1 A mathematical description of the universe

The broad aim of cosmology is to build a mathematical model that describes the observed properties of the universe. Furthermore, any model must change to accommodate the growing number of observations performed upon the universe. The aim of this course is to describe the development of a mathematical model of the universe: its contents and evolution. We begin by considering the basic principles and observations that form the cornerstone of any cosmological model.

1.1 Observed properties of the universe

1. The night sky is dark. Is the universe therefore bounded in space or bounded in time?
2. General relativity describes accurately gravity on Solar System scales.
3. The local universe is expanding uniformly and isotropically according to Hubble's law. We assume that peculiar velocities can be described local gravitational effects.
4. The Hubble time, H^{-1}, the ages of the oldest stars, and the radioactive dating of terrestrial elements all approximately agree, i.e. of the order of a few billion years in each case.
5. The observed universe is isotropic on very large scales.
6. We observe a nearly isotropic background of microwave radiation with a blackbody spectrum of temperature $T = 2.73$ K. This is the cosmic microwave background (CMB).
7. The observed abundance of H, He and Li (and their isotopes) agree with the predictions of nucleosynthetic reactions occurring within the early universe.
8. The present day universe is structured (galaxies, clusters and superclusters) - note that this point does not contradict point (5) above.
9. Dynamical, X-ray and lensing studies imply that most of the matter in the universe is dark.
10. The CMB displays temperature fluctuations of order $(\Delta T/T)_{rms} \sim 10^{-5}$ on scales of order one degree.
11. Observations of supernovae type Ia (SNe Ia) and the CMB indicate that the universe is spatially flat (see later) and is expanding at an accelerating rate (see much later). Note that though many observational tests agree with SNe Ia + CMB results, the SNe Ia + CMB results are the simplest and best understood.

1.2 Basic principles

1. The observed universe is isotropic.
2. The cosmological principle (an extension of the Copernican principle) assumes that we occupy no special place in the universe.
3. The combination of observed isotropy with the cosmological principle implies that the universe is homogeneous.

1.3 General relativity and the line element (basic version)

Measuring distance and time in the universe represents a fundamental cosmological challenge. With the advent of Special Relativity (SR) in 1905 and General Relativity (GR) in 1915, time and space no longer remained independent quantities. General relativity explains gravity as a consequence of the changing geometry, or spacetime, of the universe. Geometry is itself related to the local properties of matter and energy.

This leads to the famous statement, **'curved space tells matter how to move, matter tells space how to curve'**.

GR describes the geometry of the universe as a four dimensional spacetime in terms of the Ricci curvature tensor $R_{\mu\nu}$, the curvature scalar R and the metric tensor $g_{\mu\nu}$. These terms are related to the local distribution of matter and energy via the Einstein Field Equation (EFQ),

$$R_{\mu\nu} - \frac{1}{2}\, g_{\mu\nu}\, R = -8\pi\, G\, T_{\mu\nu}, \tag{1}$$

where $T_{\mu\nu}$ is the stress–energy tensor (i.e. matter and energy) and G is Newton's constant in relativistic ($c = 1$) units. We ignore the possible effect of Λ at present.

The most basic geometrical operation is to measure the distance between nearby points via the line element, e.g. Pythagoras's theorem. Within the spacetime defined by the Riemann tensor, the line element, ds, describing the infinitesimal distance between two nearby points is

$$\mathrm{d}s^2 = g_{\mu\nu}\, \mathrm{d}x^\mu\, \mathrm{d}x^\nu. \tag{2}$$

Depending upon the exact definition of $g_{\mu\nu}$, the diagonal elements are $g_{00}, g_{11}, g_{22}, g_{33} = +1, -1, -1, -1$ for Euclidean or flat space (Index 1 refers to time and indices 2, 3 and 4 refer to space). Non–diagonal elements are zero in this case.

One of the early goals of theoretical cosmology was to develop a spacetime metric or line element that would satisfy the aforementioned cosmological principle, i.e. homogeneity and isotropy. Robertson and Walker (1935) demonstrated that in *any* homogeneous and isotropic cosmological model based upon Riemannian geometry, the most general line element is

$$\mathrm{d}s^2 = c^2\ \mathrm{d}t^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1-kr^2} + r^2\ (\mathrm{d}\theta^2 + \sin^2\theta\ \mathrm{d}\phi^2)\right], \tag{3}$$

where $a(t)$ is freely chosen and k is a constant.

As we shall see later, k is related to the radius of curvature of the pseudo–surface defined by $\mathrm{d}t = 0$. In addition, $a(t)$ will be seen to describe the time evolution of the physical scale of the universe. Its form results from the insertion of the components of the metric tensor into the EFQ to obtain the **Friedmann equation** (or, as we shall obtain it, via a Newtonian analogue employing test particles in a locally flat universe). As $a(t)$ is used to describe the time evolution of the size of the universe, the spatial coordinates r, θ and ϕ are considered as fixed, or co–moving, coordinates.

1.3.1 Understanding the curvature constant

What are the effects of curvature upon geometry? Consider a 2D surface that forms the surface of 3D sphere (Figure 1). The curvature of a 3D sphere of radius R is defined as $K \equiv 1/R^2$. If one proceeds a short distance D from a point of origin and describes a circle about this point, what is the circumference thus traced?

