Chapter 14

Relativistic cosmology II: the Robertson–Walker geometry.

14.1 The Robertson–Walker metrics as models of the Universe.

The Robertson–Walker (R–W) metrics were derived in Sec. 9.8:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{R^2(t)}{\left(1 + \frac{1}{4}kr^2\right)^2}\ \left[\mathrm{d}r^2 + r^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right)\right], \tag{14.1}$$

where $R(t)$ is a function (called **scale factor**) to be determined from the Einstein equations. The arbitrary constant k, called **curvature index**, when nonzero, can be scaled to $+1$ or -1 by the transformations $r = r'/\sqrt{|k|}$, $R(t) = \sqrt{|k|}\widetilde{R}(t)$. After this, the limiting transition $k \to 0$ becomes impossible. This can cause difficulties when comparing the three classes with each other [112, 113], see Sec. 14.6.

Other representations of the R–W metric are also in use. Let

$$r = \frac{2r'}{1 + \sqrt{1 - kr'^2}} \qquad \Longleftrightarrow \qquad r' = \frac{r}{1 + \frac{1}{4}kr^2}. \tag{14.2}$$

This changes (14.1) to

$$\mathrm{d}s^2 = \mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r'^2}{1 - kr'^2} + r'^2\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2\right)\right]. \tag{14.3}$$

If $k \neq 0$, then this can be transformed still further. With $k > 0$, the transformation:

$$r' = \frac{1}{\sqrt{k}}\ \sin\psi \tag{14.4}$$

[112] A. Krasiński and C. Hellaby, *Phys. Rev.* **D65**, 023501 (2002).

[113] A. Krasiński and C. Hellaby, in: Topics in mathematical physics, general relativity and cosmology. Proceedings of 2002 International Conference in Honor of Jerzy Plebański. Edited by H. Garcia-Compean, B. Mielnik, M. Montesinos and M. Przanowski. World Scientific, Singapore 2006, p. 279.

changes (14.3) to

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \widetilde{R}^2(t)\left[\mathrm{d}\psi^2 + \sin^2\psi\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2\right)\right], \tag{14.5}$$

and now the hypersurfaces $t =$ constant are clearly seen to be 3-dimensional spheres.[37]

With $k < 0$, the transformation

$$r = \frac{1}{\sqrt{-k}}\ \sinh\psi \tag{14.6}$$

changes (14.3) to

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \widetilde{R}^2(t)\left[\mathrm{d}\psi^2 + \sinh^2\psi\left(\mathrm{d}\vartheta^2 + \sin^2\vartheta\mathrm{d}\varphi^2\right)\right]. \tag{14.7}$$

A space $t =$ constant has constant curvature: its Riemann tensor, in the coordinates of (14.1) or (14.3), is

$$R^{IJ}{}_{KL} = \frac{k}{R^2(t)}\ \delta^{IJ}_{KL}, \qquad I, J, K, L = 1, 2, 3.$$

When $k = 0$, the hypersurfaces $t =$ constant in (14.1) and (14.3) are flat. The R–W spacetimes with $k > 0$ are colloquially called the "closed Universe", those with $k < 0$ are called the "open Universe", and the one with $k = 0$ is called the "flat Universe".

The first author to investigate these spacetimes was Aleksandr Aleksandrovich Friedmann[38] in 1922 (the metric (14.5)) and in 1924 (the case $k = -1$; in coordinates different from all those used here) [38].[39] (The case $k = 0$ was first introduced by Robertson in 1929 [39].) He solved the Einstein equations for these metrics, with dust source and with cosmological constant allowed, and treated his results as merely a mathematical curiosity. Both papers were quickly forgotten – astronomers and physicists had not yet been ready to accept their implications, while Friedmann was not a sufficiently big figure to win attention. Soon after, in 1925, he died prematurely and had no chance to claim credit for his discovery when the expansion of the Universe was accepted as a fact in 1929 [54].[40]

The case $k > 0$ was discussed by Georges Lemaître in 1927 [108]. He generalised Friedmann's solution to nonzero pressure, and was aware that it reflects some properties of the real Universe. Unfortunately, he published his papers in a local Belgian journal, in French, and, being an unknown person at that time, was also ignored. In 1930, when the expansion of the Universe was already generally known, Arthur Eddington asked whether exact models of expanding Universe might be derived from the relativity theory [115]. It

[37] Note that when $k > 0$ the transformation (14.4) maps the range $0 \le r < \infty$ onto the range $0 \le r' < 1/\sqrt{k}$, which is covered twice: as r goes from 0 to ∞, r' first increases from 0 to $1/\sqrt{k}$, achieved at $r = 2/\sqrt{k}$, and then decreases to 0 as $r \to \infty$. Consequently, the coordinates of (14.3) cover only one half of the 3-sphere. The coordinates of (14.2) and (14.5) cover the whole sphere.

[38] The correct transcription of his name from Russian is "Fridman", but following the original publications in German, the form "Friedmann" became the accepted standard.

[39] In the second paper, some conclusions about the evolution were incorrect [56].

[40] As already mentioned, Hubble did not believe that the Universe is actually expanding. He insisted that representing redshifts by velocities of recession is just a convenient mathematical trick [54, 55, 56, 114].

[114] E. P. Hubble, *The Realm of the Nebulae.* New Haven: Yale University Press, 1982, p. 207 (Republication; First published 1936).

[115] Royal Astronomical Society Discussion, *Observatory* **53**, 39 (1930).

was only then that Lemaître brought his paper from 1927 to Eddington's attention; and the English translation [116] followed.[41] Friedmann's role was recognised later [56].

The metric (14.1) was derived from symmetry assumptions by Robertson in 1933 [40] and by Walker in 1935 [41]. Friedmann and Lemaître were first to use it to derive explicit solutions of Einstein's equations. In view of this history, the spacetimes with this geometry are often called the Friedmann–Lemaître–Robertson–Walker (FLRW) models, but various subsets of this collection of four names are also in use. It seems appropriate to use the name Robertson–Walker when referring to the general metric (14.1), and the names Friedmann or Lemaître when referring to their explicit solutions.

In each of the coordinate representations, (14.1), (14.3), (14.5) and (14.7), the Einstein tensor for the R–W metric is diagonal. Consequently, if the source in the Einstein equations is a perfect fluid, then its velocity field can have only the u^0 component, and, because of $u_\alpha u^\alpha = 1$ and $g_{00} = 1$, it must be $u^\alpha \underset{*}{=} \delta^\alpha{}_0$. This, in turn, means that all those coordinate systems are comoving: each matter particle has fixed spatial coordinates, and its proper time is the time coordinate in spacetime.

14.2 Optical observations in an R–W Universe

14.2.1 The redshift

In order to use (13.23), we have to know the vector field k^α. Since every R–W spacetime is spatially homogeneous, all points within the same space t = constant are equivalent, so the result of any calculation will be independent of the spatial position of the observer. Let us then assume that the observer is at $r = 0$. A null geodesic sent off with $\dot{\vartheta}_0 = \dot{\varphi}_0 = 0$ will preserve the radial direction $\dot{\vartheta} = \dot{\varphi} = 0$ all along. Hence, such a geodesic obeys

$$0 = \mathrm{d}t^2 - \frac{R^2(t)}{\left(1 + \frac{1}{4}kr^2\right)^2}\mathrm{d}r^2. \tag{14.8}$$

For a ray proceeding *towards* the observer sitting at $r = 0$ we thus have

$$\int_{t_e}^{t_o} \frac{\mathrm{d}t}{R(t)} = -\int_{r_e}^{r_o} \frac{\mathrm{d}r}{1 + \frac{1}{4}kr^2}. \tag{14.9}$$

It can be verified that the parameter v on the geodesic defined by

$$\frac{\mathrm{d}t}{\mathrm{d}v} = \frac{1}{R(t)} \tag{14.10}$$

[116] G. Lemaître, *Mon. Not. R. Astr. Soc.* **91**, 483 (1931).

[41] The 1931 translation omitted some details of the 1927 paper. One of the omitted pieces showed that Lemaître was aware of the expansion of the Universe before Hubble announced his law of redshifts. This gave rise to a public discussion in 2011 on who was the translator of the original paper and who (and why) edited out the pieces in question. Suspicions were cast on Eddington and on Hubble. A scrutiny of documents by Mario Livio

[117] M. Livio, Mystery of the missing text solved. *Nature* **479**, 171 (2011)

showed that "it is now certain that Lemaître himself translated his article, and that he chose to delete several paragraphs and notes without any external pressure." He omitted the few pieces because he found them to be no longer sufficiently important in 1931. This story is described by J. P. Luminet [108].

is affine. In this parametrisation, the tangent vector to the geodesic in (13.23) is

$$k^\alpha = \left[\frac{1}{R}, -\frac{1}{R^2}\left(1 + \frac{1}{4}kr^2\right), 0, 0\right]. \tag{14.11}$$

The velocity field is $u^\alpha = \delta^\alpha{}_0$ everywhere, so in (13.23) we have

$$k^\alpha u_\alpha = 1/R, \tag{14.12}$$
$$1 + z = R(t_o)/R(t_e). \tag{14.13}$$

When $t_o/t_e - 1$ is small, we have

$$z \approx \left[\dot{R}(t_e)/R(t_e)\right](t_o - t_e) \approx \frac{\dot{R}}{R}\,\delta\ell, \tag{14.14}$$

since in the coordinates of (14.8) $t_o - t_e = \delta\ell$ is the distance from the light-source to the observer; $\dot{R} \overset{\text{def}}{=} \mathrm{d}R/\mathrm{d}t$. The same result follows from (13.31) when $\omega_{\alpha\beta} = \sigma_{\alpha\beta} = 0 = \dot{u}^\alpha$ and $\theta = 3\dot{R}/R$. This is the Hubble law, with the Hubble parameter $H = (c/3)\theta = c\dot{R}/R$ (the factor c appeared because the physical time τ is t/c).

