L3 Astro - Section 7 - Cosmology 1

Nick Kaiser

Contents

List of Figures

1 The Hubble law

- Hubble observed Cepheid variable stars in "spiral nebulae" (late 20's)
  - period-luminosity relation: period $\rightarrow L \rightarrow$ distance D
- large distances firmly established them to be well outside of the MW
  - external 'galaxies'
- he also obtained spectra $\rightarrow$ redshift $\rightarrow$ "recession velocity"
- he found that *the local universe is expanding*
  - recession velocity roughly proportional to distance
- Later studies extended the 'Hubble diagram' (magnitude vs log-z) to larger distances and improved precision and accuracy
  - $cz = H_0 r$ with (current) *expansion rate* $H_0 \simeq 70\mathrm{km/s/Mpc} \simeq (1.4 \times 10^{10}\mathrm{yr})^{-1}$
  - departures on small-scales – growth of structure $\rightarrow$ 'peculiar' velocities
  - and departures from linearity on very large scales $D \sim c/H$

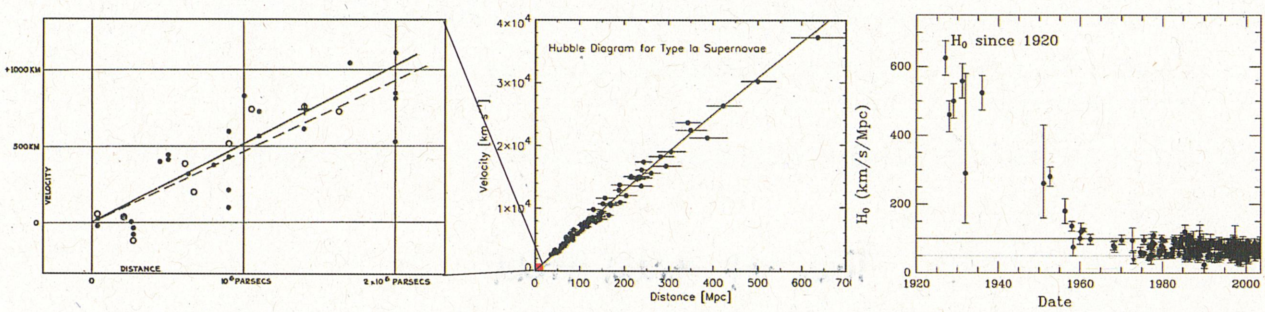


Figure 1: The Hubble diagram over the ages. On the left is Hubble's original plot from 1929. In the centre is a modern determination using supernovae. The red box at the lower left indicates the range of data that Hubble used. On the right is shown a compilation by John Huchra of historical estimates of the present-day expansion rate H_0. A common feature of such measurements is an underestimate of systematic errors that plague measurements of cosmological distances.

Hubble's observations – apparent distances as a function of recession velocity – are (and were at the time - in the 20s) interpreted in terms of *homogeneous and isotropic world-models* based on Einstein's then newly created *general theory of relativity* (GR). Let's now introduce the salient features of GR:

2 The essential features of Einstein's theory of gravity

2.1 The fundamental principles

2.1.1 The Galilean equivalence principle

In GR there is no 'force' of gravity; the fact that all objects fall the same way under the influence of gravity (the Galilean principle of equivalence) allowed Einstein to propose that gravity is curvature of space-time and that particles follow straight lines – or 'geodesics' – in space-time.

- *The curvature of space-time tells matter how to move.* (John Wheeler)

2.1.2 Space-time is everywhere locally the same as in special relativity

Just as the surface of a smooth object like an apple is locally flat and indistinguishable from Euclidean space, the so-called *'manifold'* of space-time is everywhere locally indistinguishable from Minkowskian space.

Now, living on Earth, we don't see that; we see the effect of gravity everywhere. But to see the true, locally Minkowskian, nature of the manifold of space-time, one need only jump out of the window of a tall building. Until you hit the ground, there is no gravity. According to Einstein, going into free-fall locally *'nulls-out'* the effects of gravity.

That means that there is a *light-cone structure* built into the manifold of space-time. It is an *absolute property of space-time*. The light cones allow one to categorise 4-vectors – the separation between neighbouring two events being the prototypical vector – into 3 classes:

1. **time-like:** a pair of events that a massive particle – or *'observer'* – can travel between
2. **space-like:** a pair of events that no observer can travel between
3. **null:** a pair of events that can be linked by a photon – whose path lies in the light cone

and all freely falling observers, no matter how they are moving, agree on this categorisation.

Another absolute property of space-time is that, while there are a family of possible freely falling observers at any point in space-time – depending on how they are moving and how they are oriented – *they can sense if they are rotating*. There is, again locally, a *non-rotating frame* that is somehow 'built-in' to the manifold.

2.1.3 Matter controls the curvature of space-time via its stress-energy tensor

- In special relativity all physical laws (electromagnetism, for example) are expressed in terms of 4-vectors and tensors.
- In Newton's theory, the gravitational potential ϕ is generated by the density of mass ρ
  - via *Poisson's equation:*
  - $\boxed{\nabla^2\phi = 4\pi G\rho}$ Trace of tidal field
- in special relativity there is a rank-two symmetric tensor $\mathbf{T}$, whose 10 components $T^{\mu\nu}$ are
  - T^{00} the energy density (equal to ρc^2)
  - T^{0i} the momentum density vector
  - T^{i0} the energy flux density vector (equal to the momentum density)
  - T^{ij} the stress or pressure 3-tensor (the flux density of momentum)
- for slowly moving masses only T^{00} is important (the others are suppressed by powers of v/c or v^2/c^2) and – aside from the factor c^2 – is the same as the mass density
- Einstein realised that there is an essentially unique rank-2 tensor determined from the curvature of space-time – we call it the *Einstein tensor* and denote it by $\mathbf{G}$ – that is driven (or 'sourced') by the stress-energy tensor
  - via *Einstein's equation:*
  - $\boxed{\mathbf{G} = 8\pi\kappa\mathbf{T}}$
  - this equation expressed the way that, in the second part of John Wheeler's beautiful aphorism,
    - *"matter tells space-time how to curve"*
  - single free parameter: $\kappa = G/c^4$

Spacetime is locally flat

$$ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta = \Lambda^\alpha{}_{\alpha'} \Lambda^\beta{}_{\beta'}\, g_{\alpha\beta}\, dx^{\alpha'} dx^{\beta'}$$

$$dx^{\alpha'} = \Lambda^{\alpha'}{}_\alpha dx^\alpha$$

2.2 The machinery of general relativity: differential geometry

2.2.1 The metric

The geometry is described by the *metric*. But this also encodes the *coordinate system* – which is arbitrary in GR (much of GR is concerned with *"generalised covariance"*; how to do physics in arbitrary coordinates).

Given set of *events* (points) on region or patch of manifold that is covered by some coordinate system x^α we can measure distances dl_{ij} between and then fit for a function $g_{\alpha\beta}(\vec{x})$ that gives the so-called 'line-element':

$$dl^2 = g_{\alpha\beta} dx^\alpha dx^\beta \tag{1}$$

where we use the *Einstein summation convention*.

The metric is a distillation of measurements of physical distances and how they relate to coordinate intervals.

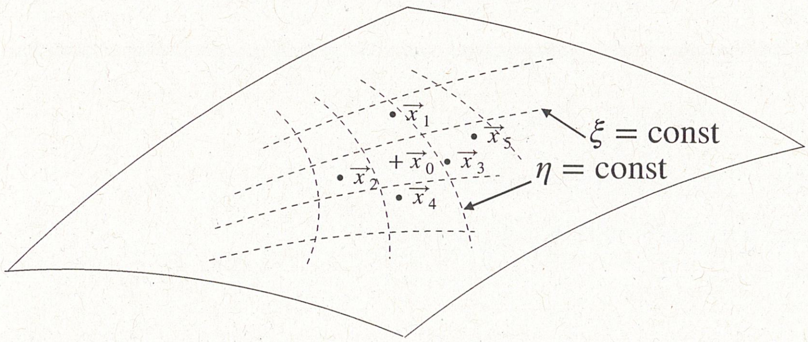


Figure 2: A portion of a 2-dimensional manifold on which lines of (unprimed) coordinates $x^\alpha = (\xi, \eta)$ are shown along with a set of points (events in 4D). Distances – in the manifold itself, not taking a 'short-cut' through the fictitious embedding space – can be measured using pieces of string or rulers. This allows the determination of the metric; which we can think of as a distillation of such distance measurements.

The metric plays the role in GR of the potential in Newtonian gravity.

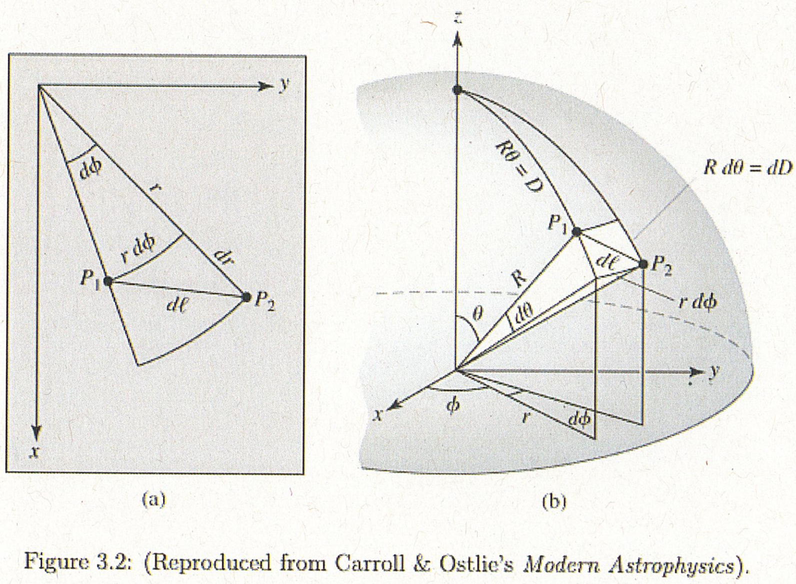


Figure 3.2: (Reproduced from Carroll & Ostlie's *Modern Astrophysics*).

Figure 3: Illustration of the line element $dl^2 = g_{\alpha\beta} dx^\alpha dx^\beta$ in polar coordinates on a plane (left) where $dl^2 = dr^2 + r^2 d\phi^2$ and on a sphere embedded in 3-D (right) where $dl^2 = R^2(d\theta^2 + \sin^2(\theta) d\phi^2)$. In homogeneous cosmological models, the metric of the spatial surfaces of constant proper time since the 'big-bang' are closely analogous: $dl^2 = a^2(d\chi^2 + S_k(\chi) d\Omega^2)$ where $a = a(\tau)$, $d\Omega^2 = d\theta^2 + \sin^2(\theta) d\phi^2$ and where the function S_k can be $\sin\chi$, χ or $\sinh\chi$. There are all spaces of constant curvature.