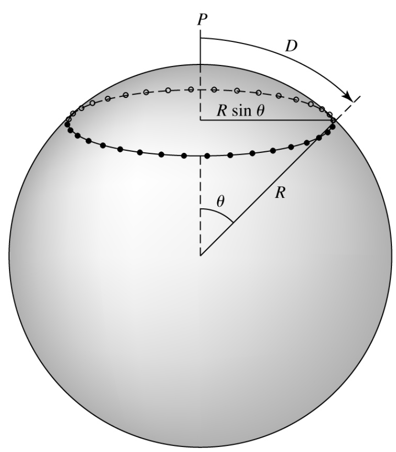

Figure 1: The geometry of a two dimensional surface considered embedded within a three dimensional space.

The measured circumference $C_{meas} = 2\pi R\ \sin\theta$ and $\theta = D/R$. Rearranging this yields

$$C_{meas} = 2\pi R \sin(D/R). \tag{4}$$

Taking a Taylor expansion in $\sin(D/R)$, one obtains

$$\begin{aligned} C_{meas} &= 2\pi R \left[\frac{D}{R} - \frac{1}{6}\left(\frac{D}{R}\right)^3 + O\left(\frac{D}{R}\right)^5 \right], \text{or} \\ C_{meas} &= 2\pi D \left(1 - \frac{1}{6}\frac{D^2}{R^2}\right) = 2\pi D \left(1 - \frac{KD^2}{6}\right). \end{aligned} \quad (5)$$

So the geometry measured on the two dimensional surface is modified by a term depending upon the inverse square of the effective radius of curvature. The circumference of a circle drawn on a surface of positive curvature is less than the circumference expected based upon Euclidean geometry (with the opposite being true for a surface of negative curvature). As the radius of curvature tends to infinity (i.e. flat space), the properties of a circle drawn upon the surface revert to the Euclidean case.

1.3.2 Determining the line element of a curved space

A fundamental operation in geometry is to measure the distance between two closely separated points P_1 and P_2. The expression for this distance is referred to as the line element. The line element for a flat, two dimensional surface may be written using Polar coordinates r and ϕ as

$$(\mathrm{d}l)^2 = (\mathrm{d}r)^2 + (r\ \mathrm{d}\phi)^2, \quad (6)$$

as shown in Figure 2a. For an example of a line element describing a curved 2D surface we again

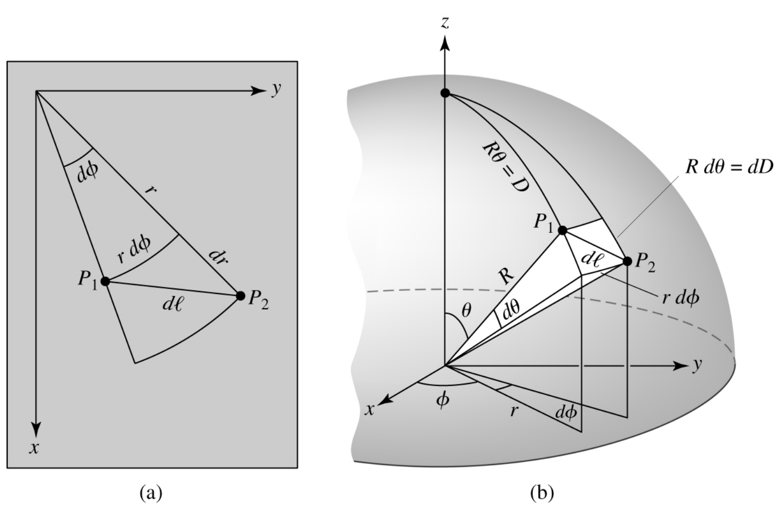

Figure 2: The line element of a curved two dimensional surface.

consider the surface of a sphere of radius R (Figure 2b). Here $K = 1/R^2$ and the distance between points P_1 and P_2 is now

$$\begin{aligned}(\mathrm{d}l)^2 &= (\mathrm{d}D)^2 + (r\ \mathrm{d}\phi)^2 \\ &= (R\ \mathrm{d}\theta)^2 + (r\ \mathrm{d}\phi)^2. \end{aligned} \tag{7}$$

We have $r = R\sin\theta$ and $\mathrm{d}r = R\cos\theta\ \mathrm{d}\theta$, therefore

$$R\ \mathrm{d}\theta = \frac{\mathrm{d}r}{\cos\theta} = \frac{R\ \mathrm{d}r}{\sqrt{R^2 - r^2}} = \frac{\mathrm{d}r}{\sqrt{1 - r^2/R^2}} \tag{8}$$

and

$$\begin{aligned}(\mathrm{d}l)^2 &= \left(\frac{\mathrm{d}r}{\sqrt{1 - r^2/R^2}}\right)^2 + (r\ \mathrm{d}\phi)^2 \\ &= \left(\frac{\mathrm{d}r}{\sqrt{1 - Kr^2}}\right)^2 + (r\ \mathrm{d}\phi)^2. \end{aligned} \tag{9}$$

The extension to a three dimensional space is straightforward

$$(\mathrm{d}l)^2 = \left(\frac{\mathrm{d}r}{\sqrt{1 - Kr^2}}\right)^2 + (r\ \mathrm{d}\theta)^2 + (r\sin\theta\ \mathrm{d}\phi)^2, \tag{10}$$

where r now measures the radial distance from the origin.