The Hubble law in fact holds exactly in the R–W models. To see it, note that

$$\ell(t) = R(t)\int_0^{r_e} \frac{\mathrm{d}r}{1 + \frac{1}{4}kr^2}, \tag{14.15}$$

is the distance of the light-source at $r = r_e$ from the observer at $r = 0$, calculated within the hypersurface of constant t. In comoving coordinates r_e = constant, and then

$$\frac{\mathrm{d}\ell}{\mathrm{d}t} = \frac{\dot{R}}{R}\,\ell = H\ell/c. \tag{14.16}$$

14.2.2 The redshift–distance relation

Assume that, this time, the origin is at the source G in Fig. 13.2. The element of surface area in the surface $\{t = t_o = \text{constant}, r = \overline{r} = \text{constant}\}$ is, in the coordinates of (14.1),

$$\delta S_o = \left[\frac{R(t_o)\,\overline{r}}{1 + \frac{1}{4}k\overline{r}^2}\right]^2 \sin\vartheta \mathrm{d}\vartheta \mathrm{d}\varphi.$$

On the other hand, close to the origin $r = 0$, the metric of the space t = constant is approximately flat, so in that space $\delta S = r^2\delta\Omega$, where $\delta\Omega = \sin\vartheta\mathrm{d}\vartheta\mathrm{d}\varphi$ is the element of solid angle. Comparing these quantities with (13.55) we see that in the R–W models

$$r_G = \frac{R_o\overline{r}}{1 + \frac{1}{4}k\overline{r}^2}, \tag{14.17}$$

where $\overline{r}$ is the r-coordinate of the observer and R_o is the value of R at the time when the ray passes the observer. By the reciprocity theorem (13.58), we then have

$$r_o = \frac{R_o\overline{r}}{(1 + z_o)\left(1 + \frac{1}{4}k\overline{r}^2\right)}. \tag{14.18}$$

This provides a relation between $\overline{r}$ and the observable quantity r_o, and (14.13) then provides a relation between R and the observable quantity z. The integral of (14.8) along a radial *outgoing* null geodesic is

$$\int_{t_e}^{t} \frac{\mathrm{d}t'}{R(t')} = \int_0^r \frac{\mathrm{d}\widetilde{r}}{1+\frac{1}{4}k\widetilde{r}^2}. \tag{14.19}$$

We can calculate this integral when $R(t)$ is given explicitly. We will do it in Section 14.6.

The quantities that are used in astronomy to characterise $R(t)$ are the **Hubble parameter** of (14.16)[42] and the **deceleration parameter** q_0:

$$q_0 \overset{\text{def}}{=} -RR,_{tt}/R,_t{}^2\big|_{\text{now}} \equiv -1 - \frac{c}{{H_0}^2}\frac{\mathrm{d}H_0}{\mathrm{d}t}. \tag{14.20}$$

With $q_0 < 0$ the expansion of the Universe accelerates ($R,_{tt} > 0$), with $q_0 > 0$ it decelerates.

14.3 The Friedmann equations

With a perfect fluid source, in a general geometry, the Einstein equations must determine six components of the metric tensor $g_{\alpha\beta}$ (the remaining four are fixed by a choice of the coordinates), the energy-density, the pressure and three components of the velocity field u^α (the fourth component is determined by $u_\alpha u^\alpha = 1$). This is 11 quantities to be found from 10 equations $G_{\alpha\beta} = \kappa T_{\alpha\beta}$, and the set is underdetermined. To make it determinate, we have to add one more equation, and usually this is the equation of state.[43]

This indeterminacy survives in the R–W metrics, in spite of their high symmetry. They automatically define a perfect fluid energy-momentum tensor, with the energy-density and pressure being functions of $R(t)$, $\mathrm{d}R/\mathrm{d}t$ and $\mathrm{d}^2R/\mathrm{d}t^2$. In order to obtain an explicit solution, we have to add an equation of state relating the pressure to the energy-density. With all functions depending only on t, this must be the barotropic equation $\epsilon = \epsilon(p)$.

What equation of state is most appropriate for describing the Universe as a whole? At present, the mass-density in the Universe is so small ($< 10^{-28}$ g/cm^3) that pressure does not influence the large-scale dynamics of matter and $p = 0$ is an acceptable hypothesis. However, to describe the evolution of the Universe in the early period of high density, pressure must be taken into account. We shall consider here the later periods and, following Friedmann (1922), we assume $p = 0$.

In the orthonormal basis for the metric (14.1) or (14.3),[44] the Einstein equations are

$$G_{00} = \frac{3k}{R^2} + 3\frac{\dot{R}^2}{R^2} = \kappa\epsilon - \Lambda, \tag{14.21}$$

$$G_{11} = G_{22} = G_{33} = -\left(\frac{k}{R^2} + \frac{\dot{R}^2}{R^2} + 2\frac{\ddot{R}}{R}\right) = \kappa p + \Lambda. \tag{14.22}$$

[42] It is most often denoted H_0, to stress that its current value is meant.

[43] In vacuum, because of $G^{\alpha\beta}{}_{;\beta} = 0$, only six of the Einstein equations are independent ([9], p. 160).

[44] Since the G_{ij} in (14.21) – (14.22) are scalars independent of r, they come out the same no matter which of the two coordinate representations is used. However, they would be different for (14.5) and (14.7) because there, the constant k is scaled to $+1$ or -1, respectively.

The equations of motion $T^{\alpha\beta}{}_{;\beta} = 0$ reduce here to the single equation

$$\left(\epsilon R^3\right)^{\cdot} + 3R^2 \dot{R} p = 0. \tag{14.23}$$

With $p = 0$, this becomes the mass conservation equation

$$\rho R^3 = \text{const} \stackrel{\text{def}}{=} 3\mathcal{M}/(4\pi), \tag{14.24}$$

where $\rho = \epsilon/c^2$ is the mass-density. If $\dot{R} \neq 0$, then (14.22) follows from (14.21) and (14.23). If $\dot{R} = 0$, then the solution of (14.21) – (14.22) is the 'Einstein Universe' (10.74). With $p = 0$ and $\dot{R} \neq 0$ only (14.21) must be solved. In consequence of (14.24) it becomes

$$R_{,t}{}^2 = \frac{2G\mathcal{M}}{c^2 R} - k - \frac{1}{3}\Lambda R^2. \tag{14.25}$$

With $\dot{R}/R = H/c$ – the Hubble 'constant', we obtain from (14.24) – (14.25)

$$\rho = \frac{3H^2}{8\pi G} + \frac{c^2}{8\pi G}\left(\frac{3k}{R^2} + \Lambda\right). \tag{14.26}$$

The quantities ρ and H are *in principle* measurable.[45] Also in principle, the value of Λ may be deduced from observations. Hence, (14.26) gives us a possibility to determine the sign of k for the real Universe – *assuming that it has a Friedmann geometry.*

The recently published results of the Planck satellite mission [85] imply $-\Lambda \approx 10^{-46}$ km^{-2}. A sphere of this curvature would have the radius of ≈ 109 Mpc.

14.4 The ΛCDM model

If $k = 0 = \Lambda$, then we obtain from (14.26) the **critical density**

$$\rho_{cr} \stackrel{\text{def}}{=} \frac{3H^2}{8\pi G}. \tag{14.27}$$

This notion played an important role in cosmology in those years (approx. 1929 – 1999) when the majority view was that $\Lambda = 0$. The challenge was then to decide whether the mass-density in the Universe is smaller than ρ_{cr}, larger than ρ_{cr} or equal to ρ_{cr} (which would imply, via (14.26), $k < 0$, $k > 0$ or $k = 0$, respectively). The current obligatory view in astronomy is that the Universe is expanding at an accelerated rate. This is possible only when Λ, or something that imitates its action, is nonzero. (All those "somethings" are collectively called "dark energy".) This is seen from (14.25) differentiated by t:

$$R_{,tt} = -\frac{G\mathcal{M}}{(cR)^2} - \frac{1}{3}\,\Lambda R. \tag{14.28}$$

[45] The ucertain element is measuring the distances. All methods to measure intergalactic distances rely on theoretical assumptions that are difficult to verify, and they can be used only for nearby galaxies. For distant galaxies, the Hubble law is *assumed* to hold exactly, and then used as a measure of distance.

If $\Lambda \geq 0$, then $R_{,tt} < 0$, i.e., the expansion of the Universe is decelerated by gravitational attraction. With $\Lambda < 0$, the accelerated expansion sets in when $R = R_i = \left[-3G\mathcal{M}/\left(c^2\Lambda\right)\right]^{1/3}$. The time of that inflection point can be calculated once we know the solution of (14.25).

The conclusion about accelerated expansion follows only when the observations of supernovae of type Ia are interpreted using the models of the R–W class. With inhomogeneous models, one may account for these observations by inhomogeneities in mass distribution and expansion rate, without introducing Λ – see Chapter 15.

The current conventional wisdom in astronomy says still more: the values of H_0 and Λ are such that they imply $k = 0$ in (14.26), so the Universe is not only expanding with acceleration but remains spatially flat. Such a model is called ΛCDM, the CDM meaning "cold dark matter". The solution of (14.25) with $k = 0$ is

$$R(t) = \left(\frac{6M_0}{(-\Lambda)}\right)^{1/3} \sinh^{2/3}\left[\frac{\sqrt{-3\Lambda}}{2}\left(t - t_{B\Lambda}\right)\right], \tag{14.29}$$

where $t_{B\Lambda}$ is the Big Bang time and $M_0 \overset{\text{def}}{=} G\mathcal{M}/c^2$. The inflection $R_{,tt} = 0$ occurs at

$$t_i - t_{B\Lambda} = \frac{1}{\sqrt{-3\Lambda}} \ln\left(\frac{\sqrt{3}+1}{\sqrt{3}-1}\right). \tag{14.30}$$

Taking $-\Lambda = 10^{-46}\ \mathrm{km}^{-2}$ [85] and remembering that the physical time is t/c we obtain from (14.30) $\left(t_i - t_{B\Lambda}\right)/c = 8 \times 10^9$ years. The current estimate of the age of the Universe is $\approx 13.67 \times 10^9$ years [85]. So, the acceleration began at somewhat earlier than 2/3 of the present age of the Universe.