2.2.2 Parallel transport

Given a vector $\vec{\Delta x}$ at a point in a manifold one can *parallel transport* it along a line.

- if you drive over a line of wet paint, the vectors between the splotches of paint are parallel transported

See figure 5).

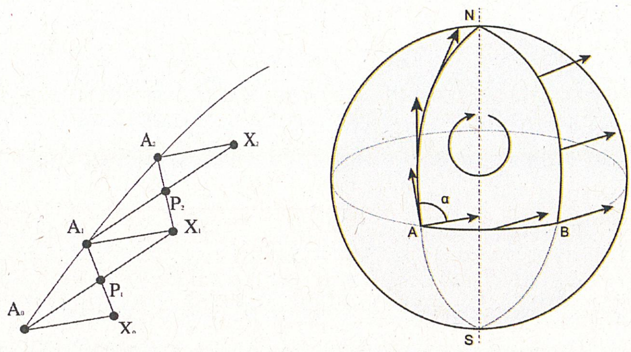


Figure 4: Left: 'Schild's ladder'. If we have a vector (here $\vec{X}_0$) at some point $\vec{x} = \vec{A}_0$, and a path $\vec{x}(\lambda)$ through that point, we can make a parallel transported copy $\vec{X}_1$ of $\vec{X}_0$ at $\vec{A}_1$ by erecting a vector $\vec{A}_1 - (\vec{A}_0 + \vec{X}_0)$. We then construct a vector from $\vec{A}_0$ to the mid point of that vector and extend it the same distance. This gives the end point of the vector $\vec{X}_1$. If we make the steps of the ladder smaller we obtain in the limit the (scaled down) parallel transported vector along the path. Right: the result of parallel transporting a vector on a curved manifold depends not just on the end points of the path but on the path taken. The difference is proportional to the area enclosed between the paths.

2.2.3 Covariant differentiation

The *covariant derivative* $\nabla\vec{V}$ of a vector *field* $\vec{V}(\vec{x})$ answers the question:

- *how does the field change with respect to a parallel transported copy of itself?*

It is well defined as a limit; but the change for a finite path is path dependent.

2.2.4 The connection

Both parallel transport and covariant differentiation involve the *connection* or the *Christoffel symbols* which can be computed from the metric:

$$\nabla\vec{V} \to V^\alpha{}_{;\beta} \equiv \underbrace{V^\alpha{}_{,\beta}}_{\partial V^\alpha/\partial x^\beta} + \Gamma^\alpha{}_{\mu\beta} V^\mu \tag{2}$$

where

$$\Gamma^\alpha{}_{\mu\nu} = \underbrace{g^{\alpha\beta}}_{\text{inverse of } g_{\alpha\beta}} \left(g_{\beta\mu,\nu} + g_{\beta\nu,\mu} - g_{\mu\nu,\beta}\right) \tag{3}$$

The connection plays the role in GR of the gravitational acceleration vector $\mathbf{g} \equiv -\boldsymbol{\nabla}\phi$.

2.2.5 Curvature

The curvature tensor $\vec{R}$ can be defined as the change $\Delta\vec{V}$ obtained after transporting a vector $\vec{V}$ around a parallelogram defined by two vectors $\vec{a}$ and $\vec{b}$ (see figure 5).

We can write – in geometric notation –

$$\Delta\vec{V} = \mathbf{R}(\ ,\vec{V},\vec{a},\vec{b}) \tag{4}$$

or – in component notation –

$$\Delta V^\alpha = R^\alpha{}_{\beta\mu\nu} V^\beta a^\mu b^\nu \tag{5}$$

where the *Riemann curvature tensor* is expressed in terms of the connection by

$$\boxed{R^\alpha{}_{\beta\mu\nu} \equiv \Gamma^\alpha{}_{\beta\nu,\mu} - \Gamma^\alpha{}_{\gamma\nu}\Gamma^\gamma{}_{\beta\mu} - \{\nu \Leftrightarrow \mu\}} \tag{6}$$

which evidently contains both derivatives and products of the Christoffel symbols.

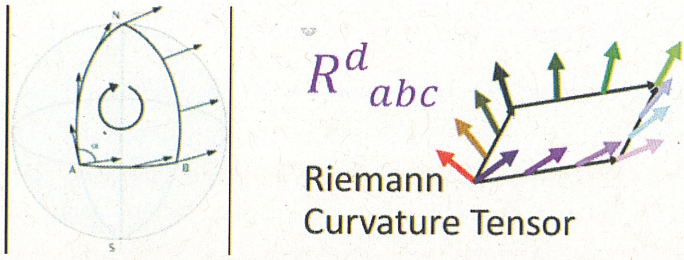


Figure 5: Curvature us defined in terms of parallel transport of a vector around a loop. The change in the vector depends on the original vector and the 2 vectors defining the path in a linear manner: $\Delta\vec{V} = \mathbf{R}(\ ,\vec{V},\vec{a},\vec{b})$, which is the equivalent, in GR, of the Newtonian (tidal) gravitational field. It is computable from the connection, which in turn derives from the metric, and when contracted gives the Einstein tensor on the LHS of his field equations.

2.2.6 Geodesic deviation

Curvature plays a key role in the *equation of geodesic deviation.* (see figure 6):

$$\left(\frac{d^2\vec{\xi}}{d\lambda^2}\right)^{\alpha} = \underbrace{\left(R^{\alpha}{}_{\mu\beta\nu}p^{\mu}p^{\nu}\right)}_{\text{tidal field tensor}}\xi^{\beta} \tag{7}$$

whereby the curvature controls the focussing – or defocussing – of neighbouring geodesics. In Newtonian gravity the same role is played by the gravitational *tidal field*:

$$\frac{d^2\boldsymbol{\xi}}{dt^2} = -\boldsymbol{\xi}\cdot\underbrace{\boldsymbol{\nabla}\boldsymbol{\nabla}\phi}_{\text{tidal field}}\ . \tag{8}$$

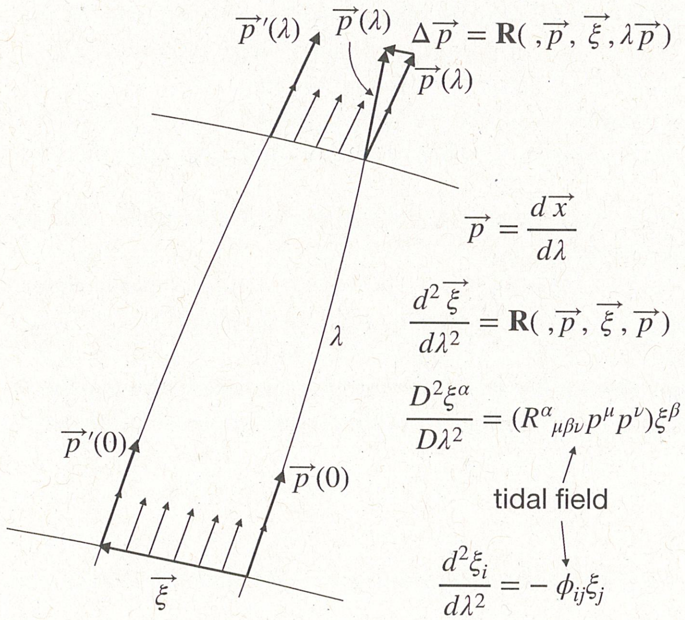


Figure 6: Geodesic deviation. This is a space-time diagram; time increases vertically. We start with a particle with 4-momentum $\vec{p}(0)$ at the bottom right. We can choose a parameter λ along the path – proportional to the proper time on a clock carried by the particle such that $\vec{p} = d\vec{x}/d\lambda$. Clone the particle and parallel transport it's momentum along the *separation vector* $\vec{\xi}$ to make $\vec{p}\,'(0)$. We have two particles with parallel momenta. Advance them forward in time (or λ) to make $\vec{p}(\lambda)$ and $\vec{p}\,'(\lambda)$ which will no longer be parallel (unless there is no curvature of space-time). Transport $\vec{p}\,'(\lambda)$ back to the location of the original particle and subtract the momenta. The result $\Delta\vec{p}$ is, from the definition of curvature, as indicated. But since $\vec{p} = d\vec{x}/d\lambda$, the rate of change of $\Delta\vec{p}$ is the same as $d^2\vec{\xi}/d\lambda^2$. This gives the *geodesic deviation equation.*

2.2.7 Einstein's equations

The geodesic deviation equation allows one to identify certain components of the Riemann tensor with the corresponding components of $\partial^2\phi/\partial x_i\partial x_j$. This ties down the single free parameter κ in Einstein's gravity to be $\kappa = G_{\rm N}/c^4$ so Einstein's equation becomes

$$\mathbf{G} = 8\pi(G_{\rm N}/c^4)\mathbf{T} \tag{9}$$

where, just as it is a *contraction* of the tidal field tensor – its trace ϕ_{ii} – that appears in Poisson's equation

$$\nabla^2\phi = 4\pi G_{\rm N}\rho \tag{10}$$

the *Einstein tensor* $\mathbf{G}$ is a certain contraction of the curvature tensor:

$$G_{\mu\nu} = R_{\mu\nu} - \tfrac{1}{2} g_{\mu\nu} R \tag{11}$$

where $R_{\mu\nu} = R^{\alpha}{}_{\mu\alpha\nu}$ and $R = g^{\mu\nu} R_{\mu\nu}$ are known respectively as the *Ricci tensor* and *Ricci scalar*.

2.2.8 Raychaudhuri's equation

A related useful result is *Raychaudhuri's equation.* While the relative acceleration of a pair of particles depends on both the tidal focussing caused by the local matter and the tidal field coming from distant matter. But the latter does not change the volume occupied by a collection of particles. If one takes the inverse cube root of the volume: $a = \sqrt[3]{V}$, this obeys the equation

$$\ddot{a} = -\frac{4\pi}{3} G(\rho + 3P/c^2)a \tag{12}$$

This is exactly as in Newtonian gravity except that pressure appears along with the density. This is one of the fundamental equations of cosmology.

2.2.9 The cosmological constant

Einstein obtained the field equations $\mathbf{G} = 8\pi\kappa\mathbf{T}$, which are the simplest compatible with Newtonian gravity with ρ replaced by $\mathbf{T}$ as the source, in 1915. In 1917, in order to allow static (non-expanding) solutions with vanishing pressure, he proposed the modification

$$\mathbf{G} - \Lambda\mathbf{g} = 8\pi\kappa\mathbf{T} \tag{13}$$

which introduces a new constant of nature Λ with units of inverse length squared[1].

