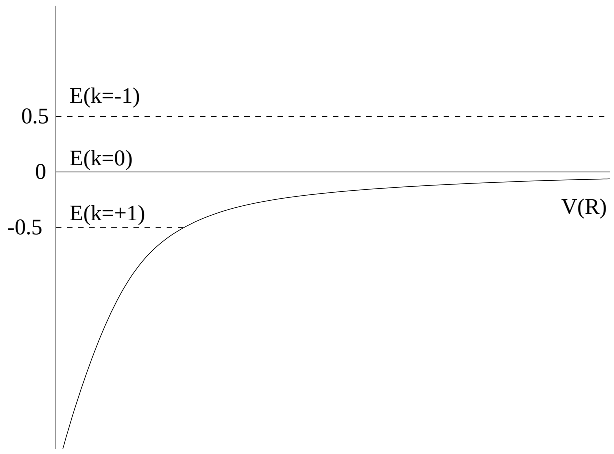
$$K = 1/2\dot{a}^2 \quad \text{Kinetic Energy} \tag{3.63}$$

$$V(a) = -\left(\frac{H_0^2 \Omega_0 a_0^3}{2}\right) a^{-1} \quad \text{Potential Energy} \tag{3.64}$$

$$E = \frac{-k}{2} \quad \text{Total Energy} \tag{3.65}$$

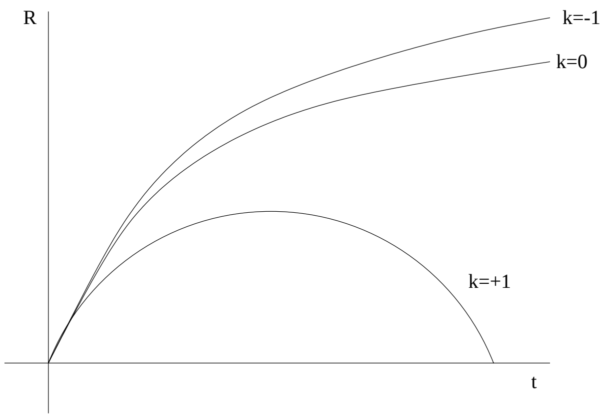
Then the evolution equation resembles a simple energy equation in mechanics:

$$K + V = E. \tag{3.66}$$



Now, if $K = E - V > 0$ then $\dot{a}^2 > 0$, so the universe is expanding (or contracting). Meanwhile, $E = V$ implies $\dot{a} = 0$, so that the expansion has stopped (perhaps only at one moment). Furthermore $E - V < 0$ is impossible, as K cannot be negative. For $k = -1$ the energy is always greater than $V(a)$, so the universe expands forever. For $k = 0$ the universe also expands forever, but the expansion asymptotically slows to a stop as $a \to \infty$.

For the $k = +1$ universe, the universe reaches a maximum radius where $E = V(a_{\text{max}})$. However, the second derivative $\ddot{a} < 0$ (see equation (3.31)) so the universe begins to contract, eventually reaching $a = 0$ in the *big crunch*.



3.6.2 The Flat Universe Solution

For the flat universe model, $k = 0$ and $\Omega_0 = 1$, so

$$\dot{a}^2 = (H_0^2 a_0^3) a^{-1} \tag{3.67}$$

If we let $X = a/a_0$, then

$$\boxed{\dot{X} = H_0 X^{-1/2}}. \tag{3.68}$$

This equation is integrable:

$$
\begin{aligned}
X^{1/2}\,\mathrm{d}X &= H_0\,\mathrm{d}t && (3.69)\\
\frac{2}{3}X^{3/2} &= H_0 t && (3.70)\\
\Rightarrow t &= \left(\frac{2}{3H_0}\right)X^{3/2} && (3.71)
\end{aligned}
$$

or

$$\boxed{t = \left(\frac{2}{3H_0}\right)\left(\frac{a}{a_0}\right)^{3/2}}, \tag{3.72}$$

and

$$\boxed{a = a_0\left(\frac{3H_0 t}{2}\right)^{2/3}}. \tag{3.73}$$

Evaluate at $t = t_0$ and $a = a_0$:

$$\boxed{t_0 = \frac{2}{3H_0}} \tag{3.74}$$

A measurement of H_0 can tell us the age of the universe! It is usual to write

$$H_0 = h(100\,\mathrm{km\,s^{-1}}Mpc^{-1}) \tag{3.75}$$

where h is dimensionless. We can write H_0^{-1} in terms of years (y):

$$
\begin{aligned}
100\,\mathrm{km\,s^{-1}}Mpc^{-1} &= \frac{100\,\mathrm{km\,s^{-1}}}{3.26\times 10^6\,\ell\,\mathrm{y}}\\
&= \frac{100\,\mathrm{km\,s^{-1}}}{3.26\times 10^6(3\times 10^5\,\mathrm{km\,s^{-1}\,y})}\\
&= \frac{100}{(3.26\times 10^6)(3\times 10^5\,\mathrm{y})}\\
&= \frac{1}{9.78\times 10^9\,\mathrm{y}}
\end{aligned}
$$

so within 2.2%:

$$\boxed{H_0^{-1} = h^{-1}10^{10}\text{years}}. \tag{3.76}$$

Recent estimates give $h \approx 0.67 \approx 2/3$. For this value, $t_0 \approx 10$ eons or 10 billion years. As the oldest stars are thought to be some 12-13 billion years old, something is wrong with our model! Either $k \neq 0$ or the universe is not matter-dominated, or both.

3.7 The development Angle

In order to solve the $k = \pm 1$ cases, we will need a change of variable. Let

$$\xi \equiv \int_0^t \frac{dt'}{a(t')} \tag{3.77}$$

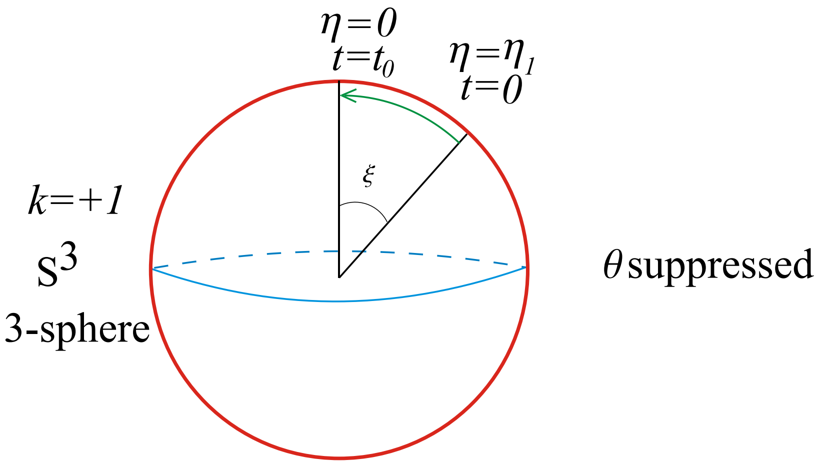
be the *horizon coordinate*, also known as the *development angle*. Note that

$$\mathrm{d}\xi = \frac{\mathrm{d}t}{a(t)} \tag{3.78}$$

$$\Rightarrow a(t) = \frac{\mathrm{d}t}{\mathrm{d}\xi}. \tag{3.79}$$

Physical interpretation

The development angle tells us the net coordinate distance in η that a photon has travelled since the big bang. Consider a photon emitted at radial coordinate η_1 and time coordinate $t_1 = 0$. The photon travels toward us in the negative η direction, until it is observed in our telescopes at $\eta = 0$ and $t = t_0$.



Photons have zero proper time, so

$$0 = \mathrm{d}\tau^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}\eta^2 \tag{3.80}$$

$$\Rightarrow \mathrm{d}t = \pm a(t)\,\mathrm{d}\eta. \tag{3.81}$$

The sign depends on whether the photon is increasing or decreasing its η coordinate. For a photon moving toward us, $\mathrm{d}\eta < 0$ so $\mathrm{d}t = -a(t)\,\mathrm{d}\eta$. Thus

$$\boxed{\mathrm{d}\eta = -\frac{\mathrm{d}t}{a(t)} = -\,\mathrm{d}\xi}. \tag{3.82}$$

Let us integrate this equation from emission ($\eta = \eta_1$, $t = 0$) to absorption ($\eta = 0$, $t = t_0$):

$$-\eta_1 = \int_{\eta_1}^{0} \mathrm{d}\eta = -\int_0^{t_0} \frac{\mathrm{d}t}{a(t)} = -\xi(t_0). \tag{3.83}$$

In other words

$$\boxed{\eta_1 = \xi_0}. \tag{3.84}$$

This demonstrates that ξ_0 gives the η coordinate of an object so far away that its light has taken the entire age of the universe to reach us. For a three sphere S^3 the coordinate η is an angle, hence the term development angle. But also, ξ_0 tells us the coordinate of the horizon: objects with $\eta > \xi_0$ are too far away for us to see them; possibly our descendants may see them at some future time.

Note also that we can invert the relation between ξ and time:

$$t(\xi) = \int_0^{\xi} a \mathrm{d}\xi'. \tag{3.85}$$

3.8 The $k = +1$ Universe

$$\dot{a}^2 + k = H_0^2 \Omega_0 \left(\frac{\rho}{\rho_0}\right) a^2, \qquad \rho = \rho_0 (a/a_0)^{-3} \tag{3.86}$$

$$= (H_0^2 \Omega_0 a_0^3) a^{-1}. \tag{3.87}$$

Let $k = +1$, and gather together the constants in C, where by equation (**??**),

$$C = H_0^2 \Omega_0 a_0^3 = \frac{\Omega_0}{H_0 |\Omega_0 - 1|^{3/2}}. \tag{3.88}$$

$$\boxed{\dot{a}^2 + 1 = \frac{C}{a}}. \tag{3.89}$$

Now,

$$\dot{a} = \frac{\mathrm{d}a}{\mathrm{d}t} = \frac{\mathrm{d}a}{\mathrm{d}\xi}\frac{\mathrm{d}\xi}{\mathrm{d}t} \tag{3.90}$$

$$= \frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}\xi} = \frac{a'}{a}, \tag{3.91}$$

where a prime denotes differentiation with respect to ξ: $a' \equiv \mathrm{d}a/\mathrm{d}\xi$.. The evolution equation becomes

$$\frac{a'^2}{a^2} + 1 = \frac{C}{a}, \tag{3.92}$$

or

$$a'^2 + a^2 = C a. \tag{3.93}$$

This nonlinear first order equation can be made linear by differentiation:

$$2a'a'' + 2aa' = Ca' \tag{3.94}$$

$$\Rightarrow a'' + a = \frac{C}{2}. \tag{3.95}$$

The complementary functions are Sin and Cosine, with particular integral $C/2$. The general solution is:

$$a(\xi) = \frac{C}{2} + A\cos\xi + B\sin\xi. \tag{3.96}$$

We need initial conditions to determine A and B. First, at the big bang, $a(0) = 0$, so

$$A = -\frac{C}{2} = -\frac{\Omega_0}{2H_0|\Omega_0 - 1|^{3/2}}. \tag{3.97}$$

Next, for the second initial condition we must go back to the first order equation, $a'^2 + a^2 = Ca$. At $\xi = 0$,

$$a'^2(0) + 0 = 0 \tag{3.98}$$

$$\Rightarrow B = 0 \tag{3.99}$$

$$\Rightarrow a(\xi) = \frac{C}{2}(1 - \cos\xi). \tag{3.100}$$

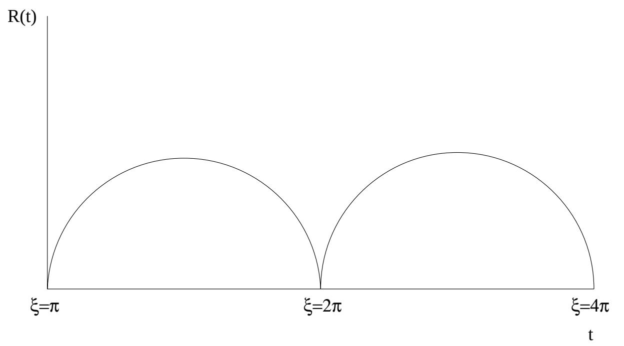
We now have our solution:

$$\boxed{a(\xi) = \frac{\Omega_0}{2H_0|\Omega_0 - 1|^{3/2}}(1 - \cos\xi)}. \tag{3.101}$$

Note that $a(2\pi) = 0$! Thus when $\xi = 2\pi$ the universe has collapsed in a big crunch. Recall that $\xi(t)$ gives the net change in η along a photon path. For a photon traversing the three-sphere, η goes from 0 to π and back again to 0 (just like a geodesic on a two-sphere going from the North pole ($\theta = 0$ to the South pole and back again). Thus a photon can only circumnavigate the universe once.

Note that a reaches its max value at $\xi = \pi$:

$$a_{\max} = \frac{\Omega_0}{H_0|\Omega_0 - 1|^{3/2}}. \tag{3.102}$$



To find $t(\xi)$ requires a simple integration:

$$t(\xi) = \int a\,\mathrm{d}\xi = \frac{\Omega_0}{2H_0|\Omega_0 - 1|^{3/2}}(\xi - \sin\xi). \tag{3.103}$$

In order to find the age of the universe t_0, we need to find ξ_0. To begin, we have *two* expressions for a_0 (we drop the absolute value signs on $\Omega_0 - 1$ since $\Omega_0 > 1$:

$$a_0 = \frac{1}{H_0\left(\Omega_0 - 1\right)^{1/2}} \tag{3.104}$$

$$= \frac{\Omega_0}{2H_0(\Omega_0 - 1)^{3/2}}\left(1 - \cos\xi_0\right). \tag{3.105}$$

or

$$(\Omega_0 - 1) = \frac{\Omega_0}{2}\left(1 - \cos\xi_0\right). \tag{3.106}$$

Solving for $\cos\xi_0$ gives

$$\cos\xi_0 = \frac{2 - \Omega_0}{\Omega_0}. \tag{3.107}$$

Example $\Omega_0 = 2$. Here $\cos\xi_0 = 0$ so $\xi_0 = \pi/2$ or $3\pi/2$. The latter corresponds to the collapsing phase of the universe. For $\xi_0 = \pi/2$ we are halfway (in ξ) to the maximum size of the universe. The calculation of t_0 gives

$$t_0 = \frac{1}{H_0}\left(\frac{\pi}{2} - \sin\frac{\pi}{2}\right) = 0.57H_0^{-1} \approx 5.7h \times 10^9\,\mathrm{y}. \tag{3.108}$$

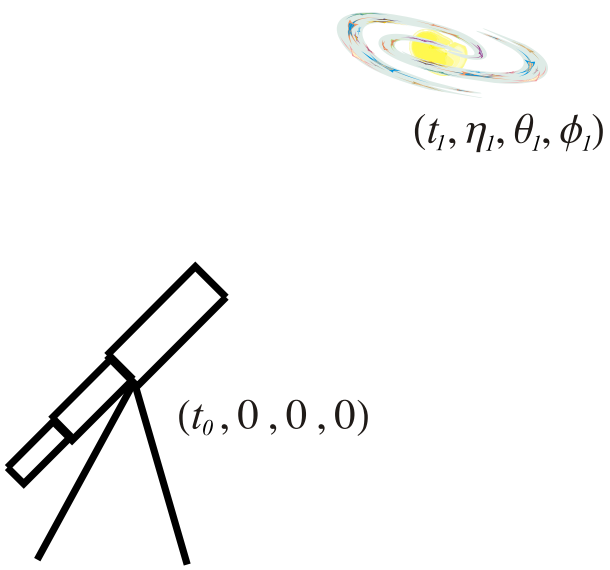
Thus if $h = 2/3$ then $t_0 \approx 8.5$ billion years, which is certainly too small.

Exercise 3.2 *Find $a(\xi)$ and $t(\xi)$ for a $k = -1$ matter-dominated universe with parameters H_0 and Ω_0.*

Chapter 4

Observations in an Expanding Universe

4.1 Redshift



For photons

$$\begin{aligned} 0 = \mathrm{d}\tau^2 &= \mathrm{d}t^2 - a(t)\,\mathrm{d}\eta^2 && (4.1) \\ \Rightarrow \mathrm{d}\eta &= -\frac{\mathrm{d}t}{a(t)}. && (4.2) \end{aligned}$$

The negative sign arises because the photons are moving in the negative η direction.