14.5 The redshift drift – a test for accelerated expansion

The Hubble parameter H depends on the cosmic time coordinate. The first author to consider the possibility of measuring the rate of change of H, called **redshift drift**, was A. Sandage [118]. He estimated that, with the technology of then, $\approx 10^7$ years of monitoring the value of H would be required to detect its changes [119]. By a more recent estimate, a few decades would be sufficient [120]. We will show how the hypothesis of accelerated expansion of the Universe can be confirmed or refuted by detecting just the sign of $\mathrm{d}H/\mathrm{d}t$.

For galaxies, t_e increases when t_o increases. From (13.21) and (13.23), $\mathrm{d}t_o/\mathrm{d}t_e = 1 + z$. Using this, and also $\mathrm{d}R(t_e)/\mathrm{d}t_o = \mathrm{d}R(t_e)/\mathrm{d}t_e \times \mathrm{d}t_e/\mathrm{d}t_o = (\mathrm{d}R(t_e)/\mathrm{d}t_e)/(1 + z)$, we obtain

$$\frac{\mathrm{d}z}{\mathrm{d}t_o} = \frac{1}{R(t_e)}\left[\frac{\mathrm{d}R(t_o)}{\mathrm{d}t_o} - \frac{\mathrm{d}R(t_e)}{\mathrm{d}t_e}\right]. \tag{14.31}$$

[118] A. Sandage, *Astrophys. J.* **136**, 319 (1962).

[119] R. Lazkoz, I. Leanizbarrutia and V. Salzano, *European Phys. J. C*, **78**, 11 (2018). It is funny to note that the authors cite Allan Sandage as S. Allan.

[120] A. Loeb, *Astrophys. J.* **499**, L111 (1998).

So, $\mathrm{d}z/\mathrm{d}t > 0$ when $\mathrm{d}R/\mathrm{d}t$ has increased between t_e and t_o, and $\mathrm{d}z/\mathrm{d}t < 0$ in the opposite case. Thus, if $\mathrm{d}z/\mathrm{d}t_o > 0$ for some observed objects, then the expansion of the Universe had been accelerating for some time, but if $\mathrm{d}z/\mathrm{d}t_o < 0$ for all observed objects, then no accelerated expansion has ever occurred.

14.6 The Friedmann solutions with $\Lambda = 0$.

With $\Lambda = 0$ and $\mathrm{d}R/\mathrm{d}t > 0$, eq. (14.25) can be written as follows:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = \left(\frac{2G\mathcal{M}}{c^2 R} - k\right)^{1/2} \overset{\text{def}}{=} \left(\frac{2\alpha}{R} - k\right)^{1/2}. \tag{14.32}$$

The solutions of (14.32) are best represented in a parametric way. For $k < 0$ we obtain

$$R = -\frac{\alpha}{k}(\cosh\omega - 1), \qquad t - t_B = \frac{\alpha}{(-k)^{3/2}}(\sinh\omega - \omega), \tag{14.33}$$

where ω is the parameter, and t_B is an arbitrary constant. Eqs. (14.33) can be solved for $t = t(R)$, but the $R(t)$ defined by (14.33) is not an elementary function.

For $k = 0$ the solution is

$$R = \left[\frac{9}{2}\alpha\left(t - t_B\right)^2\right]^{1/3}. \tag{14.34}$$

In this case the constant α can be scaled by transformations of r, so its actual value merely defines the unit of distance and has no physical meaning.

Finally, for $k > 0$, the solution of (14.32) is

$$R = \frac{\alpha}{k}(1 - \cos\omega), \qquad t - t_B = \frac{\alpha}{k^{3/2}}(\omega - \sin\omega). \tag{14.35}$$

In all three cases we have taken into account the observed fact that at present $\dot{R} > 0$. As we predicted in Sec. 12.3, each of these solutions has a singularity at $t \to t_B$, where $R \to 0$ and $\rho \to \infty$. At $t = t_B$, all matter of the model would be condensed in one point. The last model has a second singularity, in the future at

$$t = t_{FS} = t_B + \frac{2\pi\alpha}{k^{3/2}} = t_B + \frac{2\pi GM}{k^{3/2}c^2} = t_B + \frac{8\pi^2 G\rho_o {R_o}^3}{3k^{3/2}c^2}. \tag{14.36}$$

The models (14.33) and (14.34) will continue to expand forever. The model (14.35) has a finite lifetime; at $t = t_{FS}$ its existence is terminated. The graphs of $R(t)$ corresponding to different values of k are shown in Fig. 14.1.

Formally, eqs. (14.35) describe a cycloid,[46] on which t runs through an infinite range, while R oscillates between 0 and $R_{\max} = 2\alpha/k = 2GM/\left(kc^2\right)$. However, $\mathrm{d}R/\mathrm{d}t \to \infty$ as $t \to t_{FS}$, and the integration of (14.32) through this point is not possible. This is why the $k > 0$ Friedmann Universe has a finite time of existence equal to $(t_{FS} - t_B)$.

[46] It is strictly a cycloid when $k = 1$. With $k < 1$ the cycloid is flattened in the R-direction, with $k > 1$ it is squeezed in the t-direction.

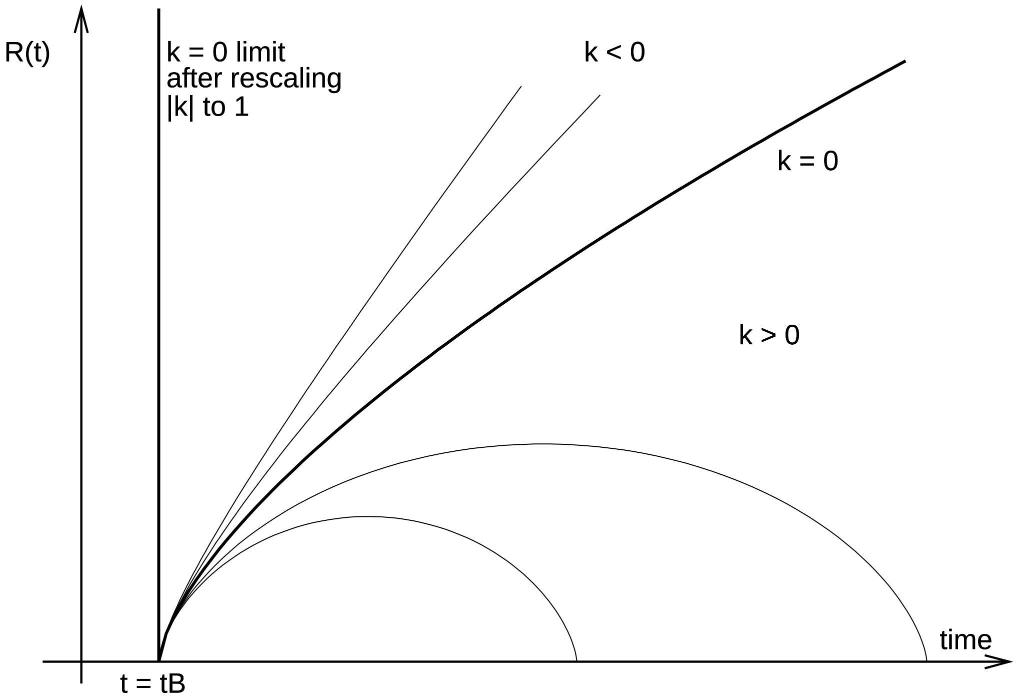

Figure 14.1: The functions $R(t)$ corresponding to Friedmann models with different values of k. When $k \to 0$, the rescaling of R required to achieve $|k| = 1$ maps the $k = 0$ graph into the vertical straight line at $t = t_B$, and no place is left for the $k < 0$ models. Note that if $k > 0$, then the value of k determines the lifetime of the model, so *it has physical meaning*. Similarly, for models with $k < 0$, the value of $-k$ determines the asymptotic value of the expansion velocity, $\lim_{t\to\infty} \dot{R}$, and also has a physical meaning. Thus, rescaling k results in tampering with physical parameters.

14.6.1 The redshift – distance relation in the $\Lambda = 0$ Friedmann models

Let $\mathcal{H}_0 \overset{\text{def}}{=} H_0/c$. From (14.32) $R,_{tt} = -\alpha/R^2$, and from (14.20) $R,_{tt} = -q_0 R,_t{}^2/R \equiv -q_0 R \mathcal{H}_0{}^2$, so $R = \left[\alpha/\left(q_0 \mathcal{H}_0{}^2\right)\right]^{1/3}$. Using this, (14.32) can be rewritten as

$$k = \frac{2\alpha}{R} - (\mathcal{H}_0 R)^2 = (2q_0 - 1)\left(\frac{\alpha \mathcal{H}_0}{q_0}\right)^{2/3}. \tag{14.37}$$

Thus the signs of k and of $(2q_0 - 1)$ are the same, and $q_0 = 1/2$ when $k = 0$. With $k \neq 0$,

$$\alpha = \left(\frac{k}{2q_0 - 1}\right)^{3/2} \frac{q_0}{\mathcal{H}_0}, \qquad R = \frac{1}{\mathcal{H}_0}\sqrt{\frac{k}{2q_0 - 1}}. \tag{14.38}$$

The integrals in (14.19) are different for each k. For $k > 0$ the result is

$$\arctan\left(\frac{\sqrt{k}r}{2}\right) = \frac{1}{2}\left[\omega(t) - \omega\left(t_e\right)\right]. \tag{14.39}$$

But, from (14.35), $\omega(t) = \arccos(1 - kR/\alpha)$ and $\omega(t_e) = \arccos\{1 - kR/[(1+z)\alpha]\}$. Substituting for α and R from (14.38) we obtain from (14.18) the **Mattig formula** [121].

$$r_o = \frac{q_0 z + (q_0 - 1)\left(\sqrt{2q_0 z + 1} - 1\right)}{\mathcal{H}_0 {q_0}^2 (1+z)^2}. \tag{14.40}$$

With $k < 0$, we again obtain (14.40) from (14.19), even though the intermediate formulae are different. When $k = 0$, we substitute $q_0 = 1/2$ in (14.40) and obtain

$$r_o = 2\,\frac{\sqrt{1+z} - 1}{\mathcal{H}_0 (1+z)^{3/2}}. \tag{14.41}$$

The limit of (14.40) at $q_0 \to 0$ is $r_o = [1/(2\mathcal{H}_0)]\,[1 - 1/(1+z)^2]$. This corresponds to $\alpha = 0$, i.e. $\mathcal{M} = 0$, which is the Minkowski spacetime in coordinates connected with an expanding family of timelike straight lines.