If we move this over to the right-hand side of the field equations, we see that a positive Λ would correspond to a matter source with $T^{\mu\nu} = (\Lambda/8\pi\kappa)\mathrm{diag}(1, -1, -1, -1)$, or, equivalently, to matter with positive density $\rho = T^{00}/c^2$ and strong, but negative, pressure $P = -\rho c^2$. I.e. matter with strong *tension.*

This means that a positive Λ causes a pair of test particles to accelerate away from one another – because the pressure terms in $\rho + 3P/c^2$ outweigh the effect of the density – rather as would a negative mass density without pressure or tension.

With the discovery of the expansion of the universe, the cosmological constant was discarded. But it has re-emerged recently with vigour, and in two different situations.

First, in the theory of *inflation* it is assumed that the dynamics of the universe was at early times dominated by the effect of a nearly spatially constant and slowly time-varying relativistic scalar field dubbed the 'inflaton' (a cousin of the Higgs field, if you like).

Second, in 1999, it was shown convincingly[2] that the universe is now entering an accelerating phase. This is naturally interpreted as the effect of the cosmological constant or, perhaps, the influence of another scalar field analogous to the inflaton.

2.3 What is *the* gravitational field?

In Newtonian gravity we have the potential ϕ, the gravitational acceleration $\mathbf{g} = -\boldsymbol{\nabla}\phi$ and the tide $\phi_{ij} = \partial^2\phi/\partial x_i \partial x_j$.

In Einstein's gravity we have, analogously, the metric $g_{\mu\nu}$, the connection $\Gamma^{\alpha}{}_{\mu\nu}$ and the curvature $R^{\alpha}{}_{\beta\gamma\delta}$. Which of these deserves to be called *the* gravitational field?

The answer is the *curvature*. In many situations, one can have non-trivial metric tensors – i.e. coordinate systems in which $g_{\mu\nu} \neq \mathrm{diag}(-1, 1, 1, 1)$ – and non-vanishing connection also. These include simply working in curvilinear coordinate systems, such as spherical coordinates, but can also be used include affects of acceleration, as in a rocket, or on a rotating roundabout. But they have nothing to do with gravity *per se.*

[1] The minus sign is conventional and tied up with the convention for the sign of the metric. We are using the convention that the metric has 'signature' (sum of signs of eigenvalues) +2.

[2] A critical addition to the Hubble diagram for supernovae was the evidence from the CMB that the spatial curvature of the universe must be very close to being spatially flat, so this required something to augment the rather low measured matter density.

The curvature, on the other hand, vanishes in such situations, and it would be zero everywhere and always in a world where Einstein's constant κ (or Newton's $G_{\rm N}$) were zero. The curvature, or tidal field, considered as a geometric entity $\mathbf{R}$ is generated by the presence of matter $\mathbf{T}$ (along with boundary conditions; the field equations, like Poisson's equations, not providing explicitly all the components of the curvature). If it vanishes – and being a tensor, its vanishing or not-vanishing is an absolute fact – then there is no gravity. In this sense, there is nothing 'relative' about general relativity.

This makes nonsense of the oft stated form of 'Einstein's principle of equivalence' (EEP): *that gravity and acceleration are indistinguishable.*. If you are being accelerated in a rocket out in empty space then there is no curvature; the gravitational field vanishes. Adopting this view, we would have to say that Pound and Rebka did not measure the *gravitational redshift* in their celebrated experiment in 1959. What they measured was merely the affect of their apparatus being accelerated.

A better way to state the EEP is to say that, if you are freely falling, you will not see any effect of gravity (locally). Note that this is very different from the situation with e.g. electromagnetism, where one can measure e.g. the electric field $\mathbf{E}$ locally. Fundamentally that is because, in electromagnetism, particles generally have different charge-to-mass ratios, whereas in gravity, the Galilean equivalence principle – that all particles fall the same way under the influence of a gravitating body – means that all objects have the same effective charge-to-mass ratio.

3 Homogeneous Expanding Universe Models

3.1 The Friedmann, Lemaitre, Robertson & Walker (FLRW) model

3.1.1 The cosmological principle

Observations – particularly of the *cosmic microwave background* (CMB) – show our universe to be highly *isotropic* on large scales.

Establishing that the Universe is *spatially homogeneous* is more difficult, but there is good evidence for that as well.

At the time of Hubble there was very little firm evidence for either. But, perhaps presciently, or perhaps because the only solutions of Einstein's equations were those of very high symmetry, cosmologists adopted the *cosmological principle* that the Universe is spatially homogeneous (i.e. that, aside from local irregularities, it looks the same from all points in space – at a given 'cosmic-time').

A subset of cosmologists went further than this and argued for what is called the *perfect cosmological principle:* that not only does the Universe appear the same at all points in space but also at all times. This was not widely adopted[3], but, interestingly it is now believed by most cosmologists that both during *inflation* in the very early universe and in the distant future as well, the Universe will become time invariant and obey the perfect cosmological principle.

3.1.2 The metric and the line element

The cosmological principle tightly constrains the form of the metric of space-time. There is still some coordinate freedom, but one allowed form is

$$ds^2 = -c^2 d\tau^2 + a^2(\tau)(d\chi^2 + S_k(\chi)^2(d\theta^2 + \sin^2\theta d\phi^2)) \tag{14}$$

where it is important to note that e.g. $d\tau^2 \equiv (d\tau)^2$ *not* $d\tau^2 = 2\tau d\tau$.

This gives the (squared) interval of proper separation between neighbouring 'events' where, as in *special relativity* (Minkowski space-time)

$$ds^2 \text{ may be } \begin{cases} \text{negative} \\ \text{zero} \\ \text{positive} \end{cases} \text{ for } \begin{cases} \text{timelike} \\ \text{null} \\ \text{spacelike} \end{cases} \text{ intervals} \tag{15}$$

The *spatial coordinates* (χ, θ, ϕ) are dimensionless fixed labels attached to *fundamental observers* (FOs) who carry clocks measuring proper time τ.

[3] In 1953, Herbert Dingle, who was president of the Royal Astronomical Society said 'Since it causes me considerable discomfort to use names that are clearly misleading, I shall refer to the "cosmological principle" as the *cosmological assumption* and to the "perfect cosmological principle" as the *cosmological presumption*'

The *spatial geometry* – i.e. the geometry of *hypersurfaces* of constant time τ – may be spherical, flat (Euclidean) or hyperbolic (saddle-like), depending on the *curvature eigenvalue* k, according to which the function

$$S_k(\chi) = \begin{cases} \sin\chi \\ \chi \\ \sinh\chi \end{cases} \quad \text{for} \quad k = \begin{cases} +1 \\ 0 \\ -1 \end{cases} \tag{16}$$

A simple proof that S_k can only be one of these functions is given in figure 7 which shows that, were we to replace $S_k(\chi)$ by an arbitrary function $f(\chi)$ then that the spatial curvature of the equatorial plane $\theta = 0$ is proportional to f''/f (where prime denotes derivative with respect to χ). In a spatially homogeneous universe then

$$f'' = -k \times f \tag{17}$$

where k is the *curvature eigenvalue*, and whose solutions are the functions $S_k(\chi)$.

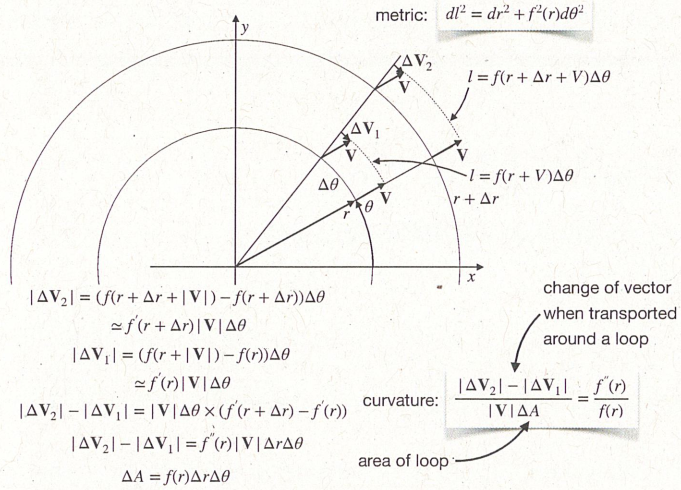


Figure 7: In 2-dimensions, the intrinsic curvature of space is defined by a single number; how much a vector changes if you carry it around a loop keeping it parallel to itself divided by the area of the loop (and the length of the vector). It has units of inverse area. It's illustrated here on a plane. For a Euclidean space – one for which $dl^2 = dr^2 + r^2 d\theta^2$ there is no change in the vector. But for a more general circularly symmetric space – or the equatorial slice through a 3-dimensional spherically symmetric space – with $dl^2 = dr^2 + f^2(r)d\theta^2$ – the curvature is non-zero and is equal to f''/f. Homogeneous spaces must have $f''/f =$ constant, the possible solutions of which are $f(r) = \sinh(r)$, $f(r) = r$ or $f(r) = \sinh(r)$.

The geometry of the equatorial slice through the FLRW spatial geometry is illustrated in figure 8.

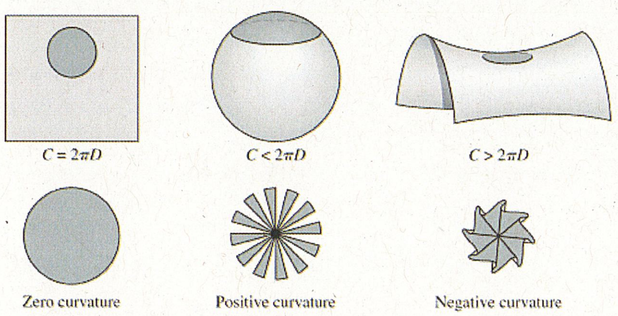


Figure 8: In 2-dimensions there are 3 choices of geometry that are homogeneous and isotropic: the sphere, the plane and the saddle. They are distinguished by whether the circumference of a circle is less than, equal to, or greater than 2π times the radius. An alternative way to determine the curvature locally is to take a vector and 'parallel transport' it around a small closed loop and differencing the transported copy from the original (see figure 5).

Another useful form for the metric is obtained if we define a new time coordinate η such that $d\tau = a(\tau)d\eta$:

$$ds^2 = a^2(\eta)\left(-d\eta^2 + d\chi^2 + S_k^2(\chi)(d\theta^2 + \sin^2\theta d\phi^2)\right) \tag{18}$$

where $a(\eta) = a(\tau)$, and which is said to be a 'conformal transformation' of the simper metric without the scale factor. This leads to the terminology of τ being called 'conformal time' and χ being called the 'conformal distance'.

