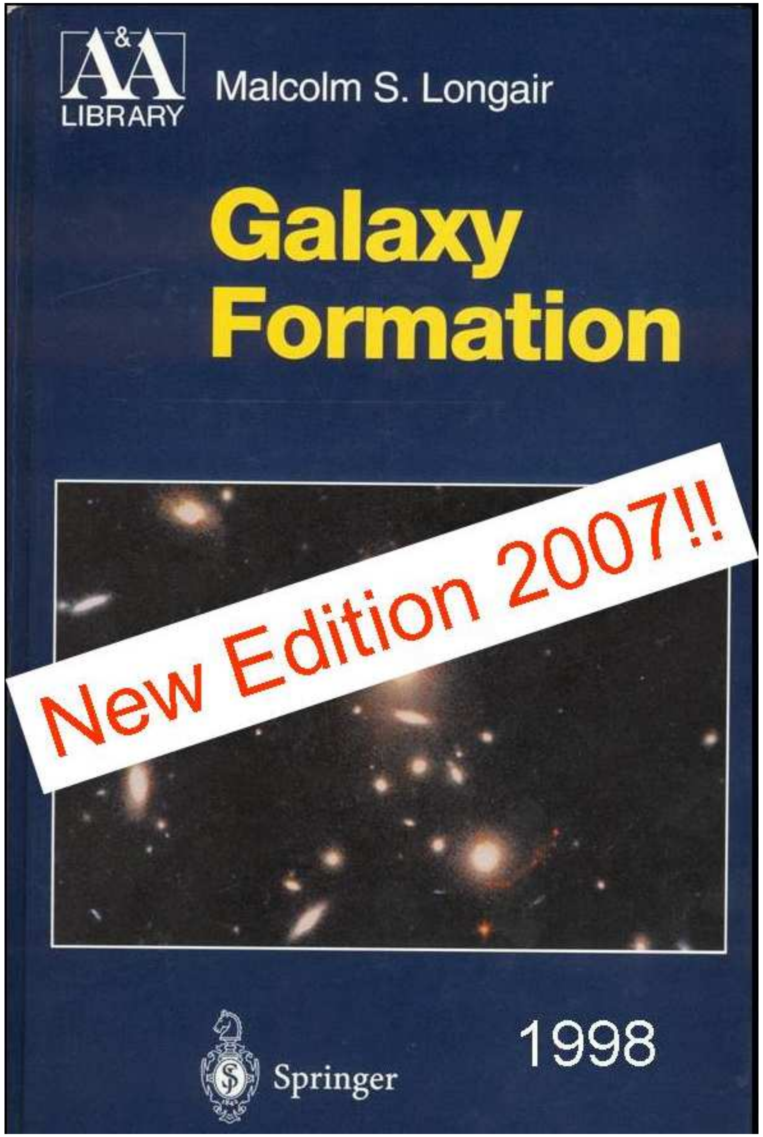
The Basic Structure of the Standard Cosmological Models

- Isotropic Curved Spaces
- Derivation of the Robertson-Walker Metric
- Results which follow from the R-W Metric
- Space-time Diagrams for the Standard World Models
- Conformal Diagrams

The Book of the Lectures

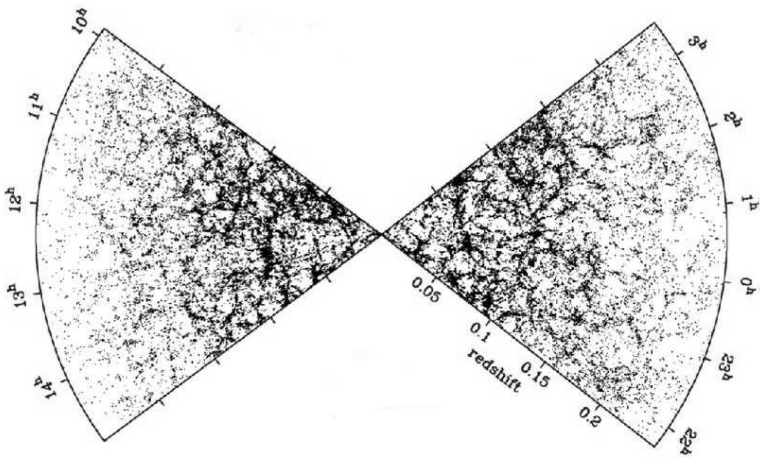


The results I will present are all derived in the new edition of Galaxy Formation

The Isotropy and Homogeneity of the Universe

COBE

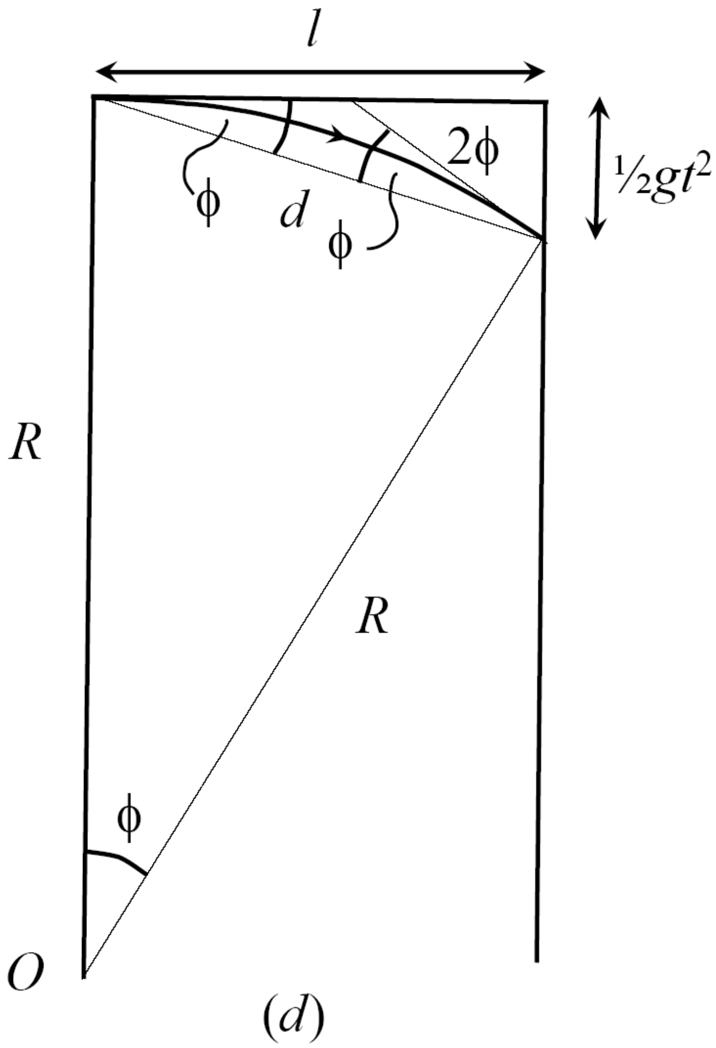
2dF Galaxy Survey



In the first Lecture, we discussed the evidence for the overall isotropy and homogeneity of the Universe.

- Away from the Galactic plane, the radiation is isotropic to better than one part in 10^5. At this level, significant temperature fluctuations $\Delta T/T \approx 10^{-5}$ were detected on scales $\theta \geq 10^\circ$.
- Although the Universe is inhomogeneous on smaller scales, the galaxies display the same degree of inhomogeneity throughout the Universe at the present epoch.
- Therefore, in the first approximation, we can take the Universe to be completely isotropic and homeogeneous.

Isotropic Curved Spaces



A lift accelerated upwards.

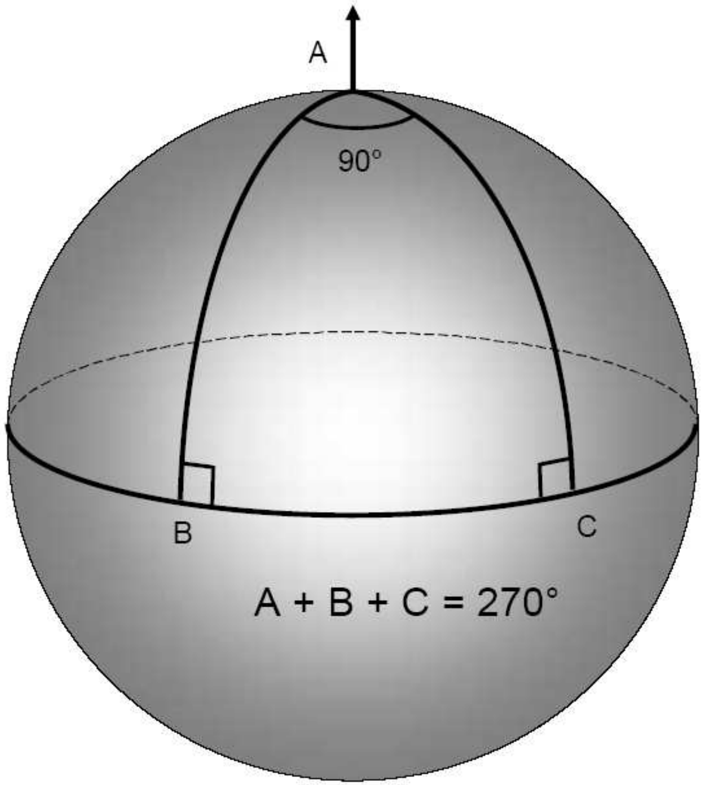
Einstein's great insight in developing General Relativity was that space-time could not be flat if the Principle of Equivalence was to hold good.

In the time the light ray propagates across the lift, a distance l, the lift moves upwards a distance $\frac{1}{2}|\boldsymbol{g}|t^2$. Therefore, in the frame of reference of the accelerated lift, and also in the stationary frame in the gravitational field, the light ray follows a parabolic path. Approximating the light path by a circular arc of radius R, it is straightforward to show that

$$R = \frac{2l^2}{|\boldsymbol{g}|t^2} = \frac{2c^2}{|\boldsymbol{g}|} \; . \qquad (1)$$

The radius of curvature of the path of the light ray depends only upon the local gravitational acceleration $|\boldsymbol{g}|$.

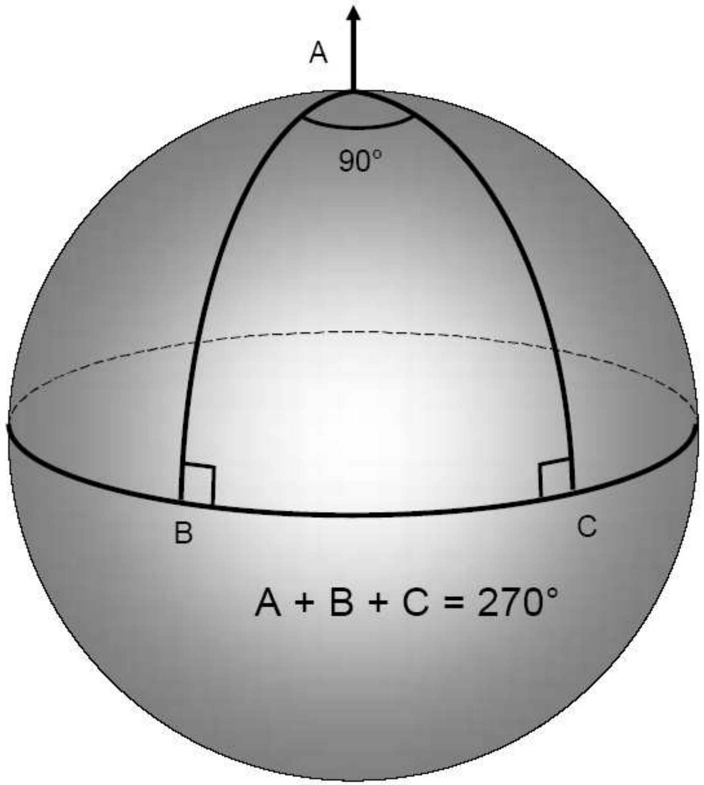
Isotropic Curved Spaces



Consider first the simplest two-dimensional curved geometry, the surface of a sphere. The three sides of this triangle are all segments of great circles on the sphere and so are the shortest distances between the three corners of the triangle. The three lines are geodesics in the curved geometry.

