Physics Before and After Einstein
M. Mamone Capria (Ed.)
IOS Press, 2005
© 2005 The authors

Chapter 6

The Rebirth of Cosmology: From the Static to the Expanding Universe

Marco Mamone Capria

Among the reasons for the entrance of Einstein's relativity into the scientific folklore of his and our age one of the most important has been his fresh and bold approach to the cosmological problem, and the mysterious, if not paradoxical concept of the universe as a three-dimensional sphere. Einstein is often credited with having led cosmology from philosophy to science: according to this view, he made it possible to discuss in the progressive way typical of science what had been up to his time not less "a field of endless struggles" than metaphysics in Kant's phrase.

It is interesting to remember in this connection that the German philosopher, in his *Critique of Pure Reason*, had famously argued that cosmology was beyond the scope of science (in the widest sense), being fraught with unsolvable contradictions, inherent in the very way our reason functions. Of course not everybody had been impressed by this argument, and several nineteenth-century scientists had tried to work out a viable image of the universe and its ultimate destiny, using Newtonian mechanics and the principles of thermodynamics. However, this approach did not give unambiguous answers either; for instance, the first principle of thermodynamics was invoked to deny that the universe could have been born at a certain moment in the past, and the second principle to deny that its past could be infinite.

Einstein's contribution to cosmology surely strengthened confidence in the power of human reason to investigate every conceivable subject-matter, including the most ambitious ones. It is from Einstein's cosmological writings that physicists drew the kind of boldness that decades later enabled some of them to author books with such titles as *A Brief History of Time* and *The Mind of God*. A different question is whether the road opened by Einstein led in fact to that consensus that most people consider as a trademark of the scientific endeavour. As we shall see (and as will be even more clear from Chap. 13) this is at least doubtful.

The starting point of relativistic cosmology was Einstein's 1917 paper, "Cosmological considerations on the general theory of relativity" [17] where, just one year after the apparently definitive formulation of general relativity, a novelty was introduced at the very core of the theory. Some years later, Einstein somewhat repented of this step, but by that time it had taken on a life of its own. In any case, as we shall see, the development of relativistic cosmology was an essential factor in the progress in the understanding of the power and limits of general relativity itself.

1. Einstein and Newton's Universe

In tackling in 1917 the problem of the structure of the universe, "this fundamentally important question", Einstein started by pointing out that in Newtonian physics the equivalence between the law of universal attraction and Poisson's equation is not strict, but requires suitable conditions at infinity, which in general relativity translate into the assumption that "it is possible to select a system of reference so that at spatial infinity all the gravitational potentials $g_{\mu\nu}$ become constant" [17, p. 177].

Einstein tried to show that Newtonian physics is unable to deal properly with the cosmological problem for an infinite universe. If matter is uniformly distributed in such a universe, then the gravitational potential should be *divergent* at infinity, rather than tend to a finite limit (this is a version of the so-called *gravitational paradox* of Newtonian physics). The only wayout – Einstein went on – is to assume that the mass density is not constant, but goes to zero faster than $1/r^2$ as the distance r from a given 'centre' goes to infinity. So Newton's *material* universe turns out to be essentially finite, and with a centre, though it is contained in an infinite (Euclidean) space.

Interestingly, Newton had already discussed this topic, if from a different angle, and drawn the *opposite* conclusion. In the first of his four letters to the Reverend Bentley (dated 10 December 1692), an important source for the study of his cosmological thought, Newton explained how the different celestial structures which we observe could well have arisen from an initially uniform, *infinite* distribution of matter, while from a finite distribution this development would have been very unlikely:

> As to your first Query, it seems to me that if the Matter of our Sun and Planets, and all the Matter of the Universe, were evenly scattered throughout all the Heavens, and every Particle had an innate Gravity towards all the rest, and all the whole Space, throughout which this Matter was scattered, was but finite; the Matter on the outside of this Space would by its Gravity tend towards all the Matter on the inside, and by consequence *fall down into the middle of the whole Space, and there compose one great spherical Mass*. But if the Matter was *evenly disposed throughout an infinite Space*, it could never convene into one Mass, but some of it would convene into one Mass and some into another, so as to make an infinite Number of great Masses, scattered from one to another throughout all that infinite Space [[5, pp. 281–2]; italics added].

In other words, with an infinite distribution Newton thought that it was possible to avoid a different difficulty facing cosmology, namely that of all matter collapsing into a single big mass – a destiny apparently inconsistent with astronomical observation. Thus Newton was advancing the idea that in an infinite material universe a dynamic equilibrium is possible; that is, it is possible for the gravitational field to vanish everywhere in the large, while locally what is known today as 'gravitational instability' can give rise to stars and stellar systems.

Einstein had also other reasons to reject a Newtonian finite universe. He argued, by means of statistical considerations, that one cannot have a plausible Newtonian cosmology even by assuming conditions at infinity for the gravitational potential, since in that case one has to accept the possibility that some heavenly body could escape to infinity, unless the difference of potentials between 'here' and 'infinity' was big enough. The last possibility Einstein rejected, as inconsistent with the then available observational evidence of *small values for the "stellar velocities"*. Another statistical difficulty of Newtonian cosmology arose if Boltzmann's distribution for the molecules of a gas in thermal equilibrium was applied to the stars.

To solve these difficulties within Newtonian physics, Einstein advanced a suggestion, "which does not in itself claim to be taken seriously", but was meant to introduce a change in his field equations which, on the contrary, he wished to advance quite seriously. If at the left-hand side of Poisson's equation:

$$\Delta\phi = 4\pi G\rho, \tag{1}$$

a term is added in the form:

$$\Delta\phi - \lambda\phi = 4\pi G\rho, \tag{2}$$

where λ is a "universal constant", then the new equation clearly admits a solution ϕ corresponding to a uniform mass distributions ρ_0: namely, the constant potential function $\phi = -4\pi G\rho_0/\lambda$. This change eliminates the need of enforcing conditions at infinity.

The idea of (2) was probably inspired by the modified gravitational potential ϕ_N built out of the mass-point potential $Ae^{-r\sqrt{\lambda}}/r$ (A is a constant), as proposed by the German theoretical physicist Carl Neumann in 1896. Neumann introduced this form of the potential in order to solve the gravitational paradox in the form of the impossibility to assign, *at any point*, a finite value to the gravitational potential corresponding to a uniform infinite mass distribution. In fact ϕ_N goes to zero at infinity even with such a mass distribution, and (2) is precisely the equation which it satisfies.

2. Conditions at Infinity and Mach's Principle

The hypothesis that at *spatial* infinity the $g_{\mu\nu}$ should be Minkowskian in some suitable coordinate system was the very hypothesis Einstein had adopted in his previous papers when dealing with the astronomical problems of the solar system (Chap. 5, §11), but for the universe as a whole such an hypothesis was "by no means evident *a priori*", as he said – for a number of reasons.

First, in relativity a condition at spatial infinity is necessarily dependent on the choice of a coordinate system, and so it appears to conflict with the principle of general covariance. Second, such conditions all contradict the principle of relativity of inertia (or Mach's principle), according to which: "In a consistent theory of relativity there can be no inertia *relatively to 'space'*, but only an inertia of masses *relatively to one another*" [17, p. 180]. In fact under the stated conditions at infinity, the $g_{\mu\nu}$ would only slightly differ from the Minkowskian values, with the result that "inertia would indeed be *influenced*, but would not be *conditioned* by matter (present in finite space)" [p. 183]; and, as we know, this requirement is an essential part of Mach's principle. Third, statistical objections can be raised in relativistic cosmology just as in the Newtonian case.

Of course, another option would be to abstain from fixing any conditions at infinity for the universe as a whole, and instead to fix such conditions in a case by case manner, for space–times modelling particular physical systems (as was done by Schwarzschild when deriving his solution).[1] However, this agnostic position was not to Einstein's liking:

[1]Note, however, that the derivation of the Schwarzschild solution does not require that the metric is asymptotically Minkowskian, since this (as well as staticity) happens to be a *consequence* of spherical symmetry.

> This is an incontestable position, which is taken up at the present time by de Sitter. But I must confess that such a complete resignation in this fundamental question is for me a difficult thing. I should not make up my mind to it until every effort to make headway toward a satisfactory view had proved to be vain [p. 182].

We shall hear again of Willem de Sitter, a professor of astronomy at the university of Leiden, who published between 1916 and 1917 three influential papers on general relativity, which, among other virtues, were responsible for awakening Arthur Eddington's fateful interest in the theory. Thus Einstein decided to *remove* the problem of imposing conditions at infinity, by using the radical remedy of postulating that space is a "self-contained continuum of finite spatial (three-dimensional) volume". In fact, if no 'infinity' (and no boundary) exists, then, obviously, no 'conditions at infinity' (or boundary conditions) are needed any more.

3. Einstein's Static Universe

The idea that an infinite, Euclidean space is not the only conceivable geometric model for the universe, largely pre-dates Einstein's cosmological paper. Indeed, it has been convincingly argued that the way Dante Alighieri in his *Divina Commedia* (fourteenth century) manages the system of "heavens" of his Paradise implies that he thought of them as if they were two-dimensional analogues of the parallels of an ordinary sphere: in other words, that he re-intepreted the Tolemaic cosmos as a three-dimensional sphere *ante litteram* (cf. [67,14]). However, it is to the nineteenth century that we owe an explicit and mathematically precise conjecture to this effect, and first of all to Riemann's generalization to higher dimension of Gauss' differential geometry of surfaces (Chap. 5, §7). Following in Riemann's footsteps, Clifford had written:

> We may postulate that the portion of space of which we are cognizant is practically homaloidal [i.e. flat], but we have clearly no right to dogmatically extend this postulate to *all* space. A constant curvature, imperceptible for that portion of space upon which we can experiment, or even a curvature which may vary in an almost imperceptible manner with the time, would seem to satisfy all that experience has taught us to be true of the space in which we dwell [4, p. 201].

Clifford had in fact a preference for the hypothesis that the universe, except for small local variations, was a space with constant positive curvature. Now this was precisely one of the hypotheses endorsed by Einstein in his 1917 paper [17]; let us list them, each one followed by a short commentary.

(I) *There is a privileged, time-orthogonal coordinate system, such that, with respect to the time coordinate matter looks "permanently at rest"* (staticity).

As we shall see, this hypothesis was weakened in subsequent theories, as regards the staticity of the matter distribution, but not in the all-important assumption of the existence of a privileged time coordinate (which will be called *cosmic time*), which was clearly a robust injection of Newtonianism into 'general relativity'.

(II) *The universe is a 3-sphere*, that is, it can be represented (except for local variations) as the set of points $(\xi^1, \xi^2, \xi^3, \xi^4)$ in 4-dimensional Euclidean space satisfying the equation:

$$\left(\xi^1\right)^2 + \left(\xi^2\right)^2 + \left(\xi^3\right)^2 + \left(\xi^4\right)^2 = R_0^2. \tag{3}$$

This can be seen as a special application of what will be named the *cosmological principle*, which assumes that the universe is both homogeneous and isotropic, as regards both geometry and mass-energy content.[2] Versions of the cosmological principle occur in most cosmological theories.

The 3-sphere is the only (up to isometry) *compact* 3-manifold with constant curvature which satisfies the further topological condition of being simply-connected (the property that every loop can be continuously deformed to a point). These conditions cannot be satisfied with a nonpositive constant curvature, though – it must be added – the condition of simply-connectedness has no obvious physical justification. It is worth pointing out that at the time the classification of 3-spaces with constant curvature was still at a very primitive stage, and topological considerations entered the discussion in a rather naive form; practically the only topological discussions had to do with the choice between the 3-sphere and the projective 3-space (also called "elliptic space": this is the simplest instance of a multiply-connected 3-space with constant positive curvature). In a letter to Weyl of June 1918 Einstein wrote that his preference for the 3-sphere was rooted in an "obscure feeling" [34, p. 78].

(III) *On a large scale, the matter content of the universe can be modelled as a pressure-less, incoherent fluid* (a dust, cf. Chap. 5, §10), so that in the privileged system the only nonvanishing component is the matter density ρ – constant both in space, because of (II), and in time, because of (I).

Hypothesis (III) has been considerably modified in other cosmologies, as different matter models have been advanced, the most popular being that of a perfect fluid, with nonvanishing pressure. What has hardly ever been modified is the fluidodynamical description itself (cf. [63, Note IX, pp. 415–6]).