1.3.3 A relativistic line element

Our geometric model of the universe requires a relativistic line element describing a four dimensional spacetime of general curvature. We measure distance now as the "proper distance" between neighbouring events. We denote the time component of the proper distance to be $c\ \mathrm{d}t$ and the line element is written as

$$(\mathrm{d}s)^2 = (c\ \mathrm{d}t)^2 - \left(\frac{\mathrm{d}r}{\sqrt{1 - Kr^2}}\right)^2 - (r\ \mathrm{d}\theta)^2 - (r\sin\theta\ \mathrm{d}\phi)^2. \tag{11}$$

We note that the physical separation $(\mathrm{d}l)^2 = -(\mathrm{d}s)^2$ for $\mathrm{d}t = 0$.

We would further like to label stationary objects within this spacetime using a fixed or co-moving coordinate and to describe the varying scale of the spacetime via a dimensionless relative scale factor $a(t)$, i.e.

$$r(t) = a(t)\ \omega. \tag{12}$$

However, as the expansion of the universe will affect all of its geometric properties, including curvature, we define the time dependent curvature in terms of a time independent curvature constant, i.e.

$$K(t) \equiv \frac{k}{a^2(t)}. \tag{13}$$

Substituting for r and K in Equation 11 we obtain

$$(\mathrm{d}s)^2 = (c\ \mathrm{d}t)^2 - a^2(t)\left[\left(\frac{\mathrm{d}\omega}{\sqrt{1-k\omega^2}}\right)^2 + (\omega\ \mathrm{d}\theta)^2 + (\omega \sin\theta\ \mathrm{d}\phi)^2\right], \tag{14}$$

which is the Robertson-Walker line element presented in Equation 3 (note the slight change in variable representation).

1.4 A historical interlude

1912: Spectroscopic observations lead Vesto Slipher to obtain an apparent velocity of +300 kms^{-1} for M31. By 1924 he had painstakingly obtained apparent velocities for 41 objects. Of these, 31 were negative, i.e. receding.

1920: The Shapley–Curtis debate. **Shapley argued** correctly that the sky distribution of globular clusters indicated that the Sun was located far from the centre of the Galaxy. However, he incorrectly believed (using an erroneous result from van Maanen) that the spiral nebulae displayed observable rotational proper motions and were thus located within the (albeit very large) Galaxy. **Curtis argued** that the spiral nebulae were extra-galactic, mainly by assuming that novae observed in the nebulae were similar to novae observed in our galaxy and then following distance dimming arguments. However, he mistrusted Shapley's Cepheid-based distance to the globular clusters and favoured a smaller, Kapteyn Galaxy (based upon star counts) with the Sun close to the centre.

1925: Using Cepheid variable stars Hubble demonstrates that M31 and M33 lie at 285 kpc from the Galaxy and that NGC 6822 lies at > 214 kpc (note that these distance estimates were much too low due to a problem with Cepheid brightness calibration). All appear to be extragalactic, isolated systems.

1929: Using Cepheids to calibrate a secondary distance indicator (the brightest star in a galaxy), Hubble extended his distance–recession velocity sample and reported the relation $v = H\ d$ where $H \approx 500$ kms^{-1} Mpc^{-1} (the value for H is so high compared to the modern value due to the aforementioned problem with Cepheid calibration – if d is biased low, H will be biased high).

1.5 An examination of Hubble's law using the RW line element

Hubble found that $v = H\ d$. Milne examined this apparent expansion in terms of a universe consisting of a mesh of interlocking triangles. Consider three neighbouring galaxies located at the vertices of a particular triangle (Figure 3). This particular universe is homogeneous and is isotropic on large scales. If this universe is expanding uniformly then at some later time these galaxies will have expanded to form a new triangle similar to the first but with all distances scaled by some factor,

$$l(t) = a(t)\ l_0. \tag{15}$$

Would an observer studying the recession velocities of galaxies within the homogeneous, isotropic and uniformly expanding universe discover Hubble's law? Differentiating $l(t)$ one obtains

$$\frac{\mathrm{d}l}{\mathrm{d}t} = \frac{\mathrm{d}a(t)}{\mathrm{d}t}\ l_0 = \frac{\mathrm{d}a(t)}{\mathrm{d}t}\ \frac{l(t)}{a(t)}$$

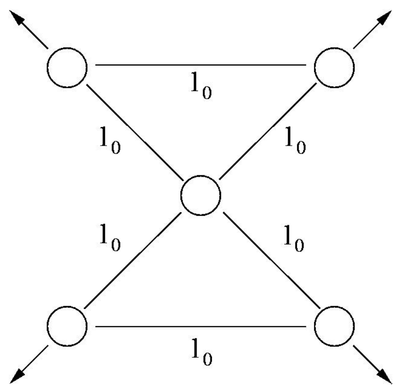

Figure 3: Milne's model of uniform expansion.

$$\begin{aligned} v &= \frac{\mathrm{d}l}{\mathrm{d}t} = \frac{\dot{a}}{a}\, l(t) \\ H(t) &= \frac{\dot{a}}{a} \end{aligned} \tag{16}$$

