Arch. Hist. Exact Sci. 55 (2001) 189–220. © Springer-Verlag 2001

Lambda: The Constant That Refuses to Die

JOHN EARMAN

Communicated by J.Z. BUCHWALD

1. Introduction

Einstein's general theory of relativity (GTR) was only a little over a year old when he added a "cosmological term" to his gravitational field equations. From the beginning, Einstein was not happy with this term because, it seemed to him, it marred the beauty of his theory, and over the years he found its presence increasingly distasteful, eventually referring to it as his "biggest blunder."[1] But although the cosmological term has, from time to time, been perceived to give off a bad odor, it has stubbornly refused to go away – partly because (Einstein notwithstanding) there is a logical place for it in GTR, partly because it has always had its champions, and partly because it gets periodically called upon to solve problems and resolve crises in cosmology. And as I write, the observational evidence is mounting that the actual universe is characterized by a positive cosmological constant (Λ), or something that acts like one. Thus, the time seems ripe to review the history of the oft reviled but never defeated Λ.[2]

The plan of the paper is as follows. Section 2 provides a taste of the nineteenth century antecedents of Λ. Section 3 discusses Einstein's (1917a) "Cosmological Considerations on the General Theory of Relativity" where Λ makes its debut. Einstein put Λ on stage both to make room in his GTR for a static cosmological model and to harmonize the GTR with Mach's principle. The physical implications of Λ as a cosmic repulsive force are explained in Sect. 4. Three developments that led Einstein to reject the cosmological term are discussed in Sect. 5. After the rejection, Einstein had a "bad conscience" about having introduced Λ in the first place and would have nothing further to do with it. But as will be seen in Sect. 6, other physicists thought that once the genie had been let out of the bottle it was not so easy to put it back again. Both R. C. Tolman and Georges Lemaître argued that Λ has a natural place in GTR and that only observation can decide

[1] Attributed to Einstein by Gamow (1958, 66–67). Note that Gamow does not purport to be quoting Einstein directly. He says: "Einstein remarked to me many years ago that the cosmic repulsion idea was the biggest blunder he had made in his entire life." The account given in Gamow's autobiography *My World Line* (1970, 44) is similar. The existing evidence is insufficient to decide whether Einstein himself used the word "blunder" or whether this was Gamow's embellishment.

[2] Much relevant historical information on Λ is found in North (1965) and Kragh (1996). There are a number of good review articles on the physics of Λ. Especially recommended are Carroll, Press, and Turner (1992), Cohn (1998), and Sahni (1999).

whether in fact its value is zero. As recounted in Sect. 7, there were also more positive reasons for keeping Λ alive. First, a positive Λ offered a solution to the apparent paradox that the earth and stars are older than the expansion age of the universe, which is bounded from above by the Hubble time. Second, Lemaître championed a cosmological model with $\Lambda > 0$ that promised to explain cosmological structure formation. But as time wore on and estimates of the Hubble time went up, the age problem seemed less pressing, and more detailed investigations of Lemaître's model deflated the promise to explain structure formation. Λ was ready to be put back on the shelf. However, as described in Sect. 8, the ghost of Λ was kept alive in steady state cosmology. And, as detailed in Sect. 9, in the late 1960s Λ was explicitly invoked once more, this time to explain the predominance of $z = 2$ redshift measurements for quasi-stellar objects. This revival lasted only a few years. However, particle physicists began to argue that it is difficult if not impossible to put Λ back on the shelf since it is to be interpreted as the vacuum energy density of quantum fields. The controversy about how to calculate the value of this energy density is critically examined in Sect. 10. Section 11 takes a brief look at inflationary cosmology which provides for an effective cosmological constant in the very early universe of such a large value that the resulting accelerated expansion (or inflation) produces observed features of the cosmos that are otherwise puzzling in the standard big bang model. However, simple versions of inflation run into problems unless Λ is present to make up for the "missing mass" in the universe. Evidence that Λ is indeed with us is reviewed in Sect. 12. Recent redshift measurements for supernovae indicate that the rate of expansion of the universe is increasing, which requires either a $\Lambda > 0$ or else a strange form of matter (now called "quintessence") that mimics some of the features of Λ. Conclusions are presented in Sect. 13.

2. Prehistory of Λ

In the 1890's both von Seeliger and Neumann realized that standard Newtonian cosmology did not provide for a consistent model of a cosmos with an infinite space filled with a static uniform distribution of matter [see von Seeliger (1895, 1896) and Neumann (1896)]. For in such a model the integrals corresponding to the Newtonian gravitational potential at a point and the Newtonian gravitational force at a point are divergent.[3] Two options were contemplated. The first was to maintain standard Newtonian gravitational theory and conclude that a static homogeneous cosmos is forbidden. What is permitted, for example, is an island universe in which there is a kind of center where the mass density ρ reaches a maximum and then falls off more rapidly than $1/r^2$ as $r \to \infty$ (where r measures the distance from the center). Such a model allows the relevant integrals to converge. The second option provides for a static homogeneous cosmology by modifying Newton's theory. In particular, Neumann proposed that the Newtonian gravitational potential $\varphi(r) = -\kappa m/r$ for a point mass m be multiplied by the factor $\exp(-\sqrt{\Lambda}r)$, where Λ is a positive constant, while von Seeliger proposed to

[3] For a detailed discussion of this and other problems in Newtonian cosmology, see Norton (1999).

add this term to Newton's inverse square force law. Such a modification of Newton's law of gravitation had already been proposed a number of times for various reasons [see North (1965, 17)]. By making Λ sufficiently small, deviations from standard Newtonian predictions will be indistinguishable except at large distances. Von Seeliger's version of the modification carries with it a corresponding modification of Poisson's equation

$$\nabla^2 \varphi = 4\pi\kappa\rho \tag{1}$$

to

$$\nabla^2 \varphi - \Lambda\varphi = 4\pi\kappa\rho. \tag{2}$$

Although Einstein does not refer to either von Seeliger or Neumann in the paper that introduces the cosmological constant into GTR, it is certain that he knew about the work of the former since it gets a mention in his *Über die Spezielle und die allgemeine Relativitätstheorie* (1917b), which is dated December 1916, two months before he communicated his cosmological paper (1917a) to the Berlin Academy.

3. Einstein's embrace of Λ

In "Cosmological Considerations on the General Theory of Relativity" Einstein (1917a) used the problem which had exercised von Seeliger and Neumann as a springboard to launch the cosmological term. He rejected the option of an island universe on the grounds that statistical mechanical considerations showed that, if the island is treated as a gas whose molecules are stars, the system would depopulate.[4] So, following von Seeliger, Einstein chose the second option of modifying Newton's law of gravitation. Replacing (1) by (2), Einstein noted, allows a solution of the form

$$\varphi = -\frac{4\pi\kappa\rho_o}{\Lambda}, \quad \rho_o = const. \tag{3}$$

i.e. a static uniform distribution of matter throughout space with a mean density of ρ_o.

Einstein then argued that, by analogy, the original gravitational field equations[5]

$$G_{\mu\nu} := R_{\mu\upsilon} - \frac{1}{2} g_{\mu\upsilon} R = 8\pi\kappa T_{\mu\nu} \tag{4}$$

can be generalized to

$$R_{\mu\upsilon} - \frac{1}{2} g_{\mu\upsilon} R + \Lambda g_{\mu\upsilon} = 8\pi\kappa T_{\mu\nu}. \tag{5}$$

The seeming analogy between the move from (1) to (2) and the move from (4) to (5) is enhanced if one operates with Einstein's sign conventions, on which the cosmological term in (5) is prefixed by a minus rather than a plus sign. In any case, this bit of reasoning

[4] This argument is questionable; see Norton (1999).