Note that $\mathrm{d}r_o/\mathrm{d}z = 1/\mathcal{H}_0 > 0$ at $z = 0$ and becomes negative at large z. Thus the function $r_o(z)$ is increasing at small z and decreasing at large z, achieving a maximum in between. This means that there is refocussing in the Friedmann models with $\Lambda = 0$. This is true for every k, including $k = 0$, for which the maximal r_o occurs at $z = 5/4$.

14.7 The Newtonian cosmology.

We will now use Newton's theory to describe the motion of a cloud of particles that interact only by gravitation. We will make the same assumptions that had led us to the Robertson–Walker models in relativity: a homogeneous and isotropic matter distribution, $\rho = \rho(t)$, and a spherically symmetric initial distribution of velocities:

$$v_i(t_0) = v(t_0, r)\frac{x_i}{r}\ . \tag{14.42}$$

Let us consider the motion of the particle on the sphere of radius $r(t)$. Since the distribution of matter is spherically symmetric, the force exerted on this particle by matter that lies outside this sphere is zero (see Exercise 3). The particle thus moves in the potential

$$V(r) = -\frac{GM}{r} = -\frac{4}{3}\pi G\rho(t) r^2(t). \tag{14.43}$$

Consequently, the equation of motion of the particle is

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} \equiv \frac{\partial v_i}{\partial t} + v_j\frac{\partial v_i}{\partial x_j} = -\frac{GM}{r^3}x_i = -\frac{4}{3}\pi G\rho(t) x_i(t). \tag{14.44}$$

Using eq. (14.44) one can easily verify that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{v_i}{v}\right) = 0, \tag{14.45}$$

[121] W. Mattig, *Astron. Nachr.* **284**, 109 (1958).

where $v = \sqrt{v_i v_i}$, i.e. the direction of the velocity vector at every point remains the same for all time. Therefore, using (14.42) we can rewrite (14.44) as follows

$$\frac{\mathrm{d}v}{\mathrm{d}t} \equiv \frac{\partial v}{\partial t} + v\frac{\partial v}{\partial r} = -\frac{GM}{r^2} = -\frac{4}{3}\pi G\rho(t)r. \tag{14.46}$$

This is the equation of motion. The equation of continuity $\partial\rho/\partial t + (\rho v_i),_i = 0$, in consequence of $\rho = \rho(t)$, (14.42) and (14.45) takes the form

$$\frac{1}{\rho}\,\frac{\partial\rho}{\partial t} + v,_r + \frac{2}{r}\;v = 0. \tag{14.47}$$

We denote:

$$\frac{1}{\rho}\frac{\partial\rho}{\partial t} \overset{\text{def}}{=} -\,3F(t). \tag{14.48}$$

Then we have in (14.47) $(r^2 v)\,,_r = 3F(t)r^2$, which is integrated to give

$$v(t,r) = F(t)r + \frac{\mathcal{K}(t)}{r^2}. \tag{14.49}$$

We substitute this in (14.46) and obtain

$$\dot{F}r + \frac{\dot{\mathcal{K}}}{r^2} + \left(Fr + \frac{\mathcal{K}}{r^2}\right)\left(F - \frac{2\mathcal{K}}{r^3}\right) = -\frac{4}{3}\pi G\rho(t)r. \tag{14.50}$$

This is an algebraic equation in r whose coefficients are functions of t. The coefficients of different powers of r have to vanish separately, hence

$$\mathcal{K}(t) = 0, \qquad \dot{F} + F^2 = -\frac{4}{3}\pi G\rho(t). \tag{14.51}$$

Consequently, (14.49) becomes

$$v(t,r) = F(t)r, \tag{14.52}$$

and from (14.42) we have $v_i = F(t)x_i$, i.e. the Hubble law.

Equation (14.52) can be integrated again since $v = \mathrm{d}r/\mathrm{d}t$. We substitute this and (14.48) in (14.52) and obtain $(1/r)\mathrm{d}r/\mathrm{d}t = -[1/(3\rho)]\partial\rho/\partial t$, which is integrated with the result $r(t) = (A/\rho)^{1/3}$, where A is an arbitrary constant. Hence, for each particle separately

$$\rho r^3 = A = \text{constant} \overset{\text{def}}{=} \frac{3M}{4\pi}. \tag{14.53}$$

This is an analogue of (14.24). Substituting (14.53) in (14.48), and then substituting the result in (14.51) gives $\ddot{r}/r = -GM/r^3$. Consequently,

$$\dot{r}^2 = \frac{2GM}{r} - 2K, \tag{14.54}$$

where K = constant. This equation is the same as (14.32), so it has the same 3 types of solutions. In relativity, (14.32) defined the scale factor $R(t)$ that determined the changes of distance (see (14.15)). Here, (14.54) determines the distance of an arbitrary particle from the origin $r = 0$. In the R–W models, k determined the sign of the spatial curvature. Here, multiplying (14.54) by m, the mass of the fluid particle, we see that $(-mK)$ is the total energy of the particle. If $K > 0$, then the energy is negative; then at $r \to GM/K$ the velocity goes down to zero; the particle turns back towards the origin and falls onto it in a finite time. When $K = 0$, the particle can escape infinitely far from the origin, but its velocity tends to zero as $r \to \infty$. If $K < 0$, then $\dot{r} \underset{r\to\infty}{\longrightarrow} \dot{r}_0 > 0$.

Eq. (14.54) was found by E. A. Milne and W. H. McCrea in 1934 [57], and the authors were surprised that this result had not been known earlier. It could have been found still in the 18th century, if only anybody allowed for the *possibility* that the Universe is not static. However, up until 1929 everybody *knew* that the Universe was unchanging in time.

Although the Milne–McCrea solutions are formally identical with those of Friedmann, their physical interpretation is different. The solutions of (14.54) contain typical Newtonian notions such as absolute space and absolute velocity of matter (relative to space points). In contrast to relativity, they do not imply any law of propagation of light. Making the Milne–McCrea models consistent with special relativity requires the extension of Newton's theory by new postulates. Such a theory (called "kinematical relativity") was devised by Milne [122], but it did not gain acceptance because, unlike relativity, it was well-suited only to this particular cosmological model.

14.8 The Friedmann solutions with the cosmological constant.

Now we will discuss the full Friedmann equation (14.25):

$$\dot{R}^2 = \frac{2G\mathcal{M}}{c^2 R} - k + \frac{1}{3}\lambda R^2, \tag{14.55}$$

where $\Lambda \overset{\text{def}}{=} -\lambda$. The method of discussion used above was first presented by A. Friedmann [38]; it can also be found in the book by Robertson and Noonan [123] and, in a more extended form, in Rindler's book [124]. Friedmann's original discussion was incomplete because he considered only the case $k = +1$, and did not know the case $k = 0$.

Exemplary graphs of the function $\lambda(R)$ determined by $\dot{R}^2 = 0$ are shown in Fig. 14.2. When $k \leq 0$, $\lambda(R)$ is monotonic and negative for all R. When $k > 0$, $\lambda(R)$ increases from $-\infty$ at $R = 0$ through 0 (attained at $R = 2G\mathcal{M}/(c^2k)$) to the maximum $\lambda_E = c^4k^3/\left(9G^2\mathcal{M}^2\right)$, attained at $R = 3G\mathcal{M}/(c^2k) \overset{\text{def}}{=} R_E$, and then monotonically decreases to zero as $R \to \infty$. Since λ is a universal constant, R can vary only along straight lines parallel to the R-axis. Since λ calculated from (14.55) is never smaller than the $\lambda(R)$ determined by $\dot{R} = 0$, the allowed area for R-values is above the corresponding $\dot{R} = 0$ curve, and the extrema of R lie on $\dot{R} = 0$. By definition, $R(t) > 0$. This implies the following:

[122] E. A. Milne, *Kinematic Relativity.* Clarendon Press, Oxford 1948.

[123] H. P. Robertson and T. W. Noonan, *Relativity and cosmology.* W. B. Saunders Company, Philadelphia – London – Toronto 1968, p. 374 – 378.

[124] W. Rindler, *Essential relativity. Special, General and Cosmological.* Revised 2nd ed. Springer, Berlin 1980.

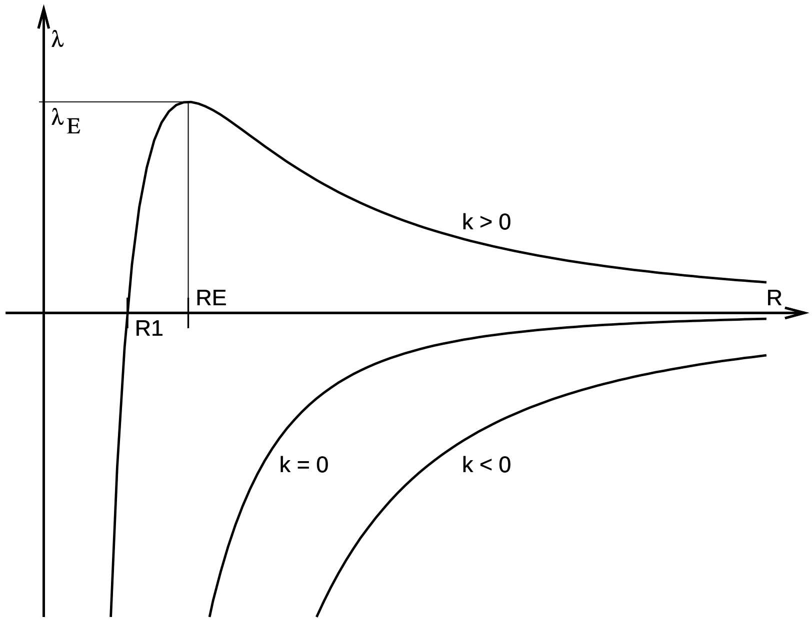

Figure 14.2: The curves $\dot{R}^2 = 0$ on the (R, λ) plane. The area accessible for evolution is above each curve. $R_1 = R_{\max} = 2G\mathcal{M}/(c^2 k)$ is the size of the Universe at maximal expansion when $\lambda = 0 < k$. $\lambda_E = c^4 k^3/(3G\mathcal{M})^2$ is the minimal value of λ at which R changes between 0 and ∞; it corresponds to the maximum of the $k > 0$ curve that occurs at $R = R_E = 3G\mathcal{M}/(c^2 k)$.