We need a procedure for working out how non-Euclidean the curved geometry is. The way this is done in general is by the procedure known as the parallel displacement or parallel transport of a vector on making a complete circuit around a closed figure such as the triangle. The total rotation of the vector is 270°. Clearly, the surface of the sphere is a non-Euclidean space.

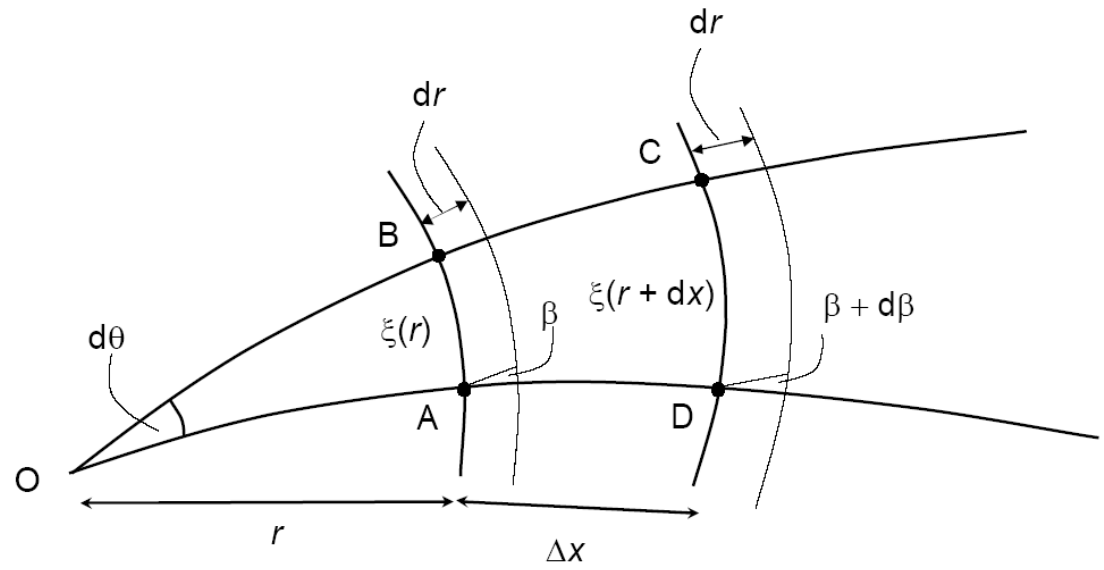
Isotropic Curved Spaces



This procedure illustrates how we can work out the geometrical properties of any two-space, entirely by making measurements within the two-space. Suppose the angle at A was not 90° but some arbitrary angle θ. Then, if the radius of the sphere is R_{C}, the surface area of the triangle ABC is $A = \theta R_{\mathrm{C}}^2$. Thus, if $\theta = 90^\circ$, the area is $\pi R_{\mathrm{C}}^2/2$ and the sum of the angles of the triangle is 270°; if $\theta = 0^\circ$, the area is zero and the sum of the angles of the triangle is 180°. The difference of the sum of the angles of the triangle from 180° is proportional to the area of the triangle, that is

$$(\text{Sum of angles of triangle} - 180^\circ) \propto (\text{Area of triangle})\ .$$

Isotropic Curved Spaces

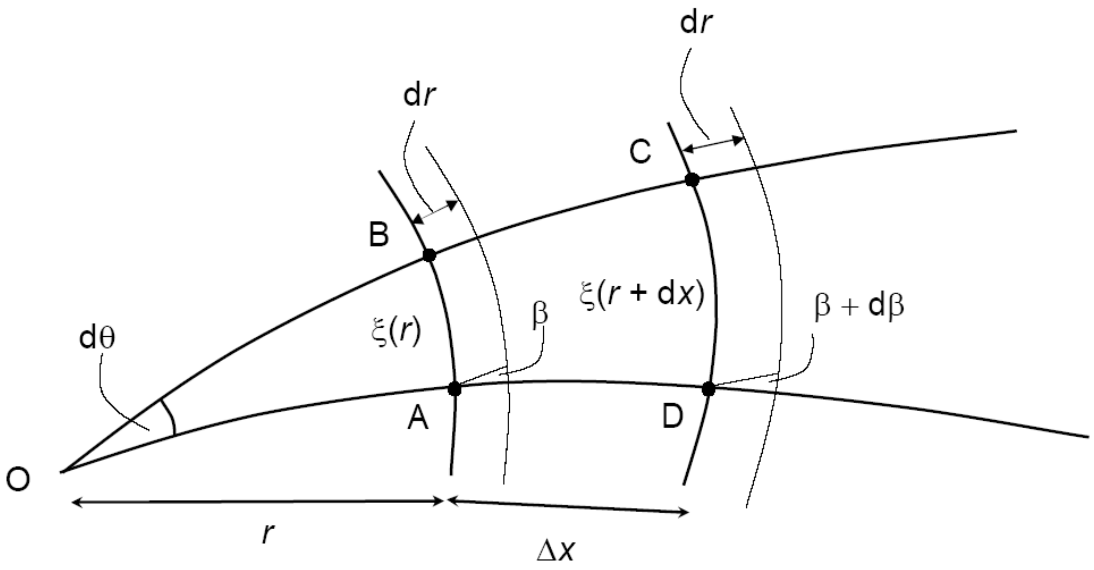


We now work out the sum of the angles round a closed figure in an isotropic curved space. The procedure is shown schematically in the diagram which shows two geodesics from the origin at O being crossed by another pair of geodesics at distances r and $r + \Delta x$ from the origin. The angle $\mathrm{d}\theta$ between the geodesics at O is assumed to be small. In Euclidean space, the length of the segment of the geodesic AB would be $\xi = r\,\mathrm{d}\theta$. However, this is no longer true in non-Euclidean space and instead, we write

$$\xi(r) = f(r)\,\mathrm{d}\theta\ . \qquad (2)$$

We now to work out the angle between the diverging geodesics at distance r from the origin.

Isotropic Curved Spaces



It can be seen that the angle between the geodesics is

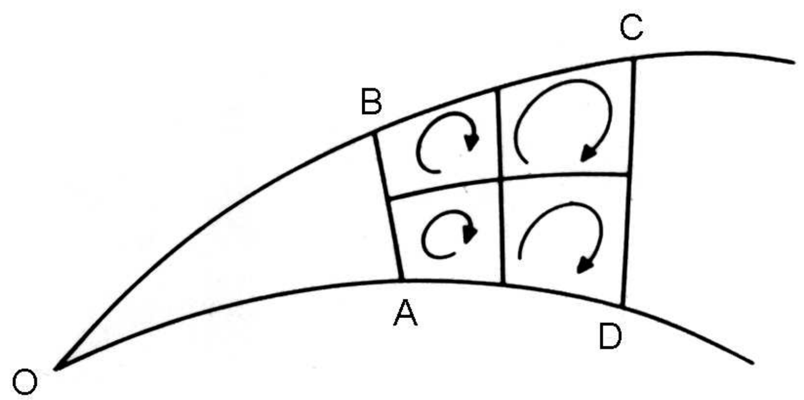
$$\beta = \frac{\xi(r + \mathrm{d}r) - \xi(r)}{\mathrm{d}r} = \frac{\mathrm{d}\xi(r)}{\mathrm{d}r} = \mathrm{d}\theta \frac{\mathrm{d}f(r)}{\mathrm{d}r} \,. \tag{3}$$

Now move a distance Δx further along the geodesics. The change in the angle β, $\Delta\beta$ is

$$\Delta\beta = \frac{\mathrm{d}\xi(r + \Delta x)}{\mathrm{d}r} - \frac{\mathrm{d}\xi(r)}{\mathrm{d}r} = \frac{\mathrm{d}^2\xi(r)}{\mathrm{d}r^2}\Delta x = \frac{\mathrm{d}^2 f(r)}{\mathrm{d}r^2}\Delta x \,\mathrm{d}\theta \,. \tag{4}$$

Isotropic Curved Spaces

In Euclidean space, $\xi(r) = f(r)\,\mathrm{d}\theta = r\,\mathrm{d}\theta$, $f(r) = r$ and hence (4) becomes $\beta = \mathrm{d}\theta$. Furthermore, in Euclidean space, $\mathrm{d}^2 f(r)/\mathrm{d}r^2 = 0$ and so $\Delta\beta = 0$, $\beta = \mathrm{d}\theta$ remains true for all values of r.



Now, the rotation of the vector $\mathrm{d}\beta$ depends upon the area of the quadrilateral ABCD. In the case of an isotropic space, we should obtain the same rotation wherever we place the loop in the two-space. Furthermore, if we were to split the loop up into a number of sub-loops, the rotations around the separate sub-loops must add up linearly to the total rotation $\mathrm{d}\beta$. Thus, in an isotropic two-space, the rotation $\mathrm{d}\beta$ should be proportional to the area of the loop ABCD and must be a constant everywhere in the two-space, just as we found in the particular case of a spherical surface.

Isotropic Curved Spaces

The area of the loop is $\mathrm{d}A = \xi(r)\Delta x = f(r)\Delta x\,\mathrm{d}\theta$, and so we can write

$$\frac{\mathrm{d}^2 f(r)}{\mathrm{d}r^2} = -\kappa\, f(r)\ , \tag{5}$$

where κ is a constant, the minus sign being chosen for convenience. This is the equation of simple harmonic motion which has solution

$$f(r) = A \sin \kappa^{1/2} r\ . \tag{6}$$

We find the value of A from the expression for $\xi(r)$ for small values of r, which must reduce to the Euclidean expression $\mathrm{d}\theta = \xi / r$. Therefore, $A = \kappa^{-1/2}$ and

$$\boxed{f(r) = \frac{\sin \kappa^{1/2} r}{\kappa^{1/2}}\ .} \tag{7}$$

κ is the curvature of the two-space and can be positive, negative or zero. If it is negative, we can write $\kappa = -\kappa'$, where κ' is positive and then the circular functions become hyperbolic functions

$$\boxed{f(r) = \frac{\sinh \kappa'^{1/2} r}{\kappa'^{1/2}}\ .} \tag{8}$$

Isotropic Curved Spaces

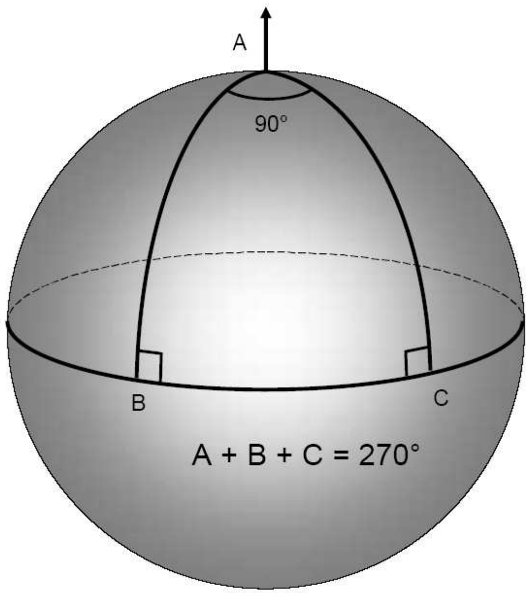
In the Euclidean case, $\mathrm{d}^2 f(r)/\mathrm{d}r^2 = 0$ and so $\kappa = 0$.