[5] Here $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\upsilon}$ is the Ricci tensor, $g_{\mu\upsilon}$ is the metric tensor, R is the Ricci curvature scalar, and $T_{\mu\nu}$ is the stress energy tensor. The signature convention for the metric is chosen to be $(+ + + -)$.

shows how slippery analogies are to handle, for the Newtonian limit of (5) is *not* (2) but rather

$$\nabla^2\varphi + \Lambda = 4\pi\kappa\rho \tag{2'}$$

[see Trautman (1965)].[6] But however shaky the analogical reasoning, Einstein arrived at the "right" result (5).[7]

What did the generalized field equations (5) accomplish for Einstein? Assume a homogeneous, isotropic spacetime. The most general line element takes the form

$$ds^2 = a^2(t)d\sigma^2 - dt^2 \tag{6}$$

where $a(t)$ is the scale factor, sometimes called the "radius of the universe" and thus often denoted by $R(t)$, and $d\sigma^2$ is the line element of a Riemann space of constant curvature $k = 0$ (flat space), $k = +1$ (closed space), or $k = -1$ (open elliptic space). In such a setting the stress-energy tensor must take the form of a perfect fluid

$$T^{\mu\nu} = (\rho + p)V^\mu V^\nu + pg^{\mu\nu} \tag{7}$$

where p is the pressure of the fluid and V^μ is the normed four-velocity of the fluid. The original form of the gravitational field equations (4) imply that

$$\ddot{a} = -\frac{4\pi\kappa}{3}(\rho + 3p)a \tag{8}$$

and

$$\dot{a}^2 = \frac{8\pi\kappa}{3}\rho a^2 - k. \tag{9}$$

The pair (8) and (9) are commonly referred to as the *Friedmann equations*. They do not appear in Einstein's (1917a), but the anachronism of using these equations can, perhaps, be excused because it makes for a much cleaner and more perspicuous discussion.[8] Conservation of energy $T^{\mu\nu}{}_{;\upsilon} = 0$, which is a consequence of (4), adds the equation

$$\dot{\rho} = -3(\rho + p)\frac{\dot{a}}{a}. \tag{10}$$

From (9) we see that a static solution ($\dot{a} = 0 = \ddot{a}$) requires that $\rho = const.$ – which is no surprise – while (8) implies that $\rho = -3p$. If $\rho > 0$, the pressure p must be negative, which we may surmise would have been unattractive to Einstein in 1917.

---

[6] Trautman states but does not derive (2′). A derivation can be given on the basis of the following assumptions. First, the metric can be written as $g_{\mu\upsilon} = \eta_{\mu\nu} + \epsilon\gamma_{\mu\upsilon}$, where $\eta_{\mu\nu}$ is the Minkowski metric and $\gamma_{\mu\upsilon}$ is assumed to be time independent. Next, in calculating the Ricci tensor it is assumed that terms of second order in $\epsilon\gamma_{\mu\upsilon}$ can be neglected. As for matter, it is assumed that the proper density ρ is low and that in the expression for the stress-energy tensor terms of order $(v/c)^2$, $\rho(v/c)$, and $\epsilon\rho$ can be neglected. And finally, it is assumed that terms of the order $\epsilon\Lambda$ are to be neglected.

[7] The sense in which (5) is the correct generalization of the original Einstein gravitational field equations will be discussed below.

[8] The Friedmann equations were derived in Friedmann (1922) and Lemaître (1927).

In any case, in 1917 he was assuming a dust model, i.e. $p = 0$, which means that a static solution requires $\rho = 0$. That is, the only static dust model allowed by the original gravitational field equations is the trivial empty one. And by (9) this trivial model cannot be closed ($k = +1$).

For the generalized field equations (5), the conservation law $T^{\mu\nu}{}_{;\upsilon} = 0$ continues to hold, and (10) remains the same. However, (8) and (9) are replaced respectively by

$$\ddot{a} = -\frac{4\pi\kappa}{3}(\rho + 3p)a + \frac{\Lambda}{3}a \tag{8$'$}$$

and

$$\dot{a}^2 = \frac{8\pi\kappa}{3}\rho a^2 - k + \frac{\Lambda}{3}a^2. \tag{9$'$}$$

As before, a static solution $a = a_E = const.$ requires $\rho = const.$ And for dust matter, (8′) requires

$$\Lambda = 4\pi\kappa\rho. \tag{11}$$

Together (11) and (9′) give

$$\Lambda = \frac{k}{a_E^2}. \tag{12}$$

Since ρ is assumed to be positive, Λ must also be positive. Thus, by (12), $k > 0$ and Einstein's static universe is spatially closed.

The upshot is that introducing the cosmological term into his gravitational field equations allowed Einstein to achieve a static cosmological model, which he believed in 1917 – on the basis of very little evidence – was needed to represent the actual cosmos. But Λ also apparently allowed him to fulfill two other closely related items on his agenda: the "relativity principle" and "Mach's principle." The fact that Einstein's static universe is spatially closed obviated the need to specify any boundary conditions at infinity, a feature which Einstein found to be attractive: "[B]oundary conditions presuppose a definite choice of the system of reference, which is contrary to the spirit of relativity" (1917a, 183). The version of Mach's principle of which Einstein was then enamored held that the distribution of matter energy should determine the metric structure of spacetime. This has a corollary – never explicitly stated by Einstein but surely understood – that the correct gravitational field equations should not admit a solution when $T^{\mu\nu}$vanishes identically. But the original field equations (4) admit empty Minkowski spacetime as a solution. The modified field equations (5) do not admit this solution when $\Lambda \neq 0$. This fact is not mentioned in Einstein's (1917a), but it is made explicit in Pauli's (1921) influential encyclopedia article on relativity theory.

In sum, the introduction of Λ appeared in 1917 to be a master stroke: a simple modification of the original gravitational field equations for GTR made room for the desired static cosmology and at the same time seemed to harmonize GTR with the sorts of overarching principles that Einstein had used to get from special to general relativity. Whatever satisfaction Einstein derived from this apparent triumph was soon to evaporate. But before continuing with the historical narrative, I will pause to consider the physical effects of a positive cosmological constant.

4. The physical meaning of Λ

A glance at Eq. (8′) shows that in expanding Friedmann models ($\dot{a} > 0$), the rate of expansion is slowed by the presence of matter and (positive) pressure, whereas the rate of expansion is speeded up by the presence of a positive cosmological constant. In this sense the factor $\frac{\Lambda a}{3}$ represents a cosmic repulsive force when $\Lambda > 0$. This conclusion is reinforced when one examines the details of solutions to the gravitational field equations (5) with cosmological term. Consider, for example, empty space solutions ($T^{\mu\nu} = 0$). In such cases the Ricci curvature scalar R is equal to 4Λ, and the field equations reduce to $R_{\mu\upsilon} = \Lambda g_{\mu\upsilon}$. When $\Lambda = 0$ the unique static spherically symmetric exterior field solution is that first given by Schwarzschild:

$$ds^2 = \frac{dr^2}{1 - \frac{2m}{r}} + d\Omega^2 - \left(1 - \frac{2m}{r}\right) dt^2 \tag{13}$$

where $d\Omega^2 = (d\theta^2 + \sin^2\theta\, d\phi^2)$ and m is the mass of the central body. The generalization of the Schwarzschild solution for $\Lambda > 0$ is

$$ds^2 = \frac{dr^2}{1 - \frac{2m}{r} - \frac{1}{3}\Lambda r^2} + d\Omega^2 - \left(1 - \frac{2m}{r} - \frac{1}{3}\Lambda r^2\right) dt^2. \tag{14}$$

In the Newtonian weak field, slow motion approximation, the Newtonian potential φ is given by $(-g_{44} - 1)/2$. In the case of (14) we get

$$\varphi = -\frac{\kappa m}{r} - \frac{\kappa}{6}\Lambda r^2. \tag{15}$$

So even when $m = 0$, a particle moving in the field (14) feels a repulsive radial force proportional to $\frac{\Lambda r}{3}$, assuming that $\Lambda > 0$. Solar system tests of GTR based on the Schwarzschild metric are affected in principle. For instance, the presence of a positive cosmological constant induces an additional perihelion shift for Mercury of $\Lambda/(5 \times 10^{-42}\,\mathrm{cm}^{-2})$ seconds of arc per century [see Rindler (1969, 208)]. If Λ is small, this effect will be undetectable. But on a cosmic scale – that is, for r large enough – the presence of a cosmological constant will manifest itself in other effects.