- (1) For $\lambda < 0$, only models oscillating between $R = 0$ and $R = R_{\max}$ exist. The $R_{\max}$ is greater for $k < 0$ and smaller for $k > 0$. The "cosmological attraction" implied by $\lambda < 0$ will always halt and reverse the expansion of the Universe.
- (2) For $\lambda = 0$ and $k \leq 0$ the model expands monotonically from infinite density at $R = 0$ to zero density at $R \to \infty$ or contracts monotonically from $R = \infty$ to $R = 0$; such models were discussed in Sec. 14.6.
- (3) For $\lambda = 0$ and $k > 0$ the model oscillates between $R = 0$ and $R_{\max} \overset{\text{def}}{=} 2G\mathcal{M}/(c^2 k)$.
- (4) For $0 < \lambda < \lambda_E$ the following solutions exist:
  - (4a) If $k \leq 0$, then the model expands from a singularity at $R = 0$ to zero density at $R \to \infty$, or contracts from $R = \infty$ to $R = 0$. The $\dot{R}$ in (14.55) cannot change sign in this case. Moreover, if $\dot{R} > 0$, then $\dot{R}(t)$ is an increasing function at sufficiently large R; the expansion then proceeds with acceleration.
  - (4b) If $k > 0$, then we have two cases:
    * (4b$_A$) If $R < R_E \overset{\text{def}}{=} 3G\mathcal{M}/(c^2 k)$, then the model is oscillating.
    * (4b$_B$) If $R > R_E$, then the model contracts from zero density at $R = \infty$ to a finite maximal density at $R = R_{\min}$ (with $\dot{R}(R_{\min}) = 0$), and then expands again to $R \to \infty$.
- (5) For $\lambda = \lambda_E$ there are several possibilities:

– (5a) If $k \leq 0$, then the model constantly expands or contracts, like in case (4a).

– (5b) If $k > 0$, then the right-hand side of (14.55) has a double root $R = R_E = 3G\mathcal{M}/(c^2 k)$, and then the time required to reach R_E from either side,

$$t = \int_{R(t_0)}^{R_E} \left(\frac{2G\mathcal{M}}{c^2 R} - k + \frac{1}{3}\lambda_E R^2 \right)^{-1/2} \mathrm{d}R \tag{14.56}$$

is infinite. Then there exist the following subcases:

* (5b$_A$) If $R < R_E$, then the model either expands from a singularity at $R = 0$ asymptotically to $R = R_E$ at $t \to \infty$, or contracts from the asymptotic state $R \to R_E$ at $t \to -\infty$ to a singularity at $R = 0$.
* (5b$_B$) If $R > R_E$, then the model either expands from an asymptotic state $R \to R_E$ at $t \to -\infty$ to $R \to \infty$ at $t \to \infty$, or contracts from $R = \infty$ at $t = -\infty$ to $R \to R_E$ at $t \to \infty$.
* (5b$_C$) There also exists the static 'Einstein Universe' (10.74) with $R \equiv R_E$. It is unstable because a small perturbation of R will make it expand as in case (5b$_B$) or contract as in case (5b$_A$).

• (6) For $\lambda > \lambda_E$, with any sign of k, there exist only models that monotonically expand from $R = 0$ to $R \to \infty$ or monotonically contract. Just as in cases (4a) and (5b$_B$), the expansion proceeds with acceleration, since $\dot{R}(t)$ increases at sufficiently large R.

This discussion is summarised in Figs. 14.3 – 14.5. Fig. 14.3 shows all the recollapsing models. With $\lambda < 0$ but close to zero, the maximal size of the model with $k \leq 0$ can be arbitrarily large and also its lifetime can be arbitrarily long. Fig. 14.4 shows all the models for which $\lambda = \lambda_E$, and Fig. 14.5 shows the remaining ones.

With $\lambda \neq 0$, the connection between the sign of spatial curvature and the type of motion no longer holds. There are models with $k < 0$ that recollapse after a finite time (case (1)) and models with $k > 0$ that expand to reach infinite extent (cases (4b$_B$), (5b$_B$) and (6)). The latter do not always attain a singular state.

14.9 Horizons in the Robertson–Walker models.

Every observer in the Universe receives information about distant objects via light rays. In most of the R–W models there exist such objects, from which a given observer has not yet received any ray. Then, in some of the R–W models (those in which expansion proceeds with acceleration) objects exist from which a given observer has not received and will never receive any ray. The boundaries separating objects already observed from those not yet observed, and objects observable from unobservable are called *horizons*. These definitions will be made precise below. This section is based on the paper by Rindler [125].

[125] W. Rindler, *Mon. Not. Roy. Astr. Soc.* **116**, 662 (1956); reprinted in *Gen. Relativ. Gravit.* **34**, 133 (2002), with an editorial note by A. Krasiński, *Gen. Relativ. Gravit.* **34**, 131 (2002) and author's (auto)biography by W. Rindler, *Gen. Relativ. Gravit.* **34**, 132 (2002).

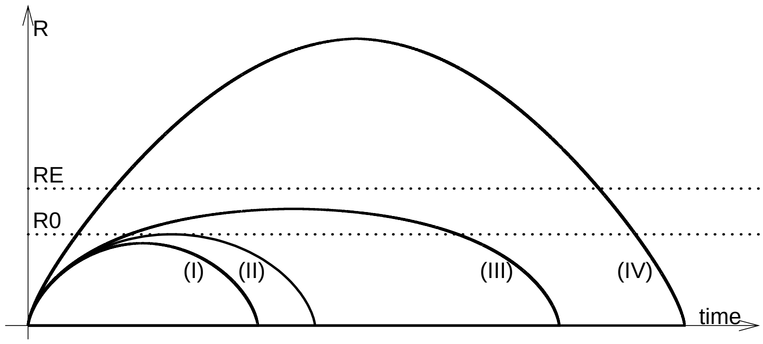

Figure 14.3: $R(t)$ for the recollapsing Friedmann models, for different values of k and of λ. Curve (I): $k > 0$, $\lambda < 0$; curve (II): $k > 0$, $\lambda = 0$ (this is one of the models considered in Sec. 14.6); curve (III): $k > 0$, $0 < \lambda < \lambda_E$; curve IV: $k \leq 0$, $\lambda < 0$. The horizontal line RE is the constant $R = R_E = 3G\mathcal{M}/c^2$ in the static Einstein Universe. The horizontal line R0 is the maximal R in the recollapsing $\lambda = 0$ model. As seen from Fig. 14.2, negative λ will always force recollapse, irrespective of the sign of k. However, with $k \leq 0$, an arbitrarily large maximal R can occur with a sufficiently small absolute value of λ. For better visualisation, the parameters of the four curves were chosen as close to each other as possible. On all curves, $G\mathcal{M}/c^2 = 1$. On curve (I) $k = +1$, $\lambda = -0.1$, on curve (II) $k = +1$, $\lambda = 0$, on curve (III) $k = +1$, $\lambda = +0.1$, on curve (IV) $k = -1$, $\lambda = -0.1$. With negative values of λ of sufficiently great absolute value, the lifetime for $k \leq 0$ models may be shorter than in any $k > 0$ model.

The **event horizon** for an observer A is the hypersurface in the spacetime that divides the collection of all events into two nonempty classes: those that have been, are or will be observed by A, and those that A has never observed and will never be able to observe. Not every R–W spacetime has event horizons.

The **particle horizon** for an observer A at the instant t_0 is a 2-dimensional surface in the space $t = t_0$ that divides all particles (i.e., world lines of matter) into two nonempty classes: those that A has seen up to $t = t_0$, and those that A has not yet seen.

There exist R–W models without any of these horizons (e.g., the Newtonian model described in Sec. 14.7). All the Friedmann solutions with $\Lambda = 0$ have particle horizons.

We will use the coordinates of (14.1). Let us define

$$\sigma(r) \overset{\text{def}}{=} \int_0^r \frac{\mathrm{d}r'}{1 + \frac{1}{4}kr'^2} = \begin{cases} \frac{2}{\sqrt{k}} \arctan\left(\frac{\sqrt{k}r}{2}\right) & \text{for} \quad k > 0, \\ r & \text{for} \quad k = 0, \\ \frac{1}{\sqrt{-k}} \ln\left(\frac{1+\frac{1}{2}\sqrt{-k}r}{1-\frac{1}{2}\sqrt{-k}r}\right) & \text{for} \quad k < 0. \end{cases} \tag{14.57}$$

This quantity is proportional to $\ell(t, r)$ – the length of the arc of the curve $\{t, \vartheta, \varphi\}$ = constant between the point $r = 0$ and the point with the running value of r.

With $k < 0$ we have $\sigma(r) \to \infty$ for $r \to 2/\sqrt{-k}$, hence the range $r \in [0, 2/\sqrt{-k})$ covers the whole (infinite) space t = constant.

