Chapter One: World-Models

Add Ch ?
p. option?

CHAPTER ONE: World-Models

1.1 Basic Observations and Assumptions

Modern cosmology rests on a small number of fundamental observations and assumptions. The observations are:

1) The sky is dark at night.

2) The velocities of recession of the galaxies increase with increasing distance of the galaxy.

3) The distribution of galaxies is statistically isotropic on linear scales greater than about 50 Mpc.

4) The Universe appears to be filled with an almost-isotropic bath of radiation whose spectrum is close to that of a black-body at a temperature $T \sim 2.7$ K.

The assumptions are:

a) Not only are we not privileged observers by result of any special location in space (the Copernican Principle), but we are also typical observers, whose data on the large-scale appearance of the Universe are at any time statistically the same as those of any other observer who moves with the mean velocity of the galaxy population in his vicinity. This more sweeping assumption is called the Cosmological Principle.

b) The known laws of physics, particularly those of General Relativity, apply throughout the Universe in all places and at all times.

1.1.1 The dark night sky

If a Universe were uniformly populated with n sources of radiation per unit volume, of average luminosity L, then the number dN of sources in a spherical shell, radius r, thickness dr, around an observer would be given by:

$$dN = n.4\pi r^2.s(r).dr$$

where s(r) is a factor which specifies the departure of geometry from Euclidean. The brightness dB at the centre of the shell due to these sources would be:

$$dB = dN.L/4\pi r^2.s(r) = nLdr$$

As this is independent of the radius of the shell, the total observed brightness B due to all the sources in shells with radii from 0 to R is

$$B = \int_{r=0}^{r=R} dB = nLR.$$

This diverges with R, i.e. the brightness of the sky in a uniform, unbounded Universe would itself be unbounded in this approximation. This conflicts violently with the observation that the night sky is dark.

If we allow for the fact that nearby sources would screen more distant ones, the expected brightness becomes that of an enclosure whose walls are the surfaces of the sources - i.e. the expected brightness is not infinite, but at visible wavelengths, for example, would be that of a typical stellar photosphere; the night sky should be as bright as the Sun, still in gross conflict with observation. If we postulate that the space between the sources is filled with absorbing matter which causes the brightness from distant sources to fall off faster than $1/r^2s(r)$ then we can limit the expected brightness only if the Universe is sufficiently young that the absorbers have not yet come to equilibrium with the radiation from the sources, but remain at a temperature much less than that of a stellar photosphere (or if the 'absorbers' are in a sense 'radiation sinks' which never come to radiative equilibrium).

The postulate of a young Universe, needed for a successful 'absorber' explanation of the dark night sky, would however be sufficient by itself to remove the problem. If the Universe is of age T, then the maximum value of R from which we receive radiation cannot exceed $R_{max}=cT$, and B is therefore small if T is sufficiently small. Alternatively, we might postulate that R is limited by a physical boundary to the Universe - that for $R>R_{max}$ there are no more sources; or, less dramatically, that n or L is a function of r which falls off something like $\exp(-r/h)$ where h is a characteristic scale height. In this case B would be limited to nLh.

The darkness of the night sky thus tells us either that a) the Universe is young, or b) the Universe is finite or c) the Universe has a radiation sink which does not come to thermodynamic equilibrium, as would a piece of dark absorbing matter, or d) that the mean properties of the Universe are changing with time so that L is greatest for the nearest sources (whose light takes the shortest time to reach us).

1.1.2 The recession of the galaxies

'Hubble's Law' is now so famous that it is easy to forget that he was not the first astronomer to obtain galaxy spectra or to correlate them with distance. By 1925 V.M.Slipher of the Lowell Observatory had determined 40 of the 45 known radial velocities of what were then called 'white spiral nebulae' from the Doppler shifts of the absorption lines in their spectra. As early as 1918, C.Wirtz attempted to use galaxy spectra to determine the motion of the Sun with respect to the reference frame of the nebulae, but abandoned this attempt because the velocities correlated less with direction in the sky than with the diameters of the galaxies. He noted that the radial velocity was greater the smaller the galaxy's angular diameter and that there was an excess of positive (recessional) radial velocities. In 1925 Lundmark proposed a velocity-distance law in the form

$$V = 513 + 10.365r - 0.047r^2 \text{ km/s}$$

where r was the galaxy's distance in units of M31 distances (this law describes a maximum V = 2250 km/s at 110 M31 distances).

In 1929 Hubble (Publications of the National Academy of Sciences, 15, 168 (1929)) established 'a roughly linear relation between velocities and distance among nebulae for which velocities have previously been established (mostly by Slipher) ... (which) appears to dominate the distribution of velocities.' In the immediately preceding paper, the cosmologist H.P.Robertson points out that a linear velocity-distance reltion is to be expected in some classes of world-model derived from General Relativity.

Hubble and M.K.Humason, working at Mount Wilson, determined the velocity-distance constant H (v = Hd) to be 500 km/s/Mpc. By 1930 Shapley's work on Cepheid variables had raised this to 558 km/s/Mpc. Baade's discovery of errors in the Cepheid calibration reduced this to 180 km/s/Mpc. In the 1950s, Sandage at Palomar showed that what Hubble had identified as the 'brightest stars' in his more distant galaxies were in fact HII regions; the consequent recalibration reduced H to 75 km/s/Mpc. By 1970, a variety of different distance indicators for the more distant galaxies then available yielded values for H between 45 and 150 km/s/Mpc, but many 'clustered' around 100 km/s/Mpc. This value has been adopted in much astronomical work as a working assumption. In 1975, Sandage and Tammann published results of an extensive study from which they obtained values of H near 50 km/s/Mpc.