A similar result is obtained working straight from the RW line element: one notes that $kr^2 = a^2r^2/R^2$ and assumes that the quantity $ar/R \ll 1$. The radial proper distance is related to the metric distance by $\mathrm{d}l(t) = \sqrt{-(\mathrm{d}s)^2} \approx a(t)\ \mathrm{d}r$. The velocity of the object is then

$$v = \frac{\mathrm{d}l}{\mathrm{d}t} \approx \dot{a}\ r = \frac{\dot{a}}{a}\ l \Rightarrow H(t) = \frac{\dot{a}}{a}. \tag{17}$$

The assumption of $ar/R \ll 1$ simply tells us that Hubble's law is only valid on scales much less than the global curvature constant. The conclusion from Milne's analysis and the analysis of the RW line element is that the apparent recession of galaxies with us located at the origin is naturally explained as a **universal** expansion of all test particles (galaxies) within a homogeneous and isotropic universe. The geometric model of the universe receives its first 'tick'.

1.6 The Friedmann equation

The Friedmann equation describes the time variation of the scale factor $a(t)$[1]. It can be derived by substituting the metric terms described by the RW line element into the EFQ. However, we will investigate a version of the Friedmann equation using a Newtonian analogue and Birkhoff's theorem. Consider a universe with coordinate distances defined by the RW line element. The universe is populated with galaxies with a space density ρ and pressure $P = 0$ (i.e. the galaxies do not interact with each other).

[1]To simplify the discussion of the scale factor and its history we (and others) set $a(t_0) = 1$ with $a(t)$ taking a value relative to the current epoch.

Consider a spherical volume of the universe of radius l and mass M. We further consider the dynamical behaviour of a test particle (a single galaxy if you like) of mass m located on the surface of this spherical shell. Birkhoff's theorem states that the mass within the sphere will act upon the test particle as if the entire mass M were concentrated at the centre of the sphere. From the Newtonian equation of motion of the test particle we therefore obtain

$$m\,\frac{\mathrm{d}^2 l}{\mathrm{d}t^2} = -\frac{GMm}{l^2}. \tag{18}$$

Multiplying the equation by $\dot{l}$ generates

$$\frac{\mathrm{d}}{\mathrm{d}t}\,\frac{\dot{l}^2}{2} = \frac{\mathrm{d}}{\mathrm{d}t}\,\frac{GM}{l}. \tag{19}$$

Integrating yields

$$\frac{\dot{l}^2}{2} - \frac{GM}{l} = E, \tag{20}$$

where the integration constant has units of energy. Note: this equation may be interpreted as a basic energy relation of the form, *Kinetic* + *Potential* = *Total*. The RW line element indicates that $l(t) = l_0\ a(t)$, where l_0 is independent of time and $a(t)$ is the time–varying, universal scale factor. Therefore, we may write

$$\begin{aligned}
\frac{\dot{l}^2}{2} - \left(\frac{G}{l}\right)\left(\frac{4\pi l^3 \rho(t)}{3}\right) &= E \\
\frac{\dot{a}^2}{2} - \frac{4\pi G \rho(t) a^2(t)}{3} &= \frac{E}{l_0^2}, \\
\dot{a}^2 - \frac{8\pi G}{3}\ \rho(t)\ a^2(t) &= \frac{2E}{l_0^2}.
\end{aligned} \tag{21}$$

The result is one form of Friedmann's equation and it has three general solutions:

E < 0 : The potential term is proportional to $1/a(t)$ and always dominates (see Equation 20). Hence $a(t)$ cannot increase without limit, instead it must reach some maximum value (at which point $\ddot{a} < 0$) and decrease.

E $= 0$: $a(t)$ increases throughout time, tending toward (but never reaching) an asymptotic maximum scale as $t \to \infty$. The critical density value corresponding to this case is

$$\rho_c(t) = \frac{3}{8\pi G}\ \left(\frac{\dot{a}}{a}\right)^2 = \frac{3H^2(t)}{8\pi G}. \tag{22}$$

The current value of the Hubble parameter is $H_0 = H(t_0) = 70 \pm 7\ \mathrm{kms^{-1}\ Mpc^{-1}}$ and the corresponding value of the critical density is $\rho_{c,0} = (9.2 \pm 1.8) \times 10^{-27}\ \mathrm{kg\,m^{-3}} = (1.4 \pm$

0.3) $\times 10^{11}$ $M_\odot$ Mpc^{-3}. Current estimates of the space density of galaxies are of order 1.4×10^{-2} Mpc^{-3}. If one assumes that the mass of a typical galaxy is $10^{11} M_\odot$, then one concludes that visible galaxies contribute about 1% of the value of $\rho_{c,0}$.

$\mathbf{E} > 0$: $a(t)$ increases monotonically for all time. The universe expands forever.