For $k > 0$ we have $\sigma \to \pi/\sqrt{k}$ for $r \to \infty$. Hence, $r = \infty$ corresponds to the point that

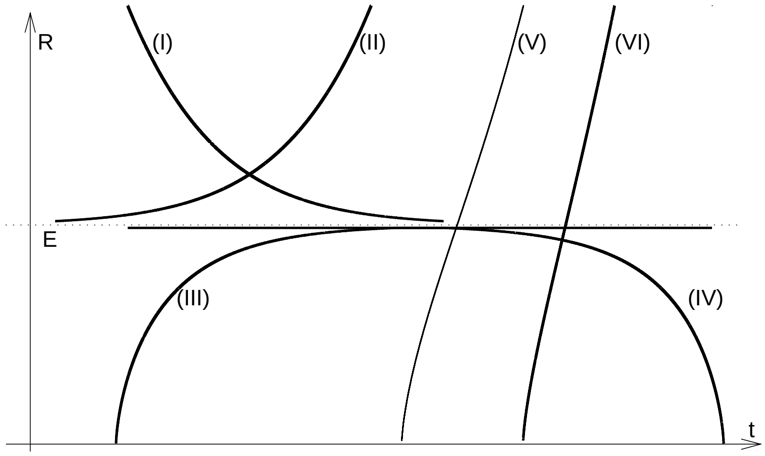

Figure 14.4: The graphs of $R(t)$ for the Friedmann models with $\lambda = \lambda_E$. The dotted horizontal line marked E is the static Einstein Universe in which $R = R_E = 3G\mathcal{M}/(c^2k) =$ constant. Curves I to IV represent models that approach the Einstein Universe asymptotically towards the future or towards the past; for all of them $k > 0$. On curve I, $R > R_E$ and $R_{,t} < 0$. On curve II, $R > R_E$ and $R_{,t} > 0$. On curve III, $R < R_E$ and $R_{,t} > 0$. On curve IV, $R < R_E$ and $R_{,t} < 0$. Curve V represents the model with $k = 0$, curve VI the model with $k < 0$; they both have inflection points at intersections with $R = R_E$. Time-reverses of curves V and VI (i.e. collapsing models) are also solutions of eq. (14.55), but are omitted for clarity. Curves III – VI represent models that have singularities in the past or in the future. Curves I and II represent models with no singularity. Note that the static Einstein Universe is unstable: an arbitrarily small perturbation that sets R off the value R_E will cause the model to expand or contract away from the initial state. The actual values of parameters in the figure are $G\mathcal{M}/c^2 = 1$ on all curves, $k = +1$ for curves I – IV and $k = -1$ for curve VI.

is antipodal to $r = 0$ on each sphere $t =$ constant. The r-curve can be continued beyond this point, but the r coordinate does not cover the extra stretch.

The continued line may wind multiply around the sphere $t =$ constant. On such multiply-wound curves we define σ to be $\sigma(r) = n\pi/\sqrt{k} + \widetilde{\sigma(r)}$, where n is the number of passages of the curve through the poles $r = 0$ and $r = \infty$, and $\widetilde{\sigma(r)}$ is calculated by (14.57) between the final r and the last pole passed. Hence, also in the model with $k > 0$, a point in the space $t =$ constant can be assigned to every value of σ.

For the observer placed at $r = 0$, the equation of motion of a particle at $r = r_1$ is

$$\ell(t, r_1) = R(t)\sigma(r_1). \tag{14.58}$$

The equation of motion of a photon emitted at $r = r_1$ and proceeding *towards* the observer at $r = 0$ is, from (14.9) and (14.57):

$$\sigma(r) = \sigma(r_1) - \int_{t_1}^{t} \frac{\mathrm{d}\tau}{R(\tau)}. \tag{14.59}$$

The following discussion is concentrated on models with infinite time of existence. Recollapsing models will be mentioned only casually.

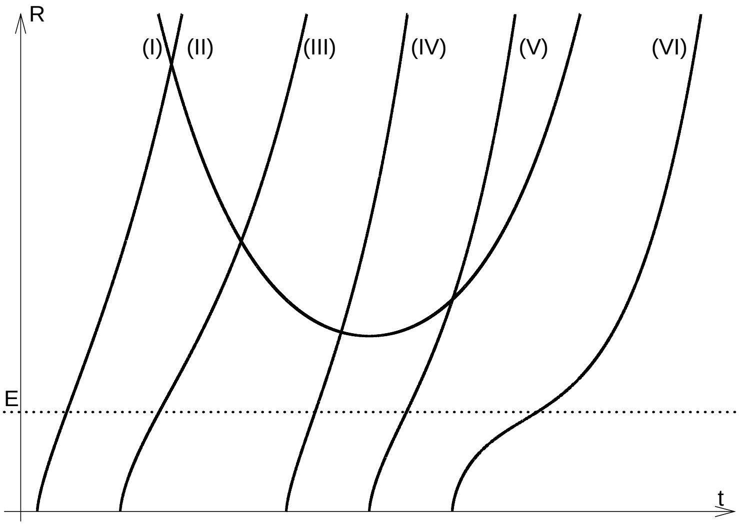

Figure 14.5: The graphs of $R(t)$ for the remaining Friedmann models, for different values of k and λ. The horizontal line E is the constant value of $R = R_E = 3G\mathcal{M}/(c^2k)$ in the static Einstein Universe. On curves I – III $0 < \lambda < \lambda_E$; on curves IV – VI $\lambda > \lambda_E$. On curve I $k > 0$ and $R > R_E$, on curve II $k < 0$, on curve III $k = 0$. On curves IV, V and VI k is, respectively, negative, zero and positive. Curves II and III have their inflection points for $R > R_E$, curves IV to VI have their inflection points for $R < R_E$. Positive λ implies cosmic repulsion that opposes gravitational attraction. Beyond the inflection point the repulsion prevails and sets the Universe into accelerated expansion. With large initial R, the repulsion prevents collapse even for $k > 0$.

When $k \leq 0$, the necessary and sufficient condition for the existence of an event horizon is the convergence of $\int_{t_1}^{t} \mathrm{d}\tau/R(\tau)$ to a finite limit at $t \to \infty$.

Then, $\sigma(r_1)$ differs from $\sigma(r)$ by a finite quantity even at infinite future. In other words, there exist particles at $r = r_H$ (with $\sigma(r_H) < \infty$) such that the photon emitted from there at $t = t_H$ will never reach the observer at $r = \sigma(0) = 0$.

From (14.59) we get for r_h (the minimal value of r_H):

$$\sigma(r_h) = \int_{t_h}^{\infty} \frac{\mathrm{d}\tau}{R(\tau)} < \infty. \tag{14.60}$$

When $k > 0$, the event horizon exists if $\sigma(r_h) \leq \pi/\sqrt{k}$. If $\sigma(r_h) > \pi/\sqrt{k}$, then each photon sent from $r = r_h$ can, in a finite time, travel further than half around the spherical space, so it will reach the observer at $r = 0$ on one route or another.

(However, for recollapsing models with $k > 0$, the event horizon may exist even if $\sigma(r_H) \overset{\text{def}}{=} \int_{t_H}^{t_{FS}} \mathrm{d}\tau/R(\tau) > \pi/\sqrt{k}$ – see the example below).

In consequence of isotropy of the R–W models, if an event horizon exists for one direction, then it exists for every direction of observation. In consequence of their homogeneity, the origin $r = 0$ may be placed at any point of the space $t =$ constant. Hence, if there exists an event horizon for one given observer, then event horizons exist for all observers.

When the event horizon exists, the events at $r > r_h$ (when $k \le 0$) or $r_h < r \le \pi/\sqrt{k}$ (when $k > 0$) will never be seen by the observer at $r = 0$. How can this be explained intuitively?

If $R(t)$ increases sufficiently rapidly, then the spatial distance between the observer at $r = 0$ and the light source at $r = r_1 > r_h$, given by (14.58), increases so fast that light cannot overcome this "swelling of space", even though it keeps going towards the observer. Eddington once explained this as follows: imagine someone who runs on an expanding race track, but the finish line moves away faster than he can run.

In the Friedmann models with $\Lambda = 0$ and $k \le 0$ event horizons do not exist: every event will eventually become visible for every observer because, from (14.33), $\int_{t_1}^{t_2} \mathrm{d}\tau/R(\tau) \underset{t_2\to\infty}{\longrightarrow} \infty$. However, with $k > 0$, the final singularity occurs at $\omega = 2\pi$, and then

$$\sigma(r_1) = \int_{t_1}^{t(2\pi)} \frac{\mathrm{d}\tau}{R(\tau)} = \int_{\omega(t_1)}^{2\pi} \frac{1}{\sqrt{k}}\,\mathrm{d}\omega = \frac{2\pi - \omega(t_1)}{\sqrt{k}}. \tag{14.61}$$

Hence, if $\omega(t_1) > \pi$ (i.e. $t_1 - t_B > \pi\alpha/k^{3/2}$), then events at $t > t_1$ and $r > r_1$ will not become visible for the observer at $r = 0$ before the final singularity occurs. The event horizon will thus exist.

From (14.59) it is also seen that if a particle was initially inside the event horizon for observer A, then it will remain visible for her for ever. This is because, if $1/R(t)$ is continuous and t_1 obeying the equation

$$0 = \sigma(r_1) - \int_{t_1}^{t_0} \frac{\mathrm{d}\tau}{R(\tau)} \tag{14.62}$$

exists for a given r_1 and $t_0 < \infty$, then it will exist for all $t > t_0$ and the same r_1.

Proof:

Since $\sigma(r_1)$ does not depend on t, the proof requires finding such an integration interval in which $\int_{t_1}^{t_0} \mathrm{d}\tau/R(\tau)$ has a given value. From Fig. 14.6 one sees that if one such interval $[t_1, t_0]$ exists, then for every $t_0' > t_0$ there will exist a $t_1' > t_1$ such that

$$\int_{t_1'}^{t_0'} \frac{\mathrm{d}\tau}{R(\tau)} = \int_{t_1}^{t_0} \frac{\mathrm{d}\tau}{R(\tau)}. \qquad \square$$

If the event horizon exists (i.e. $\sigma(r_1) < \infty$ for $t_0 \to \infty$ in (14.62)), then t_1 goes to a finite limit $t_{\rm out}$ as $t_0 \to \infty$.