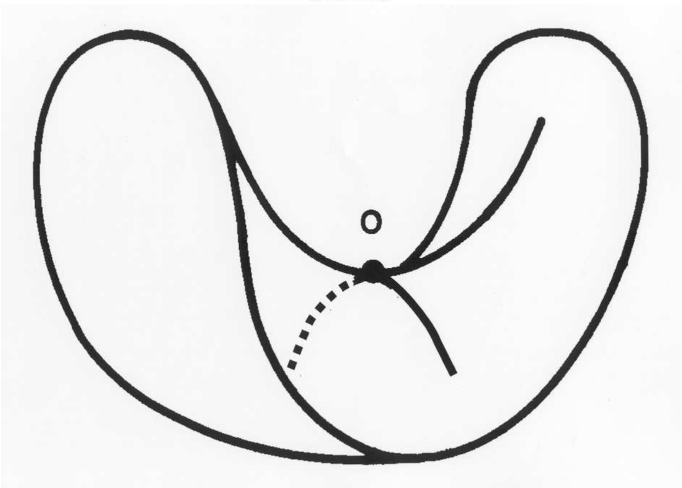
The results we have derived include all possible isotropic curved two-spaces. The constant κ can be positive, negative or zero corresponding to spherical, hyperbolic and flat spaces respectively. In geometric terms, $R_{\mathrm{c}} = \kappa^{-1/2}$ is the radius of curvature of a two-dimensional section through the isotropic curved space and has the same value at all points and in all orientations within the plane. It is often convenient to write the expression for $f(r)$ in the form

$$f(r) = R_{\mathrm{c}} \sin \frac{r}{R_{\mathrm{c}}} , \tag{9}$$

where R_{c} is real for closed spherical geometries, imaginary for open hyperbolic geometries and infinite for the case of Euclidean geometry.

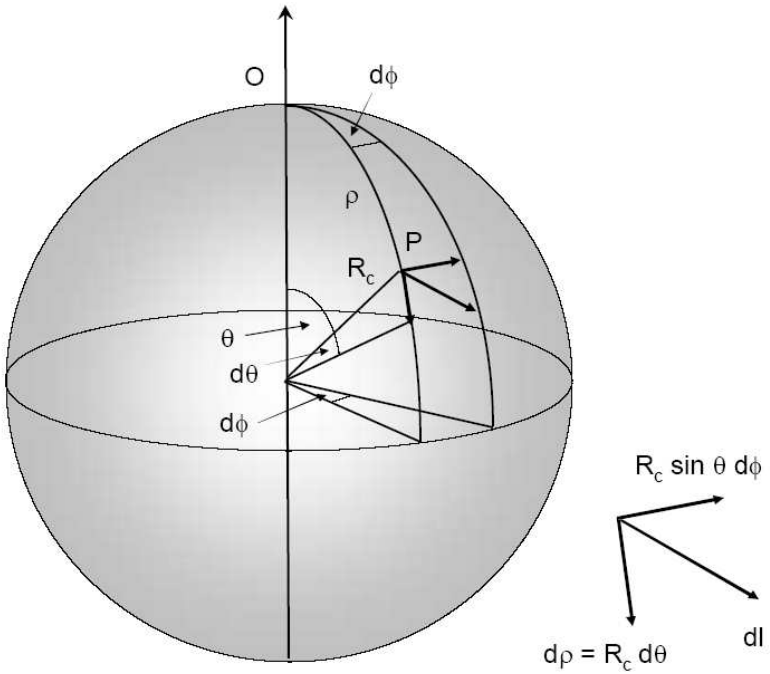
Isotropic Curved Spaces





The simplest examples of such spaces are the spherical geometries in which $R_{\rm C}$ is just the radius of the sphere. The hyperbolic spaces are more difficult to envisage. The fact that $R_{\rm C}$ is imaginary can be interpreted in terms of the principal radii of curvature of the surface having opposite sign. The geometry of a hyperbolic two-sphere can be represented by a saddle-shaped figure, just as a two-sphere provides an visualisation of the properties of a spherical two-space.

The Space-time Metric for Isotropic Curved Spaces

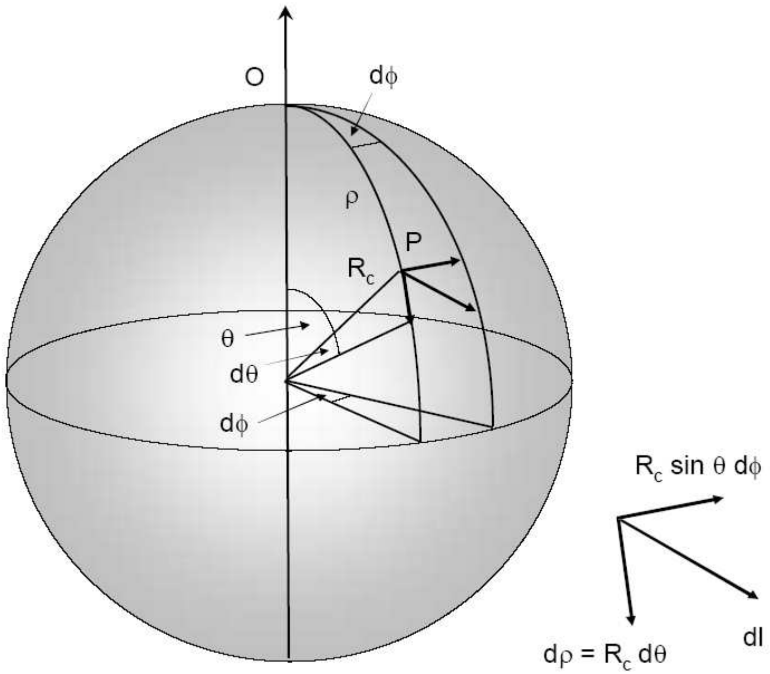


In flat space, the distance between two points separated by dx, dy, dz is

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2. \qquad (10)$$

Let us now consider the simplest example of an isotropic *two-dimensional* curved space, namely the surface of a sphere. We can set up an orthogonal frame of reference at each point locally on the surface of the sphere. It is convenient to work in spherical polar coordinates to describe positions on the surface of the sphere.

The Space-time Metric for Isotropic Curved Spaces



In this case, the orthogonal coordinates are the angular coordinates θ and ϕ, and the expression for the increment of distance $\mathrm{d}l$ between two neighbouring points on the surface can be written

$$\mathrm{d}l^2 = R_\mathrm{c}^2\,\mathrm{d}\theta^2 + R_\mathrm{c}^2 \sin^2\theta\,\mathrm{d}\phi^2, \quad (11)$$

where R_c is the radius of curvature of the two-space, which in this case is just the radius of the sphere.

The Space-time Metric for Isotropic Curved Spaces

The expression (11) is known as the *metric* of the two-dimensional surface and can be written more generally in tensor form

$$\mathrm{d}l^2 = g_{\mu\nu}\,\mathrm{d}x^\mu \mathrm{d}x^\nu. \tag{12}$$

It is a fundamental result of differential geometry that the *metric tensor* $g_{\mu\nu}$ contains all the information about the intrinsic geometry of the space. The problem is that we can set up a variety of different coordinate systems to define the coordinates of a point on any two dimensional surface. For example, in the case of a Euclidean plane, we could use rectangular *Cartesian coordinates* so that

$$\mathrm{d}l^2 = \mathrm{d}x^2 + \mathrm{d}y^2, \tag{13}$$

or we could use *polar coordinates* in which

$$\mathrm{d}l^2 = \mathrm{d}r^2 + r^2\mathrm{d}\phi^2. \tag{14}$$

The Space-time Metric for Isotropic Curved Spaces

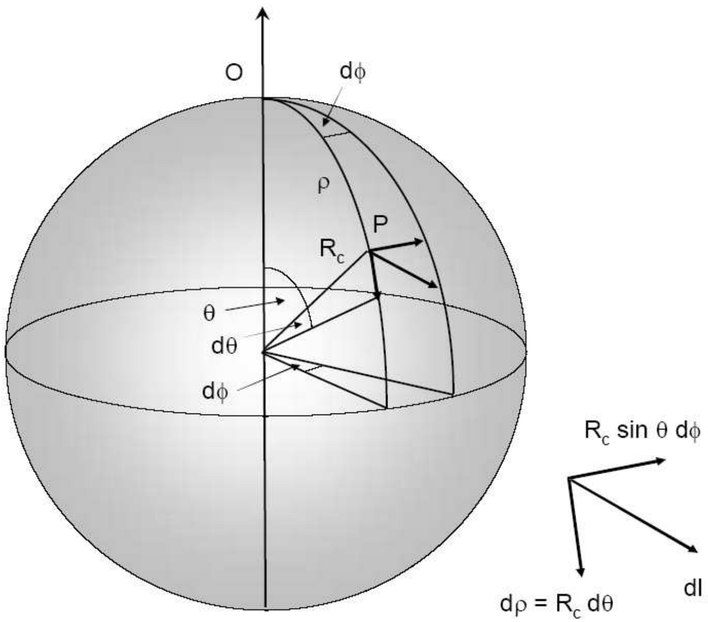
How do we determine the *intrinsic curvature* of the space in terms of the $g_{\mu\nu}$ of the metric tensor? For the case of two-dimensional metric tensors which can be reduced to diagonal form, the intrinsic curvature of the space is given by the quantity

$$\begin{aligned}\kappa &= \frac{1}{2g_{11}g_{22}}\left\{-\frac{\partial^2 g_{11}}{\partial x_2^2}-\frac{\partial^2 g_{22}}{\partial x_1^2}+\frac{1}{2g_{11}}\left[\frac{\partial g_{11}}{\partial x_1}\frac{\partial g_{22}}{\partial x_1}+\left(\frac{\partial g_{11}}{\partial x_2}\right)^2\right]\right.\\ &+ \left.\frac{1}{2g_{22}}\left[\frac{\partial g_{11}}{\partial x_2}\frac{\partial g_{22}}{\partial x_2}+\left(\frac{\partial g_{22}}{\partial x_1}\right)^2\right]\right\}. \qquad (15)\end{aligned}$$

We can use (15) to show that metrics (13) and (14) have zero curvature and that, for the surface of a sphere, the metric (11) has positive curvature with $\kappa = R_{\mathrm{c}}^{-2}$ at all points on the sphere. κ is known as the *Gaussian curvature* of the two-space and is the same as the definition of the curvature we have already introduced. In general curved spaces, the curvature κ varies from point to point in the space.

The extension to isotropic three-spaces is straightforward if we remember that any two-dimensional section through an isotropic three-space must be an isotropic two-space and we already know the metric tensor for this case.

The Space-time Metric for Isotropic Curved Spaces



The natural system of coordinates for an isotropic two-space is a spherical polar system in which a radial distance ϱ round the sphere is measured from the pole and the angle ϕ measures angular displacements at the pole.

The distance ϱ round the arc of a great circle from the point O to P is $\varrho = \theta R_c$ and so the metric can be written

$$dl^2 = d\varrho^2 + R_c^2 \sin^2\left(\frac{\varrho}{R_c}\right) d\phi^2 . \quad (16)$$

The distance ϱ is the shortest distance between O and P on the surface of the sphere since it is part of a great circle and is therefore the *geodesic distance* between O and P in the isotropic curved space. Geodesics play the role of straight lines in curved space.