5. Three blows to Λ

The first blow fell almost immediately. De Sitter (1917) produced an empty space ($T^{\mu\nu} = 0$) solution to (5) with $\Lambda > 0$. The coordinate system in which De Sitter expressed the line element gave the solution a static appearance:

$$ds^2 = dr^2 + R^2 \sin^2\left(\frac{r}{R}\right)[d\psi^2 + \sin^2(\psi)d\theta^2] - \cos^2\left(\frac{r}{R}\right) dt^2 \tag{16}$$

where R is a constant (not to be confused with the Ricci curvature scalar). Initially then, the De Sitter solution seemed to satisfy Einstein's desire for a static cosmological model. The De Sitter model, however, was violently in conflict with Einstein's understanding of Mach's principle since the spacetime of the model is curved but without any matter-energy to explain the curvature.

Einstein's first reaction was to defend Mach's principle by dismissing De Sitter's solution on the grounds that it was singular and, he suspected, the singularities were hiding the matter required by his version of Mach's principle; in particular, Einstein saw a singularity at the 'equator' $r = \frac{\pi R}{2}$ where the g_{44} component of the metric of (16) is 0. Writing to De Sitter on 22 July 1917, Einstein stated: "Such a singularity in a finite world [obviously] is, in my opinion, to be discounted as physically beyond consideration."[9] This criticism was repeated in print [see Einstein (1918)]. The linchpin of Einstein's original defense crumbled in the light of Felix Klein's demonstration that the singularity Einstein perceived at the equator is only a coordinate artifact. Einstein conceded in a letter to Klein dated 20 June 1918: "You are entirely right. De Sitter's world is, of itself, free of singularities . . . My critical remark about De Sitter's solution needs correction; a singularity-free solution for the gravitation equations without matter does in fact exist."[10] This concession never made it into print, and Pauli's (1921) widely cited encyclopedia article repeats Einstein's singularity charge.

Klein's technique for showing the ersatz character of the perceived singularity in the De Sitter solution was to show how to smoothly extend the De Sitter metric through the apparent singularity. What this extension revealed, however, was that De Sitter's solution is not static, the original appearance of staticity being due to the fact that the De Sitter coordinates used in (16) cover only a piece of the fully extended spacetime. Einstein seized on this fact as an alternative reason for dismissing the De Sitter solution. In the same letter where he conceded to Klein, he wrote: "Under no circumstances could this [De Sitter-Klein] world come into consideration as a physical possibility. For in this world time t cannot be defined in such a way that the three dimensional sections $t = const$ do not intersect and are equal to one another (metrically)."[11] When Einstein later decided that static cosmological models are not feasible, this second line of defense crumbled. From this time onward, his allegiance to Mach's principle gradually diminished to the point where he said that we should no longer speak of this principle in GTR.[12]

The second blow came from Eddington (1930), who noted that Einstein's static solution is unstable. With typical modesty Eddington credited the insight to Lemaître (1927, 1931a). "Although not expressly stated, it is at once apparent from his formulae that Einstein's world is unstable – an important fact which, I think, has not hitherto been appreciated in cosmological discussions" (1930, 668). Setting $p = 0$ in Eq. (8′) gives

$$3\ddot{a} = a(\Lambda - 4\pi\kappa\rho). \tag{17}$$

Thus, for a static dust solution the value of Λ must exactly cancel the contribution from matter. The slightest perturbation that would make ρ less than $\frac{\Lambda}{4\pi\kappa}$ by any amount whatsoever would start the universe expanding, and by (10) the expansion decreases ρ, making $\ddot{a}$ increase. Similarly, the slightest perturbation that makes ρ greater than $\frac{\Lambda}{4\pi\kappa}$

9 Schulmann et al. (1998, Doc. 363).
10 Schulmann et al. (1998, Doc. 567).
11 Ibid.
12 Quoted in Pais (1982, 288). More details on the Einstein-De Sitter and the Einstein-Klein interchanges can be found in Earman and Eisenstaedt (1999).

by any amount whatsoever will start a contraction that continually increases.[13] There is no evidence that Einstein read Eddington's (1930), but he had read Friedmann's (1922), which showed that the Einstein field equations, with or without cosmological term, admit non-static solutions with a homogeneous and isotropic matter distribution. (Initially Einstein thought that Friedmann had made an error in his calculations [see Einstein (1922)], but later he decided that Friedmann was correct [see Einstein (1923)]). And in 1931 Einstein noted that the instability of his static solution follows from the Friedmann equations [see Einstein (1931]).

But for Einstein the killing blow to Λ was Hubble's (1929) redshift observations which indicated that the universe is not static but expanding. Already in 1923 Einstein was prepared to jettison the cosmological term. As he wrote to Weyl: "If there is no quasi-static world, then away with the cosmological term."[14] As shown by the work of Friedmann (1922) – and by Lemaître who is not cited in Einstein's paper – Hubble's observations can be accounted for by the GTR in an "unforced manner," namely without the Λ-term (1931, 237). So away with Λ.

In hindsight we can see that some evidence for an expanding universe existed already in the early 1920s. In the first edition of *The Mathematical Theory of Relativity* (1923) Eddington opined that "One of the most perplexing problems of cosmogony is the great speed of the spiral nebulae. Their radial velocities average about 600 km. per second and there is a great preponderance of velocities of recession from the solar system" (161). Eddington goes on to report measurements of radial velocities of spiral nebulae made by V. M. Silpher of the Lowell Observatory, noting that "the great preponderance of positive (receding) velocities is very striking" (162). But Eddington was cautious in drawing conclusions both because of the "lack of observations of southern nebulae" and because two of Silpher's data points were from nebulae that are approaching the solar system at high speeds. And in any case, the expanding universe models of Friedmann (1922) were apparently unknown to Eddington in 1923. He therefore tried to explain Silpher's data on spiral nebulae by means of De Sitter's model.

6. Einstein's denunciation of Λ

Physicists in Einstein's orbit tended to follow his lead in rejecting the cosmological term. A silent but eloquent affirmation of Einstein's attitude is found in Peter Bergmann's *Introduction to the Theory of Relativity* (1942), which was read and approved by Einstein: the book contains not one mention of Λ. In 1958 Pauli added a footnote to his *Theory of Relativity*: "Einstein . . . *completely rejected the cosmological term* as superfluous and no longer justified. I fully accept this new standpoint of Einstein's" (1958, 220). We have seen that from the beginning Einstein was uneasy about Λ because it

[13] Since (17) applies only to a homogeneous universe, the instability in question arises only if God created a homogeneous universe with a bit too much or a bit too little matter. The real worry (as John Norton kindly reminded me) concerns the *local* instability of the matter distribution, and Eddington's analysis did not resolve this issue.

[14] Quoted in Pais (1982, 286).

was "gravely detrimental to the formal beauty of the theory [GTR]" (1919, 193). After Hubble's observations indicating that the universe is expanding he seems to have felt that Λ was tainted with the original sin of a faulty motivation. Both of these feelings are expressed in a letter of 26 September 1947 to Lemaître:

> Since I have introduced this [Λ] term, I had always a bad conscience. But at that time I could see no other possibility to deal with the fact of the existence of a finite mean density of matter. I found it very ugly indeed that the field law of gravitation should be composed of two logically independent terms which are connected by addition. About the justification of such feelings concerning logical simplicity it is difficult to argue. I cannot help to feel it strongly and I am unable to believe that such an ugly thing should be realized in nature.[15]

It is indeed difficult to argue about simplicity and beauty.[16] But the question of whether or not Einstein should have had a bad conscience for introducing Λ is irrelevant to the issue of whether Λ has a natural place in GTR. And it indisputably does. On the right hand side of the original gravitational field equations stands the stress-energy tensor $T^{\mu\nu}$. Thus, in order to have conservation of energy-momentum in the form $T^{\mu\nu}{}_{;\upsilon} = 0$ it is necessary that the left hand side of the equation have vanishing covariant divergence. So one can look for the most general symmetric second rank tensor $X^{\mu\nu}$, concocted from $g_{\mu\upsilon}$ and its derivatives, such that $X^{\mu\nu}{}_{;\upsilon} = 0$. If, as is suggested by the Newtonian limit, $X^{\mu\nu}$ is a functional of $g_{\mu\upsilon}$ and its first and second derivatives, and if it is linear in the second derivatives, then $X_{\mu\upsilon}$ must be equal to $R_{\mu\upsilon} - \frac{1}{2}g_{\mu\upsilon}R + \Lambda g_{\mu\upsilon}$. It is then an empirical question as to whether or not Λ is zero.