Integrate:

$$-\int_{\eta}^{0} \mathrm{d}\eta = \int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)} \tag{4.3}$$

$$\Rightarrow \eta_1 = \int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)}. \tag{4.4}$$

Suppose the earth and the galaxy are at rest w.r.t. the cosmic frame, so we stay at $\eta = 0$, and the galaxy stays at $\eta = \eta_1$. Also suppose another photon is emitted after a period of oscillation δt_1 ($\delta t_1 = 1/\nu_1$). Then

$$\eta_1 = \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{\mathrm{d}t}{a(t)}. \tag{4.5}$$

Take the difference between the two expressions for η_1:

$$0 = \int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)} - \int_{t_1+\delta t_1}^{t_0+\delta t_0} \frac{\mathrm{d}t}{a}. \tag{4.6}$$

$$= \int_{t_1}^{t_1+\delta t_1} \frac{dt}{a} - \int_{t_0}^{t_0+\delta t_0} \frac{dt}{a}. \tag{4.7}$$

For visible light δt_0, $\delta t_1 \sim 10^{-14}$ s. The universe does not expand much in such a small time interval, so we can treat a as a constant in the integral. Thus

$$0 = \frac{1}{a(t)}\delta t_1 - \frac{1}{a(t)}\delta t_0. \tag{4.8}$$

We can now state how expansion affects the properties of the light wave:

- Period:
$$\frac{\delta t_0}{\delta t_1} = \frac{a_0}{a_1}. \tag{4.9}$$

- Frequency ($\nu = 1/\delta t$):
$$\frac{\nu_0}{\nu_1} = \frac{a_1}{a_0}. \tag{4.10}$$

- Wavelength ($\lambda = c\delta t$):
$$\frac{\lambda_0}{\lambda_1} = \frac{a_0}{a_1}. \tag{4.11}$$

We define *redshift*, z, by

$$\boxed{z_1 = \frac{\lambda_0 - \lambda_1}{\lambda_1}} \tag{4.12}$$

so

$$\boxed{z = \frac{a_0}{a_1} - 1}. \tag{4.13}$$

Exercise 4.1 *Suppose two galaxies, at redshifts z_1 and z_2, are in the same line of sight as seen from the Earth. Light observed from the galaxies was emitted at times t_1 and t_2, where $t_1 > t_2$. Consider the creatures who lived in galaxy 1 at time t_1. At what redshift z_{12} did they observe galaxy 2?*

For nearby galaxies redshift resembles a Döppler shift. In special relativity the Döppler formula for an emitter moving at velocity $\overrightarrow{\mathbf{V}}$, radial velocity V_r is

$$\frac{\nu_{observed}}{\nu_{emitted}} = \frac{(1-V^2)^{1/2}}{1+V_r}. \tag{4.14}$$

For small radial velocities ($V \ll 1$, $\overrightarrow{\mathbf{V}} = V_r \hat{r}$) we have

$$(1-V^2)^{1/2} \approx 1 - \frac{1}{2}V^2 \tag{4.15}$$

$$\Rightarrow \frac{\nu_{obs}}{\nu_{emitted}} \approx \frac{1-\frac{1}{2}V^2}{1+V} \tag{4.16}$$

$$\approx (1-\frac{1}{2}V^2)(1-V) \tag{4.17}$$

$$\approx 1 - V \qquad \text{to 1st order} \tag{4.18}$$

$$\frac{\lambda_{obs}}{\lambda_{emitted}} \approx \frac{1}{1-V} \tag{4.19}$$

$$\approx 1 + V \tag{4.20}$$

$$\Rightarrow z = \frac{\lambda_{obs} - \lambda_{emitted}}{\lambda_{emitted}} \approx V. \tag{4.21}$$

Thus z can be interpreted as measuring a Döppler shift of a nearby galaxy, moving away with speed $V = z(z \ll 1)$. (Careful, there are LOTS of approximations here). Also, we (and the galaxy) are moving with respect to the cosmic rest frame. (Motion with respect to the cosmic frame is called *proper motion*. So there is a small real Döppler shift, which also appears in the observed z (i.e there are 2 components to z, the expansion and proper motion).

Let's look at this in more detail. Consider a galaxy at co-ordinate η_1. We can define a spatial distance $d_R(t, \eta_1)$ measured at one single cosmic time t from Earth at $\eta = 0$ to the galaxy at η_1:

$$d_R(t, \eta_1) \equiv \int_0^{\eta_1} \sqrt{g_{11}(t)}\, d\eta \tag{4.22}$$

$$= \int_0^{\eta_1} a(t)\, d\eta = a(t)\eta_1 \tag{4.23}$$

If we measure this distance at the present time t_0, we have

$$\boxed{d_R(t_0, \eta_1) = a_0\eta_1}. \tag{4.24}$$

We can also define a *recession velocity* as the time derivative of this distance:

$$\boxed{V_R(t, \eta_1) = \frac{\mathrm{d}}{\mathrm{d}t}\left[d_R(t, \eta_1)\right] = \dot{a}(t)\eta_1}. \tag{4.25}$$

The ratio of recession velocity to spatial distance gives the Hubble parameter:

$$\frac{V_R(t_0, \eta_1)}{d_R(t_0, \eta_1)} = \frac{\dot{a}_0}{a_0} = H_0. \tag{4.26}$$

Next, we compare this with z:

$$1 + z_1 = \frac{a_0}{a_1}. \tag{4.27}$$

For $z_1 \ll 1$, use a Taylor expansion for a_1:

$$a_1 = a_0 + \dot{a}(t_0)(t_1 - t_0)\ldots \tag{4.28}$$
$$\Rightarrow 1 + z_1 \approx \frac{a_0}{a_0 + \dot{a}_0(t_1 - t_0)} \tag{4.29}$$
$$\Rightarrow 1 + z_1 \approx \frac{1}{1 + \frac{\dot{a}_0}{a_0}(t_1 - t_0)} \tag{4.30}$$
$$= \frac{1}{1 - H_0(t_0 - t_1)} \tag{4.31}$$
$$\approx 1 + H_0(t_0 - t_1) \tag{4.32}$$

Thus

$$\boxed{z_1 \approx H_0(t_0 - t_1)}. \tag{4.33}$$

Now, light from the galaxy has travelled a time $(t_0 - t_1)$. This corresponds to a distance $c(t_0 - t_1)$, which should be approximately some average distance between $d_R(t_1, \eta_1)$ and $d_R(t_0, \eta_1)$. If $t_0 - t_1 \ll t_0$, these two distances will be similar: the universe will not have expanded much in between t_0 and t_1. So

$$c(t_0 - t_1) \approx d_R(t_0, \eta_1). \tag{4.34}$$

But $c = 1$ so, by equation (4.26)

$$z_1 \approx H_0 d_R(t_0, \eta_1) \tag{4.35}$$
$$= V_R(t_0, \eta_1). \tag{4.36}$$

Thus z resembles a recession velocity.

4.2 Observable quantities

The definition of distance $d_R(t_0, \eta_1)$ is not observable, since it requires a measurement done at once over a distance of millions of light years. Some more practical observables are:

- z redshift.
- δ Angular diameter of object in sky.
- μ Angular velocity of moving object in sky.
- ℓ Apparent luminosity.

In order to use these observables, we will need to know or guess some *intrinsic quantities*:

- D True diameter.
- $V_\perp$ Perpendicular velocity (to line of sight).
- L Absolute luminosity – total radiated power.

The relation between the intrinsic and observable quantities is affected by the cosmic expansion, and thus by H_0, Ω_0, q_0. So, knowing (guessing?) H_0, Ω_0, q_0, and measuring the observables, we can get the intrinsic quantities. (Can do this the other way round; By making assumptions about the intrinsic quantities, and measuring the observables, we can get values for H_0, Ω_0, q_0)

4.3 Radial Coordinate

Let us change from angle η to the radial coordinate r: Let

$$r = F(\eta) = \begin{cases} \sin\eta & k = +1 \\ \eta & k = 0 \\ \sinh\eta & k = -1 \end{cases} . \tag{4.37}$$

The differentials become

$$dr = \begin{cases} \cos\eta \, \mathrm{d}\eta \\ \mathrm{d}\eta \\ \cosh\eta \, \mathrm{d}\eta \end{cases} = \begin{cases} \sqrt{1-r^2}\, \mathrm{d}\eta \\ \mathrm{d}\eta \\ \sqrt{1+r^2}\, \mathrm{d}\eta \end{cases} . \tag{4.38}$$

Turning this around,

$$\mathrm{d}\eta^2 = \begin{cases} \frac{1}{1-r^2}\, \mathrm{d}r^2 & k = +1 \\ \mathrm{d}r^2 & k = 0 \\ \frac{1}{1+r^2}\, \mathrm{d}r^2 & k = -1 \end{cases} , \tag{4.39}$$

or, more compactly,

$$\boxed{d\eta^2 = \frac{1}{1-kr^2}dr^2}. \tag{4.40}$$

4.4 Angular Diameter Distance

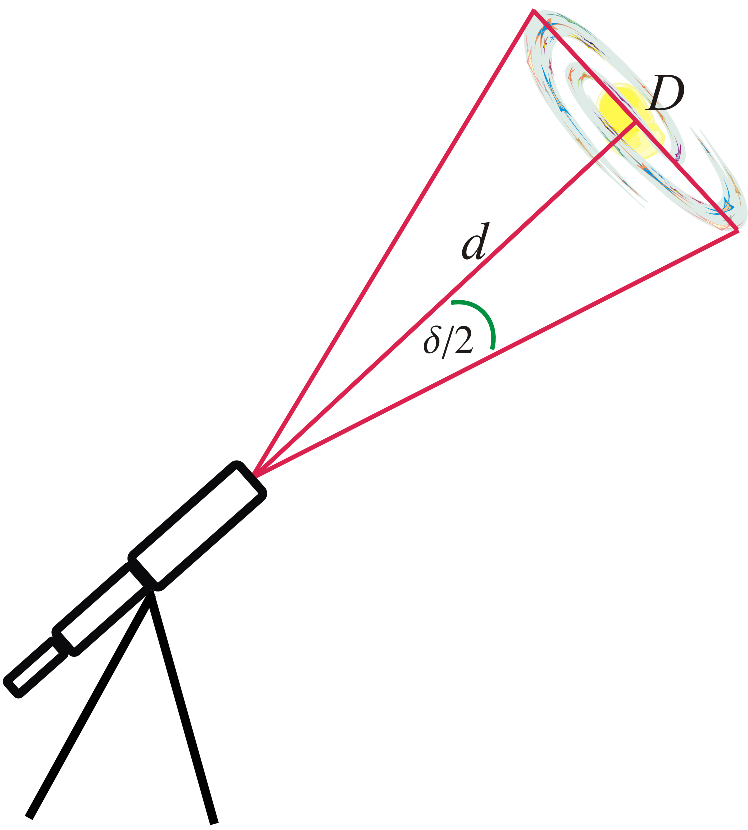
Suppose we are observing an extended object such as a galaxy, or a region of the sky with a slightly higher than average microwave background temperature. We will combine an observed quantity (the angular diameter of the object on the sky) with an intrinsic quantity (the actual size of the object perpendicular to our line of sight):

- *Observed Quantities*

  Angular diameter: δ

- *Intrinsic Quantities*

  Diameter D.



4.4.1 Euclidean Definition

$$\frac{D}{2d} = \tan\frac{\delta}{2} \approx \frac{\delta}{2}, \tag{4.41}$$

thus

$$\delta = \frac{D}{d}, \tag{4.42}$$

and we can define the angular diameter distance to be

$$\boxed{d_A = \frac{D}{\delta}}. \tag{4.43}$$

4.4.2 Cosmological Effects

Consider a galaxy which had diameter D when the light we see from the galaxy was emitted. Now, we always need to be precise about what "distance" means in cosmology, because the universe is expanding, and what we see of distant galaxies actually occurred long ago when the universe was smaller. First, D measures the distance (measured at emission time t_1) between the edges of the object (say at $\theta = \frac{\delta}{2}, \theta = -\frac{\delta}{2}$). In this situation, we integrate the Robertson–Walker metric line-element (using the r coordinate rather than η) on a path across the galaxy. But if the path just goes in the θ direction, then $\mathrm{d}t = \mathrm{d}r = \mathrm{d}\phi = 0$. So the spatial part of the line element is $\mathrm{d}s^2 = g_{\theta\theta}\mathrm{d}\theta^2$, where $g_{\theta\theta} = a(t_1)r_1 = a_1 r_1$. This means that

$$\begin{aligned} D &= \int_{-\delta/2}^{\delta/2} \sqrt{g_{22}} d\theta & (4.44) \\ &= \int_{-\delta/2}^{\delta/2} a_1 r_1 d\theta. & (4.45) \end{aligned}$$

or

$$\Rightarrow \boxed{D = a_1 r_1 \delta}. \tag{4.46}$$

So

$$d_A = \frac{D}{\delta} = a_1 r_1. \tag{4.47}$$

Recall that $1 + z_1 = a_0/a_1$. So we can also write

$$\boxed{d_A = (1+z_1)^{-1} a_0 r_1}. \tag{4.48}$$

If we think we know the diameter D of a galaxy observed with redshift z_1, then measuring δ gives us d_A, and hence $a_0 r_1$. We still need to relate this to cosmological parameters. A general method will be given below. For now, let us try the flat ($k = 0$) matter-dominated model:

$$\begin{aligned} \frac{a(t)}{a_0} &= \left(\frac{3H_0 t}{2}\right)^{2/3} = \left(\frac{t}{t_0}\right)^{2/3} & (4.49) \\ &= w^{2/3}, & (4.50) \end{aligned}$$

where $w = t/t_0$.

Now for the photons coming from the galaxy, $0 = \mathrm{d}\tau^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}r^2$, so

$$\mathrm{d}r = -\frac{\mathrm{d}t}{a(t)}. \tag{4.51}$$

Integrate to find $a_0 r_1$:

$$\begin{aligned} a_0 r_1 = a_0 \int_0^{r_1} \mathrm{d}r &= -\int_{t_0}^{t_1} \frac{a_0}{a(t)} \,\mathrm{d}t && (4.52)\\ &= \int_{t_1}^{t_0} \frac{a_0}{a(t)} \,\mathrm{d}t && (4.53)\\ &= \int_{w=t_1/t_0}^{w=1} w^{-2/3} t_0 \,\mathrm{d}w && (4.54)\\ &= 3t_0 \left(1 - \left(\frac{t_1}{t_0}\right)^{1/3}\right). && (4.55) \end{aligned}$$

From 4.49,

$$\begin{aligned} \left(\frac{t_1}{t_0}\right)^{1/3} &= \left(\frac{a_1}{a_0}\right)^{1/2} && (4.56)\\ &= (1+z_1)^{-1/2}. && (4.57) \end{aligned}$$

Thus

$$a_0 r_1 = 3t_o \left(1 - (1+z_1)^{-1/2}\right) = \frac{2}{H_0}\left(1 - (1+z_1)^{-1/2}\right) \tag{4.58}$$

We finally have the angular diameter distance in terms of H_0 and z:

$$d_A = \frac{2}{H_0}\left((1+z_1)^{-1} - (1+z_1)^{-3/2}\right). \tag{4.59}$$

4.4.3 How does angular diameter vary with redshift?

From $\delta = D/d_A$, we have

$$\begin{aligned} \delta &= \frac{H_0 D}{2} f(z_1); && (4.60)\\ f(z_1) &= \frac{(1+z_1)^{3/2}}{(1+z_1)^{1/2} - 1}. && (4.61) \end{aligned}$$

Does f(z) have a minimum? If so, it will be found where $\mathrm{d}f/\mathrm{d}z = 0$. Let $\xi = (1+z)^{1/2}$. Since this is a monotonic function of z we can look for the point where $\mathrm{d}f/\mathrm{d}\xi = 0$. Here

$$f = \frac{\xi^3}{\xi - 1} \tag{4.62}$$

and

$$\frac{\mathrm{d}f}{\mathrm{d}\xi} = \frac{3\xi^2(\xi-1)-\xi^3}{(\xi-1)^2} \tag{4.63}$$

$$= \frac{\xi^2}{(\xi-1)^2}(2\xi-3) \tag{4.64}$$

$$= 0 \qquad \text{at} \qquad \xi_{\min} = \frac{3}{2}. \tag{4.65}$$

This corresponds to

$$z_{\min} = \frac{5}{4}, \qquad f(z_{\min}) = \frac{27}{4}. \tag{4.66}$$

The minimum angular diameter of the galaxy is

$$\delta_{\min} = \frac{27H_0D}{8}. \tag{4.67}$$

How big is this in the sky? Say $h = 2/3$ and the diameter of the galaxy $D \approx 10^5$ light years:

$$\delta = \frac{27}{8}\left(\frac{2}{3}\right)(10^{10}\,\mathrm{y})^{-1}(10^5\,\ell\,\mathrm{y}) \tag{4.68}$$

$$= \frac{9}{4} \times 10^{-5} \text{ radians} \tag{4.69}$$

$$\approx 5 \text{ arc seconds.} \tag{4.70}$$

This is well within the capability of even modest telescopes – provided one can gather enough light to make a decent image!