From (I) and (II) it follows that the space–time metric in the privileged coordinate system has the form:

$$ds^2 = -R_0^2\, d\sigma^2 + c^2\, dt^2 \tag{4}$$

where $d\sigma^2$ is the metric of the unit 3-sphere S^3, that is:

$$d\sigma^2 = \sum_{a,b=1}^{3} \left(\delta_{ab} + \frac{x^a x^b}{R_0^2 - ((x^1)^2 + (x^2)^2 + (x^3)^2)}\right) dx^a\, dx^b,$$

or just

$$d\sigma^2 = \left(\delta_{ab} + \frac{x^a x^b}{R_0^2 - ((x^1)^2 + (x^2)^2 + (x^3)^2)}\right) dx^a\, dx^b, \tag{5}$$

as we may write using Einstein's convention on indices (cf. Chap. 5, §7). The coordinate system in which the metric is expressed in (5) is obtained by parallel projection along the ξ^4-axis in $\mathbb{R}^4$; it covers the upper (or the lower) hemisphere, except for the equatorial 2-sphere ($\xi^4 = 0$). It is important to remember that for 'most' manifolds (the n-sphere

[2] The first explicit statement of this principle is said to have been made in 1933 by E.A. Milne [63, pp. 156–8].

being a classic example) no single coordinate system exists covering the whole space, so all explicit coordinate representations of the metric can only describe what happens in a subspace. On the other hand, given the homogeneity of the 3-sphere, the expression (5) can be thought of as being valid in the neighbourhood of any given point.

Topologically speaking, Einstein's space–time is $S^3 \times \mathbb{R}$, and can be seen as the 4-dimensional cylinder in Minkowskian $\mathbb{R}^5$, that is, the set of all points $(x^0, x^1, x^2, x^3, x^4)$ in $\mathbb{R}^5$ satisfying:

$$\left(x^0\right)^2 + \left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 = R_0^2. \tag{6}$$

By applying the field equations to these data, one gets, after some computations, two equalities:

$$\frac{1}{R_0^2} g_{ab} = 0, \qquad -\frac{3}{R_0^2} = -\kappa c^2 \rho.$$

Now this result hardly makes sense: in fact, from the first equality one has that R_0 should be infinite, that is, that the 3-sphere should degenerate into ordinary Euclidean 3-space; but then the second equality implies that the matter density should be *null*, which means that the universe is *empty*! To put it in spatio-temporal terms, one is back to Minkowski space–time, with no matter in it.

4. The Cosmological Constant

Which assumption was to blame for this unseemly conclusion? Einstein's guess was that the culprit is not to be found among (I)–(III), but that it lies hidden in his very field equations. He proposed to change them into the following variant, which he claimed was "perfectly analogous to the extension of Poisson's equation given by (2)"

(IV) *Modified field equations*

$$R_{\mu\nu} - \lambda g_{\mu\nu} = -\kappa\left(T_{\mu\nu} - \frac{1}{2} T g_{\mu\nu}\right), \tag{7}$$

where λ is a constant, the so-called *cosmological constant*. Using (7) one obtains the following equations:

$$\frac{1}{R_0^2} g_{ab} = \lambda g_{ab}, \qquad -\frac{3}{R_0^2} + \lambda = -\kappa c^2 \rho,$$

whence it immediately follows $\lambda = 1/R_0^2 = \kappa c^2 \rho/2$. These equalities establish a connection between the constant λ and the total mass M of the spherical universe. In fact, since the volume of a 3-sphere of radius R_0 is $2\pi^2 R_0^3$, the total mass of Einstein's universe is:

$$M = 2\pi^2 R_0^3 \rho = \frac{4\pi^2}{\kappa c^2} \lambda^{-1/2}.$$

As Hermann Weyl commented, "this obviously makes great demands on our credulity" [79, p. 279]. Notice, however, that now there is no absurdity. By the introduction of λ it is possible to solve the (new) field Eqs (7) with a space–time and a matter model obeying conditions (I)–(III).

It is worth pointing out that Einstein's claim that (7) is the general relativistic version of (2) is wrong. It seems that the first scientist to realize it (a quarter of century later, in 1942), or at least to publish it, was the German cosmologist Otto Heckmann [39]. In fact it is easy to prove that the Newtonian limit (Chap. 5, §12) of (7) is *not* (2), but:

$$\Delta\phi + \lambda c^2 = \frac{\kappa\rho c^4}{2},$$

which can be rewritten as

$$\Delta\phi = \frac{\kappa c^4}{2}\left(\rho - \frac{2\lambda}{\kappa c^2}\right). \tag{8}$$

This equation is clearly satisfied by a constant potential in case one puts $\lambda = \kappa c^2\rho/2$ (note that this is the same equality which must be satisfied by Einstein's solution). One interesting consequence is that by this change in Poisson's equation Newton's insight into the behaviour of matter "evenly disposed throughout an infinite Space" (§1) happens to be fully vindicated: the net gravitational field at any point is zero, and this ensures that a Newtonian infinite material universe can indeed be stationary. In other words, the way Einstein modified the field equations to solve the cosmological problem closely resembles Newton's suggestion to Bentley on how to deal with an infinite universe![3]

Even more remarkable, perhaps, is that a long sequence of eminent authors missed this basic point and blindly endorsed Einstein's stated analogy between (2) and (7) ("generations of physicists have parroted this nonsense", as is said in [39, p. 723]). Clearly most working scientists are just too anxious to publish some 'new' piece of research of theirs to spend a sufficient amount of time reviewing the foundations of their disciplines; so they frequently end up by relying on authority much more than on rational belief, in contrast with the scientific ethos as ordinarily proclaimed.

At the end of his paper Einstein pointed out that the cosmological term was needed "only for the purpose of making possible a quasi-static distribution of matter, as required *by the fact of the small velocities of the stars*" ([17, p. 188]; italics added). From the following of the story it will be clear that Einstein also expected the introduction of λ to ensure that *only* a quasi-static (i.e. static on a sufficiently large scale) distribution was possible. This conjecture received a first blow a few months later.

5. De Sitter's Space–Time (1917)

As is clear, the Eqs (7) are also verified by any *empty* (i.e. $T_{\mu\nu} = 0$) space–time with a metric $g_{\mu\nu}$ such that:

$$R_{\mu\nu} = \lambda g_{\mu\nu}. \tag{9}$$

Any space–time satisfying this relation generates a difficulty for the Machian interpretation of general relativity, because it is a space–time with no matter content and yet

[3] This is in agreement with the 'genuinely Newtonian' cosmology described in [59], though this author, too, considers (2) as the legitimate Newtonian limit of (7).

with a metric solving the (vacuum) field equations. We recall that the original field equations had been conceived as describing how matter determines 'inertia', which means, under the assumption of the principle of equivalence, the behaviour of particles subjected to gravitation only (this behaviour being in turn described by the geodesics of the space–time). On the other hand, the λ-term makes it easier, in principle, to have plenty of formally acceptable space–times *with no matter content at all*.[4] The Minkowski and Schwarzschild space–times satisfy this condition with $\lambda = 0$ (that is, the original vacuum field equations), but what about a nonzero λ, such as the one needed for Einstein's static space–time?

The Dutch astronomer whom Einstein had cited in his first cosmological paper, de Sitter, published a few months later a solution of the modified field equations which was to prove one of the most important in the history of cosmology, not only of *relativistic* cosmology [8].

The idea (suggested to de Sitter by Paul Ehrenfest) is to obtain an homogeneous and isotropic space–time by adapting condition (3), which defines a spherical 3-*space*, so that it defines a 'spherical' *space–time*. Consider all points in $\mathbb{R}^5$ – regarded as the 5-dimensional Minkowski space – at space-like distance R_0 from the origin:

$$-\left(x^0\right)^2 - \left(x^1\right)^2 - \left(x^2\right)^2 - \left(x^3\right)^2 + c^2\left(x^4\right)^2 = -R_0^2. \tag{10}$$

This is the four-dimensional version of a familiar quadric surface, the one-sheet hyperboloid; in order to make it into a *space–time* one has to consider the metric induced on it from the *Lorentzian* structure of $\mathbb{R}^5$. Using the coordinate system obtained by parallel projection along the x^0-axis, one obtains the following expression for the metric, which closely resembles (5) (remember that for all repeated Greek indices, sum from 1 to 4 is understood):

$$ds^2 = \left(\eta_{\mu\nu} - \frac{\eta_{\mu\rho}\eta_{\nu\sigma}x^\rho x^\sigma}{R_0^2 + \eta_{\rho\sigma}x^\rho x^\sigma}\right)dx^\mu\, dx^\nu. \tag{11}$$

where $\eta_{\mu\nu}$ is the matrix of the usual Minkowski metric in $\mathbb{R}^4$ (Chap. 5, §4). This four-dimensional manifold is like the 3-sphere (or the n-sphere, for that matter, $n \geq 2$) in also having constant curvature (equal, with our conventions, to $-1/R_0^2$). By using the formula for the Ricci tensor given in Chap. 5 and with computations very similar to those that are made for the Einstein's metric, one finds:

$$R_{\mu\nu} = \frac{3}{R_0^2} g_{\mu\nu}, \tag{12}$$

so that the left-hand side of the modified field equations is

$$R_{\mu\nu} - \frac{1}{2}Sg_{\mu\nu} + \lambda g_{\mu\nu} = \left(-\frac{3}{R_0^2} + \lambda\right)g_{\mu\nu}.$$

It follows that if $\lambda = 3/R_0^2$, then (7) is satisfied *with a vanishing stress-energy tensor*. This is one part of the trouble: the λ-term makes it possible to have new space–times

[4]Ironically, all pseudo-Riemannian manifolds (of any signature) satisfying (9) are today called *Einstein manifolds*. For an encyclopaedic survey see [1].

with nontrivial inertia (i.e. different from Minkowski space–time) and empty. Indeed, it can be proven (for instance by using (24) and (25) below) that the only possibility for de Sitter space–time to be filled with a perfect fluid with density ρ, is for its pressure p to verify the equation $\rho + p/c^2 = 0$, which is a condition on the signs of both density and pressure which is difficult to reconcile with a realistic physical model unless one take both ρ and p equal to zero.[5]

Here comes a further twist. Equation (10) can be rewritten, by putting $t = x^4$, in the form:

$$\left(x^0\right)^2 + \left(x^1\right)^2 + \left(x^2\right)^2 + \left(x^3\right)^2 = R_0^2 + c^2t^2 \tag{13}$$

which is easily interpreted, geometrically, as a family of 3-spheres parametrized by t and of radius

$$R(t) = \sqrt{R_0^2 + c^2t^2}. \tag{14}$$

Clearly, $R(t)$ has a minimum (for $t = 0$) and no maximum; this means that de Sitter space–time describes a spherical universe with no beginning and no end, with a curvature which increases until it reaches the value $1/R_0^2$ and then decreases indefinitely. One may have different opinions on how to define ‘staticity’, but surely such a space–time cannot be rated ‘static’! By introducing a new time coordinate, namely $\tau = (R_0/c)\sinh^{-1}(ct/R_0)$, one gets the following expression for the de Sitter metric:

$$ds^2 = -R_0^2 \cosh^2\left(\frac{c\tau}{R_0}\right) d\sigma^2 + c^2\, d\tau^2. \tag{15}$$

This form of the metric was first published in 1922 by Cornelius Lanczos [56], a correspondent and later collaborator of Einstein [65, p. 491]. By comparison of (15) with the Einstein static metric (4) the nonstaticity of the de Sitter universe is even more evident. Thus de Sitter space–time can be taken to describe a changing, though paradoxically empty, spherical universe satisfying the modified field Eqs (7).

6. The “Discontinuities” of de Sitter Metric

Expression (15) seemed to refute Einstein’s opinion that the cosmological term was effective in ruling out non-static space–times. However, de Sitter’s own view was that his space–time *was* static, although not globally so. To explain this strange-looking but in fact quite sound opinion, one must take full account of the rôle played by the choice of the coordinate system, and in particular of the time coordinate. The great historical importance of de Sitter’s solution lies, first of all, in providing a spectacular demonstration of the wide-ranging consequences of such a choice. In fact, if one consider the subspace of de Sitter space–time contained in $x^3 + ct > 0$, $x^3 - ct > 0$ (this is a solid five-dimensional wedge in $\mathbb{R}^5$), then one can define on it two new coordinates θ and u:

[5] The function ρ can hardly make sense as a *negative* physical quantity, while a negative pressure does make sense (it would represent a cohesive force in the ‘cosmic fluid’), but the value $p = -\rho c^2$ “could not be even remotely approached by any known material”, Tolman wrote in 1934 [77, p. 348].

$$x^3 = R_0 \cos\theta \cosh u, \qquad t = \frac{R_0 \cos\theta}{c} \sinh\left(\frac{cu}{R_0}\right), \tag{16}$$

and the de Sitter metric acquires the following appearance:

$$ds^2 = -R_0^2\, d\sigma^2 + c^2 \cos^2\theta\, du^2. \tag{17}$$

It is clear that, in this coordinate system, the metric looks static indeed, as no one of its coefficients depends on t.[6] De Sitter described his solution as "the general solution for the case of a static and isotropic gravitational field in the absence of matter" (cit. in [63, p. 88]).