Although 'best estimates' of H have decreased by a factor of more than ten in the last fifty years, and there is still controversy over the value of H in the range 50-100 km/s/Mpc (which we will review later), it is important to realise that all the uncertainty is in the scale of d. The local linearity of the v-d relation is now well established, over a range of distances approaching 100 times that in Hubble's (Slipher's) original data. The galaxies are definitely engaged in a systematic motion of mutual recession, whose linearity with distance negates any notion that our Galaxy has special status - linear expansions appear centred on a co-moving observer wherever he is located in them. This beautiful result is the backbone of modern cosmology.

1.1.3 The isotropy of the galaxy distribution

Some of the more conspicuous galaxies, notably M31 and M33, are much closer to the Milky Way than are other relatively nearby galaxies. The Milky

Way is situated near one edge of a slightly flattened loose cluster of galaxies about 800 kpc across called the Local Group. About 20 Mpc away in the direction of the constellation Virgo is a cluster of thousands of galaxies, of which the Local Group may be an outlying satellite cluster. The distribution of galaxies on the sky out to distances of order 50 Mpc is strongly influenced by such clustering, and by possible super-clustering, but on larger scales the distribution of galaxies appears statistically isotropic, once allowance is made for the effects of interstellar obscuration in the Milky Way. The statistics of galaxy counts to given levels of faintness in different directions cannot yet be said to demand large-scale isotropy in the Universe, but they are certainly compatible with it, as are the statistics of counts of clusters of galaxies, and of the very luminous and distant radio-galaxies and quasars. As we shall see in some detail later, the microwave background ('relic') radiation is also highly isotropic. Various avenues of astronomical data thus agree that the large-scale Universe is probably isotropic.

1.1.4 The microwave black-body background radiation

After accounting for all local and discrete astronomical sources of microwave radiation, there is a residual smooth 'background' of radiation whose spectrum is that of a black body at a temperature near 2.7 K. This remarkable fact, discovered in 1964 by A.Penzias and R.Wilson of Bell Labs in Holmdel, New Jersey, earned its discoverers the Nobel Prize for Physics in 1978 on the grounds of its cosmological importance. It shows that, in all directions through our present Universe, we encounter radiation that has been equilibrated with matter at the same apparent temperature - i.e. that the past thermal history of matter has a common feature in all directions. It does not by itself tell us the state of the matter, when or at what local (proper) temperature the radiation came to equilibrium, but it does strongly favour those classes of world-model which impose a common thermal history on all matter on the large scales. The exploding 'big-bang' models are of this kind. As we shall see later in Chapter Four, the presence of the microwave radiation background does not prove that 'big-bang' models are correct, but it is readily consistent with them.

1.1.5 The cosmological assumptions

The two basic assumptions referred to earlier both embody matters of philosophy rather than matters of science that are open to experimental verification. Since the Copernican Revolution we have realised that the Earth is not spatially privileged in the Solar System, nor is the Solar System in the

Milky Way galaxy, or the Milky Way in the Local Group or the Virgo Supercluster. This lack of privilege does not however justify the assumption that all cosmologically-relevant observations that we make are statistically typical of those of hypothetical observers who move with the velocities of the Hubble flow. We refer to such hypothetical co-moving observers as Fundamental Observers (FOs); the Cosmological Principle asserts that there should be an infinite set of FOs, whose data on the large-scale properties of the Universe are statistically identical at any given time.

There are in fact some observations, to be described later, which hint that the motion of our galaxy relative to the local Hubble flow is not quite negligible. We really have no alternative to assumption (a) however, as we have only one Universe to observe and one place and time to observe it from. We can only hope that if we are not strictly ideal FOs, then the deviations of our data from those of FOs are merely random perturbations that will not lead us to serious misinterpretations of the large-scale order around us.

Assumption (b) - the universality of our physics - has obvious dangers. We have established our detailed 'laws of physics' from experiments whose space-time limits are usually only a tiny fraction of the space-time domain we wish to model. The alternatives to assumption (b) are paralysis and speculation, however. The best we can do is to attempt to interpret the largest-scale phenomena using known physics until forced by the data to resign this effort in favour of new ad hoc physical constructs. At the same time, we must watch for ways in which our conclusions might be invalidated by effects that we cannot gauge locally - e.g. weak but long-range forces other than gravity, or the existence of subatomic particles that we have not yet encountered in terrestrial experiments.

1.1.6 Isotropy and homogeneity

In its fullest statement, the Cosmological Principle asserts that there can be an infinite set of possible FOs, all of whom would observe a statistically isotropic and homogeneous Universe. Isotropy means that the Universe will appear statistically the same to any FO whichever direction in his sky he looks, while homogeneity means that every FO will construct the same time-history of the Universe. Homogeneity implies the existence of a Universal scale of cosmic time, on which every observer could characterise large-scale evolution using the same functions f(t) to describe the same universal variables (e.g. the mean number density of galaxies). All FOs could in principle synchronise their clocks by sequencing distinguishable large-scale events in the Universe to a common time scale.