These three cases are shown in Figure 4. Returning to the form of Friedmann's equation, one may

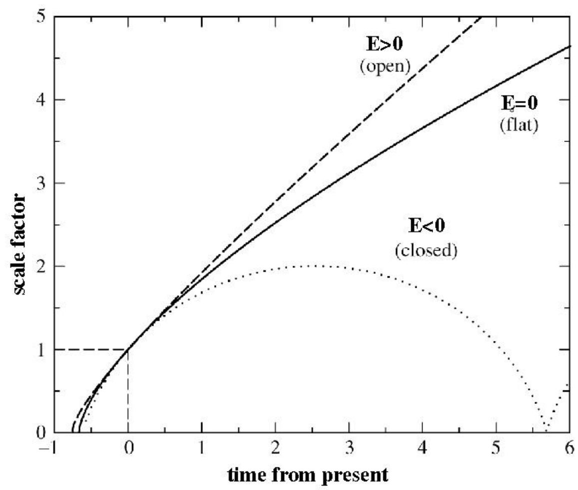

Figure 4: The evolution of the scale factor in Friedmann universes.

redefine the constant of integration to be

$$\frac{2E}{l_0^2} = -kc^2. \tag{23}$$

Re-writing the Friedmann equation yields

$$\left(\frac{\dot{a}}{a}\right)^2 - \frac{8\pi G}{3}\rho(t) = -\frac{kc^2}{a^2}. \tag{24}$$

Note that this form of the equation was that obtained by Friedmann (and re–discovered by Lemaitre) by inserting the corresponding elements of the metric tensor applicable to the RW line element into the EFQ. The parameter k takes one of three values

$$\begin{aligned} E < 0 &\Rightarrow k = +1 \\ E = 0 &\Rightarrow k = 0 \\ E > 0 &\Rightarrow k = -1. \end{aligned} \tag{25}$$

We have now obtained the complete link between the time dependence of the universal scale factor $a(t)$ and the curvature constant k first encountered in the RW line element. Proceeding further, one may redefine the Friedmann equation in terms of a new dimensionless density parameter

$$\Omega(t) = \frac{\rho(t)}{\rho_c(t)}. \tag{26}$$

If we divide the Friedmann equation (Equation 24) by $H(t)^2$ and use the above identity, we obtain

$$\begin{aligned} 1 - \Omega(t) &= -\frac{kc^2}{H^2a^2}, \quad \text{or} \\ \frac{kc^2}{H^2a^2} &= \Omega(t) - 1. \end{aligned} \tag{27}$$

This expression has the following consequences

$$\begin{aligned} k = +1 &\Rightarrow \Omega(t) > 1 \\ k = 0 &\Rightarrow \Omega(t) = 1 \quad \text{for all times,} \\ k = -1 &\Rightarrow \Omega(t) < 1 \end{aligned} \tag{28}$$

where

$$\begin{aligned} \Omega(t) > 1 &\Rightarrow \quad \text{a closed universe,} \\ \Omega(t) = 1 &\Rightarrow \quad \text{a spatially flat universe,} \\ \Omega(t) < 1 &\Rightarrow \quad \text{an open or hyperbolic universe.} \end{aligned} \tag{29}$$

Therefore, the total matter/energy content of the universe determines the overall spatial geometry, the time variation of $a(t)$ and the ultimate fate of the universe[2] Importantly, these equations indicate that the universe has a well–defined characteristic matter density that marks the limit of each of the above cases. Therefore, the determination of the total matter content of the universe became an immediate challenge for early observational cosmologists. However, to determine this critical density, one requires $H(t)$ or H_0 – the present day value of the Hubble parameter.

1.6.1 Alternative forms of the Friedmann equation

The form of the Friedmann equation given in Equation 24 may be modified to take account of the fact that the universe contains more than matter in the form of stars, gas and dust (which we might label ρ_m). The universe contains relativistic particles (photons and neutrinos) that contribute an

[2] Given the time variation of $\Omega(t)$, one question that vexed cosmologists in the past was, "If Ω_0 is not equal to 1, why should it be anywhere close to 1 at the present epoch?". If $\Omega \neq 1$, why do we appear to be living in a special epoch?

energy density u_{rel} to the universal mass-energy budget. The effective matter density of relativistic particles is $\rho_{rel} = u_{rel}/c^2$. In addition, observations of SNe Ia indicate that the expansion of the universe is accelerating. One explanation of this effect may be the presence of a "Cosmological Constant" referred to as Λ. Although these considerations will be described in later chapters, we note here the modified form of the Friedmann equation required to incorporate these effects into our model of the universe:

$$\left(\frac{\dot{a}}{a}\right)^2 - \frac{8\pi G}{3}[\rho_m(t) + \rho_{rel}(t)] - \frac{1}{3}\Lambda c^2 = -\frac{kc^2}{a^2}. \tag{30}$$

1.7 The Fluid and Acceleration equations

At present we have one equation – the Friedmann equation – and two unknowns, $a(t)$ and $\rho(t)$. The Friedmann equation can be thought of as an expression of Newtonian energy conservation for a test mass in an expanding/contracting universe. Another approach is to consider a thermodynamic expression of energy conservation. From the first law of thermodynamics we have

$$dQ = dE + PdV, \tag{31}$$

where dQ is the heat flow into or out of a region, dE is the change in internal energy, P is the pressure and dV is the change in volume of the region. In a homogeneous universe we must have $dQ = 0$, i.e. no bulk flow of heat. Processes for which $dQ = 0$ are known as *adiabatic*. Therefore, applying the first law of thermodynamics to an expanding universe, we may write

$$\dot{E} + P\dot{V} = 0. \tag{32}$$

We return to our previous example of an expanding spherical region of the universe where the physical radius $l(t) = l_0\ a(t)$. The volume of this region is

$$V(t) = \frac{4\pi}{3} l_0^3 a^3(t), \tag{33}$$

and the rate of change of volume is

$$\dot{V} = \frac{4\pi}{3} l_0^3 (3a^2\dot{a}) = V\left(3\frac{\dot{a}}{a}\right). \tag{34}$$