A space–time satisfying the modified field equations, static and empty could not but disturb Einstein. As he commented in his public reply to de Sitter [18], in case (17) were to be shown a legitimate solution of the modified field equations, this would indicate that the λ-term "does not fulfil the purpose I intended":

> Because, in my opinion, the general theory of relativity is a satisfying system only if it shows that the physical qualities of space are *completely* determined by matter alone. Therefore, no $g_{\mu\nu}$-field must exist (that is, no space–time continuum is possible) without matter that generates it.

In other words, de Sitter's solution violates Mach's principle, and yet the cosmological constant had been introduced to circumvent, in a static background, this blemish of the original field equations. So Einstein resorted to the argument that the coordinate system used in (17) was not legitimate. His point was that if $\theta = \pi/2$, then the metric degenerates (i.e. $\det(g_{\mu\nu}) = 0$), and yet "it seems that no choice of coordinates can remove this discontinuity":

> Until the opposite is proven, we have to assume that the de Sitter solution has a genuine singularity on the surface [$\theta = \pi/2$] in the finite domain; i.e., it does not satisfy the field Eqs (7) for any choice of coordinates.

Thus Einstein argued that de Sitter's space–time is not really "a world free of matter, but rather like a world whose matter is concentrated on the surface [$\theta = \pi/2$]", as is the case for "the immediate neighbourhood of gravitating mass points".

Einstein's refutation of de Sitter's counterexample is remarkably weak. Equation (17) is nothing but the expression in a certain coordinate system of metric (15); and metric (15) is indeed defined on *all* of de Sitter space–time, and it *does* satisfy (7) everywhere! There is no "discontinuity" to "remove" at $\theta = \pi/2$: one only need consider in the neighbourhood of any point p such that $\theta = \pi/2$ the analogue of the coordinate system giving (17) (in this new coordinate system, the angular variable will be a $\overline{\theta}$, with, say, $\overline{\theta}(p) = 0$). Such coordinate systems certainly exist, given the homogeneity of de Sitter's space–time. It follows, with no need for further inquiry, that this space–time has no "genuine singularity".[7]

As shown in private correspondence at the end of 1918, Einstein eventually recognized that de Sitter's solution was perfectly regular; nonetheless, he did not publish a retractation of his previous criticism [50].

---

[6] Notice that the first coordinate θ is just one angular coordinate of the 3-sphere, while the time coordinate u is given explicitly by $u = \tanh^{-1}(\frac{ct}{x^3}) = \frac{1}{2}\log\frac{x^3+ct}{x^3-ct}$.

[7] This point was well explained by Eddington in his *Report* of the same year [11, p. 88].

7. Weyl Enters the Controversy (1922–1923)

In the fourth edition of his treatise *Space–Time–Matter* Hermann Weyl discussed the contrast between the cosmological views of Einstein and de Sitter. He put it in the form as to which coordinate system – the one leading to (15) or the one leading to (17) – was more adequate to "represent the whole world in a regular manner":

> In the former case the world would not be static as a whole, and the absence of matter in it would be in agreement with physical laws; de Sitter argues from this assumption [...]. In the latter case we have a static world that cannot exist without a mass-horizon; this assumption, which we have treated more fully, is favoured by Einstein [79, p. 282].

In other words, Weyl supported Einstein's view that de Sitter's solution was not really empty of matter, but that the matter was, so to speak, 'hidden' behind the boundary of the coordinatized region.

However, in 1923 Weyl modified his views considerably ([80,82]; cf. [83]), and in so doing he posed a foundational stone of relativistic cosmology for decades to come.

In the 4-dimensional hyperboloid representing de Sitter's space–time, any of the lines obtained as the intersection with a 2-plane containing the line

$$x^0 = x^1 = x^2 = 0, \qquad x^3 + cx^4 = 0,$$

is a timelike geodesic Γ – a generating line of the hyperboloid, in fact. Now, the set of all points of space–time which can be influenced by such a worldline (its *range of influence* [*Wirkungsbereich*], as Weyl calls it)[8] is given by one 'half' of the hyperboloid: to be explicit, in the case of the line Γ_0 defined by the 2-plane $x^0 = x^1 = x^2 = 0$, this 'half' is the set M of points belonging to the half-space $x^3 + cx^4 > 0$. The important fact is that M is the range of influence not only of Γ_0, but of *the whole 3-parameters family of geodesics* Γ *obtained by the same procedure*.

Weyl's proposal was to interpret each Γ as the worldline of a "star", and, given that all these "stars" are causally connected since the "infinitely distant past" and that no influence can come from outside, to take M as the *whole* of our space–time. This assumption suggests that we select and use a coordinate system which (1) covers the whole of M (and not just the 'wedge' region); and (2) represents the Γ by parallel lines. Such a coordinate system indeed exists, and in it the metric acquires the form:

$$ds^2 = -e^{2cT/R_0}\left(dx^2 + dy^2 + dz^2\right) + c^2\, dT^2. \tag{18}$$

Note that this form of the metric is invariant for a translation of the spatial coordinates $(x, y, z) \mapsto (x + x_0,\ y + y_0,\ z + z_0)$ and also for the composite of a time coordinate translation $T \mapsto T + T_0$ with the spatial dilation $(x, y, z) \mapsto e^{-cT_0/R_0}(x, y, z)$, so that all 'fundamental observers' can be considered as equivalent. Their common origin, Weyl wrote in the German fifth edition of *Space–Time–Matter*, "makes understandable the small velocities of the stars" [80, p. 285].

Another consequence pointed out by Weyl is that even if the spatial coordinates of every star are constant, this does not mean that the distance between two stars is also constant. In fact if $D(0)$ is the distance at $T = 0$ of two stars, then the distance at time T is $D(T) = D_0 e^{cT/R_0}$.

[8] In modern notation [40] this is denoted as $J^+(\Gamma)$.

Weyl derived a crucial formula, giving the change of frequency of the light emitted by a star, as received by another star:

$$\frac{\Delta\nu}{\nu} \approx -\sin\frac{D}{R_0} \approx -\frac{D}{R_0}. \tag{19}$$

This means that the spectral lines of stars are shifted towards lower frequencies (*red-shifted*) when they are received by another star, and the redshift z is approximately proportional to the distance. This is one of the first appearances of what is known as "Hubble's law"; as we shall see this property of de Sitter's space–time will be important in directing Edwin Hubble's attention to this kind of relation in his data.

Note that, from the formula giving the relative distance of two stars, it follows that their relative speed is:

$$v = \frac{c}{R_0} D, \tag{20}$$

that is, *the speed is (exactly) proportional to the distance*, and therefore v turns out to be approximately proportional to the redshift. So, for all practical purposes, the relationship between redshift and distance can be interpreted as a Doppler effect.

The assumption that a cosmological space–time must admit a 3-parameter family of 'synchronized' timelike geodesics diverging from the past and hypersurface-orthogonal,[9] representing the fundamental cosmological objects ("stars" or, a few years later, galaxies and, still later, clusters of galaxies), survived the demise of de Sitter space–time as the centrepiece of relativistic cosmology, and became known as *Weyl's principle*.

This principle in fact singles out a class of privileged observers: once it is accepted, the static version of de Sitter's metric (17) cannot be considered as physically relevant, because the observers having u as their proper time *are not geodesic*. In relativistic cosmology, Weyl's fundamental observers came to play the rôle of the 'inertial systems' of Newtonian physics, just as the 'cosmic' (or 'universal') time took the place of the supposedly dead absolute time. This circuitous return of the basic Newtonian concepts, as we shall see (§16), will not please a great philosopher-scientist such as Kurt Gödel.

Moreover, the requirement of 'divergence from the past' rules out (15), which otherwise would satisfy the principle. In fact this condition implies that in the past the 'stars' were closer to each other than they are now, and this establishes a cosmological asymmetry between past and future which, as we shall see, was a critical assumption in the development of cosmology.

Note, finally, that the form (18) of the de Sitter metric describes an obviously non-static universe again, but with a surprise: the universe is no more a 3-sphere as in (15), but a *Euclidean 3-space*. In other words, the *space* curvature (as opposed to the *space–time* curvature, which of course is always $-1/R_0^2$, since it does not depend on the choice of the coordinate system) is zero. The moral is that, by introducing suitable coordinate systems, we can have *radically different space geometries* (Chap. 5, §5). This is one of the most striking illustrations of the fact that the general covariance principle, if not suitably restricted, leads to a radical relativization of the physical geometry of *space* (as opposed to space–time). In this sense one can say that this vindicates Poincaré's

[9] This means that the universe at a certain cosmic time is represented by a space-like 3-submanifold, which is crossed orthogonally by the fundamental worldlines.

'conventionalism' as regards physical geometry, though not his prediction that physicists would never abandon Euclidean geometry (Chap. 5, §9), as will be even more clear from the following.

8. Friedmann's Evolutionary Universes (1922, 1924)

While de Sitter, Einstein, Weyl and others were struggling about how to correctly interpret de Sitter's solution, a more radical development was taking place at Petrograd on the Baltic sea. The de Sitter metric, as displayed in formulas (15) and (18), had been the first example of a metric of the type

$$ds^2 = -R(t)^2\, d\sigma^2 + c^2\, dt^2, \tag{21}$$

which must be thought of as defined on a product manifold $\Sigma \times I$, where $(\Sigma, d\sigma^2)$ is a Riemannian 3-manifold with normalized constant curvature (i.e. the curvature K is 1 or 0 or -1), I is an open interval of the real line $\mathbb{R}$, and $R : I \to]0, +\infty[$ is a function which can be interpreted as the "radius" of the universe (or, more generally, its *scale factor*). Using formula 20 from Chap. 5, §8, we can express all these metrics in the form:

$$ds^2 = -R(t)^2\left[\frac{dr^2}{1-Kr^2} + r^2\big(d\phi^2 + \sin^2\phi\, d\chi^2\big)\right] + c^2\, dt^2, \tag{22}$$

where $K = 0, 1, -1$. Notice that (21) (or (22)) embodies both the cosmological principle (§3, II) and Weyl's principle (apart from the past-future asymmetry).

Of course also Einstein's metric falls into this class (for $R(t) \equiv R_0$). So it was a natural question to investigate whether other space–times of this type existed, apart from de Sitter's and Einstein's, either with $\lambda = 0$ or with $\lambda \neq 0$. Einstein had somewhat preempted this line of research, by implying that his static space–time was the *only* non-empty solution allowed by the modified field Eqs (7); in particular, he thought that all physically meaningful solutions of (7) were static.

So it took a Russian meteorologist to challenge Einstein's authority and proceed to an inquiry of the *non-stationary* solutions. His name was Aleksandr Aleksandrovich Friedmann; he made his major contributions to relativistic cosmology between 1922 and 1924.

In his 1922 paper, "On the curvature of space" [32], Friedmann considered the metrics of the form

$$ds^2 = -R(t)^2\, d\sigma^2 + M(x)\, dt^2, \tag{23}$$

satisfying (7), under the assumption that Σ is a 3-sphere and matter a dust, as for Einstein's space–time. He showed that they divided into two classes, according to whether $\dot{R} = 0$, i.e. $R(t)$ is a constant function (stationary case), or $\dot{R} \neq 0$ (non-stationary case), the first class comprising only Einstein's and de Sitter's metrics (the latter in the form (17)).[10] In 1924 he published a second paper, "On the possibility of a world with a constant negative curvature of space" [33], dealing with the case that Σ is not a sphere but a space with $K = -1$ (a *hyperbolic* 3-space). In fact the main argument is the same

[10] If we start from (21), 'static' and 'stationary' can be used equivalently.

with any curvature $K = 1, 0, -1$, the non-stationary case always reducing to metrics of the form (21).

The mass density ρ is supposed to depend on t only; the stress-energy tensor is given, in the coordinate system where (21) holds, by a $T_{\mu\nu}$ having only one nonzero component: $T_{44} = \rho c^4$.

Putting all these data into (7) and indicating by dots the derivatives with respect to t, one obtains just two scalar equations:

$$\frac{\dot{R}^2 + 2R\ddot{R} + c^2 K}{c^2 R^2} - \lambda = 0, \tag{24}$$

$$\frac{3(\dot{R}^2 + c^2 K)}{c^2 R^2} - \lambda = \kappa c^2 \rho. \tag{25}$$

After a few elementary transformations on (24), one obtains the (now) famous *Friedmann's equation*:

$$\dot{R}^2 = \frac{\lambda c^2 R^2}{3} + \frac{A}{R} - c^2 K, \tag{26}$$

where A is a constant which can be expressed, by exploiting (25), as:

$$A = \frac{\kappa c^4 \rho R^3}{3}. \tag{27}$$

Equation (26) describes the behaviour of the squared time derivative of the radius of the spherical universe, and as a consequence it permits one to say how the radius itself $R(t)$ behaves in time.