1.2 Cosmic Kinematics and Dynamics

1.2.1 Neo-Newtonian model-making

Cosmological models are properly developed within the framework of General Relativity (GR), which describes the observable phenomena in terms of particles with fixed co-ordinates in an expanding curved space. Milne and McCrae pointed out in two classic papers (Quarterly Journal of Mathematics, 5, 64 and 73 (1934)) that it is always open to the observer to choose to describe the observable phenomena using static Euclidean space. A description in this framework must use concepts of force and action at a distance that are avoided in GR, but such a description is not a priori less valid. GR has allowed the prediction of phenomena which are difficult to incorporate in a non-relativistic formulation, but it turns out that many of the observable cosmological phenomena are not 'relativistic' in this sense. Indeed, Milne and McCrae demonstrated that the defining equation of the GR world-models, the Friedmann-Lemaitre Equation, can be obtained from a standpoint in which space is Euclidean, time and gravitation are Newtonian, and the equivalences between observers postulated by GR are paralleled by the equivalence of the FOs postulated by the Cosmological Principle. In particular, the Newtonian 'absolute time' is paralleled even in the relativistic models by the 'cosmic time' implied by the postulate of homogeneity.

The merit of the neo-Newtonian approach is that it shows the essential features of the main classes of cosmological world-model without burying the underlying physics in the mathematical complexities of tensor calculus. Its ability to provide the basic Friedmann-Lemaitre Equation (albeit with somewhat different intepretation of the parameters) stems from the fact that the Cosmological Principle constrains even the GR models to be of a kind where the global behaviour is seen in any limited region; and in sufficiently limited regions (over which the relative velocities in the Hubble flow are small) the GR models must asymptotically be the same as the Newtonian.

1.2.2 Kinematics of the models

The Universe is to be pictured as a 'cosmic fluid' whose properties describe the spatially-averaged properties of the real Universe of galaxies, stars, etc. Each particle of the fluid has a velocity $\underline{v}(\underline{r})$; observers whose velocities are those of the fluid at their position $\underline{r}$ (i.e. observers who could be regarded as 'riding' the particles) are the FOs.

Consider one such observer, O. He constructs a co-ordinate system with himself as origin, and observes a particle P in the fluid, assigning it position vector $\underline{r}$. He determines its velocity $\underline{v}_{P/O}$ relative to himself at time t: $\underline{v}(\underline{r},t)$. He also determines the density and pressure of the fluid at P. Over a sample of points P, he can determine the functional forms for

$$\underline{v}_{P/O} = \underline{v}(\underline{r},t) \quad \rho_P = \rho(\underline{r},t) \text{ and } p_P = p(\underline{r},t)$$

A second observer, O', makes the same observations of the same point P. This second observer finds $\underline{v}'(\underline{r}',t)$, $\rho'(\underline{r}',t)$ and $p'(\underline{r}',t)$. The Cosmological Principle demands that the primed quantities are the <u>same functions</u> of their arguments as the unprimed quantities.

Now let the displacement OO' be the arbitrary vector $\underline{a}$ in O's system. Then we can write

$$v_{O/O'} = \underline{v}(\underline{a}) \text{ and } \underline{v}_{P/O'} = \underline{v}_{P/O} - \underline{v}_{O'/O}$$

$$\text{and } \underline{r}' = \underline{r} - \underline{a}$$

From this it follows that

$$\underline{v}_{P/O'} = \underline{v}'(\underline{r}')$$

$$= \underline{v}'(\underline{r} - \underline{a}) \text{ also } = \underline{v}(\underline{r}) - \underline{v}(\underline{a})$$

The Cosmological Principle requires that

$$\underline{v}'(\underline{r} - \underline{a}) = \underline{v}(\underline{r} - \underline{a}), \text{ so that}$$

$$\underline{v}(\underline{r} - \underline{a}) \text{ must } = \underline{v}(\underline{r}) - \underline{v}(\underline{a})$$

This constraint can be satisfied only if $\underline{v}(\underline{r},t) = F(t)\underline{r}$ where F(t) is a tensor whose components do not depend on $\underline{r}$. However, if the model is to be isotropic, F(t) must in fact be a scalar f(t). By similar arguments it follows that the density ρ and pressure p in the fluid cannot be functions of $\underline{r}$ at a given t.

Now concentrate on the expression

$$\underline{v} = f(t)\underline{r} \qquad (1.1)$$

This is the equation of motion of a particle in the fluid (or of an FO).

It can be solved as

$$\underline{r}(t) = \underline{r}_0 \exp \int_{t_0}^{t} f(t)dt = \underline{r}_0(R(t)/R(t_0)),$$

where R(t) is a 'scale factor' for the Universe at time t, and t_0 is an agreed-on 'scaling time' at which all co-ordinates are assigned. Clearly R(t) and f(t) are related as

$$(1.2) \qquad f(t) = \dot{R}(t)/R(t)$$

Equation (1.1) contains the first important property of all isotropic, homogeneous world-models: the only motions consistent with our postulates are those consisting of a uniform scaling of all separations by a scale factor R(t), whose logarithmic time derivative specifies the velocity-coordinate relationship (this is not quite the same as the observed velocity-distance relation, as we will see later). The form of the function R(t) will be different for different models, and our next task is to find what forms R(t) are consistent with known forces, i.e. to find the allowed dynamics of the models.

1.2.3 Classical dynamics of the models

If we fix attention on an individual particle in the fluid, the time derivative D/Dt of any parameter q at the particle is given by:

$$Dq/Dt = \partial q/\partial t + (\underline{v}.\underline{\nabla})q$$

where $\partial/\partial t$ is the Eulerian time derivative (at a fixed position in space) and $\underline{v}$ is the particle velocity. With this notation, the Equation of Continuity which expresses mass-conservation in the fluid is

$$(1.3) \qquad D\rho/Dt = -\rho\,\underline{\nabla}.\underline{v}$$

Expanding the left-hand side and substituting from (1.1) in the right-hand side, we have

$$\partial\rho/\partial t + (\underline{v}.\underline{\nabla})\rho = -\rho.3f, \text{ i.e.}$$

$$d\rho/\rho = -3\ dR/R \quad \text{(using (1.2))}$$

This integrates to

(1.4) $$\rho(t) = \rho_0 R^3(t_0)/R^3(t)$$ mass conservation

where ρ_0 is the value of ρ at the 'scaling time' t_0. Note that this relation will not hold in world-models where mass conservation is violated.