This point of view was put to Einstein by R. C. Tolman in a letter dated 14 September 1931:

> With regard to the question of setting Λ equal to zero, I think there are a number of arguments in favor of it but one fairly strong against it. On one hand, by giving Λ the definite value zero, the fundamental equations are simplified, the conclusions drawn from them are rendered less indeterminate, and it becomes no longer necessary to inquire into the significance and magnitude of what would otherwise be a new constant of nature. On the other hand, since the introduction of the Λ-term provides the most general possible expression of the second order which would have the right properties for the energy-momentum tensor, a definite assignment of $\Lambda = 0$, in the absence of experimental determination of its magnitude, seems arbitrary and not necessarily correct.[17]

---

[15] Einstein to Lemaître 26 September 1947 (Einstein Archives Doc. 15 085). Quoted by permission of the Albert Einstein Archives, the Jewish National University Library, the Hebrew University of Jerusalem, Israel.

[16] From 1915 on Einstein relied more on considerations mathematical simplicity and beauty and less on physical heuristics – with diminishing returns! See Norton (2000).

[17] Tolman to Einstein 14 September 1931 (Einstein Archives Doc. 23 031). Quoted by permission of the California Institute of Technology.

Similar sentiments are found in Lemaître's contribution to the Schilpp volume in Einstein's honor:

> Even if the introduction of the cosmological constant "has lost its sole original justification, that of leading to a natural solution to the cosmological problem,"[18] it remains true that Einstein has shown that the structure of his [gravitational field] equations quite naturally allows for the presence of a second constant besides the gravitational one. This raises a problem and opens possibilities which deserve careful discussion. The history of science provides many instances of discoveries which have been made for reasons which are no longer considered satisfactory. It may be that the discovery of the cosmological constant is such a case. (1949, 443)

Only for someone of genius (which Einstein indisputably was) or someone possessed of a monstrous ego (which Einstein wasn't) can sufficient grounds for rejecting a quantity be found in the fact that *he* introduced the quantity for reasons *he* later regretted. In any case, despite Einstein's denunciation of Λ, it refused to disappear from cosmology. It was kept alive by a number of champions, some reluctant, some enthusiastic. Before turning to these champions, it is worth noting a position that is intermediate between that of Einstein on one hand and Tolman and Lemaître on the other.

After reviewing various positions on the status of the cosmological constant, W. H. McCrea (1971) reached the following disjunctive conclusion:

> (a) If general relativity is to be treated as a self-contained theory, then the cosmical terms should be omitted; this is because the theory itself cannot assign a meaning to a fixed dimensional physical constant, while if it attempts to admit an arbitrary constant it becomes too indeterminate.
> (b) If general relativity is to be treated as only part of what is required in order to construct a theoretical model of the world of physical experience, then the cosmical terms should be retained as providing additional freedom on connecting relativity theory with other parts of physical theory. The value to be assigned to the cosmical constant may be expected to depend upon the part of the physical world being studied; this is because the theoretical model can be a model only of some idealized representation of the actual world and, in general different idealizations will have to be made in different applications. (152–153)

Attempts have been made to seize McCrea's option (b) by using Λ as a means of connecting GTR with particle physics. These attempts will be reviewed below in Sect. 10.

---

[18] This is a reference to Einstein's *The Meaning of Relativity* (1946, fn 121). The full passage reads:

> If Hubble's expansion had been discovered at the time of creation of the general theory of relativity, the cosmologic member would never have been introduced. It seems now so much less justified to introduce such a member into the field equations, since its introduction loses its sole original justification, – that of leading to a natural solution to the cosmologic problem [of allowing for a finite average density of matter].

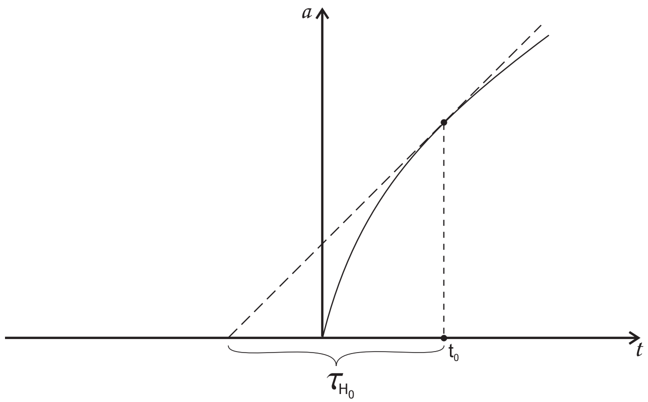

Fig. 1. A big bang model with $\Lambda = 0$

7. Keeping Λ alive

In addition to the point of view that theoretical considerations do not exclude a cosmological term from the gravitational field equations and that only observations can decide whether Λ is zero, there were reasons of a more positive sort for keeping Λ alive.

The age problem. With $\rho > 0$ and $p \geq 0$, it follows that for Friedmann cosmological models with $\Lambda = 0$, $\ddot{a} < 0$ [see Eqs. (8) and (8′)]. Redshift measurements tell us that $\dot{a}_o > 0$ (where the naught subscript on a quantity indicates its present value). Thus the graph of the scale factor $a(t)$ is concave downwards with $a(t) = 0$ at some finite time in the past, which will be designated as $t = 0$. It follows that the present age t_o of the universe is less than the Hubble time $\tau_{H_o} := H_o^{-1}$, where $H := (\dot{a}/a)$ denotes the (unconstant!) Hubble constant (see Fig. 1). Hubble's early measurements of H_o gave a value around 465 km per sec per megaparsec, which translates into a value for τ_{H_o} of two billion years. This is too short compared either to the age of the earth, as determined by radioactive decay, or to the age of stars, as determined by theories of stellar evolution. This clash is stated clearly in the second edition of Spencer Jones' *General Astronomy* (1934)[19]:

> We thus have the apparent paradox that the stars seem to be much older than the Universe. The estimate of stellar ages at which we have arrived cannot be abandoned without leaving the facts upon which it was based unexplained. On the other hand, if the Universe has been in a state of progressive expansion following upon an initial disturbance when it was much more compact than at present, the longer time scale is impossible. How the two different time scales are to be reconciled is one of the outstanding problems of astronomy at present. (414)

[19] Spencer Jones was at the time the Astronomer Royal.

This paradox was met in some quarters with a combination of obfuscation, wishful thinking, and plausible but wrong conjectures. As an apparent example of the first, I would cite De Sitter's *Kosmos* (1932):

> I am afraid that all we can do is to accept the [age] paradox and try to accommodate ourselves to it, as we have done to many paradoxes lately in modern physical theories. We shall have to get accustomed to the idea of change of the quantity R [a] commonly called the "radius of the universe," and the evolutionary changes of stars and stellar systems are two different processes, going on side by side without any apparent connection between them. After all, the "universe" is an hypothesis, like the atom, and must be allowed freedom to have properties and to do things which would be contradictory and impossible for a finite material structure. (133)

But contrary to the impression given in the semi-popular *Kosmos*, De Sitter did have a coherent proposal that fits somewhere between the categories of wishful thinking and plausible-but-false-conjecture. The idea was that when the idealization of a homogeneous and isotropic matter distribution is replaced by a more realistic assumption, the initial singularity $a(0) = 0$ of the Friedmann expanding models will be replaced by a "near approach of the galaxies during a short interval of time" near $t = 0$, from which the stars would emerge intact. De Sitter was quite confident that this idea would work:

> We must therefore accept the paradox that the stars are older than the universe, if by "the age of the universe" we mean the time elapsed since y [the scale factor a] passed through its minimum. It has been shown, however, that this minimum must not be conceived as "the beginning of the world," but as a transitory episode in the history of the universe, so that there is nothing paradoxical left of the paradox. (1933, 632)

Tolman (1934) was also hopeful that the age paradox is an artifact of the idealization of a homogeneous universe:

> [I]n connexion with this apparent difficulty as to time scales, it is to be emphasized that the highly idealized homogeneous models which we have employed can hardly be regarded as adequate for drawing any exact conclusions as to the precise state of the actual universe say 10^9 years ago. Thus, as already mentioned earlier . . . it is evident that the unique singular state at the lower limit of volume from which the expansion would appear to start in the case of certain models must be regarded as property of the homogeneous model rather than a character that could actually be found in the real universe. (1934, 485)

Here Tolman was certainly influenced by Einstein's (1931), where a similar sentiment is found, and by his correspondence with Einstein.[20] Over thirty years would elapse before a series of theorems due mainly to Roger Penrose and Stephen Hawking made it clear

[20] On 14 September 1931 Tolman wrote to Einstein:

> I was very glad to find your letter of June 27th waiting for me . . . and also to receive the reprints on the cosmological problem [Einstein (1931)] and the unified field theory.
>
> When I first saw your proposed quasi-periodic solution for the cosmological line element, I was very much troubled by the difficulties connected with the behavior of the model in the neighborhood of the points of zero proper volume. The remarks of your letter, however, pointing out that the actual inhomogeneity in the distribution of

that singularities in cosmology and gravitational collapse are not artifacts of idealized models but are generic to the solutions to Einstein's gravitational field equations.[21] In many instances there is no physically reasonable way to extend the solution through the singularities; in particular, there seems to be no physical meaning in talking about what happens "before" a big bang singularity or "after" a big crunch singularity.

Another approach to resolving the age problem involved an appeal to Λ. For a Friedmann dust universe ($p = 0$) the conservation law (10) implies that $\rho a^3 = const$. Thus for a spatially closed universe ($k = +1$), the Friedmann Eq. ($9'$) can be rewritten as

$$\dot{a}^2 = \frac{\Lambda}{3}a^2 + \frac{C}{a} - 1 \tag{18}$$

where C is a constant. Starting with (18) and setting $\Lambda = \Lambda_E := 9/4C^2$ (where the subscript E is used to denote the value of Λ in the Einstein static universe), there are three possibilities, depending on the value of the constant of integration: (i) the Einstein static universe where $a_E = \frac{2}{3}C$, (ii) A big bang model which expands from $a(0) = 0$ and asymptotically approaches the Einstein static model, and (iii) the Eddington-Lemaître model which in the future direction of time is forever expanding and which in the past direction asymptotically approaches the Einstein static model as $t \to -\infty$ (see Fig. 2). The last model "solves" the age problem all too well!

The case $\Lambda = \Lambda_E(1 + \delta)$ with $\delta > 0$ corresponds to what are called the Lemaître models, although Lemaître (1927, 1931a) himself allowed for a positive pressure. These are big bang models which expand forever but which contain a stage where the expansion rate slows down, reaching a minimum value when $a_m = (3C/2\Lambda)^{1/3}$. Not only does the expansion rate slow down, but by taking δ sufficiently small, $a(t)$ can be made to linger, for as long as one desires, close to the value at which $\dot{a}(t)$ is a minimum. Assuming that the present time t_o post-dates this "coasting phase," the longer the coasting phase, the more the age of the universe exceeds the Hubble time (see Fig. 3). Here then is a ready made solution to the age problem which does not get rid of the initial singularity $a(0) = 0$ – a feature that became highly desirable once the evidence for a big bang began to accumulate.

In the glare of hindsight wisdom it may seem strange that in the 1930s cosmologists were willing to perform theoretical contortions in order to solve the age problem but were apparently unwilling to challenge the basis of the problem: Hubble's value for H_o. But it must be remembered that at the time Hubble had the only telescope capable of fixing H_o [see North (1965, 228–229)]. Over the years estimates of H_o have dropped by a factor of 5 to 10, and the corresponding rise in the Hubble time has removed some of

---

matter might make the idealized treatment fail in that neighborhood seem to me very important . . . (Einstein Archives Doc. 23 031)

Quoted by permission of the California Institute of Technology.

[21] Introducing a positive (respectively, negative) cosmological constant tends to decrease (increase) the prevalency of singularities among the solutions of the field equations. For an account of the history of the singularity theorems in GTR, see Earman (1999).

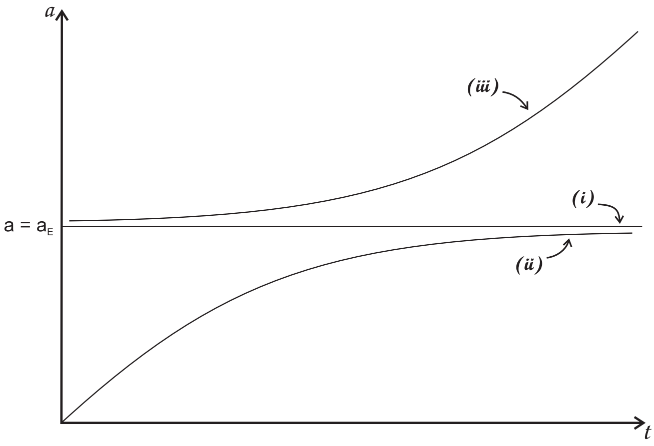

Fig. 2. $\Lambda = \Lambda_E$ models

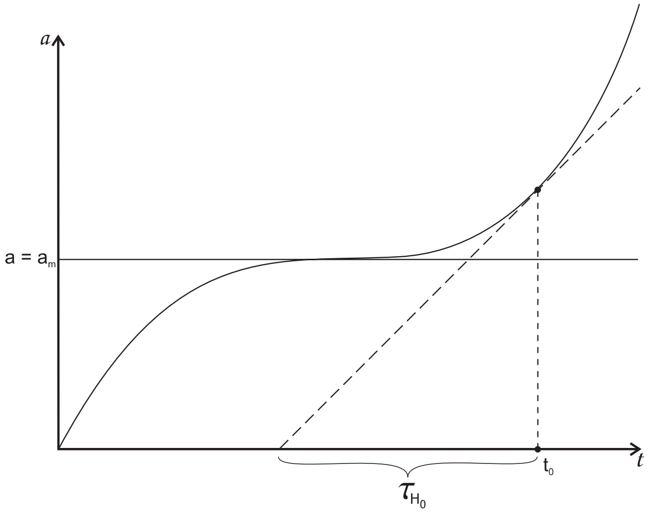

Fig. 3. A Lemaître model

urgency from the age problem. But as we will now see, the age problem wasn't the only way to motivate the Lemaître models and a positive Λ.

Structure formation. Lemaître championed Λ not only because, as embodied in the Lemaître models, it offered a way out of the age problem but also because it promised to help explain the formation of galaxies and nebulae. He assumed that prior to the coasting phase density perturbations – whose source he did not attempt to explain – left matter arranged in a kind of super-gas, the individual molecules of which form clouds of matter. During the coasting phase the near balance of the attractive force of gravity and the repulsive force due to Λ amplifies density perturbations, causing them to grow into proto-galaxies and galactic clusters. At the end of the coasting phase these galactic condensates remain condensed while the universe resumes rapid expansion, producing the presently observed recession velocities detected by Hubble.

Lemaître's own treatment of structure formation [see Lemaître (1931b, 1933a, b, 1934)] was rather qualitative and subject to various objections [see McCrea and McVittie (1931)]. But the general idea of tying structure formation to the coasting phase of the Lemaître model remained of interest through the 1950's [see Bonnor (1957)] and the 1960's [see Rawson-Harris (1969), and Brecher and Silk (1969)]. The idea seems to have been killed by Brecher and Silk (1969) who argued that only under artificial constraints can the Lemaître models escape the following dilemma: either statistical fluctuations will not produce galaxy formation, or else the formation will result in the collapse back to a singularity.