The internal energy of the sphere is

$$E(t) = V(t)\rho(t)c^2, \tag{35}$$

and the rate of change of the internal energy of the sphere is

$$\dot{E} = V\dot{\rho}c^2 + \dot{V}\rho c^2 = V\left[\dot{\rho}c^2 + 3\rho c^2\left(\frac{\dot{a}}{a}\right)\right]. \tag{36}$$

Therefore, combining equations 32, 34 and 36 we can re-write the first law of thermodynamics in the expanding region of the universe as

$$V\left(\dot{\rho}c^2 + 3\left(\frac{\dot{a}}{a}\right)\rho c^2 + 3\left(\frac{\dot{a}}{a}\right)P\right) = 0, \text{ or}$$

$$\frac{d\rho}{dt} + 3\frac{\dot{a}}{a}(\rho + P/c^2) = 0. \tag{37}$$

This is the **Fluid equation** and is the second key equation describing the expansion of the universe. Combining the Friedmann and Fluid equations generates the **Acceleration equation** providing the rate of change of universal expansion. Multiplying the Friedmann equation by a^2 generates

$$\dot{a}^2 = \frac{8\pi G}{3}\rho\, a^2 - kc^2. \tag{38}$$

Taking the time derivative of this equation we have

$$2\dot{a}\ddot{a} = \frac{8\pi G}{3}(\dot{\rho}a^2 + 2\rho\dot{a}a). \tag{39}$$

Dividing by $2\dot{a}a$ gives

$$\frac{\ddot{a}}{a} = \frac{4\pi G}{3}\left(\frac{d\rho}{dt}\frac{a}{\dot{a}} + 2\rho\right). \tag{40}$$

Substituting the Fluid equation into this term yields

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3P}{c^2}\right). \tag{41}$$

We see that a universe consisting of material with a positive energy density, i.e. $\rho c^2 > 0$, has the effect of slowing down the expansion with time. Universal components such as baryons and photons each exert positive pressure (the result of either random kinetic motion or radiation pressure). The effect of such pressure terms is also to slow down expansion. However, should a universal component contribute a negative pressure such than $P < -\rho c^2/3$, the net effect will be to increase the rate of expansion with time. Dark energy (see Lecture 8) is an example of such a universal component.

1.8 The equation of state

At present we have two independent equations (Friedmann and Fluid) and one dependent equation (Acceleration) to describe the evolution of the universe. However, we now have three unknowns: $a(t)$, $\rho(t)$ and $P(t)$. The **equation of state** describing a particular component of the universe describes the relationship between pressure and density, i.e.

$$P = P(\rho) = w\rho c^2, \tag{42}$$

for dilute cosmological "gases", i.e. a gas of matter or photon particles. Consider a low density gas of non-relativistic massive particles – such as galaxies. Non-relativistic means that the random motions of the particles are small compared to the speed of light. Such a gas obeys the perfect gas law

$$P = \frac{\rho}{\mu} kT, \tag{43}$$

where μ is the mean mass of the particles. For a non-relativistic gas, the temperature T is related to the root mean square thermal velocity $\langle v^2 \rangle$ by

$$3kT = \mu \langle v^2 \rangle. \tag{44}$$

Therefore, comparing Equation 42 to Equation 43 we see that

$$P = w\rho c^2 = \frac{\langle v^2 \rangle}{3} \rho \Rightarrow w = \frac{\langle v^2 \rangle}{3c^2} \ll 1. \tag{45}$$

Therefore, a gas of non-relativistic massive particles effectively exerts zero pressure.

1.9 H_0 and the age of the universe

In an expanding universe with low deceleration/acceleration, the quantity H^{-1} approximately defines the time taken for the distance between any two galaxies to double. A more definite relationship between the value of the scale factor and time may be obtained from the Friedmann and Fluid equation. **Note:** in the following discussion, $H_0 = H(t_0)$, where t_0 indicates the current epoch.

The Einstein–de Sitter universe (EdS): In 1932 Einstein and de Sitter postulated that, in the absence of secure observations to the contrary, the simplest assumptions governing the behaviour of the Friedmann equation should be adopted, i.e. $\Lambda = k = P = 0$. Note that these assumptions refer specifically to a universe containing a non-interacting gas of galaxies. At the time when the EdS universe was suggested a radiation dominated universe had not been considered. EdS therefore describes a spatially flat universe containing only matter, i.e. $\Omega_M = 1$. Under these assumptions, the Fluid equation states that $\rho \propto a^{-3}$, i.e.