The dynamics of the model can then be obtained using Euler's momentum equation for the fluid:

$$D\underline{v}/Dt + (1/\rho)\underline{\nabla}p - \underline{F} = 0$$

where $\underline{F}$ is the self-force per unit mass in the fluid. We use (1.1) to evaluate the first term:

$$D\underline{v}/Dt = \partial\underline{v}/\partial t + (\underline{v}.\underline{\nabla})\underline{v}$$

$$= \dot{f}\underline{r} + f^2(r\,\partial/\partial r)\underline{r}$$

$$= (\dot{f} + f^2)\underline{r}$$

$$\frac{\partial \underline{v}}{\partial t} + (\underline{v}.\underline{\nabla})\underline{v} + \frac{1}{\rho}\underline{\nabla}p - \underline{F} = 0$$

$$\underline{v} = f\underline{r} \qquad \underline{\nabla}p = 0$$

$$\underline{\nabla}.\underline{F} = -4\pi G\rho$$

As p is not a function of $\underline{r}$ (from the constraints imposed by the Cosmological Principle), $\underline{\nabla}p = 0$.

The computation of $\underline{F}$ in an infinite system cannot be made in Newtonian mechanics. We can either suppose the Universe to be finite (but very large), or we can note that Poisson's Equation gives

$$\underline{\nabla}.\underline{F} = -4\pi G\rho$$

in the Newtonian theory, so that taking the divergence of the Euler equation gives us

$$(\dot{f} + f^2)(\underline{\nabla}.\underline{r}) + 4\pi G\rho = 0$$

Substituting $f = \dot{R}/R$ from (1.2) and using $\underline{\nabla}.\underline{r} = 3$, we have

$$(\ddot{R}/R - \dot{R}^2/R^2 + \dot{R}^2/R^2) = -4\pi G\rho/3, \text{ i.e.}$$

(1.5) $$\ddot{R} = -4\pi G\rho R/3$$

Substituting for ρ in (1.5) using (1.4):

$$\ddot{R} = -4\pi G\rho_0 R^3(t_0)/3R^2$$

which can be written $\ddot{R} = - GM_0/R^2$, where

(1.6) $M_0 = 4\pi\varrho_0 R^3(t_0)/3$

is the mass inside a sphere whose radius is the 'scaling length' $R(t_0)$ at the scaling time t_0. The first integral over time (after multiplying both sides by $2\dot{R}$) is

$$\dot{R}^2 = 2GM_0/R + \text{constant}$$

With intuition born of hindsight, we will call the constant of integration $-kc^2$, to obtain finally:

(1.7) $\dot{R}^2 = 2GM_0/R - kc^2$

Equation (1.7) is identical to the corresponding equation of General Reativity, with the 'cosmological constant' set equal to zero.

1.2.4 The Cosmological Constant Λ

When Einstein began his work on cosmology, it was thought necessary to provide a static model along the lines of the Herschel/Kapteyn Universe (which was based on observations only of our local galactic neighbourhood, before Shapley's work on the globular clusters). It is clear from equations (1.5) and (1.7) that in a static universe, where $\dot{R}$ and $\ddot{R} = 0$, there can be no matter (ϱ_0 must = 0). It was therefore impossible to provide a model containing any matter, if it had to be static. This difficulty of Einstein's is in no way peculiar to GR - if gravity is the only long-range force, then the fact that it is attractive means that all static Universes will collapse unless the masses are infinitely remote from one another. If it had been known at this stage (1916) that the 'white nebulae' were galaxies outside our own, and that they were receding from one another, the 'cosmological constant' might never have been introduced.

Einstein tried to rescue the situation by introducing an additional term which permitted a static Universe to contain matter. The new term appears in GR as a constant of integration which specifies the curvature of empty space-time; relativity theory does not require this term to be zero, and it should in principle be left for experiment to determine. In the neo-Newtonian framework we can perform the equivalent of Einstein's adjustment by introducing a 'cosmic repulsive force' which can counteract gravitation to permit a static model. If we adjust Poisson's Equation to read

$$\underline{\nabla}.\underline{F} = -4\pi G\varrho + \Lambda$$

we are implying a repulsive force $\underline{F} = \Lambda\underline{r}/3$. Dimensionally the new

constant vacuum gravitation source $\varrho_{eff} = -\frac{\Lambda}{4\pi G}$, Vacuum energy density $u_{vac} = \varrho_{eff}c^2 = -\frac{\Lambda c^2}{4\pi G}$ (quantum idea?)

constant is T^{-2}. If Λ is small, the repulsion will be negligible over laboratory distances; the modification thus permits local agreement with experiment while modifying the behaviour of the large-scale Universe very substantially. It is an example of a possible lack of comprehensiveness in our local physics of the kind discussed in Section 1.1.5. It must be remembered that the form of the new term is not as arbitrary in GR as it appears to be in the adjusted Poisson's equation above.