The Space-time Metric for Isotropic Curved Spaces

We can write the metric in an alternative form if we introduce a distance measure

$$x = R_{\rm C} \sin\left(\frac{\varrho}{R_{\rm C}}\right) . \tag{17}$$

Differentiating and squaring, we find

$$dx^2 = \left[1 - \sin^2\left(\frac{\varrho}{R_{\rm C}}\right)\right] d\varrho^2 \qquad d\varrho^2 = \frac{dx^2}{1 - \kappa x^2} , \tag{18}$$

where $\kappa = 1/R_{\rm C}^2$ is the curvature of the two space.

Therefore, we can rewrite the metric in the form

$$dl^2 = \frac{dx^2}{1 - \kappa x^2} + x^2 d\phi^2 . \tag{19}$$

From the metric (19) $dl = x\, d\phi$ is a proper dimension perpendicular to the radial coordinate ϱ and that it is the correct expression for the length of a line segment which subtends the angle $d\phi$ at geodesic distance ϱ from O. It is therefore what is known as an angular diameter distance since it is guaranteed to give the correct answer for the length of a line segment perpendicular to the line of sight.

The Space-time Metric for Isotropic Curved Spaces

We can use either ϱ or x in our metric but, if we use x, the increment of geodesic distance is $\mathrm{d}\varrho = \mathrm{d}x/(1-\kappa x^2)^{1/2}$. We recall that the curvature $\kappa = 1/R_{\mathrm{c}}^2$ can be *positive* as in the spherical two-space discussed above, zero in which case we recover flat Euclidean space ($R_{\mathrm{c}} \to \infty$) and negative in which case the geometry becomes hyperbolic rather than spherical.

We now write down the expression for the spatial increment in any isotropic, three-dimensional curved space. Any two-dimensional section through an isotropic three-space must be an isotropic two-space for which the metric is (16) or (19). In spherical polar coordinates, the general angular displacement perpendicular to the radial direction is

$$\mathrm{d}\Phi^2 = \mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\,. \tag{20}$$

Note that the θs and ϕs in (20) are different from those used in the diagram. Thus, by extension of the formalism we have derived already, the spatial increment can be written

$$\mathrm{d}l^2 = \mathrm{d}\varrho^2 + R_{\mathrm{c}}^2\sin^2\left(\frac{\varrho}{R_{\mathrm{c}}}\right)[\mathrm{d}\theta^2 + \sin^2\theta\;\mathrm{d}\phi^2]\,, \tag{21}$$

in terms of the three-dimensional spherical polar coordinates ($\varrho,\ \theta,\ \phi$).

The Space-time Metric for Isotropic Curved Spaces

An exactly equivalent form is obtained if we write the spatial increment in terms of $x,\ \theta,\ \phi$ in which case we find

$$\mathrm{d}l^2 = \frac{\mathrm{d}x^2}{1-\kappa x^2} + x^2[\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2]\ . \tag{22}$$

We are now in a position to write down the *Minkowski metric* in any isotropic three-space. It is given by

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{1}{c^2}\mathrm{d}l^2\ , \tag{23}$$

where $\mathrm{d}l$ is given by either of the above forms of the spatial increment, (21) or (22). Notice that we have to be careful about the meanings of the distance coordinates – x and ϱ are equivalent but physically quite distinct distance measures. We can now proceed to derive from this metric the *Robertson–Walker metric*.

The Robertson–Walker Metric

In order to apply the metric (23) to isotropic, homogeneous world models, we need the *cosmological principle* and the concepts of *fundamental observers* and *cosmic time*.

- For uniform, isotropic world models, we define a set of *fundamental observers*, who move in such a way that the Universe always appears to be isotropic to them.
- *Cosmic time* is time measured on the clock of a fundamental observer.

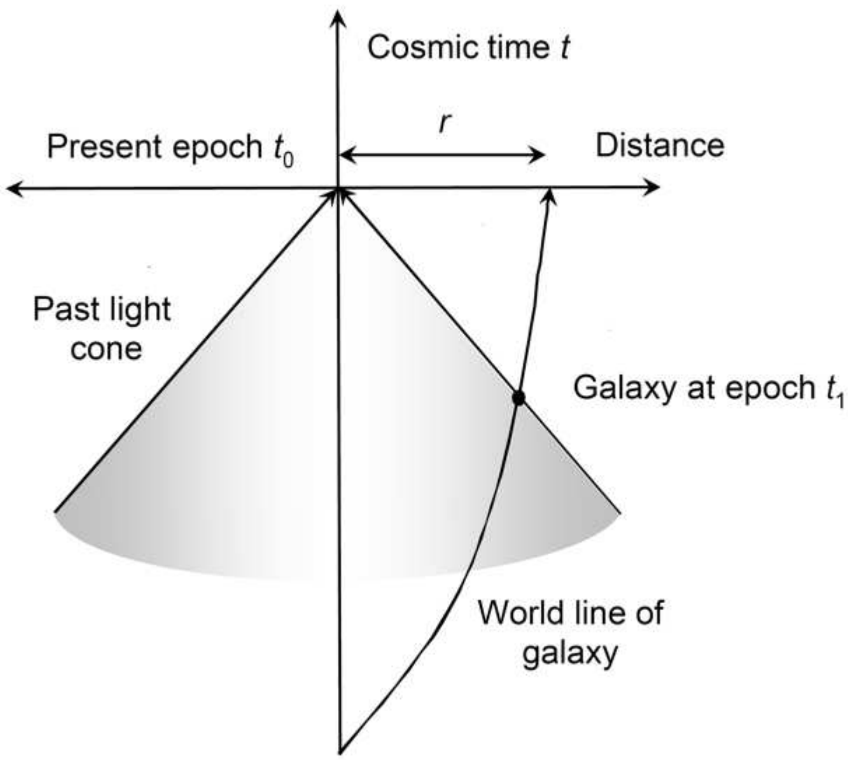
There are no problems of synchronisation of the clocks carried by the fundamental observers because, according to Weyl's postulate, the geodesics of all observers meet at one point in the past and cosmic time can be measured from that reference epoch.

From (21) and (23), the metric can be written in the form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{1}{c^2}[\mathrm{d}\varrho^2 + R_\mathrm{c}^2 \sin^2(\varrho/R_\mathrm{c})(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]\,. \tag{24}$$

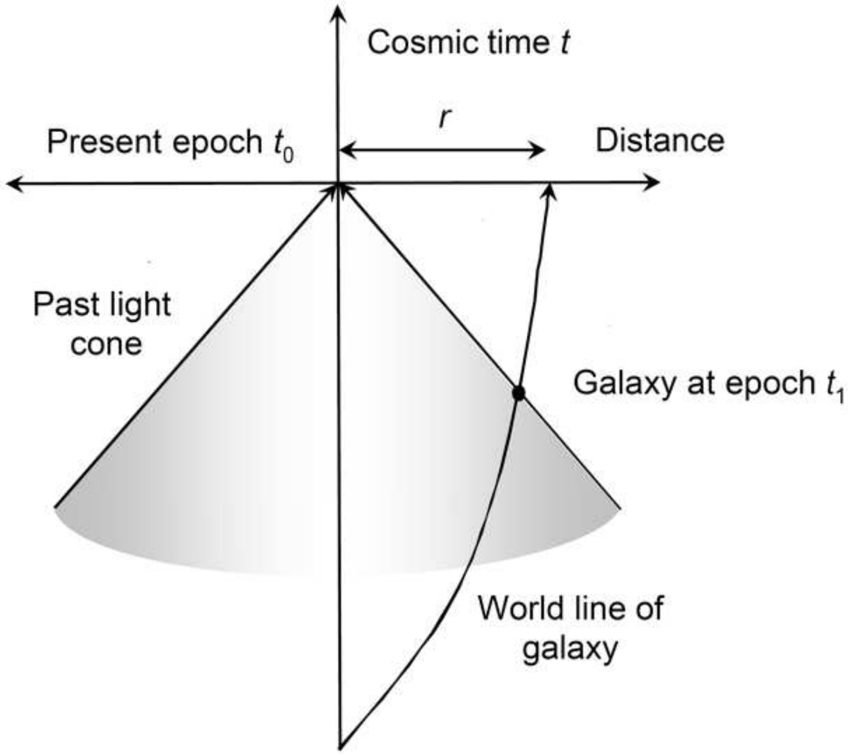
t is cosmic time and $\mathrm{d}\varrho$ is an increment of proper distance in the radial direction.

The Robertson-Walker Metric



There is a problem in applying the metric to the expanding Universe as is illustrated by the space-time diagram. Since light travels at a finite speed, we observe all astronomical objects along a *past light cone* which is centred on the Earth at the present epoch t_0. Therefore, when we observe distant objects, we do not observe them at the present epoch but rather at an earlier epoch t_1 when the distances between fundamental observers were smaller and the spatial curvature different. The problem is that we can only apply the metric (24) to an isotropic curved space defined *at a single epoch*.

The Robertson-Walker Metric



To resolve this problem, we perform the following thought experiment. To measure a proper distance which can be included in the metric (24), we line up a set of fundamental observers between the Earth and the galaxy whose distance we wish to measure. The observers are instructed to measure the distance dϱ to the next fundamental observer at a particular cosmic time t. By adding together all the dϱs, we can find a proper distance ϱ which is measured *at a single epoch* and which can be used in the metric (24). Notice that ϱ is a *fictitious distance* since we do not know how to project their positions to the present epoch until we know the kinematics of the expanding Universe. Thus, *the distance measure ϱ depends upon the choice of cosmological model*.

The Comoving Distance Coordinate

The definition of a uniform expansion is that between two cosmic epochs, t_1 and t_2, the distances of any two fundamental observers, i and j, change such that

$$\frac{\varrho_i(t_1)}{\varrho_j(t_1)} = \frac{\varrho_i(t_2)}{\varrho_j(t_2)} = \text{constant} \, , \tag{25}$$

that is,

$$\frac{\varrho_i(t_1)}{\varrho_i(t_2)} = \frac{\varrho_j(t_1)}{\varrho_j(t_2)} = \ldots = \text{constant} = \frac{a(t_1)}{a(t_2)} \, . \tag{26}$$

For isotropic world models, $a(t)$ is a universal function known as the *scale factor* which describes how the relative distances between *any* two fundamental observers change with cosmic time t. We set $a(t)$ equal to 1 at the present epoch t_0 and let the value of ϱ at the present epoch be r, that is, we can rewrite (26) as

$$\varrho(t) = a(t) r \, . \tag{27}$$

r thus becomes a *distance label* which is attached to a galaxy or fundamental observer *for all time* and the variation in proper distance in the expanding Universe is taken care of by the scale factor $a(t)$; r is called the *comoving radial distance coordinate*.

The Comoving Distance Coordinate

Proper distances perpendicular to the line of sight must also change by a factor a between the epochs t and t_0.

$$\frac{\Delta l(t)}{\Delta l(t_0)} = a(t)\ . \tag{28}$$

From the metric (24),

$$a(t) = \frac{R_{\mathrm{c}}(t)\,\sin\left[\varrho/R_{\mathrm{c}}(t)\right]\,\mathrm{d}\theta}{R_{\mathrm{c}}(t_0)\,\sin[r/R_{\mathrm{c}}(t_0)]\,\mathrm{d}\theta}\ . \tag{29}$$

Reorganising this equation and using (27),

$$\frac{R_{\mathrm{c}}(t)}{a(t)}\sin\left[\frac{a(t)r}{R_{\mathrm{c}}(t)}\right] = R_{\mathrm{c}}(t_0)\sin\left[\frac{r}{R_{\mathrm{c}}(t_0)}\right]\ . \tag{30}$$

This is only true if

$$R_{\mathrm{c}}(t) = a(t)\,R_{\mathrm{c}}(t_0)\ , \tag{31}$$

that is, the radius of curvature of the spatial sections is proportional to the scale factor $a(t)$. Thus, in order to preserve isotropy and homogeneity, *the curvature of space changes as the Universe expands as* $\kappa = R_{\mathrm{c}}^{-2} \propto a^{-2}$. κ cannot change sign and so, if the geometry of the Universe was once, say, hyperbolic, it will always remain so.