Then, although observer A will always see the particle at $r = r_1$, she will see it only up to the instant $t = t_{\rm out}$ on the particle's clock. The signal sent from $r = r_1$ at $t = t_{\rm out}$ will reach A at $t_0 \to \infty$, i.e. never.

This, in turn, means that the world line of the particle at $r = r_1$ will intersect the event horizon of A at $t = t_{\rm out}$. Moreover, a $t_{\rm out} < \infty$ exists for every $r_1 < r_h$.

Consequently, when the event horizon exists, each particle will be visible for A only for a finite period of its history and will eventually (at infinite future of A) escape from A's

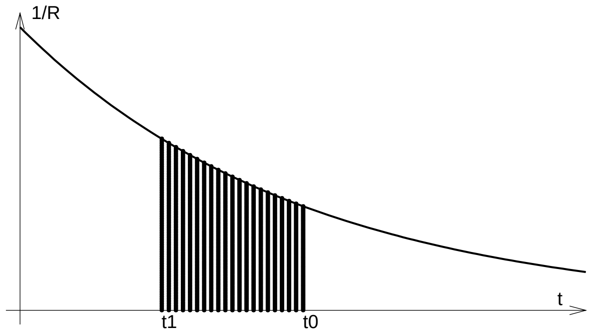

Figure 14.6: An illustration to (14.62). If the integral of the continuous function $1/R$ over the interval $[t_1, t_0]$ equals a given value $\sigma(r_1)$, then for every $t'_0 > t_0$ there exists a $t'_1 > t_1$ such that the integral of the same function over $[t'_1, t'_0]$ has the same value $\sigma(r_1)$.

field of view through the horizon. The further away the particle, the shorter the period of its history that will be visible for A.

The necessary and sufficient condition for the existence of the particle horizon is the convergence of the integral in (14.62) to a finite limit at $t_1 \to t_B$. (In those models that have no initial singularity, it is the convergence to a finite limit at $t \to -\infty$. The analysis of this case is similar and will be omitted here.)

If the integral is finite at $t_1 \to t_B$, then

$$\sigma(r_1) = \int_{t_B}^{t_0} \frac{\mathrm{d}\tau}{R(\tau)} \overset{\mathrm{def}}{=} \phi(t_0) \tag{14.63}$$

determines the farthest particles from which the observer at $r = 0$ could have received a light signal up to the instant t_0.

At fixed t_0, (14.63) is an equation of a two-dimensional sphere. Since $1/R > 0$, the function $\phi(t)$ given by (14.63) is increasing. This means that, if the particle horizon exists, then with time still more particles are within it.

Whether every particle will eventually enter the particle horizon depends on whether the integral in (14.63) has a finite limit at $t_0 \to \infty$. If it has, then, according to the previous definition, the event horizon exists and some particles will never enter A's field of view.

If $\sigma(r_1) \underset{t_1 \to t_B}{\longrightarrow} \infty$ in (14.62), then the observer at $r = 0$ could receive signals from all other particles for any t_0.

The first signal received by an observer A from each particle is the one sent out at the initial singularity at $t = t_B$. However, by eq. (14.13), that signal (for which $R(t_e) = 0$) is received with infinite redshift. This means that **in the Robertson–Walker models the initial singularity is not observable.**[47]

[47] This is not a universal property of all cosmological models. In Chapter 15 we will see that in some Lemaître–Tolman models there exist radial rays emitted from the initial singularity that would be received with an infinite blueshift by all observers. However, in the real Universe direct signals from the

In the models that contract from $R = \infty$ at $t = -\infty$ and possess a particle horizon, eq. (14.13) gives $1 + z = 0$ at $t_e = -\infty$. Since $1 + z = \lambda_o/\lambda_e$, in these models the first signal from each particle is observed with zero wavelength, i.e., is infinitely blueshifted.

In the Friedmann models with $\Lambda = 0$ particle horizons do exist.

The two kinds of horizons were defined above for observers comoving with matter. But one can also consider observers moving independently of the matter background. In some cosmological models there exist horizons also for such observers: irrespective of how fast the observer moves towards a given world line of matter, she will never see any event on that world line if the line is sufficiently far from the observer's starting point [125].

There exists one more type of horizon: the *apparent horizon* (see Secs. 11.8 and 13.4). In a Universe expanding from a singularity, it is a closed hypersurface at which both the inward- and outward-going light rays are diverging, i.e. at which the expansion scalar of any bundle of emitted light rays is positive. The calculation is simpler in the coordinates of (14.3). The tangent vector field to a bundle of null geodesics emanating radially from a sphere is found by transforming (14.11) with use of (14.2). Dropping primes, the result is

$$k^\alpha = \left[\frac{1}{R}, \frac{\varepsilon}{R^2}\ \sqrt{1 - kr^2}, 0, 0\right], \tag{14.64}$$

where $\varepsilon = +1$ for outward-going and $\varepsilon = -1$ for inward-going rays. We have

$$2\theta = k^\alpha{}_{;\alpha} = \frac{1}{\sqrt{-g}}\ \left(\sqrt{-g}k^\alpha\right)_{,\alpha}. \tag{14.65}$$

Taking $R_{,t}$ from (14.32) with the + sign (the Universe is expanding!) we find that $\theta > 0$ is equivalent to

$$\frac{r\sqrt{2\alpha/R - k}}{\sqrt{1 - kr^2}} + \varepsilon > 0. \tag{14.66}$$

This is fulfilled for an outward-going bundle ($\varepsilon = +1$), and it will be fulfilled also for an inward-going bundle ($\varepsilon = -1$) when

$$R < 2\alpha r^2 = \frac{2G\mathcal{M}}{c^2}\ r^2. \tag{14.67}$$

This is a region of past-trapped surfaces, and it always exists in a vicinity of the Big Bang. Its envelope, where $R = 2\alpha r^2$, is the past apparent horizon.

Note from (14.17) and (14.2) that, in the coordinates we are now using, rR is the source area distance from the singularity at $R = 0$. Then, note from (14.24) that $\mathcal{M}r^3$ is equal to the mass contained within the source area distance rR from the singularity.

Thus (14.67) can be written in the equivalent form $r_G < 2m$, where $m \overset{\text{def}}{=} G\mathcal{M}r^3/c^2$. This is, not accidentally, similar to the condition defining the interior of the horizon in the Schwarzschild solution. The meaning of this will become clearer in the Lemaître – Tolman model, see Section 15.8.

Big Bang cannot be received for another reason: for some time after the BB the matter in the Universe is nontransparent because photons keep interacting with other elementary particles.

14.10 The 'history of the Universe'

The idea that the Universe might have a history had been developing slowly. The discovery of the expansion of the Universe and the realisation that the Friedmann [38] – Lemaître [108] solutions of Enstein's equations account for it set this line of thinking in motion. Since the Universe had been denser in the past, it must have been hotter as well. Some time ago then, it should have been sufficiently hot that all atoms were ionised. When the temperature dropped below the ionisation temperature, radiation had to be emitted. Expansion of the Universe should have cooled the radiation. Assuming that it had a black-body spectrum all the time, the evolution of temperature can be calculated. The intensity I of the black-body radiation as a function of frequency ν is given by

$$I(\nu) = \frac{2h\nu^3}{c^2 \left\{\exp\left[h\nu/(kT)\right] - 1\right\}}, \tag{14.68}$$

where h and k are the Planck and Boltzmann constants, respectively, and T is the temperature of the radiation. The received frequencies of radiation obey (13.23), so $\nu(1+z) = \nu_e =$ constant along each ray. Consequently, to keep the form of the function $I(\nu)$ in (14.68) unchanged, we must have $T(1+z) = \text{const}$. That radiation should still be present in the Universe today. This idea first occurred to Gamow [126] and collaborators [127]. The discovery came in 1965: the radiation exists and has the temperature 2.73 K [128, 129].

The speculation went on. Atomic nuclei are built of protons and neutrons. All the known atomic nuclei could, in principle, be built by adding protons and neutrons one by one to the nucleus of hydrogen (the proton). Still further in the past, the Universe must have been hot enough to crash all heavier nuclei; only protons, neutrons and loose electrons could survive. Is it possible that matter originally consisted only of these particles, and heavier nuclei came into existence through collisions between them? This idea again occurred to Gamow [126] and Alpher and Herman [127]. Computer simulations indicated that something like this was going on, but only the first few nuclei of the Mendeleev table could be created in this way: about 25 % of the mass of the Universe would be converted into helium ^{4}He. Tiny traces of deuterium, helium ^{3}He, lithium, beryllium and boron could be created, but the falling temperature of the Universe would stop any further synthesis [130]. (Heavier elements are synthesised later, in the stars [131].) These calculated proportions of nuclides were confirmed by observations [132].

Thinking along these lines, cosmologists reconstructed the possible sequence of events in the evolution of the Universe. The conclusions mentioned above were verified obser-

[126] G. Gamow, The evolution of the Universe, *Nature* **162**, 680 (1948).

[127] R. A. Alpher and R. C. Herman, Evolution of the Universe, *Nature* **162**, 774 (1948).

[128] R. H. Dicke, P. J. E. Peebles, P. G. Roll and D. T. Wilkinson, Cosmic black-body radiation, *Astrophys. J.* **142**, 414 (1965).

[129] A. A. Penzias and R. W. Wilson, A measurement of excess antenna temperature at 4080 Mc/s, *Astrophys. J.* **142**, 419 (1965).

[130] R. V. Wagoner, W. A. Fowler and F. Hoyle, On the synthesis of elements at very high temperatures, *Astrophys. J.* **148**, 3 (1967).

[131] W. A. Fowler, *Nuclear Astrophysics*. American Philosophical Society, Philadelphia 1967.