Mathematically, the basic issue is to find under which conditions the right side vanishes, or equivalently what are the positive real roots (since R is necessarily positive) of the cubic equation $\lambda x^3/3 - Kx + A/c^2 = 0$, depending on λ and A.

If $K = 1$, it is easy to see that there is one critical value λ_c, defined by

$$\lambda_c = \left(\frac{2c^2}{3A}\right)^2,$$

and a critical interval J of real numbers such that the behaviour of $R(t)$ changes qualitatively according to where λ and $R(0)$ lie with respect to λ_c and J. An elementary computation shows that λ_c is equal to the cosmological constant of an Einstein static space–time containing the same total mass.

A summary of the multifarious non-stationary alternatives runs as follows (as stated above, Friedmann shows that stationary solutions exist only if $\lambda \neq 0$, thus proving that Einstein's argument for the cosmological constant is partly correct).

Suppose λ lies between 0 and λ_c; then if $R(t_0)$ lies on the right of J, the universe expands forever, either monotonically (if $\lambda = \lambda_c$) or after an initial contraction phase;[11] if $R(t_0)$ lies on the left of J, then the universe expands from $R = 0$ to a certain positive value of R, and then contracts: in other words the universe is *periodic*, or as it is more commonly known today, *oscillating* (although the whole of its story might well be in-

[11] Actually this last possibility, which includes the non-stationary form (15) of de Sitter metric, is not mentioned in [32].

cluded in a single cycle!).[12] Suppose λ is negative; then the universe is in all cases an oscillating one. Suppose finally that $\lambda > \lambda_c$; then the universe can only expand forever from $R = 0$.

In the case of negative constant curvature ($K = -1$) the argument follows the same pattern. Here the dividing line is whether the cosmological constant is positive (or zero), or negative; in the first case, the universe expands forever from $R = 0$; in the second case we have an oscillating universe. Friedmann did not list these different possibilities, but limited himself to pointing out that no stationary universe with a positive density can exist.

In the concluding section of [33] Friedmann tackled the problem of the finiteness of infinity of the universe, and argued, very reasonably, that "the cosmological equations, by themselves, are not sufficient to solve the problem of the finiteness of our universe [...]", and that other hypotheses, of a topological nature, were needed to reach any conclusions. In particular he emphasized that the case of negative curvature is also compatible with finiteness. In this he showed prescience, since the first examples of compact 3-manifolds with constant negative curvature were only discovered a few years later. (Today the issue of the topology of space–time is very much alive; cf. Chap. 13).

9. The Einstein–Friedmann Controversy

How did Einstein react to Friedmann's first article? It is not unreasonable to suspect that the new options crowding around his basic hypothesis of a spherical universe obeying (7) must have provoked a strong revulsion in him. To make things worse, the cosmological constant, far from being a gatekeeper stymieing cosmological anarchy, was contributing to it... Whatever the circumstantial and psychological reasons, it is a fact that he came hastily to the conclusion that Friedmann had to be wrong on a very basic point, which destroyed the main theorem of his paper, namely, that (7) admits non-stationary solutions. He wrote as much in a short note which appeared on 8 September 1922, in the same journal [21]. In it he stated that Friedmann had made a simple mathematical mistake, from which the non-constant ρ had been born. According to Einstein, from the equation $\operatorname{div} T = 0$ (cf. Chap. 5, §13) it just followed:

$$\frac{\partial \rho}{\partial t} = 0, \tag{28}$$

implying that both ρ and $R(t)$ had to be constant. Thus Friedmann's paper was useful, after this substantial correction, only insofar as it unwittingly proved that (7) could not have non-stationary solutions. In other terms, Friedmann had provided, according to Einstein, an incorrect proof of the fact that the only solution of (7), under the given assumptions, was Einstein static space–time!

Friedmann learnt about [21] by chance, from a letter of a colleague, Y.A. Krutkov, himself quite knowledgeable on relativity, who was working in Germany at the time. Thus, on 6 December of the same year, Friedmann addressed to Einstein directly a very courteous letter, where he described in detail the computation which made him doubt

[12] In his semipopular book Friedmann referred to the "cycles of existence" of the Hindu religion, and also to "creation from nothing", although just as "curiosities" [34, p. 206].

Einstein's conclusion (28). The point is a quite simple one, in fact, and can be formulated in the following few lines. In general the divergence of $T_{\mu\nu}$ is:

$$(\mathrm{div}T)_\mu = \frac{1}{|g|^{1/2}}\frac{\partial}{\partial x^\lambda}\left(g^{\lambda\nu}|g|^{1/2}T_{\mu\nu}\right) - g^{\lambda\sigma}\Gamma^{\nu}_{\lambda\mu}T_{\nu\sigma},$$

and it is easy to check that in the examined case all components of div$T$ vanish identically, except for the 4th; thus, since div$T = 0$, one must have:

$$0 = (\mathrm{div}T)_4 = \frac{\partial}{\partial t}\left(|g|^{1/2}\rho\right) = \frac{\partial}{\partial t}\left(hc\rho R^3\right),$$

where $h = \det(h_{ab})$ and $d\sigma^2 = h_{ab}\,dx^a\,dx^b$. Of course this equation *does* not imply that either ρ or R is constant, but only that the product ρR^3 is (note that the factor hc is time-independent); and this is exactly what must be expected, in agreement with (27).

Friedmann was right: Einstein, not he, had made an error – and quite an elementary error at that. So he felt safe in asking Einstein to let him know whether he agreed with the computation, and suggested that Einstein published a retractation in the same journal, perhaps together with excerpts from the letter. At the time Einstein was abroad, and only returned in Berlin in March 1923. It is not known whether he read Friedmann's letter at that time, but, if he did, it must not have shaken his faith that a non-static solution for (7) simply could not exist.

By a further lucky coincidence, Krutkov was in Leiden as a visiting scholar in May, and he met Einstein, who had travelled there to attend Lorentz's last public lecture before retirement. Krutkov insisted on reading Friedmann's article with him, and succeeded in bringing Einstein to recognize his mistake. In a letter dated May 18, two days after Einstein's return to Berlin, he wrote: "I won Einstein in his argument against Friedmann. Petrograd's honour is rescued!". In fact on 21 May Einstein did submit a retractation note [22], as short as the previous one, where he mentioned Friedmann's letter and the discussions with Krutkov, admitted a computational error of his own, and stated that Friedmann's claim, that the field equations have not only stationary, but also "dynamical (i.e. time-varying) centrally-symmetric solutions", was "correct and important".

After the publication of the second note, "everybody was much impressed by my fight with Einstein and my eventual victory", as Friedmann wrote in a letter of 13 September 1923; he added that he was glad of this, since from then on his articles would have had a smoother passage to publication in the scientific journals. In 1924, the same year that his second (and last) cosmological article [33] appeared,[13] Friedmann published a popular book on relativity, entitled *The Universe as Space and Time*, and planned with a colleague a four-volume technical treatise on relativity. Only the first volume of this major work came out, however, since he suddenly fell ill with typhus and died in 1925, aged 37.

10. What Came of Friedmann's Discovery?

In his first draft of [22], which has survived, Einstein had inserted a clause to the effect that to Friedmann's evolutionary solutions "it was hardly possible to give a physical

[13] As is clear from Friedmann's letter to Einstein quoted above, the results of [33] had essentially been obtained by him in 1922.

meaning" [34, p. 47]. Although he ultimately decided to delete it, it is clear that he was convinced that the physical universe was not subject to any large-scale motions. In any case, he did not work on the consequences of Friedmann's articles during the next few years; as a matter of fact, he devoted to cosmology only a few dozen pages in the rest of his scientific production.[14] In itself this is a curious fact.

What is definitely strange is that no one seems to have taken notice of Friedmann's results, before or after the publication of the two Einstein notes. For instance, Weyl failed to mention Friedmann's first article in the 1923 edition of his *Raum, Zeit, Materie* [80], and this is perhaps forgivable; but he did not cite [32] and [33] even in his expository 1930 article where he remained loyal to de Sitter space–time [83].[15]

In 1927 Georges Lemaître – a Belgian mathematician, astronomer, and Catholic priest (since 1923) – discussed the evolutionary solutions for a perfect fluid [58]; what he added to Friedmann's treatment, apart from the generalization to $p \neq 0$, was an examination of recent astronomical data relevant to an assessment of the cosmological models.

In his paper Lemaître did not refer to Friedmann's articles. According to his own account, it was Einstein who at the Solvay conference in Bruxelles, of the same year, mentioned them to him. It has been surmised that the reason the articles and notes of Friedmann and Einstein, published on *Zeitschrift für Physik*, "the best known journal in theoretical physics of that epoch", had completely escaped Lemaître's attention is that he could not read German [34, p. 56], but one can doubt this explanation.[16] It is somewhat amusing that in a letter of 5 April 1930, Lemaître felt he had to inform de Sitter – who certainly knew German well! – of the very *existence* of Friedmann's articles and Einstein's notes. Nevertheless, at a meeting of the British Association in 1931, de Sitter talked of the "brilliant discovery" of "the expanding universe", which had been made… by Professor Lemaître, of course, and "discovered by the scientific world about a year and a half ago, three [sic!] years after it had been published".

What these facts suggest is that there was a widespread resistance of some of the leading cosmologists to citing Friedmann's articles. It has been said that the reason was that Friedmann's articles were too mathematically oriented, but this is not only an inadequate explanation in itself,[17] but it is also to some extent misleading, given that Friedmann went to such lengths as to estimate the age (or, to be more precise, the "period") of the universe, in the last paragraph of his 1922 paper:

> It is left to remark that the "cosmological" quantity λ remains undetermined in our formulae, since it is an extra constant in the problem; possibly electrodynamical considerations can lead to its evaluation. If we set $\lambda = 0$ and $M = 5 \cdot 10^{21}$ solar masses, then the world period becomes

---

[14] In an appendix to the 1945 edition of [20], Einstein discussed the cosmological problem, giving Friedmann belated recognition.

[15] It is worth adding that in his 1950 preface to the reprint of the English translation of the fourth edition of [79], Weyl did not spend a single word on cosmology, not even to refer to the considerably revised section on cosmology in the fifth German, untranslated edition. A few years later, in the reprint of his famous 1921 *Encyclopädie* survey [66], Pauli devoted one long supplementary note (pp. 219–23) to Friedmann's and Lemaître's work.

[16] That it cannot be the full explanation is indicated by the circumstance that in [58] Lemaître cites two German articles, by Lanczos [56] and by Weyl [82].

[17] We shall dwell in the next section on the sparse and controversial observational evidence which was all astronomy provided at the time.

> of the order of 10 billion years. But these figures can surely only serve as an illustration for our calculations.[18]

The last clause seems to be rather a ground for praising Friedmann's scientific sense and moderation, than to indict the physical relevance of his results (by the way, his estimate is close to more recent estimates, cf. [72, p. 182]). Moreover, the very mathematical nature of Friedmann's contribution would only have increased its chances of being read and utilized. A plausible conjecture has to do with the sociology of small groups: the very few people working in cosmology at the time, who were virtually all committed to general relativity and personally related to Einstein, may have been loath to advertise a recent episode which did not throw a very positive light on Einstein's mathematical skill and physical insight and showed the serious lack of uniqueness of the answers given by his acclaimed theory. Also, Friedmann had gone too far for an outsider, by showing that one of the basic assumptions shared by the leaders in the field was wrong.[19]

However, in the end, the expanding universes introduced by Friedmann, and then rediscovered by Lemaître, Howard P. Robertson and Arthur G. Walker, did find their way into recognition. How the FLRW solutions (from the initials of the four people just mentioned) became the main ingredient of modern cosmology is the subject of the next sections.

11. The Astronomical Evidence

In 1905 a well-known astronomer, Agnes Mary Clerke, opened Chapter 26 of her book *The System of the Stars* with the following words:

> The question whether nebulae are external galaxies hardly any longer needs discussion. It has been answered by the progress of research. No competent thinker, with the whole of the available evidence before him, can now, it is safe to say, maintain any single nebula to be a star system of co-ordinate rank with the Milky Way [cit. in [7, p. 214]].

How persuasive. And yet, ironically, it was the very year two revolutions in physics were slowly making a major advance; surely a better time could have been chosen to be so final on a controversial topic. To partially excuse that 1905 confidence and to put in perspective the whole observational business in cosmology, it may be useful to remember that "only one per cent of the light in the night sky comes from beyond our Galaxy" [10, p. 1128].