4.5 Proper Motion Distance

Suppose a blob moves in a galactic jet at speed V. Let D be the distance moved in time Δt_1. We observe an angular change of δ in time Δt_0.

- *Intrinsic Quantities*

$$V = \frac{D}{\Delta t_1}; \tag{4.71}$$

$$V_\perp = \frac{D_\perp}{\Delta t_1} \tag{4.72}$$

  where $V_\perp$ and $D_\perp$ are measured perpendicular to the line of sight.

- *Observed Quantities*

  Angular velocity:

$$\mu = \frac{\delta}{\Delta t_0} \tag{4.73}$$

4.5.1 Euclidean Definition

From the discussion of angular diameter distance (equation (4.42)):

$$\delta = \frac{D_\perp}{d} = \frac{V_\perp \Delta t}{d} \tag{4.74}$$

Thus Euclid would give the distance d as

$$d = \frac{V_\perp \Delta t}{\delta} = \frac{V_\perp}{\mu}. \tag{4.75}$$

So we will define

$$\boxed{d_M \equiv \frac{V_\perp}{\mu}}. \tag{4.76}$$

4.5.2 Cosmological Effects

$$d_M = \frac{V_\perp \Delta t_0}{\delta} = \frac{V_\perp}{\delta}\left(\frac{\Delta t_0}{\Delta t_1}\right)\Delta t_1 \tag{4.77}$$

Now $V_\perp \Delta t_1 = D_\perp$, so

$$d_M = \frac{\Delta t_o}{\Delta t_1}\frac{D}{\delta} \tag{4.78}$$

The quantity $\frac{D_\perp}{\delta}$ is just the angular diameter distance d_A. Thus $d_M = \frac{\Delta t_0}{\Delta t_1} d_A$. Now, by time dilation

$$\frac{\Delta t_0}{\Delta t_1} = (1 + z_1) \tag{4.79}$$

so

$$\boxed{d_M = (1 + z_1) d_A = a_0 r_1}. \tag{4.80}$$

4.6 Luminosity Distance

Recall that power $P =$ energy/unit time.

- *Intrinsic Quantities*

  Absolute Luminosity L: the total power radiated by the object.

- *Observed Quantities*

  Apparent Luminosity ℓ: the power received per unit area by the telescope.

4.6.1 Euclidean Definition

The light from an object expands in a spherical wavefront. In a Euclidean universe, when the spherical wavefront has reached a distance d, the light has been spread over a sphere of area $4\pi d^2$. Thus the absolute luminosity is $4\pi d^2$ times the luminosity per unit area, i.e.

$$L = 4\pi d^2 \ell \tag{4.81}$$

We use the Euclidean result to *define* the luminosity distance

$$\boxed{d_L \equiv \sqrt{\frac{L}{4\pi\ell}}}. \tag{4.82}$$

4.6.2 Cosmological Effects

Again, we consider an object with redshift z_1, which emitted light at time t_1 and radial coordinate r_1. We need to express d_L in terms of these cosmological quantities. For this problem, we will be considering spherical wavefronts centred on the object. Thus it will be useful to change coordinates so that the object is at the origin $r = 0$, and we are at $r = r_1$. Express the Robertson–Walker metric in terms of (t, r, θ, ϕ) rather than (t, η, θ, ϕ):

$$\mathrm{d}\tau^2 = \mathrm{d}t^2 - a^2(t)\left(\frac{\mathrm{d}r^2}{1-kr^2} + r^2\mathrm{d}\Omega^2\right). \tag{4.83}$$

On the sphere at $r = r_1$, time t

$$\mathrm{d}\tau^2 = -a^2(t)\left(r_1^2\mathrm{d}\Omega^2\right). \tag{4.84}$$

So a sphere at radius r_1 and time t has effective radius $a(t)r_1$, with area $4\pi(a(t)r_1)^2$. We conduct our observations at time $t = t_0$, with scale factor $a(t_0) = a_0$. Thus the light from the object has been spread over an area

$$\boxed{\text{area} = 4\pi a_0^2 r_1^2}. \tag{4.85}$$

Let L_{received} be the total power received at the sphere with radius $a_0 r_1$. Then

$$\ell = \frac{L_{\text{received}}}{4\pi a_0^2 r_1^2}. \tag{4.86}$$

The received luminosity L_{received} at our time t_0 may not be the same as the absolute luminosity L emitted at time t_1: redshift decreases the energy of the light, and, as we shall see, time dilation slows down its delivery. The power (energy per unit time) reaching the sphere at radius $a_o r_1$ is

$$L_{\text{received}} = \frac{\Delta E_0}{\Delta t_0} \tag{4.87}$$

while $L = L_{\text{emitted}} = \frac{\Delta E_1}{\Delta t_1}$.

a. Redshift: For a photon with frequency ν and energy $h\nu$,

$$\nu_0 = (1 + z_1)^{-1} \nu_1. \tag{4.88}$$

The factor $(1 + z_1)^{-1}$ is the same at all frequencies, so energy summed over all frequencies scales by this factor as well:

$$E_0 = (1 + z)^{-1} E_1. \tag{4.89}$$

b. Time dilation: Energy received now over a time interval Δt_0 was emitted long ago over the time interval Δt_1, where

$$\frac{\Delta t_0}{\Delta t_1} = (1 + z_1). \tag{4.90}$$

Thus

$$\begin{aligned} L_{\text{received}} &= \frac{\Delta E_0}{\Delta t_0} = (1 + z_1)^{-2} \frac{\Delta E_1}{\Delta t_1} && (4.91) \\ &= (1 + z_1)^{-2} L. && (4.92) \end{aligned}$$

and

$$\ell = (1 + z_1)^{-2} \frac{L}{4\pi a_0^2 r_1^2}. \tag{4.93}$$

Now,

$$\begin{aligned} d_L &\equiv \left(\frac{L}{4\pi \ell} \right)^{1/2} && (4.94) \\ &= \sqrt{a_0^2 r_1^2 (1 + z_1)^2}. && (4.95) \end{aligned}$$

Thus

$$\boxed{d_L = (1 + z_1) a_0 r_1}. \tag{4.96}$$

4.7 summary

$$\begin{aligned} d_A &= \frac{D}{\delta} = (1 + z_1)^{-1} a_0 r_1; && (4.97) \\ d_M &= \frac{V_\perp}{\mu} = a_0 r_1; && (4.98) \\ d_L &= \sqrt{\frac{L}{4\pi \ell}} = (1 + z_1) a_0 r_1. && (4.99) \end{aligned}$$

Exercise 4.2 *For a matter dominated universe with $k \neq 0$, one can show that*

$$r_1 = \left(\frac{2|\Omega_0 - 1|^{1/2}}{\Omega_0^2}\right) \frac{\Omega_0 z_1 + (2-\Omega_0)\left(1 - \sqrt{1 + z_1\Omega_0}\right)}{(1+z_1)}. \tag{4.100}$$

Suppose $\Omega_0 = 2$. Find the Angular Diameter, Proper Motion, and Luminosity distances as functions of z_1 and H_0. Consider the angular diameter δ of an object with intrinsic diameter D. Above which value of z_1 will δ increase with redshift?

4.8 General Solution for a_0r_1

In this section we find a general expression for a_0r_1 in terms of cosmic parameters such as H_0 and Ω_0. Curiously, this involves an analysis of how the Hubble parameter $H(t)$ behaves when expressed as a function of redshift z.

For now, we will go back to using radial coordinate η rather than $r = F(\eta)$. We already have a useful expression for η_1, given by equation (4.7):

$$\eta_1 = \int_{t_1}^{t_0} \frac{\mathrm{d}t}{a(t)}. \tag{4.101}$$

The first task will be to replace the t variable by redshift z. Each redshift z has an associated cosmic time t (the time when objects observed with redshift z emitted their light), so we can write t as a function of z. Then

$$\begin{aligned} \frac{\mathrm{d}t}{\mathrm{d}z} &= \frac{\mathrm{d}t}{\mathrm{d}a}\frac{\mathrm{d}a}{\mathrm{d}z} && (4.102)\\ &= \frac{1}{\dot{a}}\frac{\mathrm{d}a}{\mathrm{d}z} && (4.103)\\ &= \frac{1}{aH}\frac{\mathrm{d}a}{\mathrm{d}z}, && (4.104) \end{aligned}$$

as $H = \dot{a}/a$. By equation (4.13),

$$a(z) = \frac{a_0}{(1+z)}, \tag{4.105}$$

and so

$$\begin{aligned} \frac{\mathrm{d}a}{\mathrm{d}z} &= -\frac{a_0}{(1+z)^2} && (4.106)\\ \Rightarrow \frac{\mathrm{d}t}{\mathrm{d}z} &= -\frac{a_0}{aH(1+z)^2} && (4.107)\\ &= -\frac{1}{(1+z)H(z)}. && (4.108) \end{aligned}$$

Define a function $E(z)$ by:

$$\boxed{H(z) = H_0 E(z)}. \tag{4.109}$$

By definition

$$\boxed{E(0) = 1}. \tag{4.110}$$

We can immediately obtain time t as an integral involving $E(z)$:

$$\begin{aligned} t_0 - t_1 &= \int_{z_1}^{0} \frac{\mathrm{d}t}{\mathrm{d}z}\,\mathrm{d}z && (4.111)\\ &= \frac{1}{H_0}\int_0^{z_1} \frac{1}{(1+z)E(z)}\,\mathrm{d}z. && (4.112)\end{aligned}$$

If we let $t_1 = 0$, $z_1 = \infty$ this gives the age of the universe:

$$\boxed{t_0 = \frac{1}{H_0}\int_0^{\infty} \frac{1}{(1+z)E(z)}\,\mathrm{d}z}. \tag{4.113}$$

Also, we can now find η_1 in terms of $E(z)$: equation (4.101) transforms to

$$\begin{aligned} \eta_1 &= \int_{z_1}^{0} \frac{1}{a(z)}\frac{\mathrm{d}t}{\mathrm{d}z}\,\mathrm{d}z && (4.114)\\ &= -\int_{z_1}^{0} \frac{1}{a(z)}\frac{1}{(1+z)H_0E(z)}\,\mathrm{d}z && (4.115)\\ &= \boxed{\frac{1}{H_0 a_0}\int_0^{z_1} \frac{1}{E(z)}\,\mathrm{d}z}. && (4.116)\end{aligned}$$

Finally, $a_0 r_1 = a_0 F(\eta_1)$. Note that for $k = 0$, we have $r = \eta$. Thus the distance measures become

- Angular diameter distance $(k = 0)$:

$$\boxed{d_A(z_1) = \frac{1}{(1+z_1)H_0}\int_0^{z_1} \frac{1}{E(z)}\,\mathrm{d}z}. \tag{4.117}$$

- Luminosity distance $(k = 0)$:

$$\boxed{d_L(z_1) = \frac{(1+z_1)}{H_0}\int_0^{z_1} \frac{1}{E(z)}\,\mathrm{d}z}. \tag{4.118}$$

4.8.1 The function $E(z)$

What is $E(z)$? Go back to the first evolution equation:

$$\dot{a}^2 + k = H_0^2 \frac{\rho}{\rho_{c0}} a^2. \tag{4.119}$$

Divide by a^2, with $H = \dot{a}/a$,

$$H^2 + \frac{k}{a^2} = H_0^2 \frac{\rho}{\rho_{c0}} \tag{4.120}$$

Thus with $H = H_0 E$,

$$E^2(z) = \frac{-k}{H_0^2 a^2} + \frac{\rho}{\rho_{c0}}. \tag{4.121}$$

Define

$$\Omega_{a0} \equiv \frac{-k}{H_0^2 a_0^2}. \tag{4.122}$$

Note that by equation (3.7), this quantity is a measure of the curvature of space:

$$\Omega_{a0} = -\frac{\mathfrak{R}(t_0)}{6H_0^2}. \tag{4.123}$$

Then, since $a^2 = a_0/(1+z)^2$,

$$E^2(z) = \Omega_{a0}(1+z)^2 + \frac{\rho}{\rho_{c0}}. \tag{4.124}$$

There may be several sources of mass-energy density in the universe. We will assume the total mass-energy density ρ consists of the matter density ρ_m (with pressure $p_m \approx 0$) and the radiation density ρ_γ ($p_\gamma = \rho_\gamma/3$). More speculatively, there may be a vacuum energy density ρ_Λ ($p_\Lambda = -\rho_\Lambda$), and/or a quintessence energy density ρ_q (in the simplest form $p_q = w\rho_q$ for some constant $w < 0$). Thus

$$\frac{\rho}{\rho_{c0}} = \frac{\rho_m}{\rho_{c0}} + \frac{\rho_\gamma}{\rho_{c0}} + \frac{\rho_\Lambda}{\rho_{c0}} + \frac{\rho_q}{\rho_{c0}}. \tag{4.125}$$

Matter $\rho_m \sim a^{-3}$, i.e.

$$\begin{aligned} \frac{\rho_m}{\rho_{c0}} &= \frac{\rho_{m0}}{\rho_{c0}} \left(\frac{a}{a_0}\right)^{-3} = \frac{\rho_{m0}}{\rho_{c0}}(1+z)^3 && (4.126) \\ &= \Omega_{m0}(1+z)^3, && (4.127) \end{aligned}$$

where $\Omega_{m0} = \rho_{m0}/\rho_{c0}$.

Radiation Here $\rho_\gamma \sim a^{-4} \sim (1+z)^4$, so

$$\frac{\rho_\gamma}{\rho_{c0}} = \Omega_{\gamma 0}(1+z)^4. \tag{4.128}$$

Vacuum The vacuum energy density is constant, so

$$\frac{\rho_\Lambda}{\rho_{c0}} = \Omega_{\Lambda 0}. \tag{4.129}$$

Quintessence Suppose $\rho_q = C\,a^a$ for some exponent a and some constant C. If $p_q = w\rho_q$ then the second evolution equation (3.28) gives

$$\frac{\mathrm{d}a^{3+a}}{\mathrm{d}a} = -3w\,a^{2+a} \qquad \Rightarrow a = -3(1+w). \tag{4.130}$$

Thus

$$\frac{\rho_q}{\rho_{c0}} = \Omega_{q0}(1+z)^{-3(1+w)}. \tag{4.131}$$

Putting these expressions together into equation (4.124) and equation (4.125),

$$\boxed{E^2(z) = \Omega_{\gamma 0}(1+z)^4 + \Omega_{m0}(1+z)^3 + \Omega_{a0}(1+z)^2 + \Omega_{\Lambda 0} + \Omega_{q0}(1+z)^{-3(1+w)}}. \tag{4.132}$$

Since $E(0) = 1$, we have

$$\boxed{\Omega_{\gamma 0} + \Omega_{m0} + \Omega_{a0} + \Omega_{\Lambda 0} + \Omega_{q0} = 1}. \tag{4.133}$$

Define the total energy density at t_0 to be

$$\Omega_{tot} \equiv \Omega_{\gamma 0} + \Omega_{m0} + \Omega_{\Lambda 0} + \Omega_{q0} = 1 - \Omega_{a0}. \tag{4.134}$$

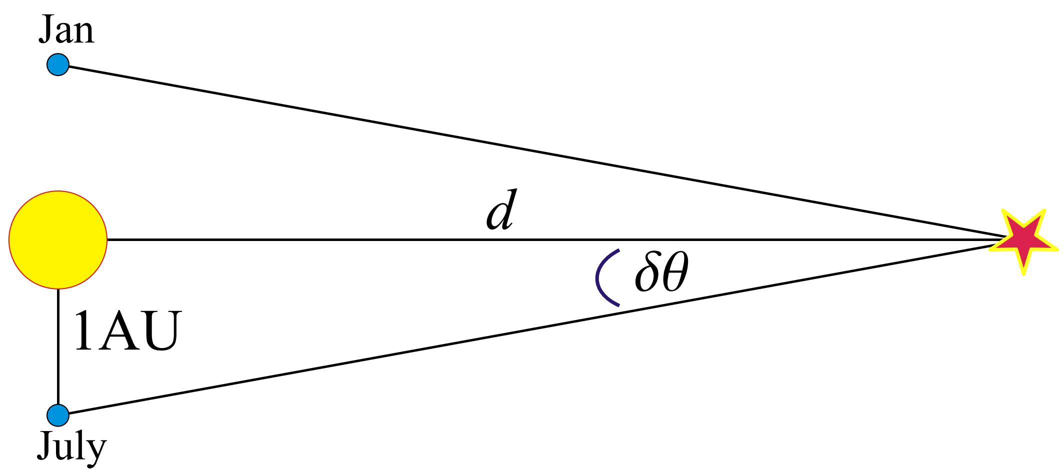
Present observations give $\Omega_{tot} = 1.02 \pm 0.02$, so the curvature term Ω_{a0} seems to be small.