Eddington's views on Λ. Eddington rivaled Lemaître as a champion of Λ, but the motivations of these two men were quite different. In *The Expanding Universe* Eddington averred that "I would as soon think of reverting to Newtonian theory as dropping the cosmological constant" (1933, 35). Eddington had at least two distinct reasons for thinking that Λ was essential to cosmological models and to spacetime theories in general. The first is based on the idea that the effective repulsive force of Λ is needed to explain the recession of the galaxies. To be sure, GTR allows for an expanding universe even when $\Lambda = 0$. But these models need to assume that large velocities of recession are built into the model *ab initio*. "This might be true," Eddington retorted, "but it can scarcely be called an *explanation* of the large velocities" (1933, 37). The notion that a scientific explanation which has to posit special initial conditions to produce the desired effect is not an acceptable explanation later helped to fuel to enthusiasm for inflationary cosmology, which eschews special initial conditions in favor of an effective cosmological constant in the very early universe (see Sect. 11). About the same time De Sitter, in a work entitled "The Expanding Universe," also endorsed the necessity of Λ for explaining expansion:

> What is it then that causes the expansion [of the universe]? Who blows up the india-rubber ball? The only possible answer is: the *lambda* does it. It is the presence of *lambda*, the "cosmological constant" of Einstein, in the equations that not only closes up the universe, making it of finite size instead of infinite, but also provides the possibility of its changing size. (1931, 9)

De Sitter knew perfectly well that GTR allowed for this possibility without the help of Λ. Perhaps he held some view of explanation similar to Eddington's.

Eddington's second reason for thinking that a positive Λ is essential to GTR derived from his idea that length is relative to a standard of comparison. But "whatever embodies this comparison unit is *ipso facto* the space of physics": "Physical space therefore cannot be featureless. As a matter of geometrical terminology features of space are described as curvatures (including hypercurvatures) . . . We have therefore no option but to look for the natural standard of length among the radii of curvature or hypercurvature of space-time" (1933, 147). It follows, on Eddington's reasoning, that Λ is indispensable:

> When once it is admitted that there exists everywhere a radius of curvature ready to serve as comparison standard, and that spatial distances are directly or indirectly expressed in terms of this standard, the law of gravitation [for empty space] ($G_{\mu\nu} = \Lambda g_{\mu\nu}$ [$-\Lambda g_{\mu\nu}$ in our sign convention]) follows without further assumption; and accordingly the existence of the cosmical constant Λ with the corresponding force of cosmical repulsion is estab-

> lished. Being in this way based on a fundamental necessity of physical space, the position of the cosmical constant seems to me impregnable; and if ever the theory of relativity falls into disrepute the cosmical constant will be the last stronghold to collapse. *To drop the cosmical constant would knock the bottom out of space.* (1933, 147–148)

In the empty space case ($T_{\mu\nu} = 0$) it follows from Einstein's field equations that the Ricci curvature R is equal to 4Λ. When matter is present, however, $R = 4\Lambda - 8\pi\kappa Tr(T_{\mu\nu})$, so that with a non-homogeneous matter distribution the radius of curvature is not everywhere of the same length. In that case, as De Sitter noted, the radius of curvature "can no longer provide the electron with the means of knowing how large it ought to be" (1932, 129). In *The Nature of the Physical World* Eddington makes it clear that his doctrine applies only to empty space. When space is not empty "the meter rod can find other lengths besides curvature to measure itself against" (1929, 153). But since space is *in fact* not empty, it hardly follows that dropping Λ would knock the bottom out of space (or spacetime), even if we accept Eddington's debatable doctrine that length is relative to a standard of comparison found in the features of spacetime.

8. Steady state cosmology and the ghost of Λ

From its inception in the late 1940s until the mid-1960s the steady state cosmology of Bondi and Gold (1948) and Hoyle (1948) represented a viable alternative to big bang cosmology.[22] The ghost of Λ haunts the steady state model because of its reliance on De Sitter spacetime. Considered as solution to Einstein's gravitational field equations, De Sitter spacetime requires either a positive Λ in an empty cosmos, or else a zero Λ in a cosmos filled with a nonstandard fluid whose density is proportional to the De Sitter value for Λ and whose pressure is proportional to the negative of this value. De Sitter chose the former alternative. The steady state model uses a representation of De Sitter spacetime discovered independently by Lemaître (1925) and Robertson (1928). In the Lemaître-Robertson coordinates the De Sitter line element takes the form

$$ds^2 = \exp(K\tilde{t})(dx^2 + dy^2 + dz^2) - d\tilde{t}^2 \tag{19}$$

where the constant K is proportional to the square root of the De Sitter value for Λ. What neither Lemaître nor Robertson – nor apparently the founders of steady state cosmology – realized was that these coordinates cover only half of the (maximally extended) De Sitter spacetime. As a result, the steady state model is geodesically incomplete in the past and, thus, it also implies a "beginning" for time, albeit of a very different kind from that of big bang models.

According to steady state cosmology, matter follows the geodesics orthogonal to the $\tilde{t} = const.$ hypersurfaces of (19). Because of the exponential expansion of the scale factor $a(\tilde{t}) = \exp(K\tilde{t})$, it would seem that the density of matter should decrease, belying the label "steady state." To compensate for the expansion, steady state theorists postulate that matter is continuously created at just the right rate to maintain constant density.

---

[22] A detailed account of the rise and fall of steady state cosmology is to be found in Kragh (1996).

Bondi and Gold (1948) did not attempt to provide field equations for their model, but Hoyle (1949) construed the steady state model as a solution to field equations which are obtained by replacing the cosmological term in the standard Einstein equations by a tensor $C_{\mu\upsilon}$ whose presence is due to the 'C'-field responsible for continuous creation of matter:

$$R_{\mu\upsilon} - \frac{1}{2}g_{\mu\upsilon}R + C_{\mu\upsilon} = 8\pi\kappa T_{\mu\nu}. \tag{20}$$

These equations allow solutions other than the steady state model, but (in modern jargon) Hoyle's claim was that the steady state model serves as an attractor for solutions to (20).

In the form (20), $C_{\mu\upsilon}$ can be seen as a replacement for Einstein's cosmological term which nevertheless produces some of the same effects as a positive Λ – in particular, exponential expansion of the universe. This aspect was somewhat disguised when McCrea (1951) proposed to move $C_{\mu\upsilon}$ to the right hand side of (20), in effect, redefining the stress-energy tensor (see also Pirani (1955) and Hoyle and Narliker (1964)). In this role $C_{\mu\upsilon}$ is the stress-energy tensor of a strange form of matter: strange in that its covariant divergence does not vanish, and strange in that it arises from a scalar field whose energy density is negative. According to this way of viewing steady state cosmology, it is this negative energy density, and not Λ, that produces a repulsive force that drives an accelerating expansion of the universe.

9. A brief revival of Λ

By the end of the 1960s it was generally agreed that Lemaître's model did not provide for a satisfactory explanation of structure formation, and with the downward trend in the estimates of the Hubble constant the problem of the age of the universe became less problematic and, thus, less of a motivation for taking the Lemaître models seriously. However, the late 1960s and early 1970s saw the emergence of yet another reason for breathing new life into these models. The redshift parameter z is defined by

$$z := \frac{\lambda_o - \lambda_1}{\lambda_1} \tag{21}$$

where λ_o is the wavelength of light observed at the present time t_o and λ_1 is the wavelength at the time t_1 of emission. In Friedmann and in Lemaître models,

$$z = \frac{a(t_o)}{a(t_1)} - 1. \tag{22}$$

The Lemaître models were invoked by Petrosian, Salpeter, and Szekeres (1967) and by Kardashev (1968) to explain the preponderance of redshifts for quasi-stellar objects (QSOs) near $z = 2$. Since during the coasting phase the scale factor remains near its value a_m for which $\dot{a}$ is a minimum, the Lemaître models would capture the redshift observations if the preponderance of QSOs were present when $a_m \sim a_o/3$. Only a few years elapsed, however, before Petrosian (1974) was ready to abandon his proposal. In particular, Petrosian and Salpeter (1970) had noted that when fitted to give a preponderance of $z \sim 2$ redshifts, the Lemaître models predict a rapid fall off of the visual intensity of QSOs with $z > 2.2$. But Petrosian (1974) reported four redshift measurements near 2.8 and 3.5, which “must be considered strong evidence against the Lemaître models unless, because of the tendency nowdays of publishing only sources with larger redshifts, these humps [at 2.8 and 3.5] are not as significant as they appear” (44).