With the new term in place, our earlier equations (1.5) and (1.7) become:

(1.8) $\ddot{R} = (\Lambda R^3/3 - GM_0)/R^2$ and

(1.9) $\dot{R}^2 = 2GM_0/R - kc^2 + \Lambda R^2/3$ Friedmann-Lemaitre

In the GR models, equation (1.9) was first given in 1922 by Alexandre Friedmann, who appears to have been the first to recognise that world models whose spatial properties depend only on the time were compatible with GR. Friedmann was a Russian whose pioneering work in this field was largely ignored in the West, and was originally sharply criticised by Einstein. The French Abbe Georges Lemaitre gave equation (1.9) for a model in which 'the radius of the Universe increases without limit' in a paper published in 1928. In modern times, it has become known as the Friedmann-Lemaitre Equation.

Given a pair of values for Λ and k, the Friedmann-Lemaitre equation specifies a form for R(t) which in turn governs the behaviour of observable quantities in the world-model. We will explore the allowed forms of R(t) and then examine ways in which we can decide which (if any of them) corresponds to the actual Universe. Before carrying out this analysis, however, it is useful to note the meaning of the terms in the Friedmann-Lemaitre equation in the context of the GR models.

1.2.5 The meaning of R and k in models based on General Relativity

In a General Relativistic (GR) formulation, the aggregate of all events constitutes the points of a Riemannian 4-space $ds^2 = g_{\mu\nu}\,dx^\mu dx^\nu$ and the constant k specifies the curvature of the spatial part of the metric. Robertson and Walker showed in 1934 that the postulates of isotropy and homogeneity in the Cosmological Principle require the $g_{\mu\nu}$ to give a metric which can be written:

(1.10) $$ds^2 = c^2dt^2 - R^2(t)\left\{d\sigma^2/(1-k\sigma^2) + \sigma^2(d\theta^2 + \sin^2\theta\, d\phi^2)\right\}$$

where σ, θ, and ϕ are spherical polar event coordinates and t is the cosmic time. The 3-spaces defined by dt=0 expand as R(t) and their curvature depends on the value of k, which in this formulation must be +1, 0 or -1. Obviously

N.B. "curvature of spacetime" is curvature of 2-space $d\theta = d\phi = 0$

$$ds^2 = c^2dt^2 - R^2(t)\,d\sigma^2/(1-k\sigma^2)$$

Gauss formula → curvature $K(t) = -\ddot{R}/R = -4\pi\rho(t)G/3 - \Lambda/3$

k=0 gives the classical flat Euclidean space, in which the proper (dt=0) distance between two FOs at $\sigma=0$ and $\sigma=\sigma_0$ is $r(t) = R(t)\sigma_0$.

The case k=+1 gives a closed space in which the coordinate σ must be in the range $0\leq\sigma<1$. The proper distance between two FOs at $\sigma=0$ and $\sigma=\sigma_0$ is, from (1.10) with dt=0 :

$$r(t) = R(t)\int_0^{\sigma_0} d\sigma/\sqrt{(1-\sigma^2)} = R(t)\sin^{-1}\sigma_0$$

This is ambiguous to within $2n\pi R(t)$, so we may regard R(t) as the radius of a closed Universe in this model.

The case k=-1 gives an open space in which σ can run from zero to infinity, but whose geometry is non-Euclidean (Bolyai-Lobatchewski space). In this case, the proper distance between observers at coordinates $\sigma=0$ and $\sigma=\sigma_0$ is:

$$r(t) = R(t)\sinh^{-1}\sigma_0$$

It will be convenient to write the above three expressions for the proper distance in the general form

(1.11) $r(t) = R(t)S_k(\sigma_0)$, where

$$S_k(\sigma) = \begin{cases} \sin^{-1}\sigma & (k=+1) \\ \sigma & (k=0) \\ \sinh^{-1}\sigma & (k=-1) \end{cases}$$

Note that the coordinates σ, θ and ϕ are fixed for any FO for all time (they are described as co-moving coordinates). They specify the relative configuration of the FOs which is maintained throughout the development of the model. The evolution of the model in time comes entirely from the variation in the scale factor R(t). Einstein's field equations applied to the Robertson-Walker metric (1.10) yield the Friedmann-Lemaitre equation.

1.2.6 The meaning of k in neo-Newtonian models

Evidently the role of R(t) in the GR models is quite similar to its role in the neo-Newtonian models of Section 1.2.3. k kas nothing to do with space geometry in the Newtonian models, however, as these models are fundamentally cast in Euclidean space. To see the significance of k in the Newtonian models, consider the case $\Lambda = 0$. If we divide equation (1.7) by $2R^2$ we obtain

(1.12) $\dot{R}^2/2R^2 = 4\pi G\rho/3 - kc^2/2R^2$

Now consider an observer at r=0 examining the motion of matter in a sphere of radius r around himself. A particle of mass m on the outside of the sphere moves with velocity (1.1)

$$\underline{v} = f(t)\underline{r}$$

relative to the observer. He will therefore gauge its kinetic energy to be

$$T = mv^2/2 = mf^2(t)r^2/2 = mr^2\dot{R}^2(t)/2R^2(t)$$

where we have used equation (1.2) to substitute for f(t). The mass M within the sphere is

$$M = 4\pi r^3\rho/3$$

and the gravitational potential energy of the mass m (as judged by this observer) will be

$$V = -GMm/r = -4\pi G\rho r^2m/3$$

We can therefore rewrite the modified Friedmann-Lemaitre equation (1.12) in the form

$$T + V = -kc^2r^2m/2R^2$$

It follows that if k is <u>positive</u>, an observer considers the mass points around him to have (T+V) <u>negative</u>, i.e. to be gravitationally bound; k negative corresponds to the masses around a given observer having more than the 'local escape velocity' and k=0 corresponds to the transition condition where T+V = 0 and the model just comes to rest with the masses infinitely separated. Thus k in the Newtonian models is a measure of the local gravitational binding, while in the GR models it specifies the type of space-geometry. These interpretations are not equivalent, and the similarity between the neo-Newtonian and GR world-models is limited by the fact that all Newtonian models are cast in k=0 <u>space</u> in the terms of the GR formalism. It turns out however that all <u>presently measurable</u> properties of the Universe depend much more on R(t) than on spatial curvature explicitly, so the Newtonian models can provide a useful framework for the visualisation of observational problems in cosmology.