The view that the galaxies, or *nebulae*, as they were called, were external to the Milky Way and in fact were systems similar to it was known as the *island universe theory*. In 1914 Eddington supported this opinion as a "working hypothesis".

---

[18] [32, p. 2000]. In the "Editor's Note" to the English translation of Friedmann's articles which we have used here, one reads that "[Friedmann] considered only a non-zero cosmological constant λ [...]" [55, p. 1987]. As the preceding quotation in the text evidently shows, this is not correct.

[19] As to mathematical skill, de Sitter had been aware, since at least 1917, that "Einstein occasionally made elementary mistakes in calculus" [52, p. 453]. Unfortunately, a reluctance to expose the failings of a genius like Einstein is conspicuous in most of the scholarly literature on him. For instance the following passage is *all* Pais has to say, in his standard 552-page biography, about the Einstein–Friedmann controversy: "*1922*. Friedmann shows that Eq. (15.20) admits nonstatic solutions with isotropic, homogeneous matter distributions, corresponding to an expanding universe [F1]. Einstein first believes the *reasoning* is incorrect [E45], then *finds an error in his own objection* [E46] and calls the new results 'clarifying'" ([65, p. 288]; italics added).

The first determination of the radial speed of a spiral galaxy, Andromeda, was made in 1912 by Vesto Slipher, of the Lowell observatory, and published in 1913. It was an *approaching* velocity of 300 km/s, not a recessional one, but as Hubble was to write at the end of his life, "the first step in a new field is the great step. Once it is taken, the way is clear and all may follow" [49]. In 1915, from the examination of the spectra of 15 spiral galaxies, Slipher concluded that they all had high radial recessional speeds, from 300 to 1100 km/s. Eddington's future collaborator in the eclipse expeditions to measure the light deflection, A.C.D. Crommelin, who in 1912 had argued against the island universe theory, expressed in 1917 the "hope" that this "sublime and magnificent" conception would "stand the test of future examination".

In 1920 what will be later called the "Grand Debate" took place between Harlow Shapley and Heber Curtis before the U.S. National Academy of Sciences (and then, in amplified form, in the published texts); among other things, the first astronomer defended the traditional view and the latter the island universe theory. One of Shapley's main pieces of evidence was provided by Adrian van Maanen, of the Mount Wilson observatory, who had detected large angular rotations in spiral galaxies, indicating that they must be rather close to the solar system. In 1924 a Cambridge authority, W.M. Smart, had confirmed van Maanen's measures, of which he extolled "the extraordinary precision", uttering another of those definitive statements:

> I do not believe that anyone would be so bold as to question the authenticity of the internal motions – regarded either as rotational or as a stream motion – found by van Maanen; in fact, the more one studies the measures, the greater is the admiration which they evoke [cit. in [41, p. 101]].

Today these 'unquestionable' results have dissolved into thin air. The story has been summed up by saying that "van Maanen had read his expectations into his data",[20] which is a reasonable hypothesis – also to be applied in other cases, as we shall see in a moment.

In his autobiography Shapley recalled: "They wonder why Shapley made this blunder. The reason he made it was that van Maanen was his friend and he believed in friends!" [7, p. 354]. In fact the event had a less edifying and gentlemanly side. Another astronomer, Knut Lundmark, from Sweden, had published in 1921 and 1922 two papers undermining van Maanen's results. Shapley's reaction had been to write a distinctly intimidating letter to Lundmark, where he qualified the latter's critical remarks of van Maanen's work as "not of significance in the larger problem", and suggested that he could find "many flaws or hasty conclusions" in Lundmark's own work. Though irritated at first, Lundmark eventually did not insist on this point, and in fact van Maanen could publish in 1923, without further ado, a paper citing Lundmark's observations in support of the soundness of his own data!

Van Maanen's colleague at Mount Wilson, Edwin Hubble, thought differently, but abstained from putting his opinion in print, and also delayed the publication (until 1925) of results on the distances of galaxies M31 (Andromeda) and M33, estimated by him to be about 930,000 light-years – a huge amount, well in agreement with the island universe theory. Mount Wilson's officials discouraged him from publishing evidence contradicting the work of another member of the same institution. Eventually, in 1935, Hubble published in the *Astrophysical Journal* his remeasurements of four of the galaxies

[20] [41, p. 110]; this well-written and scholarly small book contains a fascinating reconstruction of the case (pp. 83–110).

studied by van Maanen [48], and concluded that there was no evidence of rotations. Van Maanen, who was allowed to reply in the same issue, did not fully retract his results, but admitted that his more recent measurements "show considerably smaller values of the apparent rotational component than those first obtained", though "the persistence of the positive sign is very marked", thus requiring further investigations [78].

12. "Hubble's Law"

In 1929 Hubble stated what came to be called *Hubble's law*, namely, that redshifts and distances of galaxies are linearly related: $z = Hd$ (note that the observational counterpart of 'distance' here is 'apparent magnitude'). Several authors have pointed out that Hubble's observations suggested a quadratic rather than a linear relation (cf. [28, p. 378]). That the functional dependence had to be linear was surely inspired by de Sitter's model, and in the first place by his allegiance to general relativity, as confirmed by the following passage in his main article:

> The outstanding feature, however, is the possibility that the velocity-distance relation may represent the de Sitter effect, and hence that numerical data may be introduced into discussions of the general curvature of space [47].

Note that among the relativistic expanding universes the so-called Hubble's 'constant' is a *true* constant *only* in the case of de Sitter space–time: in the other FLRW space–times (except for the Einstein static space–time, of course, where it vanishes)[21] H is in fact a function of cosmic time. This strengthens the case for Hubble's 'reading his expectations into his data'. Nonetheless, it must be stressed that Hubble, to the end of his life [49], did not endorse the view that his discovery was decisive evidence for the expanding universe cosmology.

In the case of zero pressure and zero cosmological constant, one defines the *critical density* ρ_c as

$$\rho_c = \frac{3H^2}{\kappa c^4},$$

and so the sign of the curvature of the spatial sections is connected with the value of the Hubble's constant by the equation:

$$K = \frac{\kappa c^2 R^2}{3}(\rho - \rho_c) = \frac{\kappa c^2 R^2 \rho_c}{3}(\Omega - 1).$$

where $\Omega = \rho/\rho_c$ is called the *density parameter*. This shows the importance of evaluating H in order to determine at one stroke (under the given hypotheses) both the universe's geometry and its destiny. The positive K corresponds to $\Omega > 1$ and to a closed universe whose fate is a final collapse; the other two possibilities ($K = -1$, that is, $\Omega < 1$; and $K = 0$, that is, $\Omega = 1$) correspond to a higher or lower (respectively) rate of expansion.

It is however worth noting that the observational differences in the predictions of the three models ($K = 0$, $K = 1$, $K = -1$) are not remarkable: "all three types

[21] Notice that this is not true in the interpretation of Einstein's space–time by I. Segal in his chronometric cosmology; cf. Chap. 13.

make surprisingly similar predictions rather than grossly different ones as the trichotomy suggests", as the authors of a useful textbook in general relativity wrote in 1977 [72, p. 177].[22]

13. The Einstein–de Sitter Space–Time and the Demise of λ

It was in 1931 that Einstein abandoned for the first time the cosmological constant, in a paper where he endorsed an oscillatory model. In 1932 Einstein and de Sitter published a short joint paper, where they described a special case of Friedmann's solutions, the one with zero cosmological constant, zero spatial curvature and zero pressure. (The *Einstein–de Sitter metric*, as was to be named, had not been discussed by Friedmann, probably because his main concern was with non-Euclidean spatial geometries; however, it had been discussed in 1929 by Robertson.) The radius of the universe has the form: $R(t) = R(t_0)(t/t_0)^{2/3}$, and the space–time is topologically like $\mathbb{R}^4$. A computation shows that the scalar curvature of this space–time is $S(t) = 4/(3t^2)$. This implies that all the fundamental observers, for which t is the proper time, have a finite past: t cannot go past zero, because at zero the scalar curvature blows up to infinity. This is an example of a genuine space–time singularity, which can be physically interpreted as describing a universe expanding from a state of infinite density. In Sec. 17 we shall see how it was discovered that this property is not the awkward outcome of too special symmetry assumptions.

In their paper Einstein and de Sitter put $\lambda = 0$, but they pointed out that "an increase in the precision of data derived from observations will enable us in the future to fix its [i.e. of λ] sign and to determine its value" [26]. *This is precisely the same as Friedmann's position* (§10). According to George Gamow, Einstein once told him that the introduction of the cosmological constant had been "the greatest blunder of my life"; perhaps, talking with Gamow, who had been Friedmann's student, he was thinking also of the mathematical mistake connected with his temporary rejection of Friedmann's work (§9). It is true, however, that by 1919 Einstein had written that the λ-term was "gravely detrimental to the formal beauty of the theory" [19].

These facts may perhaps serve to explain, in part, the following amusing story by Eddington:

> Einstein came to stay with me shortly afterwards, and I took him to task about [the dropping of the Λ in [26]]. He replied: "I did not think the paper was important myself, but de Sitter was keen on it". Just after Einstein had gone, de Sitter wrote to me announcing to visit. He added: "You will have seen the paper by Einstein and myself. I do not myself consider the result of much importance, but Einstein seemed to think that it was" [cit. in [3, p. 123]].

In any case, the question of the cosmological constant cannot be dismissed so easily, as the following remarks show (more will be said in §18).

The modified field Eqs (7) can be re-written, equivalently, in the form (note the sign of λ):

[22] A sustained argument in favour of an open universe (with $\Omega \approx 0.2$) was presented 20 years later by P. Coles and G.F.R. Ellis; the authors admit that "the case [they] have made [...] though strong is by no means watertight" [6, p. 204].

$$R_{\mu\nu} - \frac{1}{2}Sg_{\mu\nu} + \lambda g_{\mu\nu} = -\kappa T_{\mu\nu}.$$

Now there is a theorem, proved by Hermann Vermeil in 1917 and included by Hermann Weyl in Appendix II of [79], that states that the only rank 2 symmetric tensor, containing the $g_{\mu\nu}$ and its first and second derivatives, and furthermore linear in the second derivatives of the $g_{\mu\nu}$, is of the form:

$$E_{\mu\nu} = c_1 R_{\mu\nu} + c_2 Sg_{\mu\nu} + c_3 g_{\mu\nu}.$$

If one requires, in addition, that $E_{\mu\nu}$ has zero divergence (that is, $g^{\lambda\nu}\nabla_\lambda E_{\mu\nu} = 0$), then

$$E_{\mu\nu} = c_1\left(R_{\mu\nu} - \frac{1}{2}Sg_{\mu\nu} + \lambda g_{\mu\nu}\right).$$

From this point of view it is clearly the *a priori* decision of neglecting λ, rather than that of including it, that needs a justification. The curious thing is that Einstein referred to a 'Vermeil-like' theorem in his 1916 outline [15, p. 144] – he did it before the theorem had been proved, and his statement there is mistaken precisely because it omits the c_3 term! It is true that this omission can be justified *a posteriori* by appeal to the Newtonian limit (§4); nonetheless the popular perception of the λ-term as a 'blunder' is certainly off the mark.[23]

It is interesting that Eddington and Lemaître both felt very strongly about the cosmological constant; the former wrote: "I would as soon think of reverting to Newtonian theory as of dropping the cosmical constant"; and the latter, when asked what had been the most important conceptual novelty introduced by general relativity, answered "without a moment's hesitation [...] 'the introduction of the cosmical constant' " [3, p. 122].

Eddington commented in 1933 that he had difficulty in discussing the proposal of Einstein and de Sitter, since he did

> not see what are 'the rules of the game'. These proposals are left as mathematical formulations, all doubtless compatible with what we observe; but there seems nothing to prevent such formulations being indefinitely multiplied [13, p. 63].

As to the singular initial state, Eddington found it unacceptable: as he wrote in 1928: "As a scientist I simply do not believe that the Universe began with a bang" – little or big, we are tempted to add [53, p. 46]. He showed in 1930 [12] that the Einstein space–time is unstable, because if it happens that ρ at some instant is different from $2\lambda/\kappa c^2$ (see §4), then from that moment on the universe will either be expanding or contracting (this can be seen as an easy consequence of (24)–(25)). Eddington wrote:

> I may mention that the proof of the instability of the Einstein configuration was the turning point in my own outlook. Previously the expanding universe (as it appeared in de Sitter's theory) had appealed to me as a highly interesting possibility, but I had no particular preference for it [13, p. 60n].