The Robertson-Walker Metric

Let us call the value of $R_\mathrm{c}(t_0)$, that is, the radius of curvature of the spatial geometry at the present epoch, $\Re$. Then

$$R_\mathrm{c}(t) = a(t)\,\Re\,. \tag{32}$$

Substituting (27) and (32) into the metric (24), we obtain

$$\boxed{\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a^2(t)}{c^2}[\mathrm{d}r^2 + \Re^2\sin^2(r/\Re)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]\,.} \tag{33}$$

This is the *Robertson–Walker metric* in the form we will use in much of our future analysis. Notice that it contains one unknown function $a(t)$, the scale factor, which describes the dynamics of the Universe and an unknown constant $\Re$ which describes the spatial curvature of the Universe at the present epoch.

The metric can be written in different ways. For example, if we use a *comoving angular diameter distance* $r_1 = \Re\sin(r/\Re)$, the metric becomes

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a^2(t)}{c^2}\left[\frac{\mathrm{d}r_1^2}{1-\kappa r_1^2} + r_1^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right], \tag{34}$$

where $\kappa = 1/\Re^2$.

The Robertson-Walker Metric

By a suitable rescaling of the r_1 coordinate $\kappa r_1^2 = r_2^2$, the metric can equally well be written

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{R_1^2(t)}{c^2}\left[\frac{\mathrm{d}r_2^2}{1 - kr_2^2} + r_2^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right], \qquad (35)$$

with $k = +1, 0$ and -1 for universes with spherical, flat and hyperbolic geometries respectively. Notice that, in this rescaling, the value of $R_1(t) = R_\mathrm{c}(t_0)a = \Re a$ and so the value of $R_1(t)$ at the present epoch is $\Re$ rather than unity. This is a popular form for the metric, but I will normally use (33) because the r coordinate has an obvious and important physical meaning.

The importance of the metrics (33), (34) and (35) is that they enable us to define the invariant interval $\mathrm{d}s^2$ between events at any epoch or location in the expanding Universe.

The Robertson-Walker Metric

To summarise, the *Robertson-Walker metric* can be written in the following form:

$$\mathrm{d}s^2 = \mathrm{d}t^2 - \frac{a^2(t)}{c^2}[\mathrm{d}r^2 + \Re^2 \sin^2(r/\Re)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]\ .$$

The metric contains one unknown function $a(t)$, the scale factor, and the constant $\Re$ which is the radius of curvature of the geometry of the Universe at the present epoch.

- t is cosmic time as measured by a clock carried by a fundamental observer;
- r is the *comoving radial distance coordinate* which is fixed to a galaxy for all time.
- $a(t)\,\mathrm{d}r$ is the element of proper (or geodesic) distance in the radial direction at the epoch t;
- $a(t)[\Re\sin(r/\Re)]\,\mathrm{d}\theta$ is the element of proper distance perpendicular to the radial direction subtended by the angle dθ at the origin;
- Similarly, $a(t)[\Re\sin(r/\Re)]\sin\theta\,\mathrm{d}\phi$ is the element of proper distance in the ϕ-direction.

The Cosmological Redshift

By cosmological redshift, we mean the shift of spectral lines to longer wavelengths associated with the isotropic expansion of the system of galaxies. If λ_{e} is the wavelength of the line as emitted and λ_0 the observed wavelength, the redshift z is defined to be

$$z = \frac{\lambda_0 - \lambda_{\mathrm{e}}}{\lambda_{\mathrm{e}}} \,. \tag{36}$$

If the redshift z were interpreted as the recession velocity v of a galaxy, these would be related by the Newtonian Doppler shift formula

$$v = cz \,. \tag{37}$$

This is the type of velocity which Hubble used in deriving the velocity–distance relation, $v = H_0 r$. It is however incorrect to use the special relativistic Doppler shift formula

$$1 + z = \left(\frac{1 + v/c}{1 - v/c}\right)^{1/2} \,, \tag{38}$$

at large redshifts. Rather, because of the requirements of isotropy and homogeneity, the relation $v \propto r$ applies at all distances, including those at which the recession velocity would exceed the speed of light.

The Real Meaning of Redshift

Consider a wave packet of frequency ν_1 emitted between cosmic times t_1 and $t_1 + \Delta t_1$ from a distant galaxy. This wave packet is received by an observer at the present epoch in the interval of cosmic time t_0 to $t_0 + \Delta t_0$. The signal propagates along null-cones, $\mathrm{d}s^2 = 0$, and so, considering radial propagation from source to observer, $\mathrm{d}\theta = 0$ and $\mathrm{d}\phi = 0$, the metric (33) gives us the relation

$$\mathrm{d}t = -\frac{a(t)}{c}\,\mathrm{d}r \qquad \frac{c\,\mathrm{d}t}{a(t)} = -\mathrm{d}r\,. \tag{39}$$

$a(t)\,\mathrm{d}r$ is simply the interval of proper distance at cosmic time t. The minus sign appears because the origin of the r coordinate is the observer at $t = t_0$. Considering first the leading edge of the wave packet, the integral of (39) is

$$\int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = -\int_r^0 \mathrm{d}r\,. \tag{40}$$

The end of the wave packet must travel the same distance in units of comoving distance coordinate since the r coordinate is fixed to the galaxy for all time. Therefore,

$$\int_{t_1+\Delta t_1}^{t_0+\Delta t_0} \frac{c\,\mathrm{d}t}{a(t)} = -\int_r^0 \mathrm{d}r\,, \tag{41}$$

The Meaning of Redshift

Therefore,

$$\int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} + \frac{c\,\Delta t_0}{a(t_0)} - \frac{c\,\Delta t_1}{a(t_1)} = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} \, . \tag{42}$$

Since $a(t_0) = 1$, we find that

$$\Delta t_0 = \Delta t_1 / a(t_1) \, . \tag{43}$$

This is the cosmological expression for the phenomenon of *time dilation*. Distant galaxies are observed at an earlier cosmic time t_1 when $a(t_1) < 1$ and so phenomena are observed to take longer in our frame of reference than in that of the source.

Expression (43) also provides an expression for the redshift. If $\Delta t_1 = \nu_1^{-1}$ is the period of the emitted waves and $\Delta t_0 = \nu_0^{-1}$ that of observed waves, $\nu_0 = \nu_1 a(t_1)$. In terms of redshift z,

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} = \frac{\lambda_0}{\lambda_e} - 1 = \frac{\nu_1}{\nu_0} - 1 \, , \tag{44}$$

$$\boxed{a(t_1) = \frac{1}{1+z} \, .} \tag{45}$$

Einstein's Field Equations

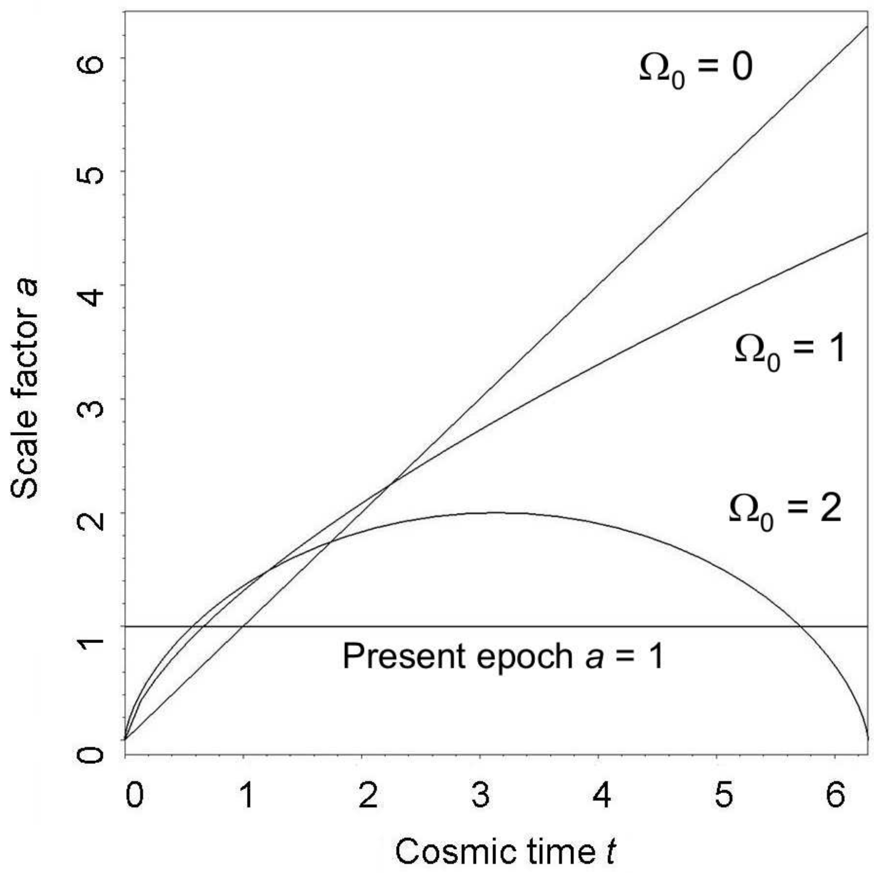
Under the assumption of isotropy and homogeneity, Einstein's field equations of General Relativity reduce to the following pair of independent equations.

$$\ddot{a} = -\frac{4\pi G}{3} a \left(\varrho + \frac{3p}{c^2} \right) + \tfrac{1}{3}\Lambda a \; ; \qquad (46)$$

$$\dot{a}^2 = \frac{8\pi G \varrho}{3} a^2 - \frac{c^2}{\Re^2} + \tfrac{1}{3}\Lambda a^2 \; . \qquad (47)$$

a is the scale factor, ϱ is the total inertial mass density of the matter and radiation content of the Universe and p the associated total pressure. $\Re$ is the radius of curvature of the spatial geometry of the world model at the present epoch and so the term $-c^2/\Re^2$ is a constant of integration. The *cosmological constant* Λ has had a chequered history since it was introduced by Einstein in 1917.