One may note that if we define $r \equiv S_k(\chi)$ we have $dr^2 = (1 - kr^2)d\chi^2$ so an alternative way to write the metric is

$$ds^2 = -c^2 d\tau^2 + a^2(\tau)\left(\frac{dr^2}{1 - kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right) \tag{19}$$

which perhaps makes it clearer that the geometry of the (2-dimensional) surfaces of constant r (or χ) is that of a sphere.

The other function $a(\tau)$ appearing in the metric is called the *cosmological scale factor*. It is not determined by symmetry considerations; rather it is determined by Einstein's field equations.

3.1.3 The Friedmann and continuity equation

In 1922, Alexander Friedmann found the equations describing an expanding homogeneous and isotropic universe.

He assumed the metric above and that the stress-energy tensor for the matter was that of an 'ideal fluid'

$$T_{\alpha'\beta'} = \begin{bmatrix} \rho c^2 & 0 & 0 & 0 \\ 0 & P & 0 & 0 \\ 0 & 0 & P & 0 \\ 0 & 0 & 0 & P \end{bmatrix} \tag{20}$$

where the primes on the indices indicate that this is in physical coordinates erected in the vicinity of a fundamental observer

As already discussed, in this tensor:

- the upper left element is the *total energy density* $\mathcal{E} = \rho c^2$
- and the lower-right sub-matrix is the diagonal stress tensor containing the *isotropic pressure* P
- three zeros in the left column are the momentum density,
- the three zeros in the top row are the energy flux density
- all these are as would be measured in the frame of a fundamental observer
- sometimes called *'comoving observers'* as they are co-moving with the cosmological fluid in that they see vanishing momentum density and energy flux in their frame of reference

Dynamical equations for $a(\tau)$ and $\rho(\tau)$ can be obtained from a combination of the Einstein field equations (by computing the connection and hence the Einstein tensor $\mathbf{G}$ and equating it to $8\pi\kappa\mathbf{T}$) and the laws of continuity of energy and momentum.

Taking any one of the spatial components of Einstein's equations gives the *Friedmann equation* for the scale factor $a(\tau)$

$$\boxed{\dot{a}^2 = (8\pi/3)G\rho a^2 - c^2 k} \tag{21}$$

where dot denotes derivative with respect to τ, so $\dot{a} \equiv da/d\tau$.

Significant features of this equation are:

- it is (if multiplied by 1/2) identical to the Newtonian expression for conservation of energy (kinetic energy + potential energy = constant) for an expanding spherical shell of radius a containing mass $M = (4\pi/3)G\rho a^3$
- dividing by a^2 the left hand side is the observable H^2 so, augmented by a measurement of the density ρ, this allows one to determine the radius of curvature a_0 and the curvature eigenvalue k
- the sign of the spatial curvature is determined by whether the kinetic energy is greater than, less than, or equal to the gravitational energy
- the pressure P does not appear – so the spatial curvature is determined by ρ alone

From continuity of energy (technically the time component of $\nabla \cdot \mathbf{T} = 0$) we get the *matter continuity equation*

$$\boxed{\dot{\rho} = -3(\dot{a}/a)(\rho + P/c^2)} \tag{22}$$

With $P = 0$, this has a solution $\rho \propto a^{-3}$ and so (defining $V \equiv (4\pi/3)a^3$) this then simply expresses conservation of mass:

$$d(\rho V)/d\tau = 0 \tag{23}$$

While for $P \neq 0$, and since $dV = 3(da/a)V$, the extra term implies

$$d(\rho V)/d\tau = -(P/c^2)dV/d\tau \tag{24}$$

which, on multiplying by c^2, we recognise as the first law of thermodynamics

$$\boxed{dE = d(\mathcal{E}V) = -PdV.} \tag{25}$$

3.1.4 The acceleration equation

The Friedmann and continuity equations are two 1st order equations. If we take the proper time derivative of the first we get

$$2\dot{a}\ddot{a} = (8\pi/3)G(\dot{\rho}a^2 + 2\rho a\dot{a}) \tag{26}$$

and use the second to eliminate $\dot{\rho}$ we get the single 2nd order *acceleration equation:*

$$\boxed{\ddot{a} = -(4\pi/3)G(\rho + 3P/c^2)a} \tag{27}$$

- which, for $P = 0$, is what Newton would have written down for an expanding sphere of dust
- but which, in general, contains an additional deceleration from the pressure
- sometimes expressed by saying 'pressure gravitates in GR'

Note that the three equations for $\dot{a}^2$, $\dot{\rho}$ and $\ddot{a}$ above are not independent, as any one of them can be obtained from the other two.

It is important to realise that the presence of the pressure here is not the effect of e.g. the kinetic energy of motion of particles in the case of a gas. Such contributions to the energy do gravitate, but are already included in the density as ρc^2 is the total energy density.

3.1.5 The 'equation of state' ∼ the rule to say 'what's pressure'

We have above either two 1st order equations (Friedmann + continuity) or 1 second order equation (acceleration) for three unknown functions of time: $a(\tau)$, $\rho(\tau)$ and $P(\tau)$.

To obtain solutions we must augment these with an *'equation of state'* giving the pressure $P = P(\rho)$. This may be

- $\boxed{P_{\rm m} = 0}$ for pressureless matter ('dust')
  - an important component at present
- $\boxed{P_{\rm r} = \rho c^2/3}$ for radiation
  - dynamically negligible now, but dominant in the past since it has $\rho_{\rm r} \propto a^{-4}$ as compared to $\rho_{\rm m} \propto a^{-3}$

Observations that the expansion of the universe is speeding up suggest we need a third component: *'dark energy'* with negative pressure to give $\ddot{a} > 0$.

Our ignorance about its equation of state is encapsulated in an unknown function of time (or redshift)

$$\boxed{\omega \equiv P/\rho c^2} \tag{28}$$

with, as an interesting special case $\omega = -1$, for which

$$\boxed{P_\Lambda = -\rho c^2} \tag{29}$$

which, as discussed, may arise from a scalar field (*'quintessence'*) or may simply represent Einstein's cosmological constant Λ.

From the continuity equation, dark energy with $\omega = -1$ has $\rho =$ constant

- and so would have become negligible in the recent past
- but will dominate in the future
- and also dominated during inflation

A note on terminology is in order. An 'equation of state' (EoS) in thermodynamics is a relation giving one thermodynamic variable in terms of 2 others (e.g. $P = P(\rho, T)$), and which may be adiabatic, isothermal etc.In cosmology people use EoS rather loosely for an expression for P in terms of density alone, and as just the ratio of pressure to density. This is either because the temperature is either assumed to be solely a function of the density – as for thermal radiation that is expanding adiabatically – or because one is not dealing with a thermodynamic system (e.g. quintessence or the inflaton) so temperature is not defined.

3.1.6 The expansion rate, critical density and the density parameters

The *expansion rate* is defined to be

$$\boxed{H \equiv \dot{a}/a} \tag{30}$$

with units of inverse time, and whose current value is measured to a few percent precision to be

$$H_0 \simeq 70\,\mathrm{km/sec/Mpc}. \tag{31}$$

The Friedmann equation can be expressed as

$$H^2 = (8\pi/3)G\rho - c^2 k/a^2 \tag{32}$$

from which we can infer that

$$k = \mathrm{sign}(\rho_0 - \rho_{\mathrm{crit}}) \tag{33}$$

where the *'critical density'* is

$$\boxed{\rho_{\mathrm{crit}} \equiv 3H_0^2/8\pi G} \tag{34}$$

which is the density the universe would have to have for the potential and kinetic terms in the Friedmann equations to balance.

We often express densities in units of the present day critical density. So, for example, the *density parameter* for the matter is defined to be

$$\boxed{\Omega_{\mathrm{m}0} \equiv \rho_{\mathrm{m}0}/\rho_{\mathrm{crit}}.} \tag{35}$$

Observations indicate that $\Omega_{\mathrm{m}0} \simeq 0.3$ (see below), so the amount of matter is about 30% of that required to 'close the universe'. If that were all there was we would conclude that the universe must have negative (hyperbolic or saddle-like) spatial curvature since $\rho_0 < \rho_{\mathrm{crit}}$, and that the current proper distance to an object at a conformal distance equal to the curvature scale ($\chi = 1$) is

$$a_0 = cH_0^{-1}|1 - \Omega_{\mathrm{m}0}|^{-1/2} \tag{36}$$

i.e. somewhat greater than the *Hubble distance* $L_{\mathrm{H}} \equiv c/H_0 \simeq 4000\mathrm{Mpc}$.

However, as already mentioned, there is good reason to think that there is also a non-negligible dark energy component with density parameter $\Omega_{\Lambda 0}$, most probably very close to $1 - \Omega_{\mathrm{m}0}$ and, in the past one had to include radiation though its current density parameter is very small $\Omega_{\mathrm{r}0} \sim 10^{-4}$.

It is usual to define $\Omega_{k0} \equiv 1 - \sum\limits_{i=\mathrm{m,r},\Lambda} \Omega_{i0}$,

- i.e. whatever would be needed to close the universe after accounting for all the matter content

using the fact that $\rho_{\mathrm{m}} \propto a^{-3}$, $\rho_{\mathrm{r}} \propto a^{-4}$, $\rho_\Lambda \propto a^0$ the Friedmann equation gives us the expansion rate when scale factor of the universe was a:

$$H = H_0[\Omega_{\mathrm{m}0}(a/a_0)^{-3} + \Omega_{\mathrm{r}0}(a/a_0)^{-4} + \Omega_\Lambda + \Omega_{k0}(a/a_0)^{-2}]^{1/2} \tag{37}$$

or, defining the *redshift* z by

$$\boxed{1 + z \equiv \frac{a_0}{a}} \tag{38}$$

we can compute the expansion rate as a function of redshift as

$$H(z) = H_0[\Omega_{\mathrm{m}0}(1+z)^3 + \Omega_{\mathrm{r}0}(1+z)^4 + \Omega_\Lambda + \Omega_{k0}(1+z)^2]^{1/2}. \tag{39}$$

3.1.7 Solutions of the Friedmann equations

If we specify type of constituents – and thus how their densities vary as a function of the size of the universe – then we can solve the Friedmann + continuity equation (or the acceleration equation) to get $a(\tau)$ and hence $\rho(\tau)$ and $P(\tau)$ also.

This requires, two boundary conditions: the present day density and expansion rate.

Unfortunately there are no analytic solutions with dark energy or pressure (except as limiting cases).