His preference was in fact for a nonsingular, spatially spherical universe asymptotically arising from Einstein static space–time and developing into de Sitter space–time (the so-called *Lemaître-Eddington model*: $K = 1$, $\lambda = \lambda_c$, with pressure). Eddington

[23] For a short history of the cosmological constant, and its recent resurrection, the note by G.F.R. Ellis [29] should be consulted.

was convinced that it was the only reasonable cosmological model available, particularly because it did not conflict with geological time estimates.[24]

Lemaître, on the other hand, gave his preference in 1931 to a singular model starting with a "primeval atom" – a first version of what is today called the "standard" cosmology.

14. Neo-Newtonian Cosmology

As we have seen, the cosmological discourse has been framed into Newtonian categories since the beginning of relativistic cosmology.

Apparently Lemaître was the first to notice, in 1931, that elementary considerations of Newtonian mechanics made it possible to derive Friedmann's Eq. (26) [34, pp. 221–2].

Three years later, the British mathematical physicists Edward Milne and William McCrea showed in detail that the Friedmann solutions are the relativistic counterpart of the possible evolutions of an homogeneous and isotropic fluid (solid) sphere kept together by gravitation only. In particular the Friedmann equations can be derived from the dynamical and the continuity equations for this fluid.

Since Newtonian space is Euclidean, the curvature of Friedmann's solutions must undergo a radical reinterpretation: the three alternatives correspond to the sphere expanding forever with exactly ($K = 0$) or more ($K = -1$) than the escape velocity, or collapsing on itself ($K = 1$). The cosmological constant can be interpreted in Newtonian terms as an additional "repulsive force proportional to distance", and Hubble's law is also valid.

Milne [60] wrote:

> It seems to have escaped previous notice that whereas the theory of the expanding universe is generally held to be one of the fruits of the theory of relativity, actually all the at-present-observable phenomena could have been predicted by the founders of mathematical hydrodynamics in the eighteenth century, or even by Newton himself.

Of course the fundamental observers in the Newtonian version cannot be fully equivalent, since they are accelerated observers, with different accelerations, so Milne's claim does not exactly hold, unless we correct Newton's physics by introducing the equivalence principle; moreover, a finite sphere has a centre and a boundary, so its points do not enjoy exactly the same view! The main interest in this revival of Newtonianism lies, apart from its paedagogical value, in the way it showed how close general relativity had come, in dealing with the cosmological problem, to classical mechanics.[25]

15. The Steady-State Theory

In 1948 two papers appeared advancing a cosmological theory which deviated from the orthodox line of using general relativity as the basis for cosmological speculation. Their authors, Hermann Bondi and Thomas Gold on one side, Fred Hoyle on the other, had

[24] "In Eddington's time most cosmological models except his own implied an 'age of the universe' less than 2×10^9 years. This was less than half the age of the oldest rocks on Earth!" (W. McCrea, in [13, p. xviii]).

[25] For a recent treatment see [74, pp. 18–28], where it is explained how one can circumvent the difficulties mentioned in the text.

extensively discussed the issue previously, so their convergence was not a case of independent discovery.[26]

Bondi and Gold adopted a simple deductive approach, assuming the relativistic space–time but not the field equations – that is, the *geometric* component of general relativity without the *dynamical* component. The main feature was the postulation of the so-called *perfect cosmological principle*, stating that the universe has to look the same not only at each point (homogeneity) and in each direction (isotropy), as in the FLRW models, but *also for each time*. Of course this principle would also be satisfied by Einstein's static universe, but the strength of the new theory was that it took into account, convincingly, those observational data that had defeated Einstein's universe, first of all the accepted redshift-distance relationship for galaxies.

Starting from (21), which is needed in order to comply with the ordinary cosmological principle, the further determination of $R(t)$ and $d\sigma^2$ is effected by exploiting the 'perfect' clause. In fact the curvature of the universe at time t is $K(t) = k/R(t)^2$, where $k = -1, 0, 1$, so for this (observable) quantity to be independent of time the only possibility (if we reject Einstein space–time) is $k = 0$, that is, the universe must be Euclidean:

$$ds^2 = -R(t)^2\big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\big) + c^2\,dt^2.$$

Moreover, the proportionality factor between velocity (observable as redshift) and distance of galaxies is given in such a universe by the Hubble constant $H(t) = \dot{R}(t)/R(t)$. The perfect cosmological principle requires that $H(t)$ be constant, say $H(t) \equiv H_0$, giving

$$R(t) = R(0)e^{H_0 t}.$$

So, if we posit $R(0) = 1$, the final expression of the metric is:

$$ds^2 = -e^{2H_0 t}\big((dx^1)^2 + (dx^2)^2 + (dx^3)^2\big) + c^2\,dt^2, \tag{29}$$

which coincides with the metric of de Sitter space–time as displayed in (18) if $R_0 = c/H_0$. In fact this last condition is not completely a matter of notation, since it can only hold if $H_0 > 0$. A further attractive feature of the steady-state approach is that it enables one to derive this inequality – that is, that the universe is expanding – from two established astronomical facts of the most elementary nature, namely: (1) that the universe is not in thermodynamic equilibrium, and (2) that the sky at night is dark.[27]

Of course an infinite and uniform universe which is forever expanding must have one disturbing peculiarity: new matter must be created continuously, in order for the matter density to remain constant (this is not required in FLRW space–times, where the matter density is time-dependent, cf. (27)). In other words, the steady-state cosmology had to introduce a systematic violation of the local energy-mass conservation (cf. Chap. 5, §13). It is true that the quantity of created matter needed is very small, nonetheless to many physicists this seemed rather hard to swallow.

As we have seen, the derivation of de Sitter metric in the steady-state approach is obtained without using Einstein's field equations; as a consequence the steady-state uni-

[26] It seems that Gold was the originator of the basic idea [53, pp. 173–9].

[27] Point (2) is enough to dismiss a static and eternal *Euclidean* universe if, furthermore, stars are assumed to be uniformly distributed (*Olbers paradox*).

verse must not be thought as empty (or filled with exotic matter), and in fact it can accommodate *any* constant mass density ρ_0. The weak side of this freedom is that, contrary to the FLRW models, there is no link between H and ρ (§11). This also depends on the fact that no alternative field equations describing the matter creation process were presented by Bondi and Gold.

It was Hoyle's paper which filled this gap. Hoyle introduced a scalar field C (that is, a real-valued function on space–time) depending only on the time coordinate, and suggested the following modification of Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2} S g_{\mu\nu} = -\kappa(T_{\mu\nu} + C_{\mu\nu}), \tag{30}$$

where $C_{\mu\nu}$ is the covariant second derivative of C (i.e. $C_{\mu\nu} = \nabla_\nu C_\mu$). It can be verified that if $T_{\mu\nu}$ is a matter tensor for a dust of density ρ and C is proportional to t, then (30) has the de Sitter metric as a solution, with $T_{\mu\nu} \neq 0$.

The steady-state theory was on several occasions defended by Bondi on methodological grounds, since it was allegedly *more* open to falsification, thus more scientific (in Popperian terms), than its rivals. The implication was that one had to prefer, at least provisionally, the steady-state theory because it was more at risk of being in conflict with observations. Given the scarcity of observational evidence at the time, the argument looked disquietingly close to claiming support for the theory from non-existent data.[28] Interestingly, Popper himself never endorsed the steady-state theory, of which he disliked the very cornerstone, the perfect cosmological principle; he also came to a rather poor opinion of the scientific status of physical cosmology as a whole [53, pp. 224–6].

Hoyle's line was different. He stressed the way the steady-state theory, with his field equations, allowed the avoidance of ad hoc initial conditions to satisfy Mach's principle; he and his collaborator Jayant V. Narlikar [44] argued that *any* solution of (30) tended asymptotically to the de Sitter universe, thus guaranteeing the link between the space–time metric and matter which is at the core of Mach's principle, and which is lost in the FLRW cosmology, as will be even more clear from the next section.

16. The Gödel's Universes

At the end of the 1940s the great logician Kurt Gödel started working on general relativity. The results of this inquiry were outlined in three short papers [36–38], which are among the most important in the history of general relativity.

As we have seen, the cosmological solutions of the field equations so far examined share two properties which are clearly inspired by the 'old' Newtonian setting: they have a privileged time coordinate $x^4 = t$, and matter defines a class of privileged, freely-falling observers (modelling the galaxies); the worldlines of these observers have a velocity vector $u^\mu = (0, 0, 0, 1)$ which is orthogonal to the level hypersurfaces $t = const$. In order for these conditions to hold for a metric $ds^2 = g_{\mu\nu}\, dx^\mu\, dx^\nu$ one must have

[28] The worthy side of the argument, that is, that the bolder theories should be given more attention since their very failures are likely to "throw light upon the state of contemporary science and may indicate where it requires supplementing", was of course not a very original idea (the quotation is by mathematician H.T.H. Piaggio, quoted approvingly by Eddington [13, p. 124]).

$$g_{a4} = 0, \qquad g_{44} = const.$$

The first condition ensures that matter is not rotating. A universe where matter has large-scale rotation clearly violates Mach's principle (cf. Chap. 5, §1).

In his first article Gödel presented explicitly a new exact solution of the modified field equations, with matter in the form of a dust, which does not admit any cosmic time and where matter (in the form of a dust with constant density) rotates. The metric, which can be considered as defined on all of $\mathbb{R}^4$, has a simple form:

$$ds^2 = a^2\left[-\left(dx^1\right)^2 - \frac{e^{2x^1}}{2}\left(dx^2\right)^2 - \left(dx^3\right)^2 + \left(e^{x^1}dx^2 + dx^4\right)^2\right], \tag{31}$$

and is a solution of (7) with a negative cosmological constant. It is easy to see that the fourth coordinate vector field is timelike; however, the fourth coordinate is not a cosmic time, i.e. a scalar function on space–time which is increasing on every future-pointing timelike curve. In fact this space–time does not admit *any* cosmic time; we shall see why after a short detour.

Gödel's model enjoys a curious property, which may be introduced by quoting a passage from Weyl's treatise [79, p. 274]:

> [...] it is not impossible for a world-line (in particular, that of my body), although it has a time-like direction at every point, to return to the neighbourhood of a point which it has already once passed through. The result would be a spectral image of the world, more fearful than anything the weird fantasy of E.T.A. Hoffmann has ever conjured up. In actual fact the very considerable fluctuations of the $[g_{\mu\nu}]$ that would be necessary to produce this effect do not occur in the region of world in which we live.

Weyl went on to remark that "there is a certain amount of interest in speculating on these possibilities inasmuch as they shed light on the philosophical problem of cosmic and phenomenal time". This is in fact what had prompted Gödel to work out his solution, and in his paper he proved that his space–time does contain closed timelike curves; indeed there are closed time-like curves through *every* point![29] This is interesting both in itself, and because it clearly shows that in such a a space–time there can be no cosmic time function. Moreover, this confirms that the mere form of the stress-energy tensor does not dispense in general relativity from appeal to specifically cosmological assumptions if one is to arrive at the FLRW space–times.

A further reason of interest is that Gödel space–time is simply-connected; this is important because it is rather easy to construct space–times with closed timelike curves (or even geodesics) if non-simply-connected topologies are allowed (e.g. one can identify points in Minkowski space–time).

Finally, as Gödel writes, his solution shows that, according to general relativity, "it is theoretically possible [...] to travel into the past, or otherwise influence the past" [36, p. 447]. Today speculation on 'time travel' is thriving, and the boundary between science and science fiction is becoming increasingly difficult to discern [76,62].

---

[29] Though not closed *geodesics*. Incidentally, in his short article Gödel never claimed that, but two authors, one being S. Chandrasekhar, 'corrected' him in 1954 of this supposed 'mistake'... What is more and worse, it seems that it took *nine* years before someone pointed out in print that Gödel was innocent of the charge! See [64].

The second paper, a very dense one, extracted a philosophical moral from the existence of such solutions. That time is not an objective entity out there has been argued by idealistic philosophers; the relativity of the time order in special relativity seemed to confirm this claim, but now general-relativistic models of the universe have reinstated a sort of absolute time under the name of cosmic time.[30] So one might argue that *even according to general relativity*, time cannot be denied an objective existence, at least on a cosmological scale. But Gödel's solutions show that a cosmology *without* cosmic time is perfectly compatible with general relativity. So after all there is still a remarkable agreement between the idealistic concept of time and modern physics: the inexorable, unique, one-way flowing of time may still be considered as a subjective illusion.[31]

Note that (31) is a stationary space–time, thus there are no cosmological redshifts in it. In his third paper Gödel stated, with little or no proofs, that rotating and non-static universes, spatially homogeneous and with closed timelike curves, exist. This seminal paper [38] was delivered as a communication at the 1950 International Congress of Mathematics, and over the next two decades it was a major impulse to research into exact solutions and topological and causality problems in general relativity.