The Critical Model and the Critical Density



The critical world model has $\Omega_0 = 1$, $\Omega_\Lambda = 0$ and separates the open from the closed models and the collapsing models from those which expand forever for the case $\Omega_\Lambda = 0$. This model is often referred to as the Einstein–de Sitter or the critical model. The velocity of expansion tends to zero as a tends to infinity. It has a particularly simple variation of $a(t)$ with cosmic epoch,

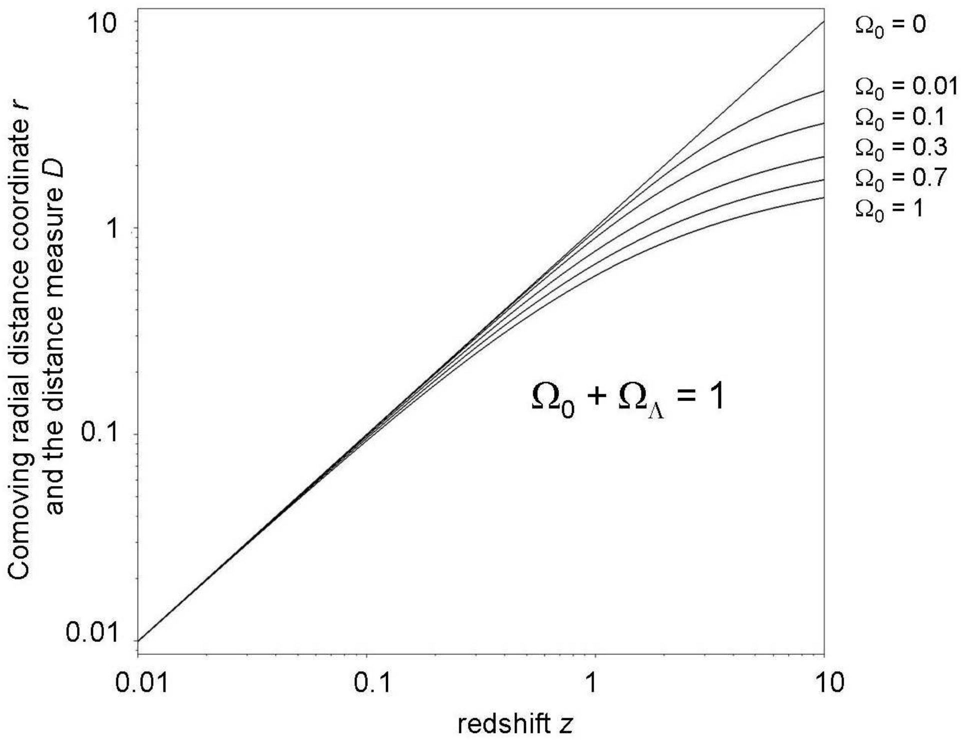
$$a = (t/t_0)^{2/3} \qquad \kappa = 0, \tag{48}$$

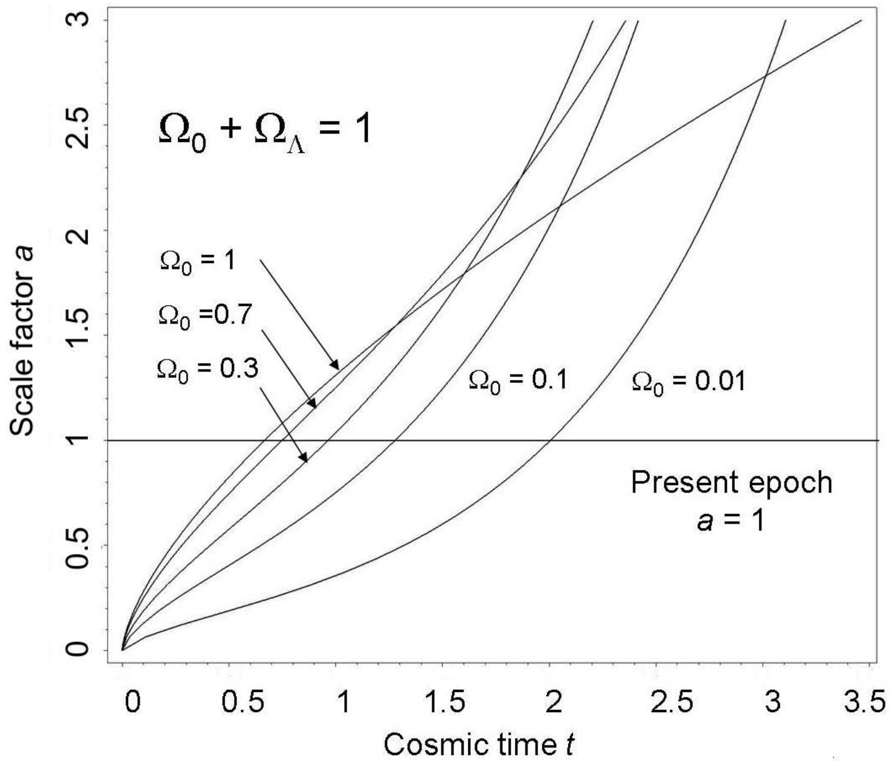
where the present age of the world model is $t_0 = (2/3)H_0^{-1}$. The density of the model $\varrho_0 = 3H_0^2/8\pi G$ is known as the critical density.

The Properties of the Concordance Model

The expression for the comoving radial distance coordinate for $\Omega_0 + \Omega_\Lambda = 1$ is

$$r = \int_{t_1}^{t_0} \frac{c\,\mathrm{d}t}{a} = -\frac{c}{H_0} \int_{\infty}^{z} \frac{\mathrm{d}z}{[\Omega_0(1+z)^3 + \Omega_\Lambda]^{1/2}} \,. \qquad (49)$$





Space-time Diagrams for the Standard World Models

Let us summarise the various times and distances used in cosmological analyses.

Comoving radial distance coordinate In terms of cosmic time and scale factor, the comoving radial distance coordinate r is defined to be

$$r = \int_t^{t_0} \frac{c\,\mathrm{d}t}{a} = \int_a^1 \frac{c\,\mathrm{d}a}{a\dot{a}} \,. \tag{50}$$

Proper radial distance coordinate We run up against the same problems we encountered in defining the comoving radial distance coordinate, in that it only makes sense to define distances at a particular cosmic epoch t. Therefore, we *define* the proper radial distance r_{prop} to be the comoving radial distance coordinate projected back to the epoch t, that is

$$r_{\mathrm{prop}} = a \int_t^{t_0} \frac{c\,\mathrm{d}t}{a} = a \int_a^1 \frac{c\,\mathrm{d}a}{a\dot{a}} \,. \tag{51}$$

Space-time Diagrams for the Standard World Models

Particle horizon The particle horizon r_{H} is defined as the maximum proper distance over which there can be causal communication at the epoch t

$$r_{\mathrm{H}} = a \int_0^t \frac{c\,\mathrm{d}t}{a} = a \int_0^a \frac{c\,\mathrm{d}a}{a\dot{a}} \,. \tag{52}$$

Event horizon The event horizon r_{E} is defined as the greatest proper radial distance an object can have if it is ever to be observable by an observer who observes the Universe at cosmic time t_1.

$$r_{\mathrm{E}} = a \int_{t_1}^{t_{\mathrm{max}}} \frac{c\,\mathrm{d}t}{a(t)} = a \int_{a_1}^{a_{\mathrm{max}}} \frac{c\,\mathrm{d}a}{a\dot{a}} \,. \tag{53}$$

Space-time Diagrams for the Standard World Models

Cosmic time Cosmic time t is defined to be time measured by a fundamental observer who reads time on a standard clock.

$$t = \int_0^t \mathrm{d}t = \int_0^a \frac{\mathrm{d}a}{\dot{a}} \,. \qquad (54)$$

Conformal time The *conformal time* τ is similar to the definition of comoving radial distance coordinate. Time intervals are projected forward to present epoch

$$\mathrm{d}t_{\mathrm{conf}} = \mathrm{d}\tau = \frac{\mathrm{d}t}{a} \,. \qquad (55)$$

Then, the Robertson-Walker metric (33) can be written in a form which makes both the space and time components of the metric change in the same way with cosmic epoch

$$\mathrm{d}s^2 = a^2(t)\left[\mathrm{d}\tau^2 - \frac{1}{c^2}[\mathrm{d}r^2 + \Re^2 \sin^2(r/\Re)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)]\right] \,. \qquad (56)$$

At any epoch, the conformal time has value

$$\tau = \int_0^t \frac{\mathrm{d}t}{a} = \int_0^a \frac{\mathrm{d}a}{a\dot{a}} \,. \qquad (57)$$

It follows from (52) and (57) that, in a space-time diagram in which comoving radial distance coordinate is plotted against conformal time, the particle horizon is a straight line with slope equal to the speed of light.

The Past Light Cone

This topic requires a little care. First, because of the assumptions of isotropy and homogeneity, Hubble's linear relation $v = H_0 r$ applies at the present epoch *to recessions speeds which exceed the speed of light*, where r is the radial comoving distance coordinate. Recall how we defined r. The fundamental observers measured increments of distance Δr at the present epoch t_0. If we consider fundamental observers who are far enough apart, this speed can exceed the speed of light. There is nothing in this argument which contradicts the special theory of relativity – it is a geometric result because of the requirements of isotropy and homogeneity.

Consider the analogue for the expanding Universe of the surface of an expanding spherical balloon. As the balloon inflates, a linear velocity-distance relation is found on the surface of the sphere, not only about any point on the sphere, but also at arbitrarily large distances on its surface. At very large distances, the speed of separation can be greater than the speed of light, but there is no causal connection between these points – they are simply partaking in the uniform expansion of the underlying space-time geometry of the Universe.

The Past Light Cone

Consider the proper distance between two fundamental observers at some epoch t

$$r_{\rm prop} = a(t)r \ , \tag{58}$$

where r is comoving radial distance. Differentiating with respect to cosmic time,

$$\frac{{\rm d}r_{\rm prop}}{{\rm d}t} = \dot{a}r + a\frac{{\rm d}r}{{\rm d}t} \ . \tag{59}$$

The first term on the right-hand side represents the motion of the substratum and, at the present epoch, becomes $H_0 r$. Consider, for example, the case of a very distant object in the critical world model, $\Omega_0 = 1, \Omega_\Lambda = 0$. As a tends to zero, the comoving radial distance coordinates tends to $r = 2c/H_0$. Therefore, the local rest frame of objects at these large distances moves at twice the speed of light relative to our local frame of reference *at the present epoch*. At the epoch at which the light signal was emitted along our past light cone, the recessional velocity of the local rest frame $v_{\rm rec} = \dot{a}r$ was greater than this value, because $\dot{a} \propto a^{-1/2}$.

The Past Light Cone

The second term on the right-hand side of (59) corresponds to the velocity of peculiar motions in the local rest frame at r, since it corresponds to changes of the comoving radial distance coordinate. The element of proper radial distance is $a\,\mathrm{d}r$ and so, if we consider a light wave travelling along our past light cone towards the observer at the origin, we find

$$\boxed{v_{\mathrm{tot}} = \dot{a}r - c\,.} \tag{60}$$

This is the key result which defines the propagation of light from the source to the observer in space-time diagrams for the expanding Universe.