However, for a universe that contains only pressure free matter there is a parametric solution (the cycloid – for a closed universe – and hyper-cycloid for an open universe)

$$\begin{aligned} a(\eta) &= A(\cosh\eta - 1) \\ \tau(\eta) &= B(\sinh\eta - \eta) \end{aligned} \tag{40}$$

where A and B are constants and η is the conformal time, with the property that $d\eta \propto d\tau/a(\tau)$, and changes in conformal 'look-back time' and conformal distance are related by $d\eta = -d\chi$. A family of such solutions is shown in figure 9.

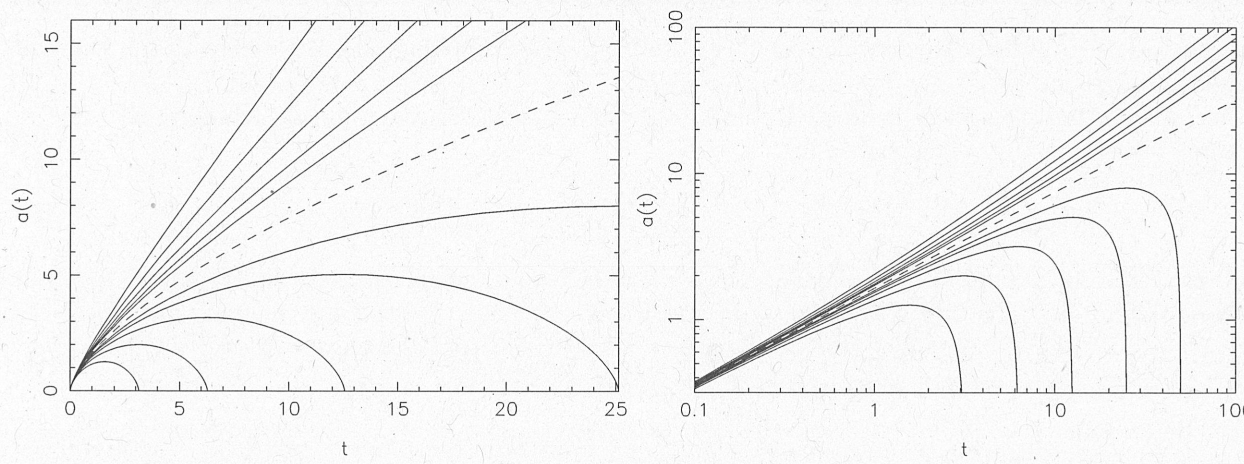


Figure 9: Cycloidal and hyper-cycloidal solutions of the Friedmann equation. These are the same as the solutions for a pebble fired upwards from a compact (Newtonian) mass. At very early times – when the kinetic energy and potential energy are both very large in relation to their difference – the solution is a power law $a \propto t^{2/3}$. The curves represent a sequence of increasing total energy.

Unfortunately, if we include dark energy one needs to solve the equations numerically.

At early times things are much simpler, because at redshifts beyond a few we can neglect the dark energy and Ω_k terms. Then

- in the *matter dominated regime* we then have power law solution
  - $\boxed{a \propto \tau^{2/3}}$
  - which can be shown, either by taking the $\eta \ll 1$ limit of the hypercycloidal solution above,
  - or, more usefully, by noting that
    - for a power law expansion $a \propto \tau^\gamma$,
    - the expansion rate goes like $H = \dot{a}/a = \gamma/\tau \propto a^{-\gamma}$
    - but, since $H^2 \propto \rho \propto a^{-3}$, so $H \propto a^{-3/2}$, this requires $\gamma = 2/3$
- and in the *radiation dominated regime* applying the argument above, but with $\rho \propto a^{-4}$ we get $H \propto \sqrt{\rho} \propto a^{-2}$ so $\gamma = -1/2$ and therefore
  - $\boxed{a \propto \tau^{1/2}}$
- we will discuss these more below

Another interesting solution emerges when the universe is strongly Λ-*dominated*, or dominated by a field with $P \simeq -\rho c^2$, which, it seems, will happen in the not so distant future, and which, it is widely suspected, happened in the distant past in the *inflationary era* that preceded the hot big bang.

For a Λ-dominated universe the universe expands exponentially with

$$\boxed{a(\tau) \propto \exp(H\tau)} \tag{41}$$

- with H asymptotically constant
- while maintaining constant energy density
- and thus creating energy out of nothing, hence *"inflation is the ultimate free lunch"* (Guth)

3.2 Interpretation of observations in FRW models

The FLRW models preceded (just), and allowed interpretation of, Hubble's observations. He was measuring recession velocities inferred from the redshift and distances obtained from flux-densities of variable stars – which he was using as *'standard-candles'*.

These models allow prediction of, in addition to flux-density, angular sizes of objects of known size – or *'standard-rulers'* as a function of redshift.

Note that, to test the model, or determine the parameters of a model, we need to have at least two ways of determining the distance to an astronomical object.

Though it should be said that, at the modest distances he was observing, relativistic effects were negligible and a Newtonian interpretation would have been adequate. But, oddly enough, Newtonian cosmology was not well developed or understood at the time.

In FLRW models, light from distant objects is focused by the gravitational lensing effect of intervening matter. This depends on how much mass there is (and the redshift depends on how the universe is expanding).

This focussing is conventionally expressed in terms of *'apparent distances'*

- $\boxed{D_L(z)}$ for flux density (or luminosity L), and
- $\boxed{D_{\mathrm{a}}(z)}$ for angular size
- both of which, in any specific cosmological model, are computable functions of redshift

these are the answers to the questions:

- *'how for away would a object of known luminosity (size) have to be <u>in an empty universe</u> in order to have the flux density (angular size) computed in the model from the metric'*

while redshift depends on how much the universe has expanded since the light we see left the objects.

These apparent-distance vs. redshift relations $D_L(z)$ and $D_{\mathrm{a}}(z)$ can be used in two ways:

1. if we assume the cosmological density parameters are known we get the intrinsic properties from observed ones
2. if we assume the intrinsic properties are known then we determine the cosmological parameters

3.2.1 The cosmological redshift - measurement

We *defined* the redshift above such that

$$1 + z \equiv a_0/a. \tag{42}$$

The reason it is called the redshift is that it is directly observable as a shift in the wavelength of spectral lines

$$\boxed{\lambda_{\mathrm{obs}}/\lambda_{\mathrm{em}} = 1 + z} \tag{43}$$

One way to understand this is by analogy with a standing wave in an expanding cavity (see left panel of figure 10).

One can also derive this by thinking about the Doppler shifts suffered by a photon bouncing repeatedly off the assumed to be steadily receding walls of an expanding cavity.

Weaknesses of this argument are:

- is radiation – e.g. that from a distant transient source – behave in the same way as a standing wave
- in reality there are no reflecting walls
- and surely the gravitational redshift has to be involved at some level
- and has been challenged (see papers on 'redshift-remapping')

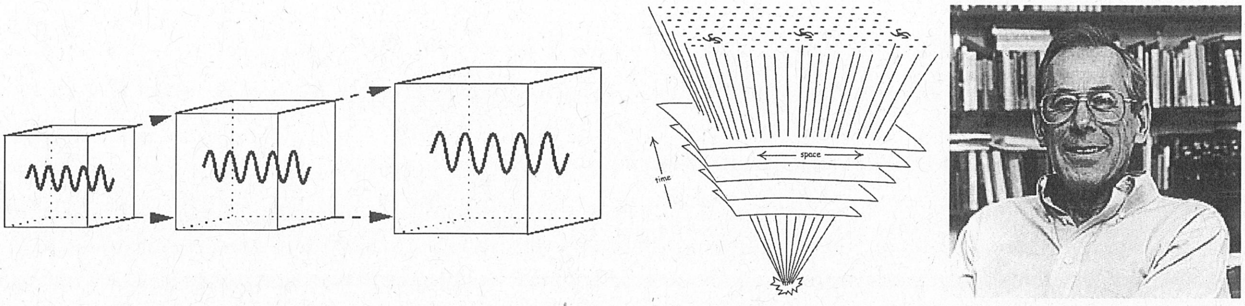
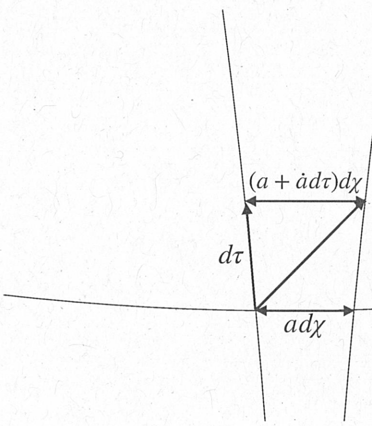


Figure 10: The cosmological redshift can be understood by analogy with standing waves in a cavity (left) for which the wavelength scales with the box size. A rigorous argument (due to Peebles - at right) is to consider the wavelength change as the product of a lot of infinitesimal shifts between a sequence of fundamental observers that the photon passes on its path.



$$\frac{d\lambda}{\lambda} = \frac{dv}{c} = \frac{\dot{a}}{a} \times a d\chi = \dot{a} d\chi$$

$$\frac{da}{a} = \frac{\dot{a} d\tau}{a} = \dot{a} d\chi$$

$$\boxed{\frac{d\lambda}{\lambda} = \frac{da}{a} \Rightarrow \lambda \propto a}$$

Figure 11: Peebles' argument for the cosmological redshift. We consider two neighbouring fundamental observers, one of whom sends a photon to the other. These observers are in free-fall, so gravity is 'transformed away' (locally at least). So the change in the frequency of the photon is just the 1st order Doppler shift. It follows that the fractional change in the wavelength is equal to the fractional change in their separation and, by extension, the fractional change in the scale factor.

- a rigorous way to prove this is illustrated in the centre panel of figure 10. One imagines a finite wavelength change $\lambda_{\rm obs}/\lambda_{\rm em}$ as being the product of a set of infinitesimal shifts.
  - by virtue of the fact that these observers are in free fall
  - and that in a locally freely falling frame the effect of gravity is 'transformed away'
  - one can be confident that the only effect is the 1st order Doppler effect $1 + \delta\lambda/\lambda = 1 + \delta v/c = 1 + H\delta x/c = 1 + H\delta\tau$ where $H = \dot{a}/a$ is the expansion rate and δt is the time elapsed as the photon makes its way between the two neighbouring observers
  - but, since $H\delta\tau = (\dot{a}/a)\delta\tau = \delta a/a$, it follows that
  - $\boxed{\delta\lambda/\lambda = \delta a/a}$
  - which we can integrate up to get $\lambda_{\rm obs}/\lambda_{\rm em} = a_{\rm obs}/a_{\rm em}$
  - this would appear to be a serious challenge to proponents of 'redshift-remapping'
- with a high resolution spectrograph, redshifts of galaxies can be measured with great precision
- and even with 'broad-band colours' one can obtain quite good accuracy (though one needs to beware of 'outliers')

- but these require that there be features in the spectrum whose 'rest-frame' wavelength is known
- a counter-example is the cosmic background radiation (CMB)
  - this has a thermal, or 'black-body', spectrum
  - which has the property that in an expanding universe it remains thermal even in the absence of interaction of matter
  - so there is nothing about the CMB that tells us at what redshift it was last interacting with matter
  - but we know that this was at $z \simeq 1100$ from Saha's equation

3.2.2 Conformal distance-redshift relation

Having established how the redshift can be measured, we now want to relate this to conformal distance. Not that this is our ultimate goal – which is to relate redshift to apparent (luminosity and angular diameter) distances.