4.9 Cosmic Distance Measures

- *Eratosthenes (∼ 240 B.C.)* : Earth's radius.
- *Hipparcos (∼ 150 B.C.)*: Distance to Moon.

4.9.1 Parallax

- *Cassini (1672)* : Distance to Mars (parallax).
- *Henderson and Bessel (1838)* : α centauri, ≈ 1.5 pc (parallax).



From the figure,

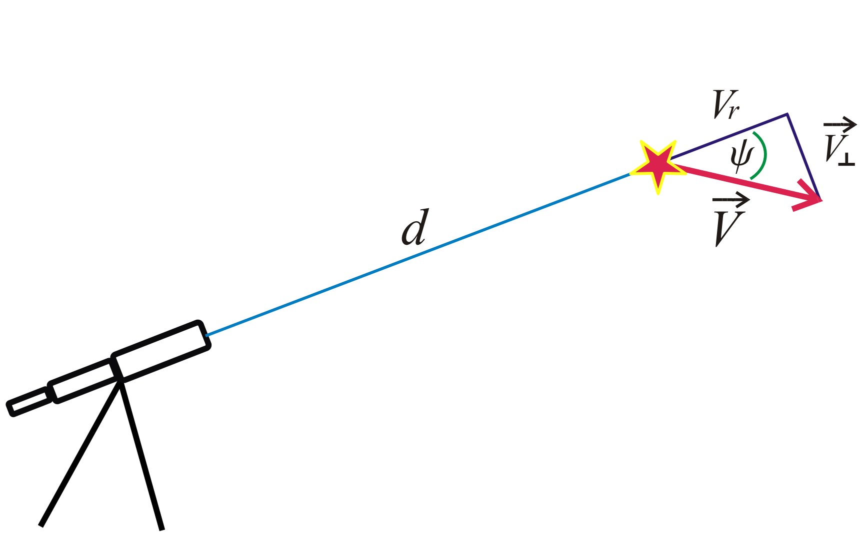
$$\boxed{d = \frac{1 A.U.}{\delta\theta}} \tag{4.135}$$

One *parsec* ($= 3.26\,\ell\,\mathrm{y}$) is defined to be the distance which gives a parallax angle $\delta\theta$ of 1 arcsecond.

- *Hipparcos satellite*: Milliarcsecond resolution $\Rightarrow$ parallax distances measured to ≈ 1000 parsecs.

4.9.2 Cluster Surveys

Consider a star moving with velocity $\overrightarrow{\mathbf{V}}$ and angle ψ relative to the line of sight.



$$\tan(\psi) = \frac{V_\perp}{V_r}. \tag{4.136}$$

Here

- V_r: measured by Döppler shift.
- ψ : unknown angle.
- $V_\perp$: $V_\perp = d\mu$, where
- $d =$ distance, and
- $\mu =$ angular velocity of observed star relative to fixed background stars.

The proper motion μ can be found by comparing archive photographs of the star. If we only knew ψ, the distance d could be found from

$$d = \frac{V_\perp}{\mu} = \boxed{\frac{V_r \tan\psi}{\mu}}. \tag{4.137}$$

If we have a cluster of stars, we can assume some probability distribution for ψ. The two simplest methods are:

a. *Moving cluster method*: (e.g. Hyades). Assume that all stars in the cluster move in the same direction. For the Hyades cluster, one obtains the distances
   - $45.5 \pm 2.5pc$ (moving cluster method)
   - $46.3 \pm 0.3pc$ (Hipparcos)

b. *Statistical cluster method*: Assume that the stars in the cluster move in all directions with equal probabilities.

4.9.3 Standard Candles

Standard Candles are objects of known intrinsic properties, for example known absolute luminosity L or period of oscillation P.

a. *Main sequence stars*

b. *Variable stars* Example: The luminosity of Cepheid stars oscillate with periods P ranging from 2 to 240 days.
   - 1912 - Leavitt finds that the apparent luminosity of Cepheid variables in the Large Magellanic clouds is proportional to their period. Since all these stars are at a similar distance to us, this suggests that the absolute luminosity is proportional also, $L(P) = CP$ for some constant C.

- 1920 - Shapley (mis)measures C.
- 1923 - Hubble observes Cepheids in Andromeda proving that spiral nebulae are outside the Milky Way.
- 1927-9 - Slipher and Hubble show distance increases with redshift. Hubble estimates $H_0 = 500\text{km s}^{-1}\text{Mpc}^{-1}$ (due to Shapley's errors). This gives an age of universe $< 2 \times 10^9$ years.
- 1952 - Baade recalibrates the distance scale, bringing H_0 below 100.

c. Brightest star in the galaxy to $\approx$ 30Mpc

d. HII regions (star forming regions like the Orion nebula) $\approx$ 60 Mpc

e. Globular clusters to $\approx$ 100 Mpc

f. Brightest galaxy in cluster to 100 Mpc

g. *Tully-Fisher relation* - For spirals we can obtain rotation rate from Döppler shifts. One finds

$$L_g = L(V_{\text{rot}}) \approx V_{\text{rot}}^4. \tag{4.138}$$

h. *Type IA Supernova* - These incredibly energetic events occur within a binary star system where one of the stars is a white dwarf (a compact star near the end of its life). The gravitational field of the white dwarf steadily sucks material away from the companion star. Eventually the white dwarf exceeds a critical mass and its outer layers become unstable, producing a tremendous explosion. The light emitted as a function of time is called the light curve. Most supernovae have similar light curves. The Peak luminosity of the curve correlates with the decay rate of luminosity. Observations can go beyond $z = 1$, i.e. several billion light years away.

Exercise 4.3

Suppose an astronomer measures both the distance and redshift of one galaxy which is part of the Virgo cluster. How does this help to determine H_0? What uncertainties are present in this determination? Suppose the astronomer then estimates the relative distance (about a factor of 5) between the Virgo and Coma clusters by comparing the apparent luminosity distribution of objects in each cluster. Why should this give a more accurate value for H_0? Why might it be better to measure the distance to more than one galaxy in Virgo?

Chapter 5

The Early Universe

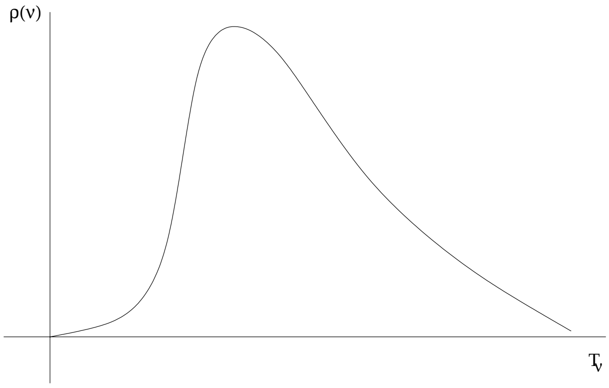
5.1 The Cosmic Microwave Background

5.1.1 Black Body Radiation

Blackbody radiation has a characteristic distribution in frequency called the *Planck spectrum*. The energy density contained in frequency interval $[\nu, \nu + d\nu]$ is

$$\rho(\nu)\,\mathrm{d}\nu = \frac{8\pi h\nu^3}{e^{h\nu/kT} - 1}\,\mathrm{d}\nu \tag{5.1}$$

where h = Planck's constant and k = Boltzmann's constant.



Suppose the universe is filled with blackbody radiation at time t_1, with temp T_1. What is the spectrum at $t_0 > t_1$? First, let $n(\nu_1)d\nu_1$ be the number of photons per unit volume in the frequency range $[\nu_1, \nu_1 + d\nu_1]$. As the energy of a photon is $h\nu$,

$$\rho(\nu_1)\mathrm{d}\nu_1 = h\nu_1 n(\nu_1)\mathrm{d}\nu_1. \tag{5.2}$$

Now, the number density ν is affected by the cosmic expansion in two ways.

- Volume increases by a factor $(a_0/a_1)^3$ between times t_1 and t_0. So the number density n should go down by a factor of $(a_1/a_0)^3$.

- The radiation redshifts to lower frequencies. A photon starting out with frequency ν_1 ends up with frequency $\nu_0 = (a_1/a_0)\,\nu_1$.

With these considerations, we can relate the number density at time t_0 to the number density at time t_1:

$$
\begin{aligned}
n(\nu_0)\,\mathrm{d}\nu_0 &= \left(\frac{a_1}{a_0}\right)^3 n(\nu_1)\,\mathrm{d}\nu_1. && (5.3)\\
&= \left(\frac{a_1}{a_0}\right)^3 \frac{8\pi\nu_1^2\,\mathrm{d}\nu_1}{e^{h\nu_1/kT_1}-1} && (5.4)\\
&= \frac{8\pi\nu_0^2\,\mathrm{d}\nu_0}{e^{h\,a_0\nu_0/a_1kT_1}-1} && (5.5)\\
&= \frac{8\pi\nu_0^2\,\mathrm{d}\nu_0}{e^{h\nu_0/kT_0}-1}, && (5.6)
\end{aligned}
$$

where

$$T_0 = \frac{a_1}{a_0}T_1. \tag{5.7}$$

Thus the spectrum is still black body at time t_0:

$$\rho(\nu_0)\mathrm{d}\nu_0 = \frac{8\pi h\nu_0^3}{e^{h\nu_0/kT_0}-1}\mathrm{d}\nu_0. \tag{5.8}$$

The total energy density in a black body spectrum is

$$\rho_\gamma(T) = \int_0^\infty \rho(\nu)\,\mathrm{d}\nu = aT^4, \tag{5.9}$$

where $a = 7.56\times10^{-15}$ erg cm^{-3}K^{-4}.

Note the scalings:

$$
\begin{aligned}
T &\sim a^{-1} \sim (1+z); && (5.10)\\
\rho_\gamma &\sim T^4 \sim a^{-4} \sim (1+z)^4. && (5.11)
\end{aligned}
$$

We could have guessed that ρ_γ would scale as a^{-4}: expansion decreases the photon number density as a^{-3}, while redshift provides a further factor of a^{-1}.

Exercise 5.1

a. *Show that for low frequencies the Planck spectrum reduces to the Rayleigh–Jeans spectrum*

$$\rho(\nu)d\nu \approx 8\pi kT\nu^2 d\nu. \tag{5.12}$$

b. *Suppose a Rayleigh–Jeans spectrum is emitted at time t_1, with temperature T_1. Show that at time t_0 the spectrum is still Rayleigh–Jeans.*

Exercise 5.2 *The number ratio of photons to baryons is*

$$n_\gamma/n_B = 0.27\sigma \tag{5.13}$$

where the entropy per baryon

$$\sigma = \frac{4aT^3}{3n_B k} = 74.0\frac{T^3}{n_B(cm^{-3})}. \tag{5.14}$$

Given the observed value of $n_\gamma/n_B \approx 10^9$, find the temperature of the universe at the time when the mass density ($n_B m_p c^2$) and radiation density of the universe were approximately equal.

5.2 The Universe before Decoupling Time

Near *decoupling time* $t_\mathrm{d} \approx 380,000$ years, the matter in the universe cooled down to a few thousand degrees. Before this time, most of the matter was ionized plasma, because temperatures were to hot for atoms to survive for very long. After t_d, typical collisions between atoms did not have sufficient energy to strip off the electrons, so neutral atoms could stay neutral. The plasma existing before t_d was opaque to the radiation, i.e. it would constantly absorb and remit the photons. Thus it would appear like a glowing fog. Near decoupling time the fog clears, leaving the glow that we see today, red-shifted all the way to microwaves.

Observations suggest the current value of Ω_0 is close to 1. As we will see, in the early universe $\Omega(t)$ was even closer to 1, and so it is an excellent approximation to set $\Omega(t) = 1$ and $k = 0$ for early universe calculations.

a. Radiation density dominated in the early universe, so $\rho \sim a^{-4}$. Thus in the first evolution equation, we can write

$$\rho = \rho_\mathrm{d}\left(\frac{a}{a_\mathrm{d}}\right)^{-4} \tag{5.15}$$

so

$$\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - k = \frac{8\pi G}{3}\rho_\mathrm{d} a_\mathrm{d}^4 a^{-2} - k. \tag{5.16}$$

Note that as $a \to 0$ the first term on the right dominates, so k can be neglected.

b. From the previous equation, $\dot{a} \to \infty$ as $a \to 0$. Recall equation (**??**):

$$\Omega(t) = 1 + \frac{k}{\dot{a}(t)^2}. \tag{5.17}$$

Thus as $t \to 0$, and hence $a \to 0$, $\Omega(t) \to 1$.

Let us now try the $k = 0$ radiation dominated solution:

$$\begin{aligned} \dot{a}^2 &= C a^{-2}, C = \frac{8\pi G \rho_\mathrm{d}}{3} a_\mathrm{d}^4 && (5.18)\\ \dot{a} &= C^{1/2} a^{-1} && (5.19)\\ a\,\mathrm{d}a &= C^{1/2}\,\mathrm{d}t, && (5.20) \end{aligned}$$

with solution

$$\frac{a^2}{2} = C^{1/2} t \tag{5.21}$$

or

$$\boxed{a(t) = \left(\frac{32\pi G \rho_\mathrm{d}}{3}\right)^{1/4} a_\mathrm{d} t^{1/2}}. \tag{5.22}$$

5.2.1 Density

Combining equations (5.15) and (5.22) gives

$$\rho(t) = \rho_\mathrm{d} \left(\left(\frac{32\pi G \rho_\mathrm{d}}{3}\right)^{1/4} t^{1/2} \right)^{-4}, \tag{5.23}$$

which leads to the simple equation

$$\boxed{\rho(t) = \left(\frac{3}{32\pi G}\right) t^{-2}}. \tag{5.24}$$

5.2.2 Temperature

Suppose the early universe contained N relativistic ($V \approx 1$) particle species. If all these particles were bosons (integer spin), then each would contribute an energy density aT^4 to the total, by equation (5.9). In this case we would have

$$\rho = N a T^4. \tag{5.25}$$

The situation is somewhat more complicated, because many particles are fermions (half-integer spin) rather than bosons. Fermions obey 'Fermi-Dirac' statistics, and so will not obey the black body distribution law, equation (5.1) (in fact their distribution just replaces the minus sign in the denominator of equation (5.1) by a plus sign). This changes the net energy density. For example, electrons have a density $\rho_{e^-} = \frac{7}{8} a T^4$. For neutrinos the energy density is half of this (because they only have one spin state compared to two for electrons): $\rho_\nu = \frac{7}{16} a T^4$.

Let $\mathcal{N}$ be the effective number of particle species, including weighting factors such as 7/8 for the electrons. Then

$$\rho = \mathcal{N} a T^4, \tag{5.26}$$

so

$$\boxed{T(t) = \left(\frac{3}{32\pi G a}\right)^{1/4} \mathcal{N}^{-1/4} t^{-1/2}}. \tag{5.27}$$

This expression shows that the effective number $\mathcal{N}$ of relativistic particles affects the temperature history of the early universe. As a consequence, $\mathcal{N}$ affects the rate of nucleosynthesis. Calculations give the following correspondence between time and temperature:

t(sec)	T
10^{-6}	10^{13}
10^{-4}	10^{12}
1	10^{10}
100	10^{9}

5.3 Nucleosynthesis

5.3.1 Pair creation of particles

A gas at temperature T consists of randomly moving particles. The particles have a range of energies, but the mean energy is $E = 3kT/2$, where k is Boltzmann's constant. Suppose $kT \approx m_e$, the mass of an electron. Then when two particles collide, there may be enough energy in the collision to create an electron-positron pair (positron = anti-electron).