We can now plot the trajectories of light rays from their source to the observer at t_0. The proper distance from the observer at $r = 0$ to the past light cone r_{PLC} is

$$r_{\mathrm{PLC}} = \int_0^t v_{\mathrm{tot}}\,\mathrm{d}t = \int_0^a \frac{v_{\mathrm{tot}}\,\mathrm{d}a}{\dot{a}}\,. \tag{61}$$

The Past Light Cone

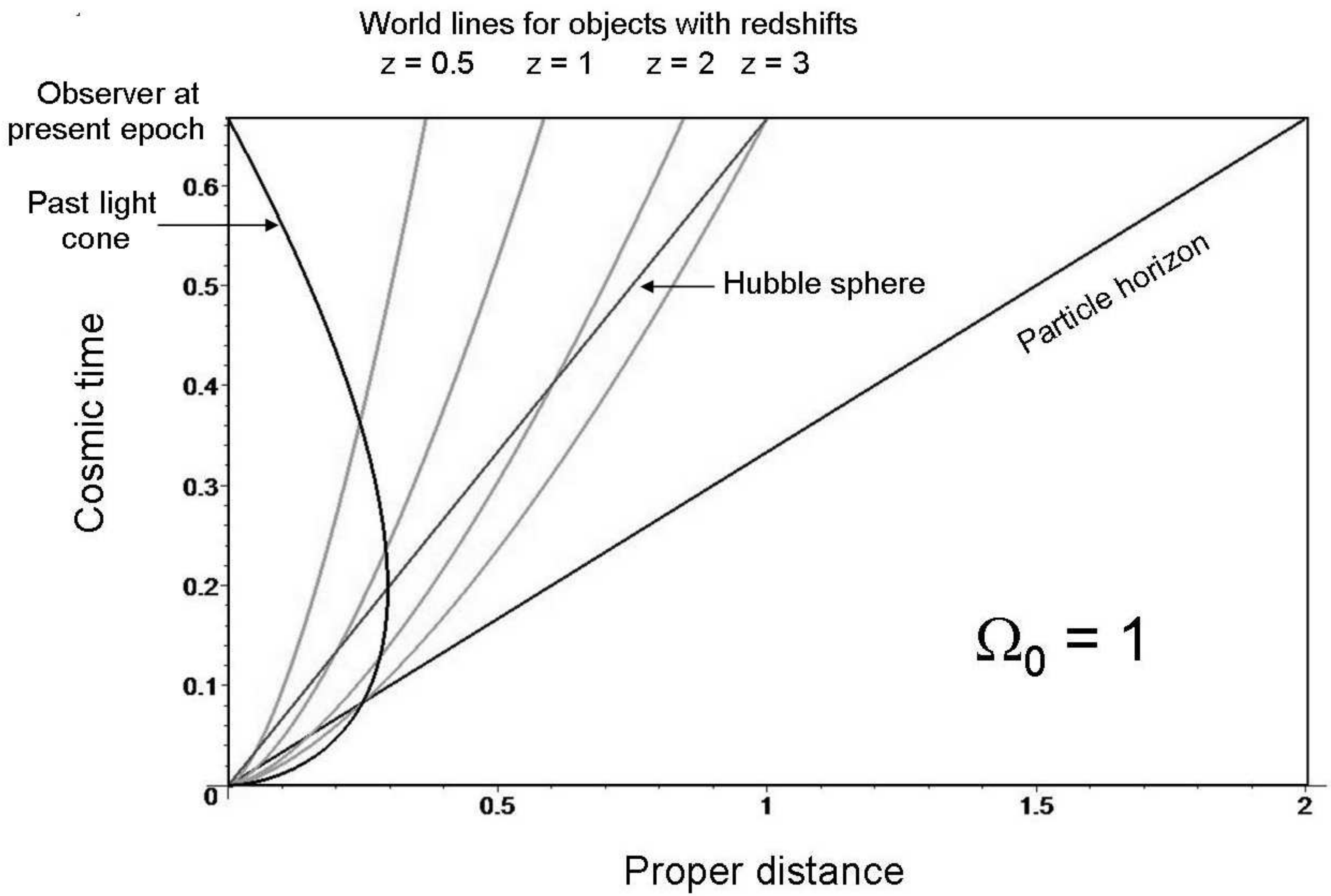
Notice that, initially the light rays from distant objects are propagating away from the observer – this is because the local isotropic cosmological rest frame is moving away from the observer at $r = 0$ at a speed greater than that of light. The light waves are propagated to the observer at the present epoch through local inertial frames which expand with progressively smaller velocities until they cross the *Hubble sphere* at which the recession velocity of the local frame of reference is the speed of light. The definition of the radius of the Hubble sphere r_{HS} at epoch t is thus given by

$$c = H(t)\, r_{\mathrm{HS}} = \frac{\dot{a}}{a} r_{\mathrm{HS}} \qquad \text{or} \qquad r_{\mathrm{HS}} = \frac{ac}{\dot{a}} \, . \tag{62}$$

Note that r_{HS} is a proper radial distance. From this epoch onwards, propagation is towards the observer until, as $t \to t_0$, the speed of propagation towards the observer is the speed of light.

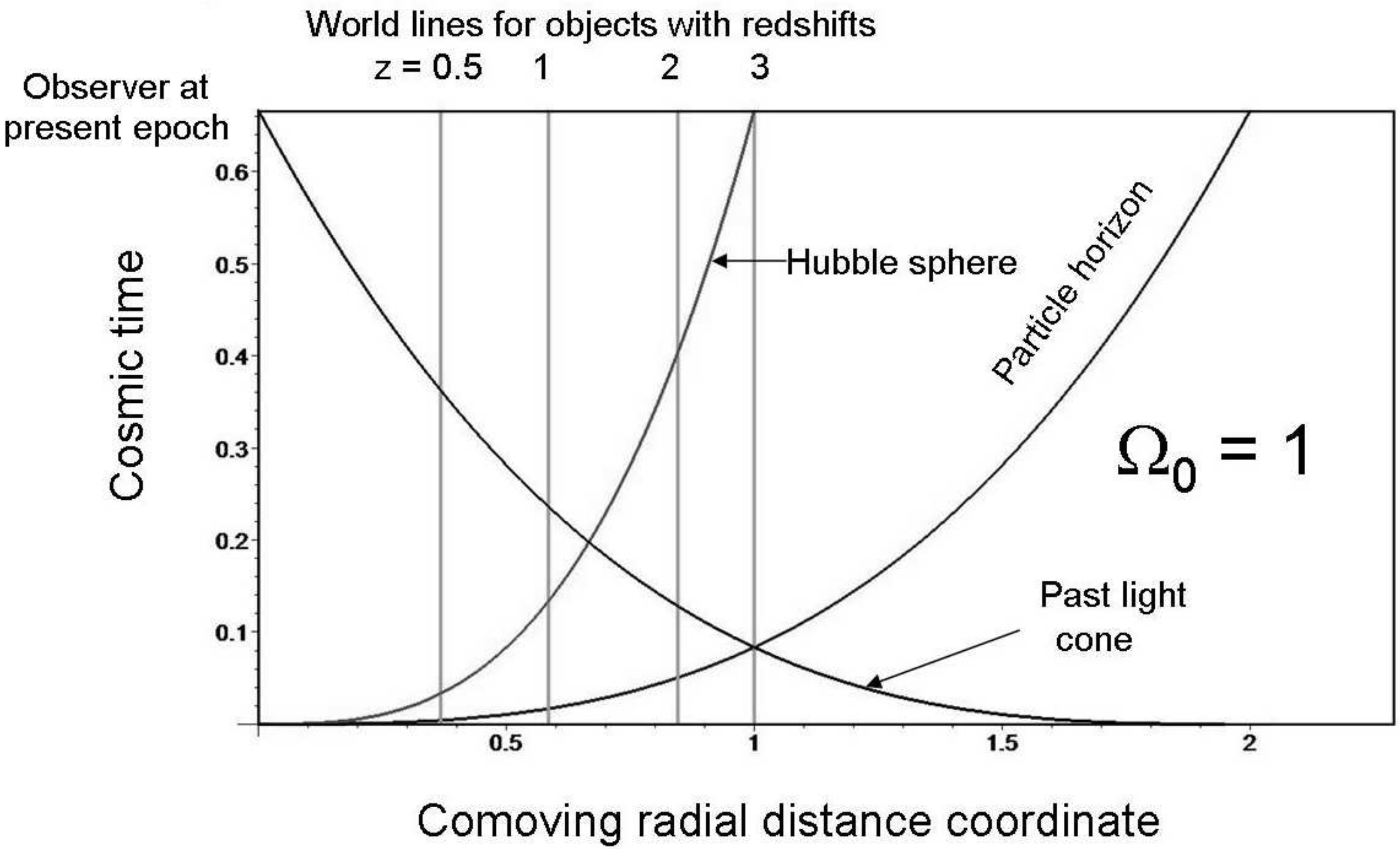
It is simplest to illustrate how the various scales change with time in specific examples of standard cosmological models. We consider first the critical world model and then our reference Λ model.

Space-Time Diagram
Cosmic Time vs. Proper Distance



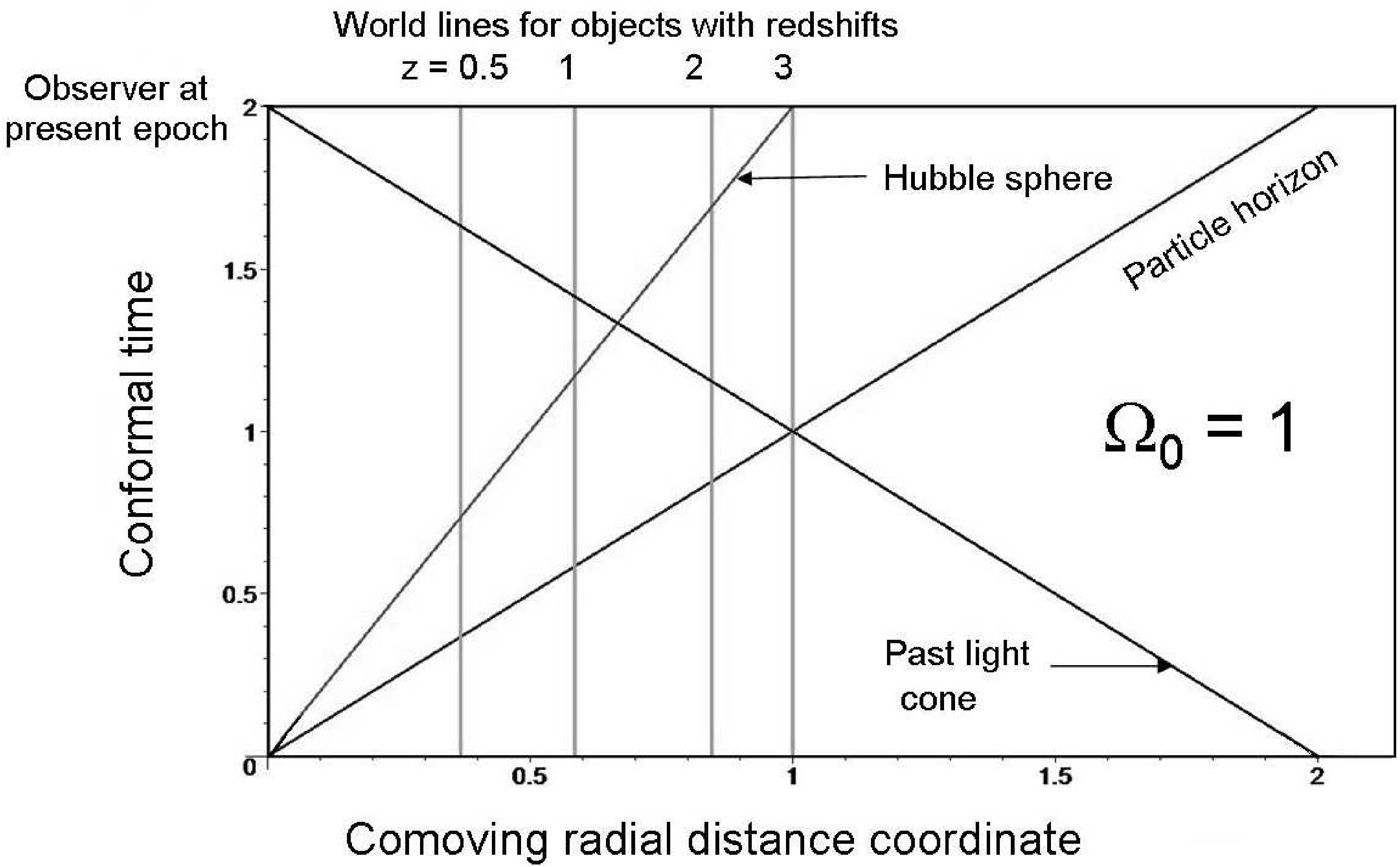
The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