Petrosian (1974) opined that although a range of Lemaître models had been killed by observations, the evidence as it then existed did not rule out all Lemaître models, nor did it force a zero value for Λ. But, he concluded, “in the absence of strong evidence in favor of Lemaître models we must once again send back the Lemaître models and along with it [sic] the cosmological constant to the shelf until their next re-appearance” (45). Petrosian may have spoken for cosmologists, but he certainly didn’t speak for particle theorists. Their collective imagination had been fired by his proposal, and once fired it focused not on cosmological observations but on how to interpret Λ as vacuum energy density.

10. Energy density of the vacuum vs. vacuum energy density

In the gravitational field equations (5), move the cosmological term $\Lambda g_{\mu\upsilon}$ to the right hand side and write $T^{tot}_{\mu\nu} = T_{\mu\nu} + T^{\Lambda}_{\mu\upsilon}$, where $T^{\Lambda}_{\mu\upsilon} := -\frac{\Lambda}{8\pi\kappa} g_{\mu\upsilon}$, making the cosmological term correspond to the stress-energy tensor of a perfect fluid of a peculiar sort: the energy density is $\rho_\Lambda = \frac{\Lambda}{8\pi\kappa}$ and the pressure is $p_\Lambda = -\rho_\Lambda$.[23] In this sense, the presence of a non-zero cosmological constant corresponds to saying that the energy-density of the vacuum ($T_{\mu\nu} = 0$) is non-zero, and if $\Lambda > 0$ (and, thus, $\rho_\Lambda > 0$) the vacuum also has associated with it a negative pressure. Note that the conservation law (10) implies that $\dot{\rho}_\Lambda = 0$, as required for the consistency of this way of speaking.

The first reference in print to Λ as representing the energy density of the vacuum is in Lemaître (1934). Perhaps confused by sign conventions, Lemaître took $\Lambda > 0$ to correspond to a negative energy density![24] Sign conventions also seem to have led Eddington (1939) to speak of Λ as “fixing the zero from which energy, momentum and stress are reckoned” (233). On his conventions, the “reckoned” energy tensor $T_{\mu\nu}$ is the difference, $G_{\mu\nu} - \Lambda g_{\mu\upsilon}$, between the actual “absolute” energy $G_{\mu\nu}$ and “absolute” energy $\Lambda g_{\mu\upsilon}$ of a standard zero condition.

Now enter from stage right the particle physicists, whose footprints on the cosmological literature are increasingly evident from the 1960’s onward. The tracks start in earnest when Zeldovich (1967, 1968) read the proposals to resurrect Λ in order to explain the $z \sim 2$ redshifts of QSOs. Though somewhat skeptical of the proposals, Zeldovich nevertheless found them stimulating because they raised for him an issue that connected cosmology to particle physics; namely, if Λ corresponds to an energy density of the

---

[23] Such a fluid violates the strong energy condition which requires that $T_{\mu\nu} V^\mu V^\nu \geq -\frac{1}{2} Tr(T_{\mu\nu})$ for every unit timelike vector V^μ. For a perfect fluid this condition requires that $\rho + 3p \geq 0$. Thus, the strong energy condition is violated by an equation of state $p = w\rho$ if $w < -\frac{1}{3}$. This is the case for “quintessence” which has been postulated to drive accelerated expansion of the universe (see Sect. 12).

[24] So many different sign conventions for the spacetime metric and the stress-energy tensor are used in the cosmological literature that the casual reader is easily confused.

vacuum, should it not be possible to calculate its value from quantum field theoretic principles as the vacuum energy density? In contrast to the classical vacuum, the vacuum in quantum field theory is an arena of seething activity, so perhaps properties of this vacuum can be used to do some real explanatory work that will have ramifications that extend from the subatomic realm to the furthermost reaches of the universe. This was the beginning of a minor industry which consists of trying to estimate the value of the vacuum energy density, finding that it is many orders of magnitude (up to 120 orders!) too large to accord with observational constraints on Λ, trying a series of increasingly desperate moves to lower the estimate, and resorting if all else fails to the anthropic principle.[25] Steven Weinberg (1989), who believes that physics thrives on crises, has been instrumental in promoting this problem to the status of a crisis for contemporary physics. I want to explain why this 'crisis' needs to be viewed with some skepticism

The problem is supposed to arise as follows. The Poincaré invariance of the Minkowski vacuum state $|0_M\rangle$ implies that the vacuum expectation value $\bar{T}_{\mu\nu} := \langle 0_M|T_{\mu\nu}|0_M\rangle$ of the stress energy tensor is Poincaré invariant. But the only (non-zero) symmetric second rank tensor $X_{\mu\nu}$ that is Poincaré invariant (in the sense that $X'_{\mu\nu} = X_{\mu\nu}$ in any two inertial frames) is proportional to the Minkowski metric, showing that $\bar{T}_{\mu\nu}$ must be of the form of a cosmological term. This reasoning is of dubious applicability in cosmology where Poincaré invariance is lost in curved spacetime models. The loss is total in some cases – in particular, the $k = +1$ Friedmann models where Lorentz invariance cannot hold even in an asymptotic sense. When taxed with this observation, Weinberg responded that although a *model* of GTR may not be Poincaré invariant, the *theory* itself is in the sense that the Poincaré group is retained as the symmetry group in local inertial frames.[26] Whatever Weinberg might mean by the claim that GTR is Poincaré invariant is irrelevant to the issue at hand. The loss of Poincaré invariance of the spacetime model means that a vacuum state cannot be picked out as the state that (among other things) is Poincaré invariant.[27] Indeed, in the cosmological setting not only cannot it not be expected that the spacetime model will admit the Poincaré group as a symmetry group, it cannot even be expected that the spacetime will admit a timelike Killing field.[28] This has the consequence that there is no natural way to identity a vacuum state for such models [see Wald (1994)]; and needless to say, an accurate portrayal of the actual universe, whose spacetime is not stationary, will belong to this class.

---

[25] The idea behind an anthropic explanation of the smallness of Λ is that such a value is forced by the requirement that the universe be such that it allows stars and observers (such as can ask why Λ has such a small value) to form. For examples of such reasoning as applied to the cosmological constant problem, see Efstathiou (1995) and Martel et al. (1998). For a skeptical look at anthropic explanations, see McMullin (1993).

[26] The issue was raised by Carlo Rovelli. See Cao (1999, 262–263).

[27] Positivity of energy is also needed to single out the Minkowski vacuum state.

[28] Intuitively the existence of such a field means that the metric is "time independent." Formally, the definition is this: a Killing field for a spacetime M, $g_{\mu\nu}$ is a vector field V^β such that $\nabla_\mu V_\nu + \nabla_\nu V_\mu = 0$, where ∇_μ is the derivative operator determined by $g_{\mu\nu}$. If there exists a timelike Killing field, then (at least locally) a coordinate system (x^i, t), $i = 1, 2, 3$, can be chosen so that the metric components are functions of the spatial coordinates x^i alone.

Whether or not the quantum state of the cosmos is a vacuum state, there remains the problem of renormalization. The quantization of a scalar field in Minkowski spacetime can be understood heuristically in terms of an assemblage of harmonic oscillators. Formally, the Hamiltonian operator has the form

$$\hat{H} = \sum_k \omega_k (\hat{a}_k^\dagger \hat{a}_k + \hat{a}_k \hat{a}_k^\dagger) \tag{23}$$

where ω_k is the frequency of wave number k and $\hat{a}_k^\dagger$ and $\hat{a}_k$ are respectively the associated creation and annihilation operators. The expectation value of $\hat{H}$ in the Minkowski vacuum state is $\frac{1}{2}\sum_k \omega_k$, which sums to ∞ if there is no high frequency cut off. This infinity can be removed by "normal ordering," i.e. rewriting $\hat{H}$ so that $\hat{a}_k$ always stands to the left of $\hat{a}_k^\dagger$. This procedure is said to be permissible in quantum field theory because "absolute energies are not measured observables; only energy differences have physical meaning" [Bjorken and Drell (1965, 30)]. However, in gravitational physics absolute energies apparently do matter since they can give rise to gravitational fields. Thus arises the worry of quantum field theorists about a vacuum energy density contribution to the gravitational field. Before addressing this worry, I need to comment on the renormalization problem in the spacetime setting.