At very high temperatures ($kT \gg m_e$) the dense gas in the early universe would have produced copious numbers of electron-positron pairs. But also pairs of electrons and positrons would crash into each other and annihilate, producing other particles or gamma rays. If the temperature had stayed constant, the creation and annihilation of pairs would have balanced. But in fact the temperature dropped quickly. As kT went below m_e pair production of electrons and positrons stopped, while annihilation continued. If there were exactly equal numbers of electrons and positrons to begin with, then we would be left with no electrons at all! However, and fortunately for life forms, a slight imbalance between matter and antimatter was present, created perhaps in the first 10^{-20} seconds of the universe. Thus the pair annihilation occurring after kT dropped below m_e, left us with a density of electrons only.

Similar processes occurred for all the other particle species. A transition happens when kT falls below the rest mass of a particle.

particle	symbol	mass	T = k/m
proton	p	1.6726×10^{-27} kg	1.1×10^{13}
electron	e	9.1094×10^{-31} kg	5.9×10^{9}

5.3.2 The neutron-proton ratio

Neutrons have a slightly higher mass than protons. In the very early universe (before 10^{-6} seconds after the big bang) the temperature was so high that the typical kinetic energy of a particle was much greater than this mass difference. Thus a particle collision could create or destroy protons and neutrons with almost equal likelihoods. As a consequence the ratio of the number density of neutrons N_n to the number density of protons N_p started out at $N_n/N_p \approx 1$.

Now, in thermal equilibrium the ratio of the number density of neutrons N_n to the number density of protons N_p will be

$$\frac{N_n}{N_p} = e^{-(M_n - M_p)/kT}. \tag{5.28}$$

For very large T this ratio goes to 1, consistent with our previous comment. On the other hand, as the temperature of the universe drops, the ratio should decrease.

However, in order to maintain thermal equilibrium, reactions must take place which convert neutrons to protons (the opposite reactions will also take place, but with less vigor). For example, a positron could smash into a neutron, giving the neutron positive charge and turning it into a proton (the positron turns into an anti-neutrino in the process).

But this picture is too simple. In the first few minutes the universe expanded and cooled too quickly for these reactions to maintain equilibrium ratios. Thus there were more neutrons than equilibrium would suggest. As a consequence, it was easier to create atomic nuclei (protons, unlike neutrons, repel each other because of their positive charge, hence making it more difficult to bind them together into nuclei).

At temperatures above about 10^{10} light nuclei form, but are quickly destroyed in collisions. As the temperature drops below 10^{10}, nucleosynthesis stops, but also several nuclei begin to survive collisions:

Deuterium	${}^2D$
Tritium	${}^3T$
Helium	${}^3He, {}^4He$
Lithium	${}^7Li$
Beryllium	${}^{11}Be$

Nucleosynthesis in stars and supernovae yield very little of these, but give us heavier elements. Observations tell us the primordial abundances of ${}^3He, {}^7Li, {}^4He, {}^2D$. This gives us valuable information about the temperature history of the early universe. For example, suppose there were N families of quarks and leptons (electrons and neutrinos). The effective number of relativistic species $\mathcal{N}$ increases with N, affecting the temperature function $T(t)$ equation (5.27).

If $\mathcal{N}$ is larger, then the universe cools more quickly, i.e. at a given time t, the temperature $T(t)$ is lower. Thus higher N implies:

- faster cooling,

- less time for $\frac{N_p}{N_n}$ to equilibrate,
- more neutrons,
- more nuclei.

Thus abundances of nuclei are affected by the number of particle families and neutrino species. 1980s Schramm : Maximum of 3 neutrino species. This cosmological result was later verified in experiments at CERN.

5.4 Galaxy Formation

5.4.1 Density Fluctuations in the Primordial Gas

Consider a fluid of density ρ, pressure p, velocity $\overrightarrow{\mathbf{V}}$, and gravitational force per unit mass $\overrightarrow{\mathbf{F}}$. We will ignore universal expansion for now. The fluid satisfies several conservation equations:

a. **Conservation of Mass** (**Equation of Continuity**)

$$\partial_t \rho + \nabla \cdot \rho \overrightarrow{\mathbf{V}} = 0. \tag{5.29}$$

(Note that the relativistic form of this is the vanishing of the 4-divergence $\partial_a(\rho u^a) = 0$).

b. **Conservation of Momentum** (**Navier-Stokes Equation**)

(This is actually the Euler equation as we will ignore viscosity).

$$\partial_t \overrightarrow{\mathbf{V}} + \overrightarrow{\mathbf{V}} \nabla \overrightarrow{\mathbf{V}} = -\frac{1}{\rho} \nabla p + \overrightarrow{\mathbf{F}}. \tag{5.30}$$

c. **Poisson Equation**

$$\nabla \cdot \overrightarrow{\mathbf{F}} = -4\pi G \rho. \tag{5.31}$$

d. **Equation of state**

$$p = p(\rho). \tag{5.32}$$

Assume the mean velocity and the mean gravitational force vanish, $\overrightarrow{\mathbf{V}} = \overrightarrow{\mathbf{F}} = 0$. The mean gravitational force is involved in the overall expansion of the universe; the effect of the expansion on density fluctuations will be discussed later. We will write our quantities in terms of a mean quantity plus a small fluctuation:

- $\rho = \bar{\rho} + \delta\rho$,
- $p = \bar{p} + \delta p$,
- $\overrightarrow{\mathbf{V}} = 0 + \delta \overrightarrow{\mathbf{V}}$,
- $\overrightarrow{\mathbf{F}} = 0 + \delta \overrightarrow{\mathbf{F}}$.

The mean quantities are assumed constant in both time and space. We can track the development of small fluctuations by linearizing the equations, i.e. by ignoring the very small terms which are quadratic or higher order in the fluctuations.

a. The mass conservation equation becomes

$$\partial_t \delta\rho + \bar{\rho} \nabla \cdot \delta \overrightarrow{\mathbf{V}} = 0. \tag{5.33}$$

b. Momentum conservation gives

$$\partial_t \delta \overrightarrow{\mathbf{V}} = -\frac{1}{\bar{\rho}} \nabla \delta p + \delta \overrightarrow{\mathbf{F}}. \tag{5.34}$$

c. The Poisson equation tells us that

$$\nabla \cdot \delta \overrightarrow{\mathbf{F}} = -4\pi G \delta \rho. \tag{5.35}$$

d. The equation of state provides us with a characteristic velocity c_s, defined by

$$c_s^2 \equiv \left(\frac{\mathrm{d}p}{\mathrm{d}\rho}\right)_{\rho=\bar{\rho}}. \tag{5.36}$$

In terms of this velocity,

$$p(\rho) = p(\bar{\rho} + \delta\rho) = p(\bar{\rho}) + \left(\frac{\mathrm{d}p}{\mathrm{d}\rho}\right)_{\rho=\bar{\rho}} \delta\rho + \ldots. \tag{5.37}$$

$$\approx \bar{p} + c_s^2 \delta\rho. \tag{5.38}$$

Thus to first order

$$\delta p = c_s^2 \delta \rho. \tag{5.39}$$

We can now eliminate δp, $\delta \overrightarrow{\mathbf{F}}$, and $\delta \overrightarrow{\mathbf{V}}$ in favor of $\delta\rho$:

a. Take the time derivative of equation (5.33).

$$\partial_t^2 \delta\rho + \bar{\rho} \partial_t \nabla \cdot \delta \overrightarrow{\mathbf{V}} = 0. \tag{5.40}$$

b. Take the gradient of equation (5.39).

$$\nabla \delta p = c_s^2 \nabla \delta \rho. \tag{5.41}$$

c. Take the divergence of equation (5.34), using equation (5.41) and equation (5.35).

$$\partial_t \nabla \cdot \delta \overrightarrow{\mathbf{V}} = -\left(\frac{c_s^2}{\bar{\rho}}\right) \nabla \cdot \nabla \delta\rho + \nabla \cdot \delta \overrightarrow{\mathbf{F}} \tag{5.42}$$

$$= -\left(\frac{c_s^2}{\bar{\rho}}\right) \nabla^2 \delta\rho - 4\pi G \delta\rho. \tag{5.43}$$

d. We can now substitute into equation (5.40),

$$\partial_t^2 \delta\rho = \bar{\rho} \left(\frac{c_s^2}{\bar{\rho}} \nabla^2 \delta\rho + 4\pi G \delta\rho\right) \tag{5.44}$$

and rearrange to find an equation for $\delta\rho$ alone:

$$\boxed{\left(\partial_t^2 - c_s^2 \nabla^2 - 4\pi G \bar{\rho}\right) \delta\rho = 0}. \tag{5.45}$$

Solution

We will try a Fourier mode with a set wavenumber $\overrightarrow{\mathbf{k}}$:

$$\delta\rho = A(\overrightarrow{\mathbf{k}})e^{i(\overrightarrow{\mathbf{k}}\cdot\overrightarrow{\mathbf{x}}-\omega t)}, \tag{5.46}$$

where $A(\overline{\mathbf{k}})$ is the amplitude. The differential equation for $\delta\rho$ gives

$$-\omega^2 + c_s^2k^2 - 4\pi G\bar{\rho} = 0. \tag{5.47}$$

Let us define the *Jeans wavenumber*

$$k_J \equiv \frac{\sqrt{4\pi G\bar{\rho}}}{c_s}. \tag{5.48}$$

Then $-\omega^2 + c_s^2(k^2 - k_J^2) = 0$, which gives us the *dispersion relation* between ω and k,

$$\boxed{\omega^2(k) = c_s^2(k^2 - k_J^2)}. \tag{5.49}$$

If $k > k_J$ then $\omega^2 > 0$. Thus $\delta\rho$ will oscillate – in fact it is modified a sound wave. If we set $k_J = 0$ then the wave travels at the sound speed c_s. For $k_J \neq 0$ the wave speed (phase velocity ω/k) is modified by gravitational forces.

However, if $k < k_J$ then $\omega^2 < 0$. In this case

$$e^{i\omega t} = e^{\pm|\omega|t}, \tag{5.50}$$

so density fluctuations grow or decay exponentially. A density fluctuation growing in amplitude ($\delta\rho > 0$ and increasing) corresponds to a clump of matter collapsing under its own gravity and becoming denser.

Inside a blob of matter there are gravitational forces attempting to collapse the blob, and fluid pressure forces which oppose this collapse. One would guess that the gravitational forces are more likely to win for large blobs of matter, because of the extra mass. In particular, define the *Jeans wavelength* to be $\lambda_J = 2\pi/k_J$. Consider a ball of radius λ_J with volume $\frac{4\pi}{3}\lambda_J^3$. Define the Jeans Mass to be the mass inside this ball:

$$M_J = \frac{4\pi}{3}\bar{\rho}\lambda_J^3. \tag{5.51}$$

Recall that collapse occurs for small k, corresponding to large λ, i.e. $\lambda > \lambda_J$, and large M, i.e. $M > M_J$.

In order to calculate M_J, we need to know the sound speed. This speed changes significantly near decoupling time.

Exercise 5.3

Suppose we keep viscosity ν in the Navier-Stokes equation, so that equation (5.30) becomes

$$\partial_t\overline{\mathbf{v}} + \overline{\mathbf{v}}\cdot\nabla\overline{\mathbf{v}} = -\frac{\nabla p}{\rho} + \delta\overline{\mathbf{F}} + \nu\nabla^2\overline{\mathbf{v}}. \tag{5.52}$$

Find a differential equation for $\delta\rho$ where terms have up to three derivatives. Assuming that solutions exist in the form

$$\delta\rho = A(\overline{\mathbf{k}}) e^{i(\overline{\mathbf{k}}\cdot\overline{\mathbf{x}}-\omega t)}, \tag{5.53}$$

find $\omega = \omega(k)$.

5.4.2 The Sound Speed in the Early Universe

The sound speed after decoupling time

In order to find the sound speed $c_s = \sqrt{\mathrm{d}p/\,\mathrm{d}\rho}$, we need to see how p varies with ρ. Consider N particles of mass m in a volume $\mathcal{V}$. The mass density $\rho = Nm/\mathcal{V}$. So we can equally use $\mathcal{V}$ as a variable in place of ρ: keeping N constant,

$$\frac{\mathrm{d}\rho}{\mathrm{d}\mathcal{V}} = -\frac{Nm}{\mathcal{V}^2} \tag{5.54}$$

$$= -\frac{\rho}{\mathcal{V}}. \tag{5.55}$$

Thus

$$c_s^2 = \frac{\mathrm{d}p}{\mathrm{d}\mathcal{V}}\frac{\mathrm{d}\mathcal{V}}{\mathrm{d}\rho} = -\frac{\mathcal{V}}{\rho}\frac{\mathrm{d}p}{\mathrm{d}\mathcal{V}}. \tag{5.56}$$

We will assume that ordinary matter and any dark matter present both obey the ideal gas law. The temperature of matter will be denoted T_m (this temperature may be different from the cosmic background radiation temperature after decoupling; after all, the matter and radiation are no longer interacting). The ideal gas law states that

$$p = \frac{NkT_m}{\mathcal{V}} \tag{5.57}$$

Thus variations in pressure p depend on variations in both T_m and $\mathcal{V}$:

$$\mathrm{d}p = \frac{Nk}{\mathcal{V}}\,\mathrm{d}T_m - \frac{NkT_m}{\mathcal{V}^2}\,\mathrm{d}\mathcal{V} = \frac{p}{T_m}\,\mathrm{d}T_m - \frac{p}{\mathcal{V}}\,\mathrm{d}\mathcal{V}. \tag{5.58}$$

We need to eliminate the temperature variable T_m in order to have $\mathrm{d}p$ just in terms of $\mathrm{d}\mathcal{V}$.

To accomplish this, we need the first law of thermodynamics ($\mathrm{d}E = -p\,\mathrm{d}\mathcal{V}$) (with no heat exchange), and the energy of a monatomic gas,

$$E = \frac{3}{2}NkT_m. \tag{5.59}$$

Thus

$$\mathrm{d}E = -p\,\mathrm{d}\mathcal{V} = \left(\frac{3}{2}Nk\right)\,\mathrm{d}T_m. \tag{5.60}$$

Using the ideal gas law to substitute for p,

$$-\left(\frac{NkT_m}{\mathcal{V}}\right)\,\mathrm{d}\mathcal{V} = \left(\frac{3}{2}Nk\right)\,\mathrm{d}T_m \tag{5.61}$$

$$\Rightarrow -\frac{\mathrm{d}\mathcal{V}}{\mathcal{V}} = \frac{3}{2}\frac{\mathrm{d}T_m}{T_m}. \tag{5.62}$$

We can now eliminate $\mathrm{d}T_m/T_m$ from equation (5.58):

$$\mathrm{d}p = -\frac{2}{3}\frac{p}{\mathcal{V}}\,\mathrm{d}\mathcal{V} - \frac{p}{\mathcal{V}}\,\mathrm{d}\mathcal{V} \tag{5.63}$$

$$= -\frac{5}{3}\frac{p}{\mathcal{V}}\,\mathrm{d}\mathcal{V}. \tag{5.64}$$

Comparing with equation (5.56) gives

$$\boxed{c_s^2 = \frac{5}{3}\frac{p}{\rho}}. \tag{5.65}$$

Matter temperature after decoupling

We can use the above treatment to find out how fast matter cooled after decoupling. First go back to equation (5.62), and integrate:

$$-\int\frac{\mathrm{d}\mathcal{V}}{\mathcal{V}} = \frac{3}{2}\int\frac{\mathrm{d}T_m}{T_m}; \tag{5.66}$$

$$-\ln\mathcal{V} = \frac{3}{2}\ln T_m + C; \tag{5.67}$$

$$T_m^{3/2} = C'\mathcal{V}^{-1}; \tag{5.68}$$

$$T_m = C'\mathcal{V}^{-2/3}, \tag{5.69}$$

where C and $C' = \exp C$ are constants of integration.