[132] A. M. Boesgaard and G. Steigman, Big bang nucleosynthesis – theories and observations, *Ann. Rev. Astron. Astrophys.* **23**, 319 (1985).

vationally, others are speculations. Details can be found in [133]. Here we list only the most important events. The leading motive was this: as we proceed backwards in time, approaching the Big Bang, the density and temperature of matter become arbitrarily high. Thus, whatever processes we know that should take place at high temperatures, must have taken place in the early Universe. Why the BB explosion occurred and what preceded it are questions that cannot be answered by means of the currently existing physics or mathematics. Thus, they are usually not asked and we take the BB as a given thing.

The natural question is, at precisely what times the consecutive stages of evolution took place. The numbers vary between sources. The selection of references given below is random; with no claim to being precise and up to date.

The current number for the time of the BB explosion is $\approx 13.67 \times 10^9$ years ago [85].

During the first 10^{-34} s, the Universe had a temperature above 10^{27} K and was supposedly described by a 'Grand Unified Theory' (GUT, still in the making) that unites the strong nuclear, weak nuclear and electromagnetic interactions ([134], Vol. II). The elementary particles we know today had not necessarily existed then, and matter might have been composed of free quarks and gluons.

Between 10^{-34} s and 10^{-32} s after the BB, inflation took place. This is a hypothetical period during which the Universe had been expanding at an exponential rate. At the end of it, the elementary particles we know today should have come into existence, together with seeds for structure formation. The reasons, mechanism and exact instant of creation of the seeds are still a matter of debate ([134], Vol. II).

About 1 s after the BB, neutrinos decoupled and should have thereafter propagated freely through space. During the next few seconds, protons, neutrons, electrons, positrons and photons existed in thermal equilibrium, at temperatures $T \geq 10^{10}$ K [52].

The formation of light atomic nuclei occurred between 2 and 1000 s after the BB [52]. At the end of this period, the temperature dropped to about 10^9 K.

Later, the Universe continued to be a radiation-dominated plasma, but the radiation mass-density ρ_r (which obeyed $\rho_r R^4$ = constant, with $R \propto 1/T$, as follows from (14.68) and (14.13)) was decreasing faster than the mass-density of massive particles, ρ_m (obeying $\rho_m R^3$ = constant). About 3×10^5 years after the BB ([134], Vol. II), and at temperatures of the order 10^3 K, ρ_r became smaller than ρ_m and radiation decoupled from matter, having too little energy to ionise the atoms that had already captured their electrons. It evolved into the CMB radiation of temperature 2.73 K that is observed today.

Still later, structures like galaxies, galaxy clusters, superclusters and larger condensations, and voids, must have formed. Even though the process of structure formation should have proceeded by gravitation and should be well within the domain of classical gravitation theory, it is still poorly understood. There exists no quantitative account of the emergence of structures, only a general firm belief that structures came about by gravitational magnification of fluctuations created very early. Models of the R–W class are

---

[133] T. Padmanabhan, *Structure Formation in the Universe*, Cambridge University Press 1993.

[134] K. Lang, *Astrophysical Formulae. Vol. I: Radiation, Gas Processes and High Energy Astrophysics; Vol. II: Space, Time, Matter and Cosmology.* Springer, Berlin – Heidelberg – New York 1999.

unable to describe structure formation – perturbations of them must be considered for this purpose. It is often claimed that the high isotropy of the CMB radiation (with maximal temperature anisotropies of $\Delta T/T \approx 2.93 \times 10^{-6}$ at the angular scale of about 0.9°; see Ref. [135]) *proves* that our Universe has a R–W geometry. Such statements are, however, meaningless without a quantitative account of interaction between the CMB radiation and inhomogeneities in matter distribution. Existing estimates show that the interaction is weak, and no temperature anisotropies larger than 10^{-5} should ever have been expected.

14.11 The redshift – distance relation in the $\Lambda \neq 0$ Friedmann models

Let us define the following three dimensionless parameters:

$$(\Omega_m, \Omega_k, \Omega_\Lambda) \overset{\text{def}}{=} \frac{1}{3{\mathcal{H}_0}^2} \left(\frac{8\pi G \rho_0}{c^2}, -\frac{3k}{{R_0}^2}, -\Lambda \right) \Bigg|_{t=t_o}, \tag{14.69}$$

where t_o is the current instant, $\rho_0 = \rho(t_o)$ is the current mean mass density in the Universe, $c\mathcal{H}_0 = H_0$ is the current value of the Hubble parameter and $R_0 = R(t_o)$. They are called, respectively, the density parameter, the curvature parameter and the cosmological constant parameter. Using them, eq. (14.26) becomes

$$\Omega_m + \Omega_k + \Omega_\Lambda = 1. \tag{14.70}$$

We assume that the observer is located at the centre of symmetry of the spacetime, so every light ray received or emitted by her will be radial. In an R–W spacetime this is no limitation of generality: because of homogeneity, every point in a space of constant t is the center of symmetry. Then, from (14.3), the definition of the observer area distance r_o (13.57), and Fig. 13.2 we see that

$$r_o = (rR)|_{\text{ray}}, \tag{14.71}$$

where $(rR)|_{\text{ray}}$ is calculated along the light ray connecting the source to the observer (to see this, take for $\delta\Omega_o$ the full solid angle 4π and calculate from (14.3) the surface area of the whole sphere with centre at the observer, $(t, r) = (t_o, 0)$, and radius $(rR)|_{\text{ray}}$). So, in order to calculate r_o we have to solve the equation of radial null geodesics in the metric (14.3). We take the initial point of the geodesic at the observer position, and follow it to the past toward the source. Then

$$\frac{1}{R}\frac{\mathrm{d}t}{\mathrm{d}r} = -\frac{1}{\sqrt{1-kr^2}}. \tag{14.72}$$

Integrating this from $r = 0$ to r_e, the position of the light source where the ray was emitted at $t = t_e$, and taking $k < 0$ for the beginning, we obtain

$$\int_{t_o}^{t_e} \frac{\mathrm{d}t}{R(t)} = -\int_0^{r_e} \frac{1}{\sqrt{1-kr^2}} = -\frac{1}{\sqrt{-k}} \operatorname{arsinh}\left(\sqrt{-k}\, r_e\right). \tag{14.73}$$

[135] W. Hu (2004), webpage:
http://background.uchicago.edu/~hu/physics/tourpage.html

To calculate $\int_{t_o}^{t_e} \mathrm{d}t/R(t)$ we first transform the variable by $\mathrm{d}t/R(t) = \mathrm{d}R/(RR,_t)$, then substitute for $R,_t > 0$ from (14.25) and for $\mathcal{M}$ at $t = t_o$ from (14.24). The result is

$$\int_{t_o}^{t_e} \frac{\mathrm{d}t}{R(t)} = \int_{R(t_o)}^{R(t_e)} \frac{\mathrm{d}R}{R\sqrt{\dfrac{8\pi G\rho_o {R_o}^3}{3c^2 R} - k - \frac{1}{3}\Lambda R^2}}. \tag{14.74}$$

Now we use $1 + z = R_o/R_e$, from (14.13). In (14.74), the running R is at the position of the emitter, and R_o is constant. So, $R = R_o/(1 + z)$ and we now change the integration variable from R to z. At the observer position $z = 0$, so (14.74) becomes

$$\int_{t_o}^{t_e} \frac{\mathrm{d}t}{R(t)} = -\int_0^z \frac{\mathrm{d}z'}{\sqrt{\dfrac{8\pi G\rho_0 {R_0}^2}{3c^2}(1+z')^3 - k(1+z')^2 - \frac{1}{3}\Lambda {R_0}^2}}. \tag{14.75}$$

This we substitute in (14.73) and use (14.69). Solving for r_e we obtain

$$r_e = \frac{1}{\sqrt{-k}} \sinh\left[\int_0^z \frac{\sqrt{\Omega_k}\mathrm{d}z'}{\sqrt{\Omega_m(1+z')^3 + \Omega_k(1+z')^2 + \Omega_\Lambda}}\right]. \tag{14.76}$$

We now use (14.71), (13.72), the definition of D_L – (13.73), we again replace the running $R = R|_{\text{ray}}$ with $R_o/(1 + z)$ and use (14.69). The final result is

$$D_L(z) = \frac{1+z}{\mathcal{H}_0\sqrt{\Omega_k}} \sinh\left[\int_0^z \frac{\sqrt{\Omega_k}\mathrm{d}z'}{\sqrt{\Omega_m(1+z')^3 + \Omega_k(1+z')^2 + \Omega_\Lambda}}\right]. \tag{14.77}$$

This is the "canonical" form of the (uncorrected) luminosity distance vs. redshift relation.

Equation (14.77) holds also for $k > 0$ if sinh and $\sqrt{\Omega_k}$ are replaced by sin and $\sqrt{|\Omega_k|}$. (This is consistent with the identity $\sinh(\mathrm{i}x) \equiv \mathrm{i}\sin x$.)

The now-standard ΛCDM model has $k = 0 = \Omega_k$. The limit of (14.77) at $\Omega_k \to 0$ is

$$D_L(z) = \frac{1+z}{\mathcal{H}_0} \int_0^z \frac{\mathrm{d}z'}{\sqrt{\Omega_m(1+z')^3 + \Omega_\Lambda}}. \tag{14.78}$$

14.12 Exercises.

1. Verify that the Einstein tensor for an R–W metric is diagonal in all the coordinate representations (14.1), (14.3), (14.5) and (14.7).

2. Prove that a null geodesic sent off radially from any point in a R–W spacetime, i.e. with $\dot{\vartheta}_0 = \dot{\varphi}_0 = 0$ at the initial point, will remain radial along its whole length.

Hint: Consider the geodesic equations.

3. Let $\rho(t, r)$ be any spherically symmetric (inhomogeneous) finite distribution of matter (i.e. $\rho(t, r) = 0$ for $r > r_0$, $r_0 < \infty$). Let A be a point inside this distribution located at $r = r_1$. Show that the total (Newtonian) gravitational force exerted on the point A by matter outside the sphere $r = r_1$ is zero.