$$\begin{aligned}
\frac{\mathrm{d}\rho}{\mathrm{d}t} + 3\,(\rho + P/c^2)\,\frac{\dot{a}}{a} &= 0 \\
\frac{\mathrm{d}\rho}{\mathrm{d}t} &= -3\,\rho\,\frac{\dot{a}}{a} \\
\frac{\mathrm{d}\rho}{\rho} &= -\frac{3}{a}\frac{\mathrm{d}a}{\mathrm{d}t}\,\mathrm{d}t \\
\ln\rho &= -3\ln a + C
\end{aligned}$$

$$\begin{aligned} \rho &\propto a^{-3} \text{ or} \\ \frac{\rho}{\rho_0} &= \left(\frac{a_0}{a}\right)^3. \end{aligned} \tag{46}$$

Note that this equation is really only telling us that the matter component of the universe is conserved. One may then re-write the Friedmann equation in the form

$$\begin{aligned} a\,\dot{a}^2 &= \frac{8}{3}\pi G\rho\, a^3 \\ a\,\dot{a}^2 &= \text{constant} \\ a^{1/2}\,\mathrm{d}a &= A\,\mathrm{d}t \\ \frac{2}{3}\,a^{3/2} &= A\,t + C \quad (C = 0 \text{ as a} \to 0 \text{ as t} \to 0) \\ \text{or}\quad a &\propto t^{2/3}. \end{aligned} \tag{47}$$

The assumption that the constant of integration is zero seems innocuous. However, it gave Lemaitre considerable trouble. With early estimates of the Hubble parameter biased high, this in turn implied a troublingly low value for the age of the universe when compared to the estimated ages of the stars. Lemaitre attempted to reconcile these two opposing age estimates by setting $C > 0$ and thus allowing the universe to exist in a static state for some arbitrary period of time prior to expanding. With hindsight this appears to be (and indeed is) a fudge – but one which at the time was forced upon Lemaitre by observations.

Note that $H = \dot{a}/a$, therefore

$$H(t) = \dot{a}/a = (2/3t^{-1/3})/t^{2/3}, \tag{48}$$

which can be rearranged as

$$t_0 = \frac{2}{3}\frac{1}{H_0}. \tag{49}$$

A further case considers an open universe ($\rho \ll \rho_c$), $\Lambda = 0$ scenario, i.e.

$$\begin{aligned} \left(\frac{\dot{a}}{a}\right)^2 &= 0 - \frac{kc^2}{a^2} \\ \dot{a}^2 &= A \end{aligned} \tag{50}$$

which leads to $a = \sqrt{A}\ t$ or, in terms of H one may write

$$\begin{aligned} H &= \dot{a}/a = 1/t \\ t_0 &\simeq 1/H_0. \end{aligned} \quad (51)$$

This is the **Hubble time** – the maximum time elapsed since $a = 0$ for a $\Lambda = 0$ universe. Note that the actual age of the universe in this case is a little bit less than $1/H_0$. The degree of inaccuracy in using a tangent to the current expansion curve will lie in the extent of the deviation of ρ from zero. Therefore, consideration of the Friedmann and Fluid equations for several basic models generates a characteristic age of the universe in terms of H_0.

1.10 Orientation and additional reading

1. Read Ryden Chapter 6 for a discussion of how $\Lambda \neq 0$ affects the analysis presented in this lecture. General, $\Lambda \neq 0$ models will be discussed in more detail later in the course.

2. The evolution of observational tests of the pre-recombination universe will be discussed in Lectures 3 (CMB) and 4 (BBN).

3. The fact that $a(t)$ is increasing with time implies that it was smaller in the past. We therefore expect the pressure term (radiation from CMB and neutrinos together with random matter velocities) to have been more important in the past – see Lectures 3 and 4.

1.11 Asides

1.11.1 The nature of redshift within the cosmological model

Redshift was originally defined in purely observational terms. However, cosmological redshift arises naturally from a consideration of the RW line element. Consider the light path of a photon travelling through the universe along a radial path (radial null geodesic), i.e. $\mathrm{d}s^2 = \mathrm{d}\theta = \mathrm{d}\phi = 0$,

$$c^2 \ \mathrm{d}t^2 = a^2(t)\left[\frac{\mathrm{d}r^2}{1-kr^2}\right]. \tag{52}$$

For a light pulse emitted at t_e and observed at t_o by an observer at a distance r_e, one may write

$$\int_{t_e}^{t_o} \frac{c \ \mathrm{d}t}{a(t)} = \int_0^{r_e} \frac{\mathrm{d}r}{\sqrt{1-kr^2}} \equiv f(r_e). \tag{53}$$

Note, that $f(r_e)$ is a fixed (or comoving) distance. If one has a good enough understanding of $a(t)$ and k one can determine the relation between distance and time in the universe. However, one may consider a second pulse of light emitted and observed at $t_e + \Delta t_e$ and $t_o + \Delta t_o$ respectively, i.e.

$$\begin{aligned}
\int_{t_e+\Delta t_e}^{t_o+\Delta t_o} \frac{c \ \mathrm{d}t}{a(t)} - \int_{t_e}^{t_o} \frac{c \ \mathrm{d}t}{a(t)} &= 0 \\
\frac{c \ \Delta t_o}{a(t_o)} - \frac{c \ \Delta t_e}{a(t_e)} &= 0 \\
\frac{c \ \Delta t_o}{c \ \Delta t_e} &= \frac{a(t_o)}{a(t_e)} \qquad (54)
\end{aligned}$$

However, $c \ \Delta t = \lambda$, and the above analysis is related to the (observationally defined) redshift by

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{\lambda_o}{\lambda_e} - 1 = \frac{a(t_o)}{a(t_e)} - 1. \tag{55}$$

This expression is normally written as $1 + z = a_0/a_e$. Cosmological redshift describes the relative expansion of the universal scale factor between the epochs of emission and observation[3]. **Note:** the observed redshift of an astrophysical source is often a combination of cosmological redshift and other physical effects, e.g. gravitational redshift, the Sachs–Wolfe effect and peculiar velocities. Peculiar velocities take alter the redshift via $cz_{obs} = cz_{cosmo} + v_{pec}(1 + z_{cosmo})$ and will discussed further in Lecture 2.