IF we now consider the expansion of the universe with scale factor $a(T)$, we have $\mathcal{V}$ scaling as $a^3(t)$, so

$$\boxed{T_m \sim a^{-2}}. \tag{5.70}$$

Thus the gas temperature T_m cools as a^{-2} as the universe expands, much more rapidly than the radiation temperature $T_\gamma \sim a^{-1}$.

The sound speed before decoupling

Before decoupling time, we need to include both radiation and matter in our thermodynamics. Here the temperature $T = T_m = T_\gamma$. Also, for early times particles move at relativistic speeds, so we must include rest mass energy in the total energy density:

$$\rho = \rho_m + aT^4 + \frac{3}{2}\frac{NkT}{\mathcal{V}}. \tag{5.71}$$

The combined pressure of the matter and radiation is

$$p = \frac{1}{3}aT_\gamma^4 + \frac{NkT}{\mathcal{V}}. \tag{5.72}$$

Let us define

$$\boxed{\sigma = \frac{4aT^3\mathcal{V}}{3Nk}}. \tag{5.73}$$

As $T \sim a^{-1}$ and $\mathcal{V} \sim a^3$, σ is a constant. In terms of σ, a long thermodynamic calculation yields

$$\boxed{c_s^2 = \frac{1}{3}\left(\frac{kT\sigma}{m + kT\sigma}\right)}. \tag{5.74}$$

5.4.3 The Effect of Cosmic Expansion on Fluctuations

We have gone to some trouble to find the Jeans mass, which tells us how large a blob of gas must be for it to collapse under its own gravity (that is, slight density enhancements collapse the blob rather than merely propagating as sound waves). However, we neglected the expansion of the universe. Cosmic expansion would be important if it proceeds as fast, or faster, than the growth of the density enhancements.

The typical growth rate for fluctuations ($\delta\rho \sim e^{-i\omega t} = e^{|\omega| t}$) is

$$\frac{1}{\delta\rho}\frac{\mathrm{d}\delta\rho}{\mathrm{d}t} = |\omega| = c_s\sqrt{k_J^2 - k^2}. \tag{5.75}$$

The fastest growth rate occurs for very small wave-numbers $k \approx 0$, i.e. very large wavelengths:

$$|\omega_{\max}| = c_s k_J = (4\pi G\bar{\rho})^{1/2}. \tag{5.76}$$

Let us compare this with the expansion rate (say for $k = 0, \Omega_0 = 1$). The expansion rate is $H_0 = \dot{a}_0/a_0$. From the first evolution equation,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\bar{\rho}, \tag{5.77}$$

so

$$\frac{\dot{a}}{a} = \sqrt{\frac{8\pi G}{3}\bar{\rho}} = \sqrt{\frac{2}{3}}|\omega|_{\max}. \tag{5.78}$$

This is almost the same as the collapse rate. Thus expansion is important! There are several techniques for including expansion in studying the growth of density fluctuations.

The simplest technique involves imagining a blob of fluid expanding into an empty universe. (This is certainly *not* the homogeneous universe we have been studying! However, it may do for a crude calculation.) The mean quantities now change in time:

$$\bar{\rho} = \rho_{\mathrm{d}} \left(\frac{a_{\mathrm{d}}}{a(t)} \right)^3; \tag{5.79}$$

$$\vec{\mathbf{V}} = \vec{\mathbf{r}} \frac{\dot{a}}{a}; \tag{5.80}$$

$$\vec{\mathbf{F}} = -\left(\frac{4\pi G}{3} \bar{\rho} a^3 \right) \vec{\mathbf{r}}. \tag{5.81}$$

The resulting calculations still give the same expressions for k_J and M_J. However, density fluctuations no longer grow exponentially, but as a power law, with $\delta\rho$ growing almost linearly with time.

5.4.4 The Later Stages of Galaxy Formation

Collapsing gas clouds eventually formed galaxies and clusters of galaxies in the first billion years of the universe. These collapsing clouds started as simply a small density fluctuation. But once $\delta\rho$ grows in a fluctuation until it is of the same magnitude as $\bar{\rho}$, the linear theory we have been considering no longer applies. In general, numerical techniques must be used to further follow the development of the proto-galaxies.

An important consideration in models of galaxy formation concerns the nature of dark matter, especially as the predominant mass in the universe may be dark. During the era of galaxy formation the dark matter may have been:

- *Cold* : Particles non-relativistic. Cold dark matter allows smaller objects (galaxies, clusters) to form more easily.

- *Hot* : Particles relativistic, i.e. $\gamma = (1 - V^2)^{-1/2} \gg 1$ or $E \gg mc^2$. Hot dark matter travelling near the speed of light tends to stream out of the gravitational well of a collapsing object. Only very large objects, in particular super-clusters of galaxies, are able to contain the hot dark matter and continue their collapse.

- *Warm* : Some models have a mixture of hot and cold dark matter.

In general, the hot dark matter models favor the initial formation of large super-clusters, which then break up to form clusters and galaxies. Meanwhile the cold dark matter models favor the initial formation of smaller objects which then merge or gather together to form the clusters and super-clusters. Present observations of the distribution and proper velocities of galaxies fit the cold and warm models better than the hot models.

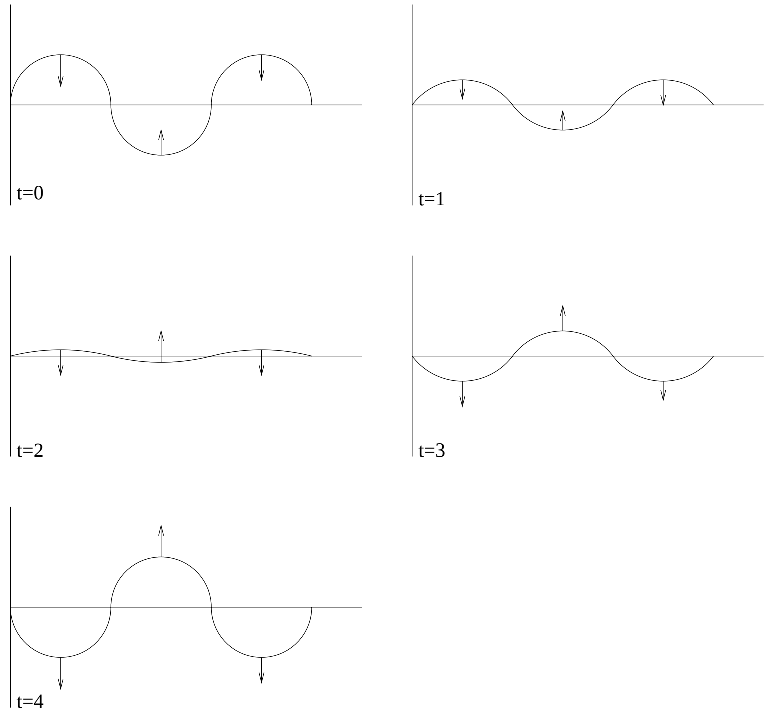
5.4.5 Fluctuations in the CMB: the Sachs–Wolfe Effect

We observe temperature fluctuations of order $\delta T/T \sim 10^{-5}$ in the microwave background. The Sachs-Wolfe effect predicts that these temperature fluctuations occur as a result of the density fluctuations at decoupling time t_d. Denser gas is hotter; on the other hand denser gas has a larger gravitational potential well – as a consequence the photons are redshifted as they struggle to pull themselves out of the well. The net effect is a reduction in energy, and hence temperature of the photons. For a density fluctuation of magnitude $\delta\rho_\mathrm{d}$ taking place with a spatial size λ_d, there will be a temperature change δT in the microwaves:

$$\frac{\delta T}{T} \approx -\frac{1}{3}\left(\frac{\lambda_\mathrm{d}}{t_\mathrm{d}}\right)\frac{\delta\rho_\mathrm{d}}{\rho_\mathrm{d}}. \tag{5.82}$$

Recall that fluctuations propagate as sound waves if they are below the Jeans mass, and collapse if they are above. Consider fluctuations that were below M_J before decoupling. These are very low frequency waves – they may have had only a few cycles between $t = 0$ and $t = t_\mathrm{d}$. This gives us an important observational tool: for some wavelengths, the wave will be peaking just at decoupling time.

To illustrate, consider a simple wave on a string:



If the wave function varies as $\cos\omega t$, then the amplitude of the wave will peak at times $t = 0, \pi/\omega, 2\pi/\omega, \ldots$. For observations of the cosmic microwave background, $t = t_\mathrm{d}$ and $\omega = \omega(k)$. Peaks occur when

$$\omega(k) = \frac{n\pi}{t_\mathrm{d}}, \qquad n = 1, 2, \ldots. \tag{5.83}$$

The lowest peak ($n = 1$) occurs at $\omega_1 = \pi/t_\mathrm{d}$, corresponding to

$$k_1 = \left(k_J^2 + \frac{\pi^2}{c_s^2 t_\mathrm{d}^2}\right)^{1/2}. \tag{5.84}$$

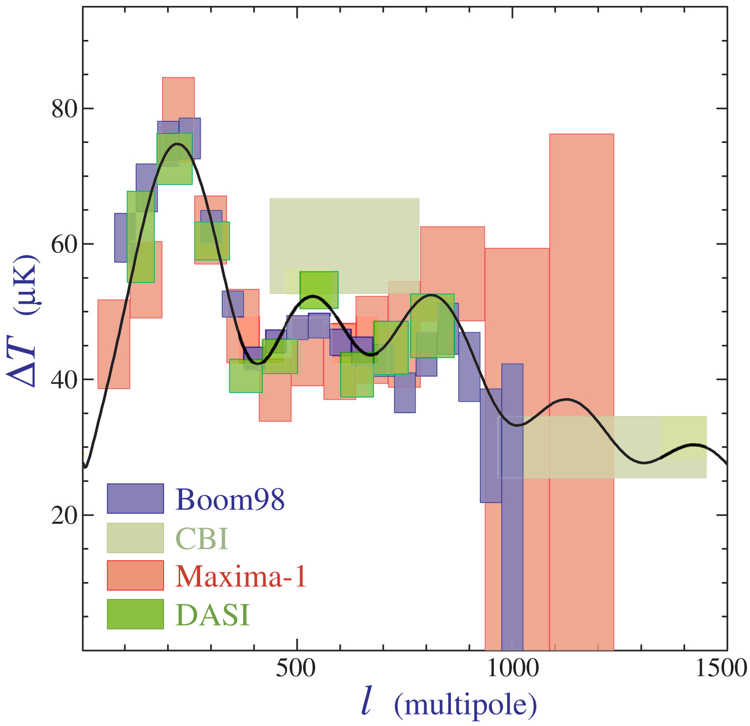


Figure 5.1: Fluctuations in μwave background temperature as a function of wave-number (spherical harmonic wave-number ℓ).

Thus a good test of cosmological theory lies in accurately predicting the value of k_1, then observing where the lowest peak is. There are some details which are beyond the scope of this course. For example, we see the sky as a sphere, so we need to express wave functions in terms of spherical harmonics rather than Fourier series. Also, intrinsic length scales at decoupling time are observed in terms of angles on the sky. The relation between apparent angular diameter and intrinsic length depends on the angular diameter distance D_A (see section 4.4).

Chapter 6

Evidence for Dark Matter

6.1 The Mass to Light Ratio

We can attempt to infer the average mass density of the universe by surveying a large region of space $\mathcal{V}$. This region will have some total luminosity L, mass M, and volume $\mathcal{V}$. We can infer the total luminosity from the apparent luminosity, using the formula $L = 4\pi d_L^2 \ell$ (equation (4.82)). Suppose the region surveyed has a fairly small redshift ($z \ll 1$). Then the astronomer may choose not to measure luminosity distance d_L directly; instead the relation between d_L and redshift z can be employed ($d_L \approx z/H_0$ for $z \ll 1$. Define the luminosity density $j_0 = L(\mathcal{V})/\mathcal{V}$, where $L(\mathcal{V})$ is the total luminosity of $\mathcal{V}$. We may use the subscript 0 on j_0 for surveys with $z \ll 1$, since they will cover galaxies seen at times close to t_0.

Let ρ_m be the matter density in the volume $\mathcal{V}$. Then $\rho_m = M(\mathcal{V})/\mathcal{V}$, so

$$\rho_m = \frac{M(\mathcal{V})}{L(\mathcal{V})} j_0. \tag{6.1}$$

If $\mathcal{V}$ is large enough, i.e. many cubic megaparsecs, then the mass to light ratio $M(\mathcal{V})/L(\mathcal{V})$ in the volume may be characteristic of the ratio M/L of the universe as a whole. In this case, we can infer the contribution Ω_{m0} of matter to the total Ω_0:

$$\begin{aligned}
\Omega_{m0} &= \frac{\rho_{m0}}{\rho_{c0}} && (6.2)\\
&= \frac{M}{L}\frac{j_0}{\rho_{c0}} && (6.3)\\
&= \frac{M}{L}\frac{8\pi G}{3H_0^2} j_0. && (6.4)
\end{aligned}$$

Davis and Huchra found

$$\boxed{j_0 \approx 2 \times 10^8 h L_\odot \mathrm{Mpc}^{-3}}, \tag{6.5}$$

where $L_\odot$ is the luminosity of the sun. This corresponds to $\frac{M}{L} \approx 1500h \frac{M_\odot}{L_\odot} \Omega_{m0}$. Thus if

$h = \frac{2}{3}$ then

$$\boxed{\frac{M}{L} \approx 1000 \frac{M_\odot}{L_\odot} \Omega_{m0}}. \tag{6.6}$$

Note that stars will only give $M/L \approx 5M_\odot/L_\odot$ – the average star has a higher mass to light ratio than the sun, but not by a factor of 1000! Old burnt out stars, interstellar clouds, gas, and dust will contribute a few more units of $M_\odot/L_\odot$ to the overall M/L, but not enough to be cosmologically significant.

Strangely, however, if we try to measure the masses of galaxies or clusters of galaxies, we get much higher values of M/L.

Exercise 6.1

Explain why a factor of h appears in the observed value of the luminosity density, equation (6.5). (e.g. why not h^2 or h^{-1}?)

6.2 Galaxy Masses

6.2.1 The Milky Way

Let us start with a basic description of our own galaxy. Let r be the distance from the galactic centre. For the sun $r_\odot \approx 8.5$ kpc. The surface brightness of the Milky Way decreases roughly exponentially away from the centre:

$$I(r) \approx I_o e^{-r/r_d}; \qquad r_d \approx 3.5 \pm 0.5\text{kpc}. \tag{6.7}$$

This implies about 70% of the luminosity of the Milky Way lies within the solar orbit. The disc brightness also varies exponentially perpendicular to the disc:

$$I(r, z) \approx I(r, 0)e^{-z/z_d}; \qquad z_d \approx 500\text{pc}. \tag{6.8}$$

The rotational velocity of the sun is $V_\odot = 220\,\text{km}\,\text{s}^{-1}$. In comparison , The Earth's orbital velocity is $V_\oplus = 30\,\text{km}\,\text{s}^{-1}$. The galactic year is about 240 million earth years. In addition to the rotation of the galaxy, the stars move randomly, with a typical velocity in the solar neighborhood of 16.5km s^{-1}.

6.2.2 Dynamics of galaxies

Escape Velocities

For a star to escape the galaxy, its total energy must be positive:

$$E = \frac{1}{2}mV^2 + m\Phi(\vec{\mathbf{x}}) > 0. \tag{6.9}$$

where $\Phi(\vec{\mathbf{x}})$ is the gravitational potential. Thus

$$\boxed{V > V_e = \sqrt{2|\Phi|}}. \tag{6.10}$$

One would expect that the fastest star in a region will have a velocity smaller than V_e (otherwise it would escape the galaxy). So we can put lower limits on the potential $\Phi(\overrightarrow{\mathbf{x}})$. This gives us information on the mass density in the galaxy, as the Poisson equation relates ϕ to ρ_m:

$$\nabla^2\Phi(\overline{\mathbf{x}}) = 4\pi G\rho_m. \tag{6.11}$$

Observations of the fastest star in the stellar neighborhood suggest that there is extra 'dark matter' that we cannot see.

Rotational Velocities

Suppose all the mass in a galaxy was concentrated in its central bulge. Then for a circular orbit the centrifugal force balances the gravitational force:

$$\frac{GM}{r^2} = \frac{V_c^2}{r} \tag{6.12}$$

where V_c is the rotational velocity and M is the mass of the galaxy. The *rotation curve* of a galaxy is the function $V_c(r)$. This is observed by looking at Doppler Shifts.