From the metric, for light, which follows null trajectories: $ds = 0$, propagating radially from the origin (i.e. with θ, ϕ fixed, so $d\theta = d\phi = 0$) we have

$$\cancel{ds}^2 = \underbrace{-c^2 d\tau^2 + a(\tau)^2(d\chi^2}_{d\chi=-cd\tau/a} + S_k(\chi)(\cancel{d\theta}^2 + \sin^2\theta\, \cancel{d\phi}^2) \tag{44}$$

and a useful chain of relations follows from the differential redshift:

$$dz = \underbrace{d(1+z) = a_0 da^{-1}}_{1+z\equiv a_0/a} = -\frac{a_0}{a^2}da = -\frac{a_0}{a^2}\dot{a}d\tau = \underbrace{-\frac{a_0}{a}H d\tau = \frac{a_0}{c}H d\chi}_{cd\tau=-ad\chi} \tag{45}$$

from which we can extract

$$\begin{aligned} d\tau &= \frac{dz}{(1+z)H(z)} \\ a_0 d\chi &= \frac{cdz}{H(z)} \end{aligned} \tag{46}$$

Integrating the latter we obtain the present epoch proper distance to redshift z

$$\boxed{a_0\chi(z) = c\int_0^z \frac{dz}{H(z)}} \tag{47}$$

or, using the expression obtained above for $H(z)$

$$\boxed{a_0\chi(z) = a_0 \int_0^{\chi(z)} d\chi = \frac{c}{H_0}\int_0^z \frac{dz}{[\Omega_{\mathrm{m}0}(1+z)^3 + \Omega_{\mathrm{r}0}(1+z)^4 + \Omega_\Lambda + \Omega_{k0}(1+z)^2]^{1/2}}} \tag{48}$$

which is a nasty integral, but something one can readily evaluate numerically given as input the values of the density parameters.

It should be noted that the integral here converges if there is any matter or radiation in the universe. In a completely empty universe (so $\Omega_{k0} = 1$) the conformal distance diverges logarithmically and in a universe with only dark energy and $\Omega_{k0} = 0$ the divergence is stronger. But these are not realistic options.

– The Einstein-de Sitter model

As an illustrative example – albeit a not an entirely realistic one – for the so-called *'Einstein-de Sitter'* model: $\Omega_m = 1$, all others zero, we obtain

$$a_0\chi(z) = (2c/H_0)[1 - (1+z)^{-1/2}] \tag{49}$$

regarding which we note the following:

- in this spatially flat model ($\Omega_k = 0$), as in any spatially flat model, the present day scale factor a_0 (being the radius of curvature) is formally infinite, while $\chi(z)$ is formally zero, but the product $a_0\chi$ is finite
- the proper distance to infinite redshift – the *horizon* – is finite
- the distance grows linearly at low-z: $a_0\chi(z \ll 1) \simeq cz/H_0$
- but this growth tapers off:
  - by $z = 3$ we are half-way to the horizon
  - by $z \simeq 1000$ (the redshift at which the photons of the CMB were last scattered) we are about 97% of the way to the horizon (see figure 12)

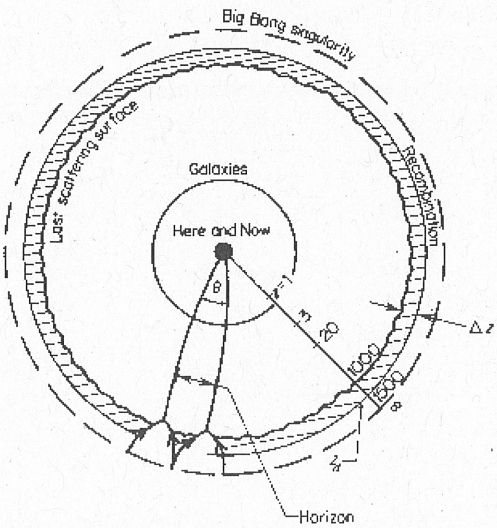


Figure 12: An equatorial slice through our universe showing the surfaces of constant redshift in a plot where radius is conformal distance χ. Galaxies can be seen out to $z \sim 10$, which is a good fraction of the volume within the entire region we can observer. Note that there is no 'edge' to the universe in this model. The density of matter, galaxies etc. is assumed to extend without limit – though we will only ever see galaxies below the redshift at which they formed (thought to be around $z \sim 20$). The horizon – the dashed circle – is simply the limit imposed by the fact that there is a maximum conformal distance that any information can have propagated in the age of the universe. The arrows labelled 'Horizon' in this plot indicate the horizon size back at $z \simeq 1000$. Parts of the sky with separation bigger than this were not, in the big-bang model, in causal contact with one-another when the photons were released, yet they have almost identical temperatures.

3.2.3 The angular diameter distance

Let's suppose we are observing a galaxy or some other extended object at redshift $z_{\rm em}$ and it subtends an angle on the sky $d\theta$

- how can we infer from this the physical linear size of the galaxy dl?

to do this, we consider two emission events that happened at the same cosmic time, so $d\tau = 0$, and at the same distance from us, so $d\chi = 0$ an let them lie in the equatorial plane, so $d\phi = 0$.

The metric then tells us that the proper size of the galaxy is $dl = ds$, where

$$ds^2 = -c^2\cancel{d\tau^2} + a(\tau)^2(\cancel{d\chi^2} + S_k(\chi)(d\theta^2 + \sin^2\theta \cancel{d\phi^2}) \tag{50}$$

so $dl = a(z_{\rm em})S_k(\chi(z_{\rm em}))d\theta$.

The factor multiplying $d\theta$ is, by definition, the angular diameter distance, so

$$\boxed{D_{\rm a}(z) = a(\tau(z))S_k(\chi(z))} \tag{51}$$

since, in empty space, an object of proper size dl subtends an angle $d\theta = dl/D$.

If we assume a spatially flat universe ($k = 0$), which is believed to be a good approximation in reality, we have $S_k(\chi) = \chi$ so the angular diameter distance is $D_{\rm a}(z) = a_{\rm em}\chi(z_{\rm em}) = (a_0/(1+z))\chi(z_{\rm em})$ or

$$\boxed{D_{\rm a}(z) = \frac{c}{1+z}\int_0^z \frac{dz}{H(z)}.} \tag{52}$$

This is a very powerful result. It turns out there is a rather accurate 'standard ruler' known as the *'baryon acoustic oscillation'* (or BAO) scale, which is a feature imprinted in the spatial distribution of galaxies and which enables a measure of $D_a(z)$ and hence a powerful test of cosmology (see figure 14). This reinforces the evidence for dark energy and helps determine the cosmological parameters.

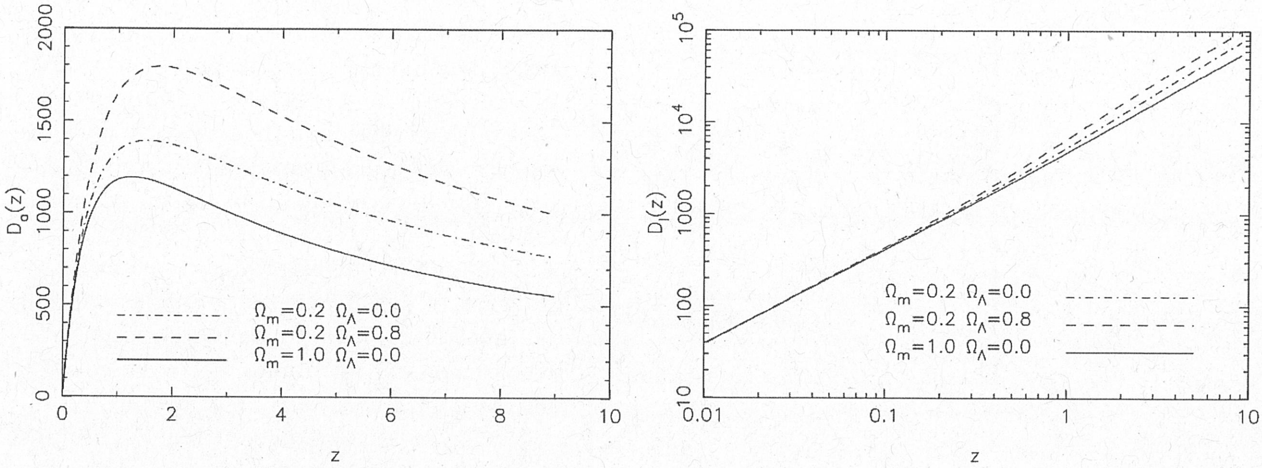


Figure 13: Angular diameter distance (left) and luminosity distance (right) as a function of redshift for various cosmological models.