1.3 Classification of World-Models

1.3.1 The static (Einstein) model

Although it is now clearly at variance with the data on the large-scale Universe, the static Einstein model is still of interest as it defines a critical value of the Λ-term. We obtain the Einstein model by putting $\ddot{R} = 0$ in equation (1.8) and $\dot{R} = 0$ in equation (1.9), noting that $R(t) = R(t_0)$ and $\rho = \rho_0$ for this static model. Λ and R must then satisfy

$$(1.13) \qquad \Lambda = 4\pi G\rho \qquad \text{and} \qquad R = c\sqrt{(k/4\pi G\rho)}$$

Note that we must have k positive for this R to be physically meaningful, which in the GR models means that k must be +1 (closed space). We can always choose the scaling factor in the Newtonian version of the model so that k=+1, to obtain correspondence with the GR model.

The values of Λ and R specified by equation (1.13) depend on the density of the model, and appear as critical parameters in some evolving models (see Sections 1.3.6 and 1.3.9 below). We will refer to condition (1.13) as defining critical values $\Lambda = \Lambda_E$ and $R = R_E$.

1.3.2 The model k=0, Λ=0 (Einstein-de Sitter model)

For this case the Friedmann-Lemaitre equation reduces to

$$\dot{R}^2 = 2GM_0/R$$

$$dR/dt = \sqrt{(2GM_0/R)}$$

and we can find a natural zero of cosmic time when R(t) = 0, in which case

$$\int_0^R \sqrt{R}\, dR = \sqrt{(2GM_0)} \int_0^t dt$$

$$2R^{3/2}/3 = t\sqrt{(2GM_0)}, \text{ i.e.}$$

$$R(t) = (2GM_0)^{1/3}(3t/2)^{2/3}$$

$$R(t) = Bt^{2/3}, \quad \text{where } B = (3/2)^{2/3}.(2GM_0)^{1/3}$$

This model 'explodes' at t=0 with finite velocity, then slows to a stop at infinite size. It is the cosmological equivalent of the parabolic-velocity condition in two-body orbits.

1.3.3 The model $k>0$, $\Lambda=0$ (Friedmann-Einstein model)

From equation (1.7), this model has $\dot{R}$ less than the corresponding value in the Einstein-de Sitter model for all $R>0$. It therefore 'explodes' at lower velocity than the Einstein-de Sitter model and its path in the R-t plane lies entirely under that of the Einstein-de Sitter. A maximum value of R is reached when $\dot{R}=0$, i.e. when

$$R = 2GM_0/kc^2$$

and the scale can always be chosen so that $k=+1$. The form of $R(t)$ can be written in parametric form

$$R = (A/2)(1-\cos\theta)\,,\quad t = (A/2)(\theta-\sin\theta)$$

which is the equation of a cycloid. Whether the cycloidal repeat of $R(t)$ corresponds to physical reality will depend on the presumed physics of the Universe near $R(t) = 0$. The model is widely known as the 'oscillating Universe'.

1.3.4 The model $k<0$, $\Lambda=0$ (Hyperbolic model)

For this model, $\dot{R}$ is everywhere greater than in the Einstein-de Sitter, and the model escapes to infinite R with $\dot{R}>0$ in finite time. At very large R, equation (1.7) becomes

$$\dot{R}^2 = -kc^2 \text{ (positive), so that}$$

$$R(t) \rightarrow ct\sqrt{(-k)}$$

This is also an 'exploding' model.

1.3.5 The deceleration parameter q

We have now seen that ALL MODELS WITH $\Lambda = 0$ IMPLY A PAST PATHOLOGICAL STATE OF THE UNIVERSE IN WHICH $R(t) \rightarrow 0$ and $\rho \rightarrow \infty$. The models differ in their rate of deceleration at given R, and can be characterised by the dimensionless <u>deceleration parameter</u>

$$(1.14)\quad q = -\ddot{R}R/\dot{R}^2 = 1/(2 - kc^2R/GM_0)$$

$$= 4\pi G\rho R^2/(8\pi G\rho R^2 - 3kc^2)$$

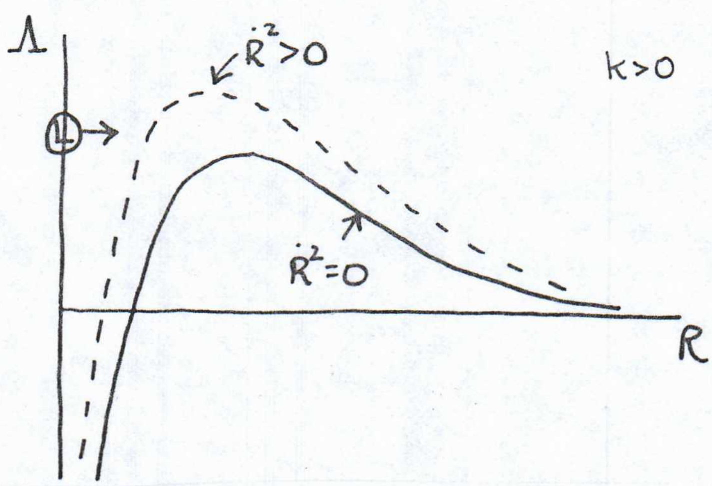
in the $\Lambda = 0$ models. Hence $q = 0.5$ at all R (all t) in the Einstein-de Sitter model, <0.5 in the hyperbolic models and >0.5 in the Friedmann-Einstein oscillating models. Evidently a reliable observation of the present value of q would establish the type of model in which we live, if we knew that Λ=0.