Space-Time Diagram
Cosmic Time vs. Comoving Distance Coordinate



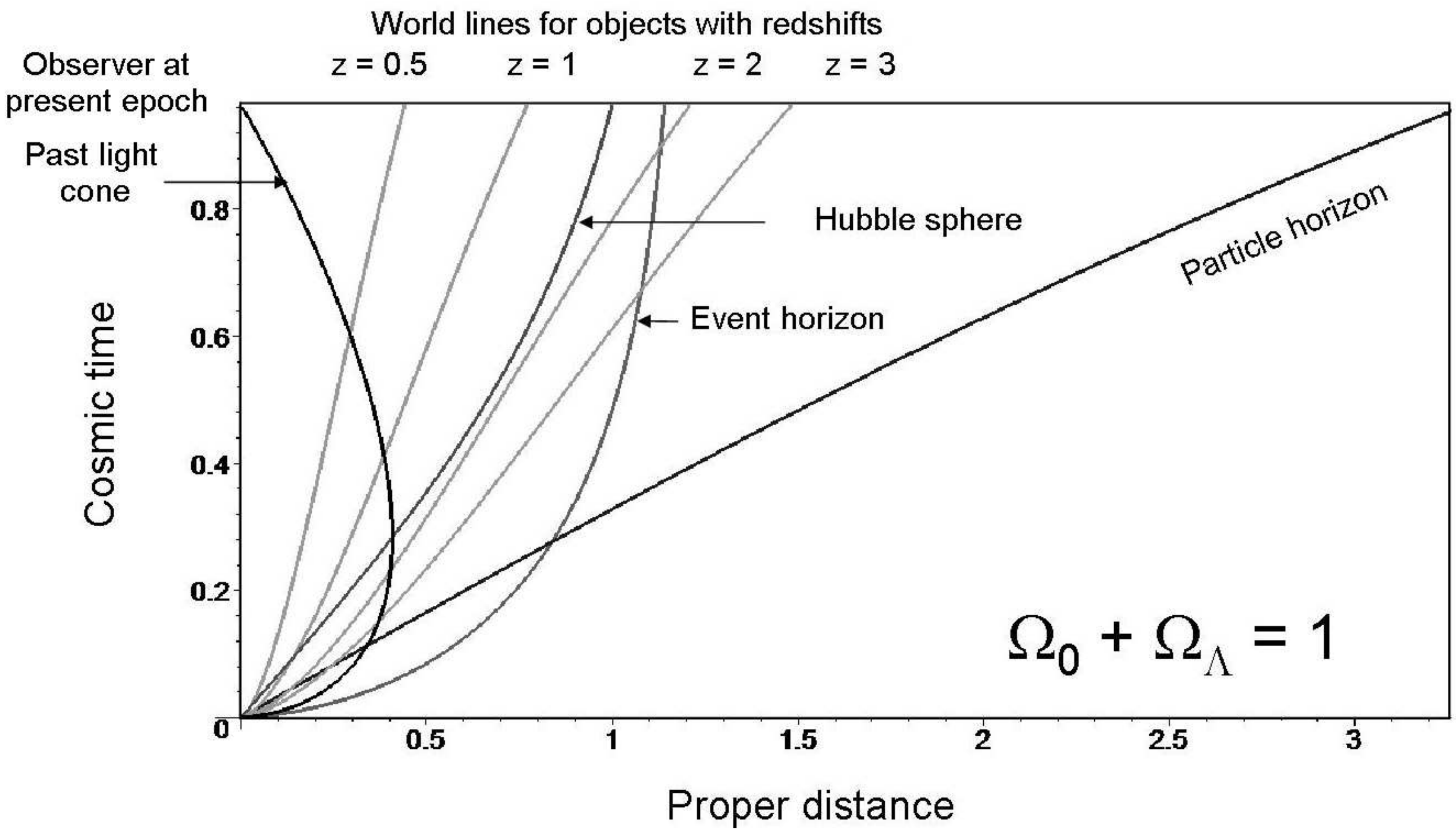
The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

Space-Time Diagram
Conformal Time vs. Comoving Distance Coordinate



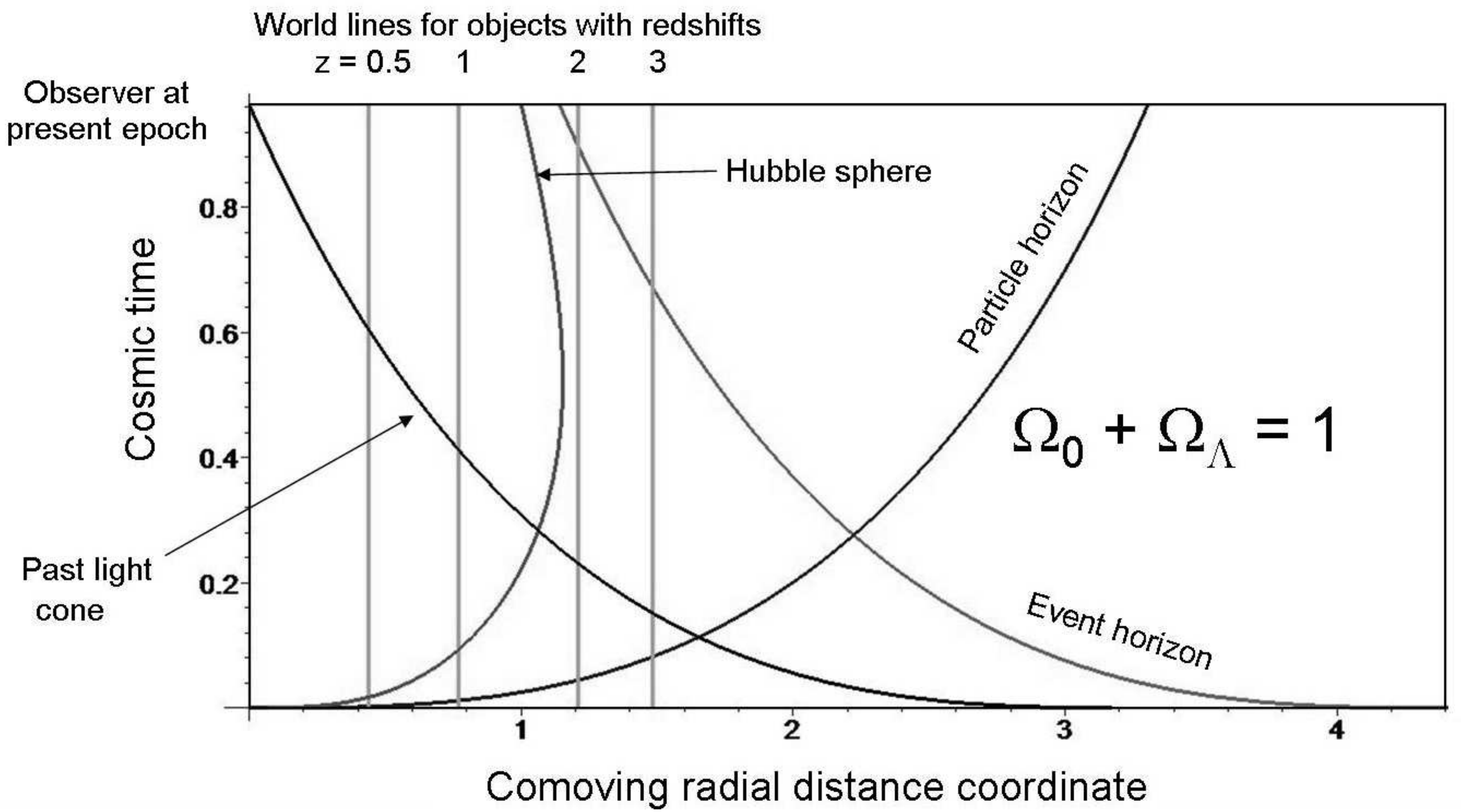
The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

Space-Time Diagram
Cosmic Time vs. Proper Distance



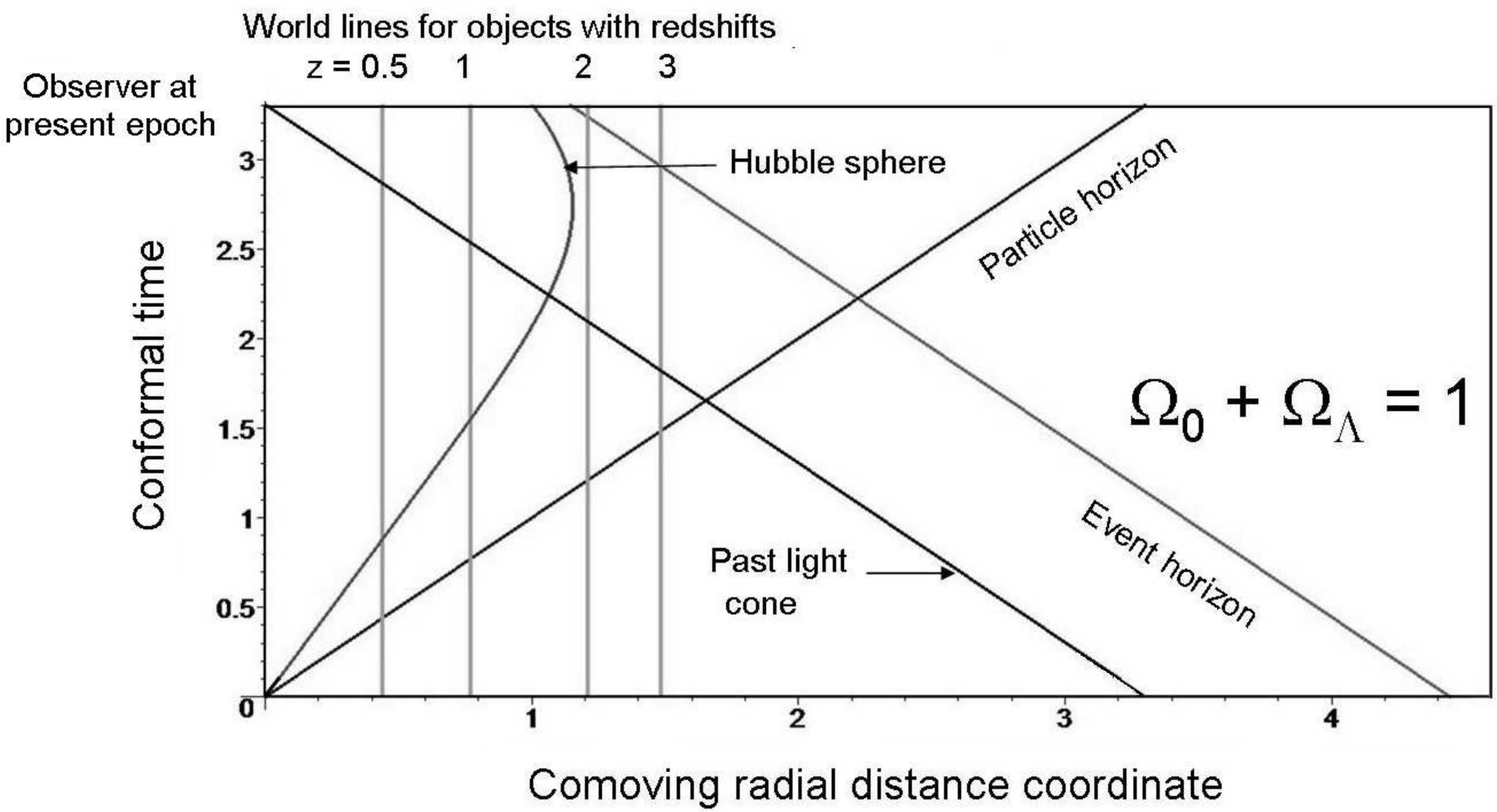
The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

Space-Time Diagram
Cosmic Time vs. Comoving Distance Coordinate



The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.

Space-Time Diagram
Conformal Time vs. Comoving Distance Coordinate



The times and distances are measured in units of H_0^{-1} and c/H_0 respectively.