Observing $V_c(r)$ gives us information about the mass distribution. For a central mass M can be obtained very simply:

$$\boxed{M = \frac{V_c^2 r}{G}}. \tag{6.13}$$

Note also that for a central mass

$$V_c(r) = \sqrt{\frac{GM}{r}}. \tag{6.14}$$

An orbit where V_c varies as $r^{-1/2}$ is called a Keplerian orbit, as the planetary orbits studied by Kepler follow this law.

While a realistic galaxy does not have all of its mass concentrated in the centre, one might nevertheless expect the outer regions of a galaxy to display a Keplerian rotation curve, as most of the mass will be inside the orbits of outer stars. However, observers find that rotation curves are almost flat! This again implies that there exists extra mass in the outer regions of galaxies that we cannot see.

Theoretical models also suggest galaxies are enclosed in a dark massive halo. Halos help stabilize the disc against warping as well as explaining the rotation curves and the observations of fast stars. Models give $\frac{M}{L} \approx 20$ for a halo the size of the visible galaxy. Extended halos give larger $\frac{M}{L}$ The observations of $V_e \approx 500$km/s near sun (fastest stars), are consistent with a halo extending to > 40 kpc.

Exercise 6.2 *Consider a spherically symmetric galaxy where the mass density $\rho = C/r$ where r is the radial coordinate and $C =$ constant. The mass density abruptly changes to 0 at some outer radius $r = r_1$. Find the circular velocity and the escape velocity as a function of r.*

6.3 The Virial Theorem

6.3.1 Statement and Proof of the Theorem

The dynamics of clusters of galaxies allows us to measure average mass density over larger distances. Consider N galaxies with:

- masses M_i;
- position $\overrightarrow{\mathbf{x}}_i$;
- Velocities $\overrightarrow{\mathbf{V}}_i = \mathrm{d}\overrightarrow{\mathbf{x}}_i/\mathrm{d}t$;
- Forces $\overrightarrow{\mathbf{F}}_i$ (grav attraction due to other galaxies)
- relative position vector $\overrightarrow{\mathbf{r}}_{ij} \equiv \overrightarrow{\mathbf{x}}_i - \overrightarrow{\mathbf{x}}_j$. This vector has magnitude $r_{ij} = |\overrightarrow{\mathbf{r}}_{ij}|$.

We assume the cluster is gravitationally bound and that the statistical properties of the galaxy have relaxed to steady values. The total energy of the cluster is a sum of two contributions:

Total Kinetic Energy

$$\boxed{K = \frac{1}{2}\sum_{i=1}^{N} M_i V_i^2}\,. \tag{6.15}$$

Total Potential Energy

$$\boxed{W = -\frac{G}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{M_i M_j}{r_{ij}}}\,. \tag{6.16}$$

Here the Newtonian potential energy for galaxies i and j is GM_iM_j/r_{ij}. The factor 1/2 appears in the sum to compensate for double counting (e.g. $i = 1$, $j = 2$ and $i = 2$, $j = 1$).

Theorem 6.1 In a statistical steady state

$$\boxed{2K + W = 0}\,. \tag{6.17}$$

Proof 6.1

$$\begin{aligned}
2K &= \sum_{i=1}^{N} M_i V_i^2 && (6.18)\\
&= \sum_{i=1}^{N} M_i \frac{\mathrm{d}\vec{\mathbf{x}}_i}{\mathrm{d}t}\cdot\frac{\mathrm{d}\vec{\mathbf{x}}_i}{\mathrm{d}t} && (6.19)\\
&= \sum_{i=1}^{N} M_i \left(\frac{\mathrm{d}}{\mathrm{d}t}\left[\vec{\mathbf{x}}_i\cdot\frac{\mathrm{d}\vec{\mathbf{x}}_i}{\mathrm{d}t}\right] - \vec{\mathbf{x}}_i\cdot\frac{\mathrm{d}^2\vec{\mathbf{x}}_i}{\mathrm{d}t^2}\right) && (6.20)\\
&= \frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{N} M_i\left(\vec{\mathbf{x}}_i\cdot\vec{\mathbf{V}}_i\right) - \sum_{i=1}^{N}\vec{\mathbf{x}}_i\cdot\vec{\mathbf{F}}_i. && (6.21)
\end{aligned}$$

The first term is the time derivative of a quantity called the 'virial', which gives the average value of $M\vec{\mathbf{x}}\cdot\vec{\mathbf{V}}$. According to our assumption of statistical steady state, the virial should not change in time. Thus we ignore the first term:

$$\begin{aligned}
2K &= -\sum_{i=1}^{N}\vec{\mathbf{x}}_i\cdot\vec{\mathbf{F}}_i && (6.22)\\
&= -\sum_{i=1}^{N}\vec{\mathbf{x}}_i\cdot\sum_{j=1}^{N} GM_iM_j\frac{\vec{\mathbf{r}}_{ij}}{r_{ij}^3} && (6.23)\\
&= -G\sum_{i=1}^{N}\sum_{j=1}^{N} M_iM_j\frac{\vec{\mathbf{x}}_i\cdot\vec{\mathbf{r}}_{ij}}{r_{ij}^3}. && (6.24)
\end{aligned}$$

Suppose we exchange the i and j labels:

$$2K = -G\sum_{j=1}^{N}\sum_{i=1}^{N} M_jM_i\frac{\vec{\mathbf{x}}_j\cdot\vec{\mathbf{r}}_{ji}}{r_{ji}^3}. \tag{6.25}$$

Now $\vec{\mathbf{r}}_{ji} = -\vec{\mathbf{r}}_{ij} = -(\vec{\mathbf{x}}_i - \vec{\mathbf{x}}_j)$. So if we sum the last two equations, we find

$$\begin{aligned}
4K &= -G\sum_{i=1}^{N}\sum_{j=1}^{N} M_iM_j\frac{(\vec{\mathbf{x}}_i-\vec{\mathbf{x}}_j)\cdot\vec{\mathbf{r}}_{ij}}{r_{ij}^3} && (6.26)\\
&= -G\sum_{i=1}^{N}\sum_{j=1}^{N} M_iM_j\frac{1}{r_{ij}} && (6.27)\\
&= -2W. && (6.28)
\end{aligned}$$

Q.E.D.

6.3.2 Application to Observations

We can use the virial theorem to estimate the typical mass to light ratio of galaxies which belong to large clusters or superclusters. We first make a few assumptions to simplify our analysis. First, in order to employ the virial theorem at all, we must assume that the cluster has achieved a statistical steady state (after a galaxy cluster forms in the early universe, it may evolve for several billion years before settling down to a steady state). Secondly, we assume that the mass to light ratio M/L is constant for all galaxies in a cluster. Then $M_i = (M/L)L_i$, and so

$$2K = \left(\frac{M}{L}\right)\sum_{i=1}^{N} L_i V_i^2 \tag{6.29}$$

$$W = -\left(\frac{M}{L}\right)^2 \frac{G}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{L_i L_j}{r_{ij}}. \tag{6.30}$$

Since K and W involve different powers of M/L, we can use the equation $2K + W = 0$ to extract M/L. However, we first need to do a bit of work to relate the quantities appearing in the expressions for K and W to observable quantities.

Given a quantity such as V_i^2 which varies from galaxy to galaxy, let

$$\langle V_i^2 \rangle \equiv \frac{1}{N}\sum_{i=1}^{N} V_i^2 \tag{6.31}$$

denote its average value. For a double sum over pairs of galaxies (with $N(N-1)/2$ pairs) the average of a quantity such as $1/r_{ij}$ becomes

$$\langle r_{ij}^{-1} \rangle \equiv \frac{1}{N(N-1)}\sum_{i=1}^{N}\sum_{j=1}^{N} r_{ij}^{-1}. \tag{6.32}$$

Then we can write equation (6.17)

$$2K = \left(\frac{M}{L}\right) N\langle L_i V_i^2 \rangle; \tag{6.33}$$

$$W = -\left(\frac{M}{L}\right)^2 \frac{G}{2} N(N-1)\langle L_i L_j r_{ij}^{-1} \rangle. \tag{6.34}$$

Velocities We can obtain information on velocities $\vec{\mathbf{V}}_i$ by observing redshifts. The cluster as a whole may have some average redshift due to the expansion of the universe (and our own motion relative to the cosmic frame), but in addition there will be extra redshifts and blueshifts due to proper galactic motions internal to the cluster. Only the line of sight (los) component of a proper motion contributes to the observed redshift, that is, we only observe one velocity component $V_{i\,\mathrm{los}}$. Fortunately, when observing many galaxies we only need one velocity component, as the other two components have the same averages.

For isotropic motions, each of the three velocity components contributes equally to $\langle L_i V_i^2 \rangle$, so

$$\langle L_i V_i^2 \rangle = 3 \langle L_i V_{\rm los}^2 \rangle. \tag{6.35}$$

Luminosities We observe the absolute luminosity L_i given the apparent luminosity and the luminosity distance to the cluster, $L_i = 4\pi d_L^2 \ell_i$. The luminosity distance in turn can be calculated from the average redshift of the cluster z and H_0 (if $z \ll 1$).

Relative Distances The relative distance r_{ij} appears in the expression for potential energy (in the denominator). But we can only observe the distance between two galaxies as it is projected on the sky (call this d_{ij}). However, we may assume random orientations for the vector $\overrightarrow{\mathbf{r}}_{ij}$. Let the angle between $\overrightarrow{\mathbf{r}}_{ij}$ and the line of sight be θ_{ij}. Then

$$d_{ij} = r_{ij} \sin \theta_{ij}. \tag{6.36}$$

Thus

$$\langle \frac{1}{d_{ij}} \rangle = \langle \frac{1}{r_{ij} \sin \theta_{ij}} \rangle \tag{6.37}$$

Now, $\overrightarrow{\mathbf{r}}_{ij}$ is uncorrelated with θ_{ij} (as θ_{ij} depends on where *we* are on Earth, rather than on how the galaxies in the cluster are situated). So

$$\langle \frac{1}{r_{ij} \sin \theta_{ij}} \rangle = \langle \frac{1}{r_{ij}} \rangle \langle \frac{1}{\sin \theta_{ij}} \rangle. \tag{6.38}$$

Warning: d_{ij} *is* correlated to θ_{ij}. Here

$$\langle \frac{1}{\sin \theta_{ij}} \rangle = \frac{\int_0^{2\pi} \int_0^{\pi} \frac{1}{\sin\theta} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi}{\int_0^{2\pi} \int_0^{\pi} \sin\theta \, \mathrm{d}\theta \, \mathrm{d}\phi} \tag{6.39}$$

$$= \frac{2\pi^2}{4\pi} = \frac{\pi}{2}. \tag{6.40}$$

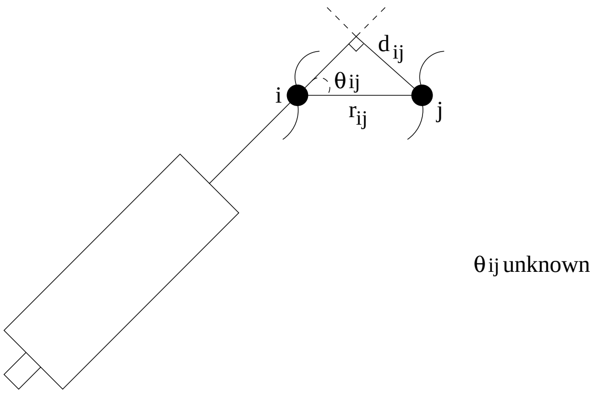
Thus

$$\langle r_{ij}^{-1} \rangle = \frac{2}{\pi} \langle d_{ij}^{-1} \rangle. \tag{6.41}$$

We will assume $L_i L_j$ is uncorrelated with r_{ij}^{-1} so

$$\langle L_i L_j r_{ij}^{-1} \rangle = \langle L_i L_j \rangle \langle r_{ij}^{-1} \rangle \tag{6.42}$$

$$= \frac{2}{\pi} \langle L_i L_j \rangle \langle d_{ij}^{-1} \rangle. \tag{6.43}$$



Putting it all together,

$$2K = \left(\frac{M}{L}\right) 3N\langle L_i V_{\rm los}^2\rangle; \tag{6.44}$$

$$W = -\left(\frac{M}{L}\right)^2 \frac{G}{\pi} N(N-1)\langle L_i L_j\rangle\langle d_{ij}^{-1}\rangle. \tag{6.45}$$

Thus the virial theorem equation (6.17) gives

$$\boxed{\left(\frac{M}{L}\right) = \frac{3\pi\langle L_i V_{\rm los}^2\rangle}{G(N-1)\langle L_i L_j\rangle\langle d_{ij}^{-1}\rangle}}. \tag{6.46}$$

Results

- Coma Cluster : $\frac{M}{L} \approx 240$.
- Perseus Cluster : $\frac{M}{L} \approx 400$.
- Median Cluster value : $\frac{M}{L} \approx 200$.

These results suggest that the matter density Ω_{m_0} lies between $\approx 0.2 - 0.4$. This is consistent with the recent WMAP microwave background estimates $\Omega_{m_0} \approx 0.27$.

Chapter 7

Cosmological Puzzles

7.1 Topology of the Universe

Recall that a *multiply connected* manifold contains curves which cannot be shrunk to a point.

k	Simple	Multiple (example)
+1	S^3 (3–Sphere)	P^3 (Projective 3-sphere)
0	E^3 (Euclidean space)	T^3 (3–torus)
-1	H^3 (Hyperbolic sphere)	Compact hyperbolic manifolds

7.1.1 Can topology be observed?

Let us consider the flat ($k = 0$) metrics. Can we distinguish the 3–torus topology from the infinite Euclidean space?

Suppose that at any one cosmic time the spatial part of the universe is the three torus T^3. We can give the three torus coordinates x, y, and z, each of which range from 0 to 1. Points with $x = 0$ and $x = 1$ are identified, just as on a circle the angles $\phi = 0$ and $\phi = 2\pi$ are identified. Similarly, we identify $y = 0$ and $y = 1$, and also identify $z = 0$ and $z = 1$.

The flat T^3 three torus cosmology has metric line element

$$\mathrm{d}s^2 = \mathrm{d}t^2 - a^2(t)(\mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2); \tag{7.1}$$

$$0 \le x < 1, \qquad 0 \le y < 1, \qquad 0 \le z < 1. \tag{7.2}$$

Here the scale factor $a(t)$ gives the periodicity length. For simplicity, we suppose this cosmology satisfies the matter-dominated $k = 0$ solution.

The simplest signature of a multiply connected space is the observation of the same object in two different places on the sky. For example, in the 3–torus cosmology we should be able to see a very distant image of the sun (if the periodicity length a_0 is less than the distance light can travel during the age of the sun). To circle the universe in the minimum possible time, the light should travel parallel to one of the axes (say the x axis). Let the sun be at $x = 0 = 1$. Sunlight comes back to our solar system when the total x interval travelled is $\Delta x = 1$. What is the time t_1 of sunlight which has circled the universe once in the x direction?

For a light beam in the x direction

$$\begin{aligned} 0 = \mathrm{d}s^2 &= \mathrm{d}t^2 - a^2(t)\,\mathrm{d}x^2. && (7.3)\\ \Rightarrow \mathrm{d}x &= \mathrm{d}t/a(t); && (7.4)\\ 1 = \int_0^1 \mathrm{d}x &= \int_{t_1}^{t_0} \mathrm{d}t/a(t). && (7.5) \end{aligned}$$

Since $k = 0$,

$$a(t) = a_0(t/t_0)^{2/3}; \qquad t_0 = \frac{2}{3H_0}. \tag{7.6}$$

Let $w = t/t_0$. Then

$$\begin{aligned} 1 &= \frac{t_0}{a_0}\int_{w_1}^1 w^{-2/3}\,\mathrm{d}w && (7.7)\\ &= 2\frac{t_0}{a_0}\left(w_1^{-1/2} - 1\right). && (7.8) \end{aligned}$$

Solving for t_1,

$$t_1 = t_0\left(1 + \frac{a_0}{2t_0}\right)^{-2}. \tag{7.9}$$

For example, if $a_0 = 10^9$ parsecs and $t_0 = 10^{10}$ y, then

$$\begin{aligned} t_1 &= t_0\left(1 + \frac{3.26 \times 10^9}{2 \times 10^10}\right)^{-2} && (7.10)\\ &= 0.74\,t_0. && (7.11) \end{aligned}$$

Thus a periodic universe would be observable if the same galaxy or quasar were seen in two different places in the sky. As the object would be seen from different sides, it would be difficult to prove that the two images were the same object. Also the two images would probably come from different times. A second possible observational signature would be periodicity in number counts of galaxies or quasars as a function of redshift. A third observational signature would involve large angle correlations in fluctuations in the microwave background.