1.3.6 Models with $k>0, \Lambda>\Lambda_E$ (Lemaitre models)

All models with non-zero Λ are best understood by considering the behaviour of $\dot{R}^2$ in a Λ-R plane. Consider the locus of all points with $\dot{R}^2=0$ in this plane when k=+1. The Friedmann-Lemaitre equation is then

$$c^2 = 2GM_0/R + \Lambda R^2/3$$

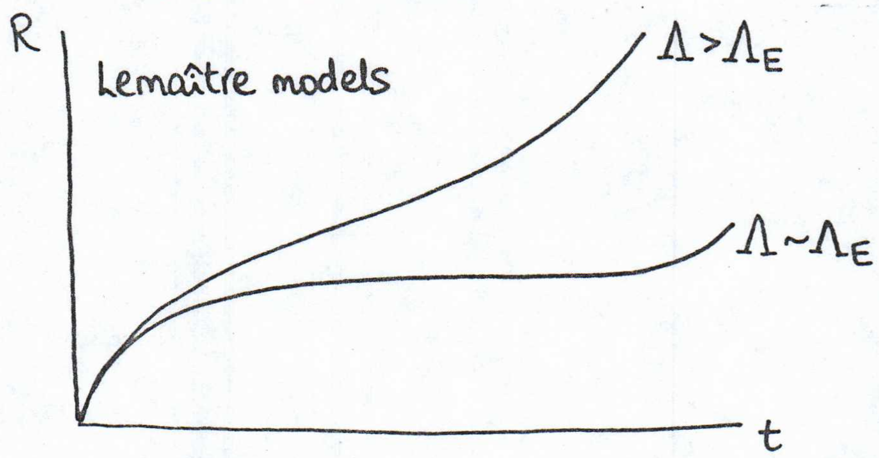
For small values of R, Λ must be large and negative, while for large values of R, Λ must be near zero but positive. It follows that the locus of $\dot{R}^2=0$ must begin at the lower left of the Λ-R plane, cross the Λ-axis at some R, pass through a maximum value of Λ, and then asymptotically approach the R-axis from above (see the diagram which follows)



From the above equation, it can be seen that the axis-crossing at Λ=0 occurs for $R=2GM_0/c^2$ and that the peak on the locus occurs at the value $\Lambda = \Lambda_E$ given by condition (1.13). To find the form of the models using this diagram, we note that the region of the diagram which corresponds to physical reality is the region where $\dot{R}^2>0$, i.e. the region above the $\dot{R}^2=0$ locus, and that a model will have a given starting value of R and $\dot{R}$, and a fixed value of Λ. The loci of the models are thus horizontal lines in the diagram.

The Lemaitre models all lie above the peak of the $\dot{R}^2=0$ locus. They explode from R=0, and so move to the right in the Λ-R plane away from their starting-point (e.g. point L). In so doing, $\dot{R}$ first decreases, reaching a minimum at the closest approach at some R near to but lower than the value of R at the peak of the $\dot{R}^2=0$ locus. After this $\dot{R}$ again increases. The general form of these models is therefore a continuously expanding R(t) which first decelerates under gravity and then accelerates under the Λ-repulsion.

A particularly interesting class of Lemaitre models is the set with Λ only slightly greater than Λ_E. For these models, the minimum value of $\dot{R}$ is very close to zero, so the models spend a very long period of time in an almost-static 'coasting' phase (see below).



1.3.7 Models with $k>0$, $0<\Lambda<\Lambda_E$

These are of two kinds. Models starting with small R are constrained to remain on the left of the $\dot{R}^2=0$ curve (remember that the region under the curve is unphysical and cannot be entered by any model). These models explode, decelerate, stop at the value of R where the horizontal line representing them in the Λ-R plane intersects the locus of $\dot{R}^2=0$, and then fall back symmetrically. They are therefore of same form as the Friedmann-Einstein models with $\Lambda=0$.

A new class of model starts at large R, to the right of the $\dot{R}^2=0$ locus, and contracts with decreasing velocity; these models stop contracting at finite R (again where their horizontal line reaches the $\dot{R}^2=0$ locus), and then re-expand symmetrically. The form of R(t) resembles a catenary, and they are variously known as 'catenary' or 'bounce' models. THEY ARE THE FIRST NON-EXPLOSIVE MODELS WE HAVE YET ENCOUNTERED.

1.3.8 Models with $k>0$, $\Lambda<0$

Negative Λ introduces a long-range attractive force which grows with distance, and all models of this kind resemble the bound Friedmann-Einstein form. They start at small R below the Λ-axis and to the left of the $\dot{R}^2=0$ locus in the Λ-R plane, decelerate and stop at a low value of R, then fall back on themselves.