3.2.4 The luminosity distance

- An analogous apparent distance D_L can be defined that uses flux density F of 'standard candles' of luminosity L
  - flat space : $F = L/4\pi D_L^2 \quad \Rightarrow \quad \boxed{D_L(z) \equiv \sqrt{L/4\pi F}}$
- to calculate $D_L(z)$ we consider a source located at $\chi = 0$
- and co-moving observers on a spherical shell at distance χ (see figure 15)

- let the source emit N photons of frequency $\nu_{\rm em}$ per (rest-frame) period $\tau_{\rm em} = 1/\nu_{\rm em}$
- that means $L_{\rm em} = Nh\nu_{\rm em}/\tau_{\rm em} = Nh\nu_{\rm em}^2$
- conservation of photons implies that N photons cross the shell per red-shifted period $\tau_{\rm obs} = (1+z)\tau_{\rm em}$
- and these photons have energy: $h\nu_{\rm obs} = h\nu_{em}/(1+z)$
- so the *energy flux* (energy per unit time) across the surface is $L_{\rm obs} = Nh\nu_{\rm obs}/\tau_{\rm obs} = Nh\nu_{\rm obs}^2$ or
  - $\boxed{L_{\rm obs} = L_{\rm em}/(1+z)^2}$
- but area is $A = 4\pi a_0^2 S_k^2(\chi) = 4\pi((1+z)a_{\rm em})^2 S_k^2(\chi)$
- so the *energy flux density* F (energy per unit time per unit area) is $F = L_{\rm obs}/A = (1+z)^{-4}/4\pi S_k^2(\chi)a(\tau_{\rm em})^2$ and hence
  - $\boxed{D_L = a_{\rm em} S_k(\chi_{\rm em})(1+z)^2}$
- or, comparing with the angular diameter distance, $D_{\rm a} = a_{\rm em} S_k(\chi_{\rm em})$ we have
  - $\boxed{D_L = D_{\rm a}(1+z)^2}$
- aside:
  - if the source has size dl then it subtends a solid angle $d\Omega \sim dl^2/D_{\rm a}^2$
  - while the intensity is $I \sim F/d\Omega \sim (L/D_L^2)/d\Omega$
  - this, for a given object (fixed L, dl) the surface brightness must vary with redshift as $I \propto D_{\rm a}^2/D_L^2 \propto (1+z)^{-4}$

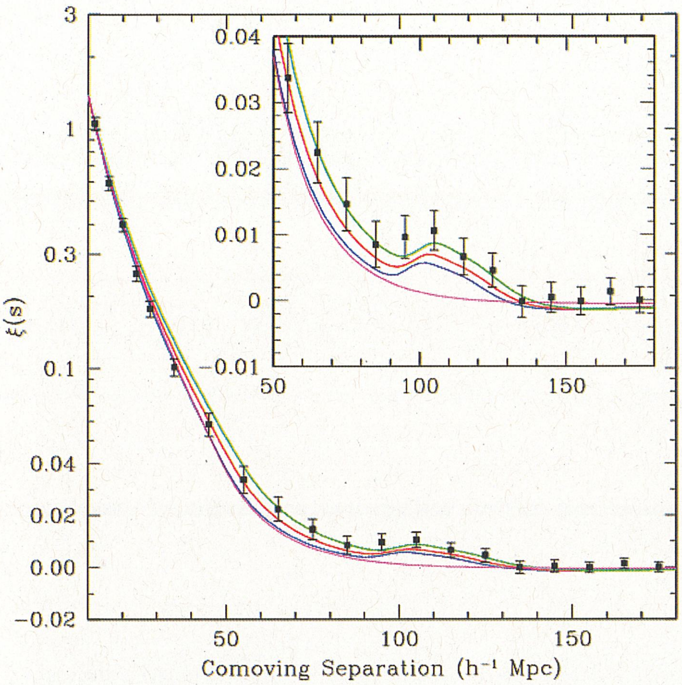


Figure 14: Baryon acoustic oscillations. In the conventional model for the origin of cosmological structure, the 'seeds' of structure were laid down during inflation in the very early universe, with a nearly scale-invariant spectrum. During the period immediately before the plasma reionized these triggered sound waves in the plasma-radiation fluid. This imprinted a feature in the spectrum of density fluctuations that emerged and later developed into fluctuations in the observed large-scale structure traced by galaxies. The feature is rather weak, but it is highly valuable as its scale is predicted from CMB observations (it is essentially the 'sound-horizon'; the product of the sound speed and the age of the universe at the decoupling epoch). Massive redshift surveys were able to measure this feature by means of the 'two-point' function characterising the galaxy distribution at left. This provides a powerful constraint on the cosmological model.

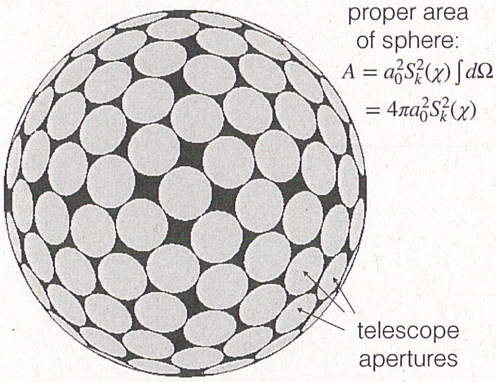


Figure 15: To calculate the luminosity distance we consider a source at the origin of our coordinate system $\chi = 0$ and we consider a sphere at distance χ covered with observers, each of whom would observe the source to have the same flux density and redshift. If the source emits N photons of frequency ν per period $\tau = \nu^{-1}$ (so the proper luminosity is $L = Nh\nu^2$) then conservation of photons implies that there must be N photons of energy $h\nu/(1+z)$ crossing the sphere per red-shifted period $\tau = (1+z)\nu^{-1}$. The energy flux density is therefore $F = (L/(1+z)^2)/A = L/4\pi a_0^2 S_k(\chi)^2(1+z)^2$ and so the luminosity distance is $D_L(z) \equiv \sqrt{L/4\pi F} = a_0 S_k(\chi(z))(1+z)$.

- so the bolometric intensity (or surface brightness) suffers a $(1+z)^{-4}$ *'surface-brightness dimming'*
- this is consistent with Liouville's theorem, which says $I_\nu/\nu^3 =$ constant along any ray since if we integrate this we get a bolometric surface brightness $I = \int d\nu I_\nu \propto \nu^4 \propto (1+z)^{-4}$

- taking the logarithm of $D_L/10\text{pc}$ (and multiplying by 5) we get the predicted *distance modulus* $m - M$ which can be compared with data for 1a SN (see figure 16)

3.2.5 The deceleration parameter

- FRW models predict linear relation $D \propto z$– i.e. Hubble's law – for $z \ll 1$
  - all observers perceive themselves to be at the 'centre of the universe'
  - analogous to ants on an expanding balloon
  - key observable is 'Hubble parameter' H_0 – essentially the inverse of the age of the universe
- going to small but finite z we start to probe departures from linearity
  - at lowest order this is parameterised by the *deceleration parameter*
    * $\boxed{q_0 \equiv (-a\ddot{a}/\dot{a}^2)_0}$
  - This led Alan Sandage, who studied with Hubble, to famously state that *'cosmology is the search for two numbers'*, these being

1. how fast the universe is expanding (H_0)
2. and how fast that expansion is slowing down (q_0)

- in Sandage's time, even the first was quite uncertain and the second largely a matter of speculation
- this changed at the end of the '90s when two groups (led by Saul Perlmutter and by Brian Schmidt) obtained the famous *'type 1a supernova Hubble diagram'* shown in figure 16
  - this leap forward was the result of careful 'standardisation' of the 'candles' in question
  - such supernovae taken as whole actually having a range of intrinsic luminosities
  - but that variation, it turns out, is a function of colour and of the duration of the supernova and so can be corrected for

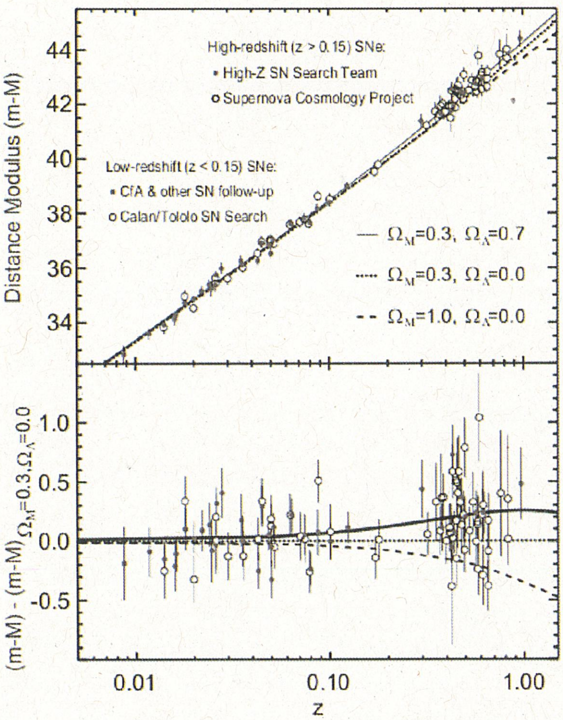


Figure 16: Hubble diagram for type 1a supernovae (contains data shown earlier, but extended here to higher redshift sources). The background to these Nobel prize winning observations is the following: By the '80s there was strong evidence that there was significant non-baryonic dark matter: much more than the roughly 5% of critical density for normal matter allowed by big-bang nucleosynthesis. At the same time, the idea that inflation predicted that the universe should have closure density and be spatially flat had firmly taken hold. For a while the CDM model with $\Omega_{\rm CDM} \simeq 1$ seemed the natural solution. But there were various problems: the predicted age of the universe was uncomfortably short; the dynamical evidence was for $\Omega_{\rm m} \simeq 0.3$, not 1; and evolution of galaxy clusters seemed slower than predicted. At the same time observations of the CMB were very hard to reconcile with an open universe with $\Omega_{\rm m} \simeq 0.3$ as the negative curvature would make the predicted scale of the ripples too small. Adding Λ to the cosmic mix, while repugnant, was becoming widely promoted. The clincher came with the high-z SN data that indicated that the universe was accelerating.

- Figure 16 shows that the supernovae at high-z have a greater apparent distance than predicted in cosmological models containing only normal matter and the observed distance is better fit by models with a cosmological constant Λ (or with 'dark energy') with $\Omega_\Lambda \simeq 0.7$.
  - the present-epoch deceleration parameter in these models is *negative*:
  - the dark energy is causing the expansion rate to *increase* with time
  - i.e. we inhabit an *accelerating universe*
- why acceleration causes an increase in apparent distance can be understood as follows
  - all distances (comoving/conformal, luminosity, angular diameter) involve $\chi(z) = (c/a_0) \int dz/H(z)$
  - and in flat models (a class to which we believe our universe closely approximates) they are simply proportional to this times factors of $1 + z$
  - imagine we compare two models
    1. a 'fiducial' model with some expansion history $H_1(z)$
    2. a 'relatively accelerating' model with idential expansion history for z below some redshift $z_\star$ but $H_2(z) < H_1(z)$ for $z > z_\star$
  - evidently, the relatively accelerating model will have greater apparent distances for sources at $z > z_\star$

3.3 The closed FLRW models

The closed model is finite, yet has no boundary. However, at least if we restrict attention to zero-pressure equation of state, we are free to take only a finite part of the total solution $\chi < \chi_{\max}$. This is a spherically symmetric mass configuration, and so should match onto the Schwarzschild solution for a point mass m, for which the space-time metric is (in units such that $c = G = 1$)

$$-d\tau^2 = -\left(1 - \frac{2m}{r}\right) dt^2 + \left(1 - \frac{2m}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2). \tag{53}$$

Comparing the angular part of the metric it is apparent that the Schwarzschild radial coordinate r and the FLRW 'development angle' χ are related by $r = a\sin\chi$. Now a particle at the edge of the FLRW model can equally be considered to be a radially moving test particle in the Schwarzschild geometry. We found that in Schwarzschild geometry, the normalisation of the 4-velocity for a radially moving particle implies

$$\left(\frac{dr}{d\tau}\right)^2 = \frac{2Gm}{r} + \text{constant}. \tag{54}$$

Compare this with the energy equation

$$\left(\frac{da}{d\tau}\right)^2 = \frac{2GM}{a} + \text{constant} \tag{55}$$

where we have defined the mass parameter $M = 4\pi\rho a^3/3$. With $r = a\sin\chi$ this implies that the Schwarzschild mass parameter is

$$m = M\sin^3\chi. \tag{56}$$

This is interesting. For $\chi \ll 1$, the mass increases as χ^3 as expected. However, the mass is maximized for a model with a development angle of $\pi/2$, or half of the complete closed model. If we take a larger development angle, and therefore include more proper-mass, the Schwarzschild mass parameter decreases. To the outside world, this positive addition of proper mass has negative total energy. This means that the negative gravitational potential energy outweighs the rest-mass energy. The gravitating mass shrinks to zero as $\chi \to \pi$. Evidently a nearly complete closed model with $\chi = \pi - \epsilon$ looks, to the outside world, like a very low mass, that of a much smaller closed model section with $\chi = \epsilon$.