7.2 Vacuum Energy Density

7.2.1 Pressure is negative!

Consider the second evolution equation with $\rho = \rho_\Lambda$ and $p = p_\Lambda$ both independent of time (and hence of $a(t)$):

$$\frac{\mathrm{d}}{\mathrm{d}a}(\rho_\Lambda a^3) = -3p_\Lambda a^2. \tag{7.12}$$

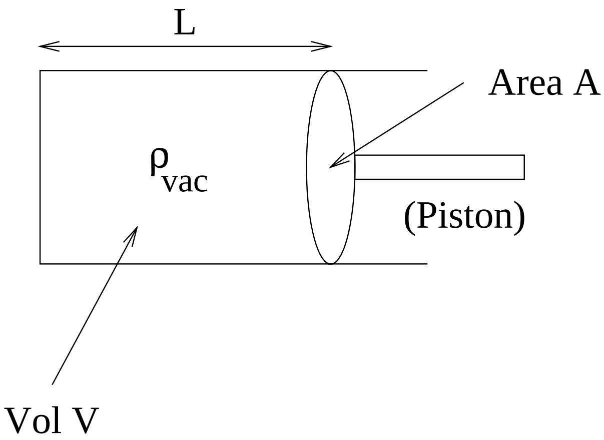
This becomes

$$\rho_\Lambda \frac{\mathrm{d}a^3}{\mathrm{d}a} = 3\rho_\Lambda a^2 = -3p_\Lambda a^2. \tag{7.13}$$

Therefore

$$\boxed{\rho_\Lambda = -p_\Lambda}. \tag{7.14}$$

7.2.2 Negative pressure from thermodynamics



Fill a piston with vacuum energy $E = \rho_\Lambda V = \rho_\Lambda LA$

$$\mathrm{d}E = \rho_\Lambda A\,\mathrm{d}l \tag{7.15}$$

The first law of thermodynamics with no heat exchange says

$$\boxed{\mathrm{d}E = -p\,\mathrm{d}V}. \tag{7.16}$$

Thus

$$-p_\Lambda\,\mathrm{d}V = \rho_\Lambda A\,\mathrm{d}l = -p_\Lambda A\,\mathrm{d}l \tag{7.17}$$

so again we have

$$\boxed{\rho_\Lambda = -p_\Lambda}. \tag{7.18}$$

7.3 Evolution Equations

With $\rho = \rho_\Lambda$ the first evolution equation reads

$$\dot{a}^2 + k = \frac{8\pi G}{3}\rho_\Lambda a^2. \tag{7.19}$$

Write this with $\Lambda \equiv 8\pi G\rho_\Lambda = \text{constant}$:

$$\dot{a}^2 + k = \frac{\Lambda}{3}a^2. \tag{7.20}$$

7.3.1 The $k = 0$ solution

$$\dot{a}^2 = \frac{\Lambda}{3} a^2 \tag{7.21}$$

$$\dot{a} = \pm\sqrt{\frac{\Lambda}{3}}\, a \tag{7.22}$$

$$\Rightarrow a(t) = Ae^{+\sqrt{\frac{\Lambda}{3}}t} + Be^{-\sqrt{\frac{\Lambda}{3}}t} \tag{7.23}$$

The first term gives exponential expansion. If the universe becomes more and more dominated by dark energy in the form of vacuum, then the expansion will become more and more exponential.

For most choices of A and B, $a(t)$ never has the value 0, so the universe would never have started with zero size. If indeed $a(0) = 0$, then $B = -A$ and $a(t) = 2A\sinh\left(\sqrt{\frac{\Lambda}{3}}t\right)$.

7.3.2 The static solution

Suppose the universe is filled with both vacuum energy and ordinary matter mass-energy. Assume $k = +1$ and, as in section 3.6.1, rewrite the first evolution equation as an energy balance equation, with

$$K = 1/2\dot{a}^2 \quad \text{Kinetic Energy} \tag{7.24}$$

$$V(a) = -\frac{\Lambda}{6}a^2 - \left(\frac{H_0^2\Omega_0 a_0^3}{2}\right)a^{-1} \quad \text{Potential Energy} \tag{7.25}$$

$$E = -\frac{1}{2} \quad \text{Total Energy.} \tag{7.26}$$

The potential energy curve has a maximum point at some expansion factor $a_{\max}$. If the universe sits precisely at this maximum, i.e. if

$$V(a_{\max}) = E = -\frac{1}{2}, \tag{7.27}$$

Then the universe will be in a static state (no expansion or contraction).

Originally, Einstein thought that this solution would save him from the absurd idea that the universe is changing its size. Less than a decade later, Hubble and Slipher found that the universe does indeed expand. Einstein then called this ‘the greatest mistake of his career’. But the real mistake was in trusting the static solution – it is an unstable equilibrium! If part of the universe has just one extra atom, then ρ_m will be slightly too high, and that part of the universe will collapse under the extra gravitational attraction. Another part of the universe with one less atom will begin to expand.

Exercise 7.1 *Suppose the universe is filled with material of energy density ρ and pressure $p = C\rho$ where $C =$ constant.*

a. *What is $\rho(a)$?*

b. *Suppose that $k = 0$. Find $a(t)$ and t_0 in terms of C. Suppose $h = 2/3$. How old is the universe (in years) if C=-1/2?*

7.4 Inflation

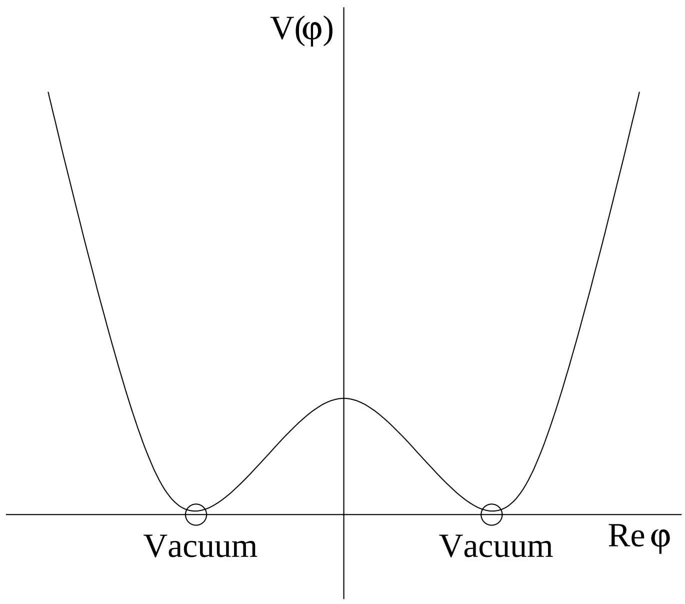
Definition : The vacuum state for a field is its minimum energy state. For electromagnetic fields, the vacuum state is $F_{ab} = 0$, i.e. $\overrightarrow{\mathbf{E}} = \overrightarrow{\mathbf{B}} = 0$. However, for some fields the vacuum state may be nonzero!

7.4.1 The Higgs effect

Consider a field ψ which has a potential energy law

$$V(\psi) = (|\psi|^2 - b^2)^2 \tag{7.28}$$

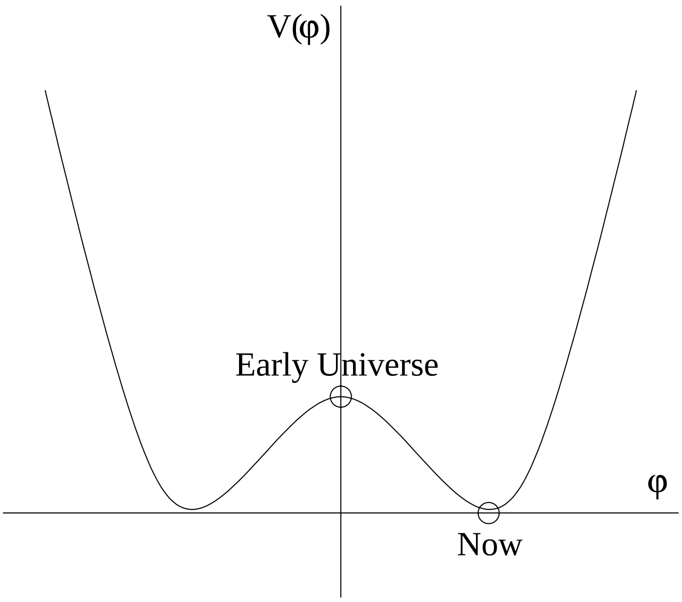
for some constant b. Then the vacuum occurs at points where $|\psi| = b$.



To make a very long story short, the weak force equations as first conceived by Yang and Mills in 1954 looked very much like electromagnetism. The equations did not quite work, as they gave a long range force and zero mass photons (just like E&M). The Higgs effect fixes these problems. This effect, applied to the Yang-Mills field by Weinberg and Salam in the late 1960's, takes into account the displacement of the vacuum due to a potential like equation (7.28). One considers a new field $h = \psi - b$, and rewrites the weak force equations in terms of h. The new equations have extra terms involving b. These extra terms modify the structure of the equations, giving $W_{\pm}, Z_0$ mass, making the weak force short range.

7.4.2 The Inflation model

Idea: In the early universe the Higgs-field (or some similar field) was not in its vacuum state.



Early means $< 10^{-33}$ seconds (or so depending on model). Consequence : a large vacuum energy associated with $V(\psi = 0)$ This sent the universe into a brief but incredibly rapid era of exponential expansion. At the end of the inflationary era the Higgs field moved to the vacuum state, releasing the old false vacuum energy to form matter and radiation.

7.5 The Flatness Problem

There are several puzzles about the beginning of the universe which the inflationary model solves. It should be emphasized that other models, such as the cyclic universe model, may also be able to solve these puzzles.
Observations give, with generous bounds:

$$0.9 \leq \Omega_0 \leq 1.1. \tag{7.29}$$

Why is Ω_0 so close to one? The question becomes more striking at early times, because for most cosmological models $\Omega(t) \to 1$ as $t \to 0$. To see this, recall from equation (3.55)

$$\boxed{\Omega(t) = 1 + \frac{k}{\dot{a}(t)^2}}. \tag{7.30}$$

so if $\dot{a} \gg 1$ then $\omega(t) \approx 1$. For both matter and radiation dominated solutions, in fact, we have

$$\begin{aligned} \lim_{t\to 0} \dot{a} &= \infty, & (7.31)\\ \Rightarrow \qquad \lim_{t\to 0} |\Omega(t) - 1| &= 0. & (7.32)\end{aligned}$$

Why did Ω start out so close to 1? This problem is misnamed the 'flatness problem' because if Ω exactly equals 1 then $k = 0$, so the universe is flat. However, for $k \neq 0$ the Ricci curvature goes as a^{-2}, so actually $k \neq 0$ universes begin with infinite curvature.

7.5.1 Inflation

Suppose at, say, $t < 10^{-35}$ sec $\Omega(t) = 0.2$ or 4.8 or anything far from 1. Then inflation occurs. As a consequence

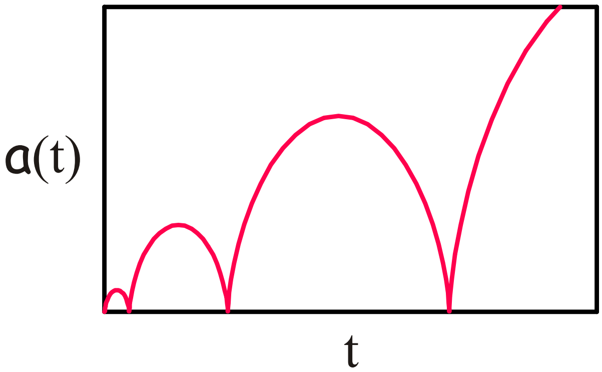
$$\begin{aligned} a(t) &= e^{\sqrt{\frac{\Lambda}{3}}\,t}, & (7.33)\\ \dot{a}(t) &= \sqrt{\frac{\Lambda}{3}}e^{\sqrt{\frac{\Lambda}{3}}\,t}, & (7.34)\end{aligned}$$

so

$$|\Omega(t) - 1| = \frac{3}{\Lambda}e^{-2\sqrt{\frac{\Lambda}{3}}\,t}. \tag{7.35}$$

Thus $|\Omega(t) - 1|$ decreases exponentially during inflation.

7.5.2 Cyclic universe



In the cyclic universe model, there have been previous universes. At the end of one universe, there is a 'big crunch', where the universe compresses to extremely high densities and bounces. Because entropy always increases, each successive cycle may be different, perhaps lasting longer. The time at which $\Omega = 1.01$ (say) will be greater in each successive cycle. We live in cycle where $\Omega_0 \approx 1.01$ at $t_0 \approx 13.8$ billion years.

7.5.3 The Anthropic principle

The anthropic principle states that we must live in a universe suitable for life as we know it. This means that many cosmological models can simply be thrown away if they lead to a dead universe. In particular, any universe with Ω very different from 1 at early times would have no observers.

If $\Omega = 1.1$ at 10^{-33} seconds then the total time between the Big Bang and Bing Crunch would be < 1 second. Thus there would not be enough time to form stars or evolve life. If $\Omega = 0.9$ then $\Omega_0 \ll 0.1$. Thus there would be not enough matter to form stars.

7.6 The Horizon Problem

7.6.1 The horizon

The *horizon* is the set of points which are just coming into view. Light from the horizon has had to travel the entire age of the universe to reach us. Equivalently, we can define the horizon to be the set of points which correspond to $z_h = \infty$, i.e. infinite redshift. The horizon coordinate η_h is equal to the development angle $\xi(t_0)$: for a photon

$$0 = \mathrm{d}\tau^2 = \mathrm{d}t^2 - a^2(t)\mathrm{d}\eta^2 \tag{7.36}$$

$$= -\frac{\mathrm{d}t}{a(t)} = -\,\mathrm{d}\xi \quad (\xi = \text{development angle}) \tag{7.37}$$

Thus

$$\eta_h = \int_0^{t_0} \frac{\mathrm{d}t}{a(t)}. \tag{7.38}$$

Galaxies with $\eta > \eta_h$ cannot yet be seen, because their light must take more than the age of the universe to reach us.

Consider two points on our horizon in opposite directions (A, B). Point B will be beyond the horizon of point A. Thus in simple cosmological models A has never communicated with B. Why, then, is the cosmic microwave background at A so similar in temperature to that at B?

7.6.2 Inflation

- Before inflation, A and B were indeed within each other horizons and exchanged heat until they were the same temperature.
- Inflation expands space between A and B by many orders of magnitude. As a consequence, they are now outside each others horizons.

7.6.3 Cyclic Universe

Thermalization occurs during previous cycles.

7.6.4 Initial Conditions

How did the universe start? It could have been created homogeneously.

7.7 Monopoles

Hypothetical particles, mass $\approx 10^{15}$ GeV ($M_p \approx 1$ Gev), with magnetic 'charge'. Some Grand Unified Theories of elementary particles suggest that these may have been created in the first 10^{-35} seconds. The puzzle is - where are they? (they should dominate the mass of the universe!).

Galaxies have magnetic fields. If lots of monopoles existed inside galaxies, they would short out the galactic magnetic fields (freely moving charges move to counteract the field). Upper bounds on monopole numbers (from the galactic magnetic field) are much less than the density predicted from the grand unified theories.

7.7.1 Inflation

Suppose monopoles were produced before inflation. Then the density of primordial monopoles would have decreased exponentially during the expansion.

7.8 Entropy puzzle

Why is the universe in a low entropy state?

Black holes are in very high entropy states (higher than a uniform gas). Gravitational collapse (even just to a star or galaxy) increases entropy. Thus the initial uniform gas in the early universe was in fact a low entropy state.

7.8.1 Inflation

Even if entropy started high, entropy/unit volume was low after inflation. Also the break down of the vacuum at the end of the inflationary era creates (almost) uniform gas.

7.8.2 Cyclic Universe

Entropy increases each cycle. But if the universe in our cycle is much larger than in the previous cycle, then the entropy/unit volume may still be lower.

7.8.3 Anthropic Principle

Evolution of life requires low entropy!