1.3.9 Models with $k>0$, $\Lambda = \Lambda_E$

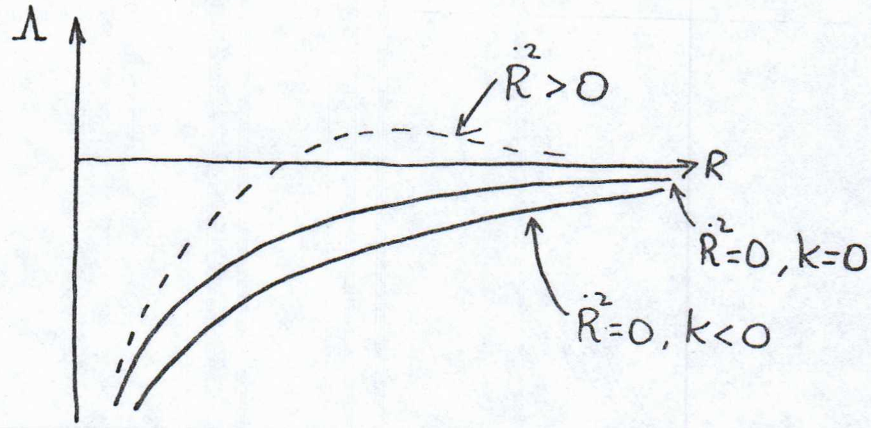
Models beginning at small R will expand at a decreasing rate and eventually stop at the peak of the $\dot{R}^2=0$ locus; at this point they become a static Einstein model. Models beginning at large R will contract at a decreasing rate and stop as a static Einstein model.

The Einstein model is itself unstable to perturbations. It is represented by the point at the peak of the $\dot{R}^2=0$ locus. Any displacement leads either to collapse under gravity or to inflation under the Λ-repulsion. The behaviour of the model after perturbation is exactly that of the two preceding models, with the time sense reversed.

An unique model in this class is the one which begins as an Einstein model but is perturbed towards larger R than its equilibrium value. This model expands at an accelerating rate, and is the only continuously expanding model which does not begin from $R(t)=0$, i.e. the only expanding model which does not originate from a 'Big Bang'. It is known as the Eddington-Lemaitre model.

1.3.10 Models with $k=0$, $\Lambda \neq 0$

With $k=0$, the locus of $\dot{R}^2=0$ in the Λ-R plane does not cross the $\Lambda=0$ axis (see diagram below).



The form of $R(t)$ is now completely specified by the sign of Λ. If $\Lambda>0$, the models are Lemaitre models; if $\Lambda<0$, they are Friedmann-Einstein models.

No 'bounce' models or Eddington-Lemaitre variants are possible.

1.3.11 Models with $k<0$, $\Lambda \neq 0$

The locus of $\dot{R}^2=0$ now lies below that for $k=0$ models, but the form of the models is still specified by the sign of Λ, as in the $k=0$ cases. The velocities are higher at any given R than they are in the $k=0$ cases, but the overall forms of the models are still those of $k=0$.

1.3.12 Summary of model types

We can now review the main forms of the allowed models, with the parameter combinations which lead to them.

a) CONTINUOUSLY EXPANDING 'BIG BANG' MODELS: The hyperbolic ($k<0, \Lambda=0$), Einstein-de Sitter ($k=0, \Lambda=0$) and Lemaitre ($k>0, \Lambda>\Lambda_E$), ($k<0$, $\Lambda>0$) models and the model ($k>0, \Lambda=\Lambda_E$ and $R(0) = 0$).

b) OSCILLATING 'BIG BANG' MODELS: Friedmann-Einstein forms ($k>0, \Lambda<\Lambda_E$), ($k\leqslant 0$, $\Lambda<0$).

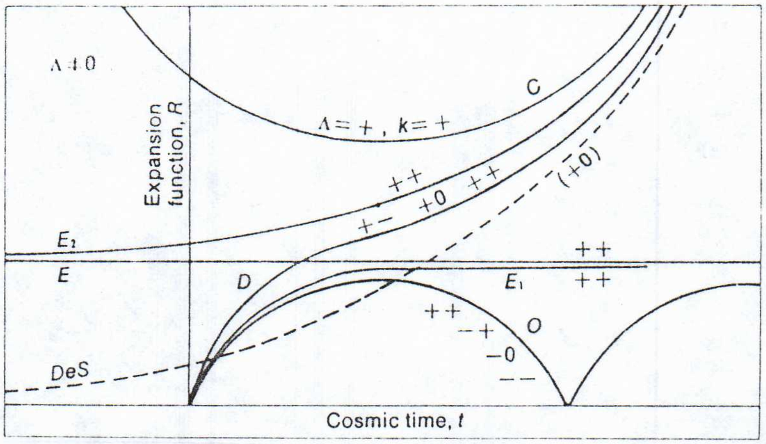
c) CONTINUOUSLY EXPANDING MODELS WITH NO BIG BANG: the Eddington-Lemaitre model ($k>0, \Lambda=\Lambda_E$, $R(0) = R_E$ perturbed outwards)

d) OSCILLATING MODELS WITH NO 'BIG BANG': 'Bounce' models ($k>0$, $\Lambda<\Lambda_E$, $R(0)$ large)

e) CONTINUOUSLY CONTRACTING MODELS: ($k>0$, $\Lambda=\Lambda_E$, $R(0)$ = large or $R(0) = R_E$ perturbed inwards)

f) STATIC MODEL: The Einstein model ($k>0, \Lambda=\Lambda_E$, $R = R_E$).

The variety of $R(t)$ forms for $\Lambda \neq 0$ is illustrated in the diagram below (taken from W.Rindler, 'Essential Relativity', p.264 (1969)).



Only cases (e) and (f) above are immediately inconsistent with the observed recession of the galaxies. To make a detailed comparison with observations, however, we must note that we are not able to observe $R(t)$

directly from an outsider's standpoint, but must extract information about R(t) and its derivatives from observations of the light received from distant objects at a single time of observation. To interpret our observations, we therefore need to examine the propagation of radiation through the different world-models. This problem is the subject of the next Chapter.