[3]Though redshift may also be defined via SR as arising from the recession velocity associated with the Hubble flow, SR is only valid in locally flat reference frames and describes the case where (effectively) $g_{\mu\nu}$ = constant.

1.11.2 Co–moving coordinates

Co–moving coordinates expand with universe. Another way of saying this is that they are fixed with respect to the **Hubble flow**. Co–moving coordinates are used to compare physical quantities derived from the metric (distance, volume, etc.) at any given epoch, to the present day (considered to be the reference epoch). For example, employing the RW line element one may express the physical distance between two galaxies (say) as $s = a(t)\ d$ where s is the physical distance and d is the co–moving distance. If no peculiar motions exist, d is constant and

$$\begin{aligned} s_0 &= a(t_0)\ d \quad \text{and} \quad s_e = a(t_e)\ d \\ \frac{s_0}{a(t_0)} &= \frac{s_e}{a(t_e)} \\ s_0 &= s_e\ \frac{a(t_0)}{a(t_e)} = s_e\ (1+z). \end{aligned} \tag{56}$$

Therefore, if two galaxies at $z = 1$ are observed to be separated by a physical distance of 1 Mpc, the co–moving separation is 2 Mpc. The same principle is applied to volumes. This definition of co–moving coordinates designed to provide a common epoch for comparing observations. The current definition of co–moving coordinates is much simpler than the 'official' relativity definition.

1.11.3 The cosmological horizon

The cosmological horizon may be defined as the maximum distance a photon could travel within the lifetime of the universe. It is a convenient definition of the largest region of the universe that could exist in causal contact at any particular epoch. The horizon is defined by considering a radial null geodesic within the RW line element, i.e.

$$\begin{aligned} c\ \mathrm{d}t &= \frac{a(t)\ \mathrm{d}r}{\sqrt{1-kr^2}} \\ \int_0^{t_0} \frac{c\ \mathrm{d}t}{a(t)} &= \int_0^{r_H} \frac{\mathrm{d}r}{\sqrt{1-kr^2}}, \end{aligned} \tag{57}$$

Where t_0 indicates the current age of the universe and r_H is the horizon distance. For an EdS universe with $k = 0$ we may write

$$\begin{aligned} r_H &= \int_0^{t_0} \frac{c\ \mathrm{d}t}{a(t)} \\ &= \int_0^{t_0} \frac{c\ \mathrm{d}t}{a_0} \frac{t_0^{2/3}}{t^{2/3}} \\ &= 3\ c\ t_0. \end{aligned} \tag{58}$$

Consideration of the cosmological horizon at the current epoch provides one method to answer **Olbers' paradox**, or "why is the night sky dark?". The universe may well be infinite in space, but it is finite in time. The combination of a finite age of the universe with a finite speed of light ensures that we can only observe photons from a finite region of the universe.

However, computation of the cosmological horizon gives rise to the **horizon problem**. When we look at the universe at the cosmological horizon, it is isotropic, i.e. large scale structure and CMB temperatures are statistically identical even though separated by 180° on the sky. However, these two regions, though in casual contact with us, are themselves causally isolated. How could the CMB and LSS have developed in exactly the same manner? The answer to the horizon problem is provided by **cosmic inflation** that postulates that the universe underwent a rapid phase of expansion at early times. Regions in the early universe were originally in causal contact and were subsequently inflated to scales much larger than the horizon during the inflationary epoch.

1.12 Summary of Lecture 1

1. Our mathematical model of the universe is based upon GR which postulates that space exists as a four dimensional spacetime defined by a global curvature constant.

2. Universal spacetime is assumed to be isotropic and homogeneous. The most general line element describing an isotropic and homogeneous spacetime is the RW line element (also referred to as a metric). The line element describes the infinitesimal distance between neighbouring coordinates.

3. The universal scale factor $a(t)$ of the RW line element evolves with time. The relationship between scale factor and time depends upon the universal mass density, pressure and presence of a Λ–term. The general form of the Friedmann equation indicates that the universe may be expanding or contracting. We shall see during our investigation of Λ that deriving a static universe is difficult.

4. The Hubble relation indicates that the universe is currently expanding. For the simple case of the EdS universe, the overall geometry (curvature) of the universe, and its eventual fate, are determined by the total matter density and the current value of the Hubble parameter.

5. Therefore, coordinated observational tests (total matter, oldest age, current expansion rate) may constrain parameter values that define the cosmological model.

   Note that such observations do not tell us whether the cosmological model as we have designed it is correct (though it is clearly a good approximation to reality), simply **which** cosmological model values **our** universe resembles.