17. The Big-Bang Cosmology 'Wins'

The discussion on the relative virtues of the steady-state cosmology and the evolutionary models – the "Big Bang" cosmology, as it was named after a scornful remark in 1950 by Fred Hoyle – during the decade starting in the mid-Fifties increasingly concentrated on astronomical issues. The most relevant results, which seemed to favour the evolutionary models, were the radio-sources counts, the redshifts of the quasars, and, perhaps most importantly, the black-body microwave radiation accidentally discovered by Arno Penzias and Robert Wilson in 1964, and interpreted as a relic of the 'big explosion' by Robert Dicke and collaborators.

One of the reasons of the success of this theory – favourably mentioned by Pope Pius XII in 1951 [53, p. 256] – was the possibility of marrying particle physics with cosmology when dealing with the first few instants of the universe after the Big Bang. The first paper discussing nucleosynthesis in the initial 'hot' stage of the universe was published by Georg Gamow in 1946.[32]

Not all competent scientists, however, have been convinced by these results that the Big Bang is the only acceptable picture of the universe (cf. also Chap. 13). In a lecture at the International Centre for Theoretical Physics (Trieste, Italy), in June 1968, Paul A.M. Dirac had this to say on cosmology [9, pp. 127–8]:

---

[30] Gödel quotes a lecture by James Jeans in 1935. In 1943 Jeans still wrote: "The hypothesis that absolute time and space do not exist brings order into man-sized physics, but seems so far to have brought something like chaos into astronomy. Thus there is some chance that the hypothesis may not be true. [...] It may be that before it can make sense, the new astronomy must find a way of determining an absolute time, which will then describe as cosmical time. The space–time unity will then be divided into space and time separately by nature itself" [51, pp. 67, 68].

[31] Einstein's reply ended in dubitative form: "It will be interesting to weigh whether these [Gödel's solutions] are not to be excluded on physical grounds" [73, p. 688].

[32] See Gamow's later survey article [35].

> One field of work in which there has been too much speculation is cosmology. There are very few hard facts to go on, but theoretical workers have been busy constructing various models for the universe, based on any assumptions that they fancy. These models are probably all wrong. It is usually assumed that the laws of nature have always been the same as they are now. There is no justification for this. The laws may be changing, and in particular quantities which are considered to be constants of nature may be varying with cosmological time. Such variations would completely upset the model makers.

As to the supposedly 'clinching' evidence of the cosmic background radiation, this was Fred Hoyle's comment thirty years after the discovery:

> How, in big-bang cosmology, is the microwave backgroud explained? Despite what supporters of big-bang cosmology claim, it is not explained. The supposed explanation is nothing but an entry in the gardener's catalogue of hypotheses that constitutes the theory. Had observation given 27 kelvins instead of 2.7 kelvins for the temperature, then 27 kelvins would have been entered in the catalogue. Or 0.27 kelvin. Or anything at all [43, p. 413].

In fact, apart from these methodological criticisms, the big-bang cosmology has to face a number of serious difficulties; some of them will be briefly mentioned in the next section.

18. Singularities and Other Open Problems

Classically, if in a theory some physical quantity becomes infinite for some value of a given parameter, this is considered a symptom that the theory reaches its limits of application near that value. In fact, since no real-world physical quantity is infinite, there must be something defective in the theory. Now one common feature of the expanding cosmological models with $\lambda = 0$ is that they all have one or two finite values of the cosmic time for which the metric degenerates. As we have seen (§6), Einstein's criterion for a singularity was the vanishing of $g = \det(g_{\mu\nu})$. In fact the problem of how to define a singularity in general relativity proved to be more tricky.

A common view about why singularities arise in the FLRW cosmological models was that they embodied unrealistic symmetry assumptions. A first step to remove this doubt was made in 1955, when an Indian theoretical physicist, Amalkumar Raychaudhuri succeeded (not very easily, in fact) in having a paper published in *Physical Review* in which he questioned this opinion [68]. Suppose that in a relativistic space–time, with matter modelled as a dust of density ρ, we have a coordinate system where the x^4-coordinate lines are geodesics which coincide with the worldlines of the particles of matter, for which $x^4 = t$ is the proper time. Then along any such worldline, say Γ, the following equation holds:

$$\frac{1}{G}\frac{\partial^2 G}{\partial t^2} = \frac{1}{3}\left(c^2\lambda - \frac{\kappa c^4 \rho}{2} - \phi^2 + 2\omega^2\right), \tag{32}$$

where $G = |g|^{1/6}$, and ω and ϕ are functions such that ω vanishes if and only if the matter worldlines are hypersurface-orthogonal, and ϕ vanishes if matter expands isotropically. From this equation it follows that if $\omega = 0$ and $\lambda \leq 0$, then G has a negative second derivative with respect to t, and this implies that at some time in the past G, and therefore g, has to vanish. Raychaudhuri's conclusion was that there was a "singularity at a finite time in the past as in isotropic models. Thus a simple change-over to anisotropy does not solve any of the difficulties" [68, p. 1125].

The argument is ingenious and correct insofar as it suggests that anisotropy is not the key, but it is inconclusive. The vanishing of g may well depend on the limits of validity of the coordinate system; for instance, it may be easily shown to occur for a suitable family of geodesic observers in (empty!) Minkowski space–time. What should be proven is that something *really* wrong, that is, something not coordinate-dependent, occurs on some timelike (or lightlike) geodesic. The simplest criterion for saying that something wrong is occurring on a timelike or lightlike geodesic is when its natural parameter (proper time, in the timelike case) cannot be extended beyond a certain finite value, either in the future or in the past. If not all inextendible timelike (resp. lightlike) geodesics are *complete* (i.e. their natural parameter ranges over all the real line), then space–time is said to be *timelike* (resp. *lightlike) incomplete*. A space–time which is incomplete in this sense and is not just a subspace of another, bigger space–time, is certainly a good candidate to be classified as genuinely *singular*.

In the mid-1960s a number of theorems were proven – by Roger Penrose, Stephen W. Hawking, Robert Geroch and others – stating that a space–time satisfying some rather broad global conditions (implying the non-existence of closed timelike curves, for instance) is singular in the sense just explained. Versions of (32) were a basic ingredient of the proofs. It is interesting to remember that for Einstein (cf. §5) the existence of a singularity in *his* sense was enough to rule out a space–time as a legitimate solution of his equations. What the singularity theorems proved was, roughly, that in relativistic cosmology singular space–times were *not* exceptional.[33]

The natural interpretation of the singularity theorem was endorsed by Hawking and George F.R. Ellis in their classic treatise *The Large-Scale Structure of Space-Time*, published in 1973:

> It seems to be a good principle that the prediction of a singularity by a physical theory indicates that the theory has broken down, i.e. it no longer provides a correct description of observations. The question is: when does General Relativity break down? [40, pp. 362–3].

Hawking and Ellis's guess is that "a breakdown occurs for lengths between 10^{-15} and 10^{-33} cm". They go on to connect the theorems with the big-bang picture:

> In any case, the singularity theorems indicate that the General Theory of Relativity predicts that gravitational fields should become extremely large. That this happened in the past is supported by the existence and black body character of the microwave background radiation, since this suggests that the universe had a very hot dense early phase.

Here "the General Theory of Relativity" is to be understood with the proviso that the cosmological term is forbidden; otherwise the singularity theorems might be seen as offering a good reason to preserve it, in the spirit of the Eddington–Lemaître space–time (§13). It is true that there are reasons suggesting that the reintroduction of the λ-term would not be enough to avoid a singularity, at least if the accepted theory of nucleosynthesis is maintained [29].

In any case, the cosmological term has other chances for a resurrection. In 1967 Russiam physicist Yakov Zel'dovich showed that in quantum field theory the stress-energy

[33] The survey [75] gives an interesting and useful discussion (pp. 796–808) of the scope and limits of the singularity theorems.

tensor of the vacuum behaves like a cosmological constant. The idea reappeared in 1980 in an influential variant of big-bang cosmology at the heart of which is the postulate that, very shortly after the big-bang (about 10^{-30} seconds later), there was a period of "inflation", that is, of very fast expansion. There are about 60 versions of "inflationary comology" on offer, to date; what is worse, the theoretical energy density of the quantum vacuum corresponds to a λ which is different from the one compatible with astronomical observations by 120 orders of magnitude (according to one estimate): this is "one of the major unsolved problems of theoretical physics" [29].

In 1990 Hawking summed up the state of the art in cosmology as regards another big open problem of standard cosmology, that of "dark matter":

> So there are two possibilities; either our understanding of the very early Universe is completely wrong, or there is some other form of matter in the Universe that we have failed to detect. The second possibility seems more likely, *but the required amount of missing 'dark' matter is enormous; it is about a hundred times the matter we can directly observe* [70, p. viii].

A rather curious situation, which will not fail to provide in the future years food for thought both for scientists and lay observers.

19. Conclusive Remarks

Cosmology is a peculiar kind of science (if it is a science at all) for the most obvious of reasons, namely that it is the study of a class of objects of which there is just one instance. In a sense it can be compared to geology or natural history, except for the worrisome circumstance that its object – space–time – is a theoretical construct in a stronger sense than can be said of the Earth or of the extinct life forms which we try to reconstruct from their relics. Though by now many more observations have been gathered than Einstein and his contemporaries could have dreamt of, cosmology cannot be but an observational, rather than an experimental, science, and moreover one whose empirical basis cannot be made sense of "without some admixture of ideology" [40, p. 134]. "*There are no purely observational facts about the heavenly bodies*", as Eddington wisely wrote in 1933 ([13, p. 17]; the italics are his).

One example is given by the cosmological principles. The standard one (§3) is often defended on grounds of Copernican humility; here is how the author of one of the classic treatises on cosmology, Richard C. Tolman, put it: "[...] it avoids the anthopocentric assignment of a unique importance to our own location in the universe", though he immediately added the somewhat inconsistent remark that it enables us to regard "the observations that we obtain as fairly representing the character of those which would be obtained from similar locations in other portions of the universe" [77, p. 363]. In other words, our position in the universe lacks "importance" only in some indefinite sense of the word, but is otherwise *very* good, since it allegedly permits us to use local observations for global purposes. The net outcome is clearly one of ('Copernican'?) scientific optimism.[34]

General relativity has surely enlarged the conceptual tool-box of the cosmologists, enabling them to discuss in mathematical terms a bewildering variety of more or less

[34] In fairness to Tolman, it must be added that a few lines below he writes that it is "most important" to emphasize that "the assumption is to be regarded merely as a working hypothesis".

well-behaved or weird possible worlds. However, it must be emphasized that even in this largest of the physical systems – the universe – relativistic cosmology has consistently relied upon Newtonian analogies. Gödel's contributions arose as a reaction to the Newtonian dominance which he saw, correctly, as entrenching upon the philosophical interpretation of cosmological models. It still remains to be seen whether the strange possibilities opened by general relativity and revealed by Gödel and others after him should be rated as a strength or a weakness of the theory.

Of all fields in which an orthodoxy can coagulate, cosmology would seem to be one of the most unlikely, given the fragility of both theory and evidence; nevertheless, even here this has happened. (Religion might be judged, on the same grounds, an even more unlikely candidate, and yet ...). The phenomenon has many reasons, like the strongly hierarchical structure of the scientific community, and (not least) the sorry tendency of scientific professionals in most fields today to overstate their case in order to seduce the public into supporting increasingly expensive research programmes. However, inflated claims may eventually backfire and cause discredit both to the individuals and to their disciplines [10].

In the history of cosmology, as in other scientific fields, we see time and again that what receives the most attention is not necessarily what is correct or even promising, but what is in agreement with the opinions, and even errors, of the scientific leaders. 'Authoritative' ideas can resist correction for decades, the more so – obviously – if there are very few people who do most of the fundamental work in the area.[35] Moreover, even competent outsiders (like Friedmann) may simply be ignored unless they succeed in connecting (like Lemaître) with the establishment. Fortunately, if we count arguments rather than heads, as we always should in all matters scientific, cosmology has never stopped being an healthily controversial field, and strictures on dissidents have never succeeded in completely stifling the debate even on the most fundamental points (cf. Chap. 13).

One last remark is in order. Substantial progress in cosmology, more than elsewhere, is a long term business. It will ultimately depend on how long humankind will resist the destructive forces which threaten to put an untimely end to its adventure. This is a most telling instance to show that scientific research – if it is meant to be more than a way of channeling one's competitive drive or of gaining material assets – cannot be separated from humanitarian concerns and participation in general cultural and political issues.

References

[1] Besse A. 1987: *Einstein Manifolds*, New York – Heidelberg – Berlin, Springer Verlag.

[2] Bondi H., Gold T. 1948: "The Steady State Theory of the Expanding Universe", *Monthly Notices of the Royal Astronomical Society*, **108**, pp. 252–70.