The total energy of a complete closed universe is therefore zero. Zel'dovich, and many others subsequently, have argued that this is therefore a natural choice of world model if, for instance, one imagines that the Universe is created by some kind of quantum mechanical tunnelling event. To be consistent with the apparent flatness of the Universe today one would need to assume that the curvature scale has been inflated to be much larger than the present apparent horizon size.

It is interesting to compare the external gravitational mass with the total proper mass. The volume element of the parallelepiped with legs $d\chi$, $d\theta$, $d\phi$ is

$$d^3x = (a d\chi) \times (a\sin\chi d\theta) \times (a\sin\chi\sin\theta d\phi), \tag{57}$$

so the total mass interior to χ is

$$M_{\text{proper}} = \rho a^3 \int_0^\chi d\chi\ \sin^2\chi \int d\theta\ \sin\theta \int d\phi = \frac{3}{2}M\left[\chi - \frac{\sin 2\chi}{2}\right] \tag{58}$$

The gravitational mass (56) and proper mass (58) are shown on the left in figure 17.

One can make an embedding diagram for this combined FLRW + Schwarzschild space-time. This is shown pm the right in figure 17.

These partial closed FRW models start from a singularity of infinite density and then expand, passing through the Schwarzschild radius $r = 2Gm/c^2$. With $r = a\sin\chi$, $m = M\sin^3\chi$, and $a = M(1 - \cos\eta)$, particles on the exterior cross the Schwarzschild radius at conformal time η when $1 - \cos\eta = 2\sin^2\chi$. For $\chi \ll 1$, this occurs when $\eta = 2\chi$. Such solutions spend the great majority of their time outside the Schwarzschild radius. For the case $\chi = \pi/2$ — i.e. half of the complete solution — the exterior particles just reach the Schwarzschild radius. It may seem strange that the matter in these models can expand from

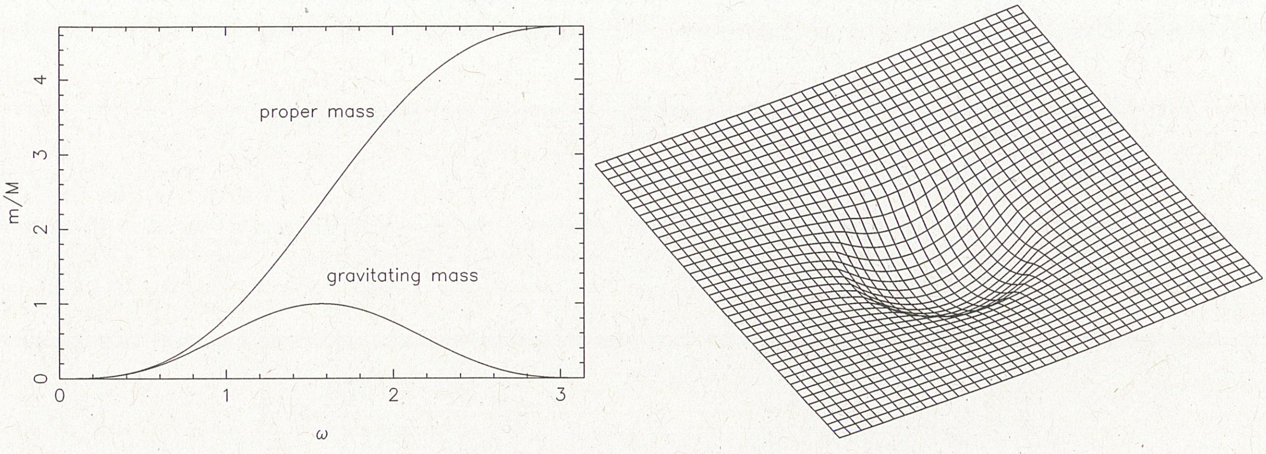


Figure 17: On the left, the proper-mass and gravitational mass for a partial closed FRW cosmology surrounded by Schwarzschild geometry are plotted against the development angle $\omega = \chi$. On the right is shown an embedding diagram. The interior is part of a sphere and the outside is like a trumpet horn. Here the 'development angle' χ is less than $\pi/2$. For $\chi > \pi/2$ we have more than half of the total FLRW model and the radius r is decreasing with increasing χ. It is still possible to match on to an exterior Schwarzschild geometry, but the embedding diagram then has a 'throat'.

Figure 18: The fact that a – potentially large – closed universe can be matched onto an external universe is the physics behind the cartoon shown at the left which is meant to indicate how in the early universe a multitude of universes could 'herniate' from a parent universe – or from each other in Andrei-Linde's 'chaotic inflation' models.

within the Schwarzschild radius, but this is indeed the case. If one considers only the collapsing phase of these models then one has the classic model for black-hole formation as developed by Oppenheimer and Snyder. The spherical mass collapses to a point, and photons leaving the surface can only escape to infinity if they embark on their journey while the radius exceeds the Schwarzschild radius. The expanding phase of these models is just the time reverse of such models; what we have is a 'white-hole' solution The initial singularity is visible to the outside world (eventually) just as photons from the outside can fall in to the final singularity.

A The Friedmann equation from the Einstein field equation

It is convenient to use the (τ, r, θ, ϕ) metric in the form

$$ds^2 = -d\tau^2 + a(\tau)^2 \left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \right) \tag{59}$$

from which we find that the non-vanishing Christoffel symbols are

$$\begin{array}{ccc} \Gamma^0{}_{jk} = \frac{\dot{a}}{a} g_{ij} & \Gamma^j{}_{0k} = \frac{\dot{a}}{a}\delta^i_k & \Gamma^r{}_{rr} = \frac{kr}{1-kr^2} \\ \Gamma^r{}_{\theta\theta} = -r(1-kr^2) & & \Gamma^r{}_{\phi\phi} = -r(1-kr^2)\sin^2\theta \\ \Gamma^\theta{}_{r\theta} = \frac{1}{r} & \Gamma^\theta{}_{\phi\phi} = \sin\theta\cos\theta & \Gamma^\phi{}_{\theta\phi} = \cot\theta \end{array} \tag{60}$$

The non-vanishing components of the Ricci tensor are then found from

$$R_{\mu\nu} = R^\alpha{}_{\mu\alpha\nu} \tag{61}$$

and

$$R^\alpha{}_{\mu\beta\nu} = \Gamma^\alpha{}_{\mu\nu,\beta} - \Gamma^\alpha{}_{\mu\beta,\nu} + \Gamma^\alpha{}_{\gamma\beta}\Gamma^\gamma{}_{\mu\nu} - \Gamma^\alpha{}_{\gamma\nu}\Gamma^\gamma{}_{\mu\beta} \tag{62}$$

to be

$$\begin{aligned} R_{\tau\tau} &= -3\frac{\ddot{a}}{a} \\ R_{rr} &= (a\ddot{a} + 2\dot{a}^2 + 2k)/(1-kr^2) \\ R_{\theta\theta} &= (a\ddot{a} + 2\dot{a}^2 + 2k)r^2 \\ R_{\phi\phi} &= (a\ddot{a} + 2\dot{a}^2 + 2k)r^2\sin^2\theta \end{aligned} \tag{63}$$

from which the Ricci scalar is

$$R = g^{\mu\nu}R_{\mu\nu} = 6(a\ddot{a} + \dot{a}^2 + k)/a^2. \tag{64}$$

The stress energy tensor for a homogeneous universe containing dust or radiation can be taken to be that of a perfect fluid

$$T_{\mu\nu} = (\rho + P)U^\mu U^\nu + g^{\mu\nu}P \tag{65}$$

where the 4-velocity is that of a *'fundamental observer'* who is moving with the fluid and has 4-velocity $U^\mu = (1,0,0,0)$ (this is properly normalised as $g_{\tau\tau} = -1$).

The Einstein field equations can be written as

$$R_{\mu\nu} = 8\pi G(T_{\mu\nu} - \tfrac{1}{2}g_{\mu\nu}T) \tag{66}$$

The $\tau\tau$ component of which is the *acceleration equation*

$$\boxed{\ddot{a} = -\frac{4\pi G}{3}(\rho + 3P)a} \tag{67}$$

which is the same as what we found for the geodesic deviation equation in the centre of a star.

Any one of the spatial equations gives $a\ddot{a} + 2\dot{a}^2 + 2k = 4\pi G(\rho - P)$ which, with the acceleration equation gives the *energy equation*

$$\boxed{\dot{a}^2 = \frac{8\pi G}{3}\rho a^2 - k.} \tag{68}$$

which are the Friedmann equations.

Differentiating the latter and combining with the former gives the *continuity equation*

$$\boxed{\dot{\rho} = -3\frac{\dot{a}}{a}(\rho + P)} \tag{69}$$

which can also be obtained as the $\mu = \tau$ component of the energy momentum conservation law $T^{\nu\mu}{}_{;\nu} = 0$ and is, in essence, the 1st law of thermodynamics as it says that the rate at which a volume element is losing mass-energy is equal to the PdV work it is doing in the process of expansion.

One could work backwards from the latter, using the acceleration equation to obtain the energy equation. But while that would tell you that $\dot{a}^2 - (8\pi G/3)\rho a^2$ is constant it would not pin down the value of k. One can, however, appeal to the Milne model, which is the limiting case of the FRW models as $\rho \to 0$ and for which $k = -1$ and $a = \tau$ to make the connection between the curvature constant and the (minus) sign of the 'total energy' $\dot{a}^2 - (8\pi G/3)\rho a^2$.

"引力红移"能"退相干"吗？

引力红移的规定 随机过程. 光子的坐标了？
路径上. "随机引力场涨落".

" decoherent time ".

麦克斯韦是对的.

辐射场

经典 辐射场 (coherent state).