[3] Chandrasekhar S. 1987: *Truth and Beauty*, Chicago – London, The University of Chicago Press.

[4] Clifford W. K. 1955: *The Common Sense of the Exact Sciences*, New York, Dover.

[35] Cf.: "On various occasions in the history of cosmology the subject has been dominated by the bandwagon effect, that is, strongly held beliefs have been widely held because they were unquestioned or fashionable, rather than because they were supported by evidence. As a result, particular theories have sometimes dominated the discussion while more convincing explanations were missed or neglected for a substantial time, even though the basis for their understanding was already present" [28, p. 367]. There is plenty of evidence to show that cosmology is not exceptional in this respect.

[5] Cohen I. B. (ed.) 1978: *Isaac Newton's papers & Letters on Natural Philosophy and related documents*, 2nd edn, Cambridge (Mass.) and London, Harvard University Press.
[6] Coles P., Ellis G. F. R. 1997: *Is the Universe Open or Closed*, Cambridge University Press.
[7] Crowe M. 1994: *Modern Theories of the Universe from Herschel to Hubble*, New York, Dover.
[8] de Sitter W. 1917: "On the relativity of inertia", *Proc. Kon. Akad. Wetensch. Amst.*, **19**, pp. 1217–25.
[9] Dirac P. A. M. 1968: "Methods in theoretical physics", pp. 125–43 of *Unification of Fundamental Forces*, Cambridge University Press, 1990.
[10] Disney M. J. 2000: "The Case Against Cosmology", *General Relativity and Gravitation*, **32**, pp. 1125–34.
[11] Eddington A. S. 1918: *Report on the Relativity Theory of Gravitation*, London, Fleetway Press.
[12] Eddington A. S. 1930: "On the instability of Einstein's spherical world", *Monthly Not. Roy. Astr. Soc.*, **91**, pp. 412–6.
[13] Eddington A. S. 1933: *The Expanding Universe*, Reprint with a preface (1987) by W. McCrea, Cambridge University Press.
[14] Egginton W. 1999: "On Dante, Hyperspheres, and the Curvature of the Medieval Cosmos", *J. Hist. Ideas*, **60**, pp. 195–216.
[15] Einstein A. 1916: "Grundlage der allgemeine Relativitätstheorie", *Ann. Phys.*, **49**, pp. 769–822.
[16] Einstein A. 1917: *Über die spezielle und die allgemeine Relativitätstheorie, Gemeinverständlich*, Braunschweig, Vieweg.
[17] Einstein A. 1917: "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie", *Preussische Akademie der Wissenschaften, Sitzungsberichte*, P. 1, pp. 142–52.
[18] Einstein A. 1918: "Kritisches zu einer von Hrn. De Sitter gegebenen Lösung der Gravitationsgleichungen", *Preussische Akademie der Wissenschaften, Sitzungsberichte*, pp. 270–2.
[19] Einstein A. 1919: "Spielen Gravitationsfelder im Aufber der materiellen Elementarteilchen eine wesentliche Rolle?", *Königlich Preussische Akademie der Wissenschaften* (Berlin), *Sitzungsberichte*, pt 1, pp. 349–56 (transl. in Einstein et al. 1932, pp. 189–98).
[20] Einstein A. 1922: *The Meaning of Relativity*, Princeton, Princeton University Press.
[21] Einstein A. 1922: "Bemerkung zu der Arbeit von A. Friedmann 'Über die Krümmung des Raumes' ", *Zeitschrift für Physik*, **21**, p. 326.
[22] Einstein A. 1923: "Notiz zu der Arbeit von A. Friedmann 'Über die Krümmung des Raumes' ", *Zeitschrift für Physik*, p. 228.
[23] Einstein A. 1931: "Zum kosmologischen Problem der allgemeinen Relativitätstheorie", *Preussische Akademie der Wissenschaften, Sitzungsberichte*, pp. 235–7.
[24] Einstein A. 1954: *Relativity. The Special and the General Theory*, 15th ed., Routledge, London and New York.
[25] Einstein A. 1994: *Ideas and Opinions* [1954], New York, The Modern Library.
[26] Einstein A., De Sitter W. 1932: "On the relation between the expansion and the mean density of the universe", *Proc. Nat. Acad. Sci.*, **18**, pp. 213–4.
[27] Einstein A., Lorentz A. A., Weyl H., Minkowski H. 1932: *The Principle of Relativity*, New York, Dover.
[28] Ellis G. F. R. 1989: "The Expanding Universe: A History of Cosmology from 1917 to 1960", pp. 367–431 of *Einstein and the History of General Relativity*, ed. by D. Howard and J. Stachel, Basel, Birkhäuser.
[29] Ellis G. F. R. 2003: "A historical review of how the cosmological constant has fared in general relativity and cosmology", *Chaos, Solitons and Fractals*, **16**, pp. 505–12.
[30] Engel A., Shucking E. (eds.) 1997: *The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings, 1914–1917*, English Translation, Princeton-Oxford, Princeton University Press.
[31] Engel A., Shucking E. (eds.) 2002: *The Collected Papers of Albert Einstein, vol. 7, The Berlin Years: Writings, 1918–1921*, English Translation, Princeton-Oxford, Princeton University Press.
[32] Friedmann A. A. 1922: "Über die Krümmung des Raumes", *Zeitschrift für Physik*, **10**, pp. 377–87. English translation in *General Relativity and Gravitation*, **31** (1999), pp. 1991–2000.
[33] Friedmann A. A. 1923: "Über die Möglichkeit einer Welt mit konstanter negativer Krümmung des Raumes", *Zeitschrift für Physik*, **21**, pp. 326–33. English translation in *General Relativity and Gravitation*, **31** (1999), pp. 2001–8.
[34] Friedmann A. A., Lemaître G. 1997: *Essais de Cosmologie*, with an introduction by J.-P. Luminet, Paris, Seuil.
[35] Gamow G. 1949: "On Relativistic Cosmogony", *Rev. Mod. Phys.*, **21**, pp. 367–73.
[36] Gödel K. 1949: "An Example of a New Type of Cosmological Solutions of Einstein's Field Equations of Gravitation", *review of Modern Physics*, **21**, pp. 447–50.

[37] Gödel K. 1949: "A remark about the relationship between relativity theory and idealistic philosophy", pp. 557–62 of [73].

[38] Gödel K. 1952: "Rotating Universes in General Relativity Theory", in L. M. Graves (ed.), *Proceedings of the International Congress of Mathematicians*, Cambridge (Mass.), vol. 1, pp. 175–81.

[39] Harvey A., Schucking E. 2000: "Einstein's mistake and the cosmological constant", *Am. J. Phys.*, **68** (8), pp. 723–7.

[40] Hawking S. W., Ellis G. F. R. 1973: *The Large-Scale Structure of Space–Time*, Cambridge University Press.

[41] Hetherington N. S. 1988: *Science and Objectivity*, Ames, Iowa State University Press.

[42] Hoyle F. 1948: "A New Model for the Expanding Universe", *Monthly Notices of the Royal Astronomical Society*, **108**, pp. 372–82.

[43] Hoyle F. 1997: *Home Is Where the Wind Blows*, Oxford University Press.

[44] Hoyle F., Narlikar J. V. 1962: "Mach's principle and the creation of matter", *Proc. Roy. Soc. London*, **A270**, pp. 334–41.

[45] Hubble E. P. 1925: "NGC 6822, a remote Stellar System", *Astrophys. J.*, **62**, pp. 409–33.

[46] Hubble E. P. 1926: "Extra-Galactic Nebulae", *Astrophys. J.*, **64**, p. 321.

[47] Hubble E. P. 1929: "A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae", *Proc. Nat. Acad. of Sciences*, **15**, pp. 168–73 (reproduced in Crowe 1994, pp. 338–45).

[48] Hubble E. P. 1935: "Angular Rotations of Spiral Nebulae", *Astrophys. J.*, **81**, pp. 334–5.

[49] Hubble E. P. 1953: "The Law of Red-Shifts", *George Darwin Lecture*, May 8, 1953, edited by A. R. Sandage.

[50] Janssen M., "The Einstein-De Sitter Debate and Its Aftermath", `www.tc.umn.edu/~janss011/pdf%20files/Einstein-De%20Sitter.pdf`.

[51] Jeans J. 1943: *Physics and Philosophy*, New York, Dover 1981.

[52] Kahn C. & F. 1975: "Letters from Einstein to de Sitter on the nature of the Universe", *Nature*, **257**, pp. 451–4.

[53] Kerzsberg P. 1989: *The Invented Universe*, Oxford University Press.

[54] Kragh H. 1996: *Cosmology and Controversy*, Princeton, Princeton University Press.

[55] Krasiński A., Ellis G. F. R. 1999: "Editor's Note", *General Relativity and Gravitation*, **31**, pp. 1985–90.

[56] Lanczos C. 1922: "Bemerkung zur De Sitterschen Welt", *Physikalische Zeitschrift*, **23**, pp. 539–43.

[57] Lemaître G. 1925: "Note on De Sitter's Universe", *J. Math. Phys.*, **4**, pp. 37–41.

[58] Lemaître G. 1927: "Un universe homogène de masse constante et de rayon croissant, rendant compte de la vitesse radiale des nébuleuses extragalactiques", *Annales de la Société scientifique de Bruxelles*, series A, **47**, pp. 29–39.

[59] Lemons D. S. 1988: "A Newtonian cosmology that Newton would understand", *Am. J. Phys.*, **44**, pp. 502–4.

[60] Milne E. A. 1934: "A Newtonian Expanding Universe", *Quart. J. Math. Oxford*, **5**, pp. 64–72.

[61] Milne E. A., McCrea W. 1934: "A Newtonian Expanding Universe", *Quart. J. Math. Oxford*, **5**, pp. 73–80.

[62] Nahin P. J. 1993: *Time Machines*, New York, American Institute of Physics.

[63] North J. D. 1990: *The Measure of the Universe* [1965], New York, Dover.

[64] Ozsváth I., Schucking E. 2003: "Gödel's trip", *Am. J. Phys.*, **71**, pp. 801–5.

[65] Pais A. 1982: *Subtle Is the Lord... The Science and Life of Albert Einstein*, New York, Oxford University Press.

[66] Pauli W. 1958: *Theory of Relativity*, New York, Dover.

[67] Peterson M. A. 1979: "Dante and the 3-sphere", *Am. J. Phys.*, **47**, pp. 1031–5.

[68] Raychaudhuri A. 1955: "Relativistic Cosmology. I", *Physical Review*, **98**, pp. 1123–6.

[69] Raychaudhuri A. 1957: "Singular State in Relativistic Cosmology", *Physical Review*, pp. 172–3.

[70] Riordan M., Schramm D. 1993: *The Shadows of Creation*, Foreword by Stephen Hawking, Oxford University Press.

[71] Robertson H. P. 1933: "Relativistic Cosmology", *Reviews of Modern Phsycis*, **5**, pp. 62–90.

[72] Sachs R. K., Wu H. 1977: *General Relativity for Mathematicians*, New York – Heildelberg – Berlin, Springer-Verlag.

[73] Schilpp P. A. (ed.) 1949: *Albert Einstein: Philosopher-Scientist*, Evanston (Ill.), Library of Living Philosophers.

[74] Sciama D. W. 1993: *Modern Cosmology and the Dark Matter Problem*, Cambridge University Press.

[75] Senovilla J. M. M. 1997: “Singularity Theorems and Their Consequences”, *General Relativity and Gravitation*, **29**, pp. 701–848.
[76] Thorne K. S. 1994: *Black Holes & Time Warps*, New York – London, Norton.
[77] Tolman R. C. 1934: *Relativity Thermodynamics and Cosmology*, New York, Dover 1987.
[78] van Maanen A. 1935: “Internal Motions in Spiral Nebulae”, *Astrophys. J.*, **81**, pp. 336–7.
[79] Weyl H. 1922: *Space–Time–Matter*, transl. of the 4th (1921) ed., with a new preface (1950), New York, Dover.
[80] Weyl H. 1923: *Raum, Zeit, Materie. Vorlesungen über allgemeine Relativitätstheorie*, 5th ed., Springer.
[81] Weyl H. 1923: “Entgegnung auf die Bemerkungen von Herrn Lanczos über die de Sittersche Welt”, *Physikalische Zeitschrift*, **24**, pp. 130–1.
[82] Weyl H. 1923: “Zur allgemeinen Relativitätstheorie”, *Physikalische Zeitschrift*, **24**, pp. 230–2.
[83] Weyl H. 1930: “Redshift and relativistic cosmology”, The Lon. Edin. Dubl. Phil. Mag. and J. of Sci., **9**, pp. 936–43.