Normal ordering corresponds in Minkowski spacetime to a point splitting procedure, i.e. taking a coincidence limit of a bi-distribution, wherein the Minkowski vacuum $|0_M\rangle$ expectation value of the bi-distribution has been subtracted off before the limit is taken. For example, calculation of the expectation value of the field operator $\hat{\phi}^2(x)$ for a state $|\xi\rangle$ involves taking the limit $x' \to x$ of $[\langle\xi|\hat{\phi}(x')\hat{\phi}(x)|\xi\rangle - \langle 0_M|\hat{\phi}(x')\hat{\phi}(x)|0_M\rangle]$. But in curved spacetime this procedure won't yield a sensible result if one tries to subtract off the bi-distribution expectation value for a fixed state; rather (as discussed in Wald (1994)), a locally constructed bi-distribution that depends on the details of the spacetime has to be subtracted off. This form of renormalization for curved spacetimes works for a large set of states with the Hadamard property which characterizes the leading singularity structure of the two-point functions. But to underscore the obvious, in this – the only existing rigorous version of renormalization for a quantum field on a curved spacetime – there is no sense in which the Minkowski vacuum state, or any fixed state, serves as a reference point for energy calculations. And there is no sense in which the renormalization procedure associates a finite non-zero energy with the Minkowski vacuum, or with any other vacuum state on a curved spacetime.

The vacuum energy enthusiasts would respond that, formal procedures notwithstanding, the vacuum energy density is real in the sense that it has measurable non-gravitational effects, and so the GTR would lead one to expect that this energy density will also produce measurable gravitational effects as well. The example cited over and over again to buttress this point of view is the Casimir effect, wherein two perfectly conducting parallel plates are attracted to one another with a force proportional to $1/d^4$, where d is the separation distance, when they are inserted into the Minkowski vacuum.[29] The usual explanation is that the insertion of the conducting plates lowers the energy density of the

[29] The $1/d^4$ dependence can be derived from considerations of the classical van der Walls force. But the numerical factor requires a quantum field theoretic explanation.

vacuum between the plates because those electromagnetic field modes with wavelengths too long to fit between the plates are excluded. To obtain a finite value for the attractive force a high frequency cutoff must be imposed, but the result is independent of the details of the cutoff because what matters is the difference between the energy density between the plates and the energy density outside, and this difference depends only on the low frequency modes. Without in any way gainsaying the brilliance of Casimir's original derivation, it can be noted that there are alternative derivations of the Casimir force which eschew vacuum energy density and rely only on the source fields of the electrons in the conducting plates. In particular, a calculation of the change in the dipole energy due to the fields of the dipoles acting back on themselves reproduces the Casimir result when normal ordering of the field operators is imposed, setting the energy density of the vacuum to zero [see Milonni (1994, Sects. 7.4–7.6)]. Unfortunately, the matter is less than clear cut since in the perturbative expansion used in quantum electrodynamics, normal ordering does not remove the vacuum energy density beyond zeroth order, and to get rid of the concept of vacuum energy altogether while retaining an explanation of the Casimir effect necessitates more drastic action such as Schwinger's source theory.[30]

Typically, enthusiasts of the interpretation of Λ as vacuum energy density cite the Casimir effect to bolster their position. But what they rarely say is that their position only makes sense if the form of the high frequency cutoff matters, whereas, as we have just seen, the Casimir effect is independent of the form of the cutoff. Returning to the case of the quantization of a scalar field of mass m in terms of an assemblage of harmonic operators, one proceeds to calculate ρ_{vac} by putting the oscillators in a box of volume L^3, imposing a cut off at $k_{\max} >> m/\hbar$, where $\omega_k^2 = k^2 + m^2/\hbar^2$, and computing the energy E_L of the box. ρ_{vac} is then defined as $\lim_{L\to\infty}(E_L/L^3)$, which turns out to have the value $\hbar k_{\max}^4/16\pi^2$. Here is a typical comment on this calculation:

> [W]e can estimate $k_{\max}$ as the energy scale at which our confidence in the formalism no longer holds. For example, it is widely believed that the Planck energy $E^* \approx 10^{19} GeV \approx 10^{16} erg$ marks the point where conventional field theory breaks down due to quantum gravitational effects. Choosing $k_{\max} \approx E^*/\hbar$ we obtain $\rho_{vac} \approx 10^{74} GeV \hbar^{-3} \approx 10^{92}$ g/cm^3. This is . . . approximately 120 orders of magnitude larger than allowed by observation. [Carroll, Press, and Turner (1992, 503)]

Granting that classical GTR and conventional quantum field theory on curved spacetimes both break down at the Planck scale where quantum gravity effects come into play, how does it follow that the correct quantum theory of gravity will yield a finite non-zero vacuum energy density, much less a density of 10^{92} g/cm^3, in some appropriate semi-classical limit?

Suppose for sake of argument that there is no non-sequitur here. Rather than concluding that there is a "cosmological constant problem," one might alternatively conclude that there is something suspect either in the very the notion of vacuum energy density or else in the notion that this energy can serve as a source for the gravitational field.[31]

---

30 As noted by Rugh et al. (1999), who provide a detailed discussion of this vexed subject.

31 As argued by Rugh and Zinkernagel (2000). See also Enz (1974) for some additional skeptical remarks about the concept of zero-point energy.

Of course, one can hold out hope that some cancellation mechanism deriving from a symmetry principle in elementary particle physics will push the value of vacuum energy density down to a tolerable range. But what is hard to understand is how such a mechanism could push the value of Λ down to the very small but positive value indicated by present observations (see Sect. 12) without cancelling it altogether.

11. Inflation

By the end of the 1960s the hot big bang cosmological model, based on the $\Lambda = 0$ Friedmann solutions to Einstein's gravitational field equations, had enjoyed several successes, including an account of primordial nucleosynthesis and the detection of the cosmic microwave background radiation. Its main rival, the steady state model, had been discredited in the eyes of most cosmologists. Thus, the hot big bang model deserved to be called the standard model for cosmology. But despite its exalted status, some theorists remained unconvinced, not because the standard model couldn't accommodate the available empirical data but because the explanations it afforded seemed to them unsatisfying. A paradigm example of this explanatory malaise is given by the so-called *horizon problem*, which can be illustrated by reference to the conformal diagram of the $k = 0$ Friedmann model (see Fig. 4). Before the time t_ℓ of last scattering, the universe is opaque to our optical and radiotelescopes. Using microwave detectors we can see the cosmic 2.7°K background radiation originating from this surface, and it looks very much the same whatever direction we point our detectors. But for directions of sufficiently great angular separation, the points of origination (such as x and y in Fig. 4) have no common causal past ($J^-(x) \cap J^-(y) = \emptyset$).[32] Thus, it seems impossible to explain in terms of causal interactions why the cosmic background radiation is uniform to one part in 10^{-4} – in the standard model, the uniformity just has to be postulated as part of the initial conditions.[33]

Alan Guth (1981) proposed that this and other perceived explanatory inadequacies of the standard big bang model could be overcome by inserting in the very early universe an "inflationary" era which lasted only from $t_i = 10^{-35}$ sec. to $t_f = 10^{-33}$ sec. but which saw such accelerated expansion that Fig. 4 is changed to Fig. 5. The overlap of the past lightcones of x and y in Fig. 5 now allows for a normal causal explanation of the uniformity of the cosmic background radiation. At the end of the inflationary era, the universe resumes its more leisurely expansion as per the standard model. In sum, Guth's scenario promised to retain many of the good features of the standard big bang model while overcoming its perceived shortcomings. Moreover, the mechanism Guth proposed as the driver of inflation promised to unite cosmology and particle physics, a promise that attracted into cosmology a number of workers trained in high energy

[32] If M, $g_{\mu\nu}$ is a spacetime and $x \in M$, the *causal past* $J^-(x)$ is defined as $\{p \in M$: there is a (possibly trivial) future directed causal curve from p to $x\}$.

[33] Compare to the complaint that led Eddington to declare that Λ is needed to give an explanation of the rapid recession of galaxies (see Sect. 7).

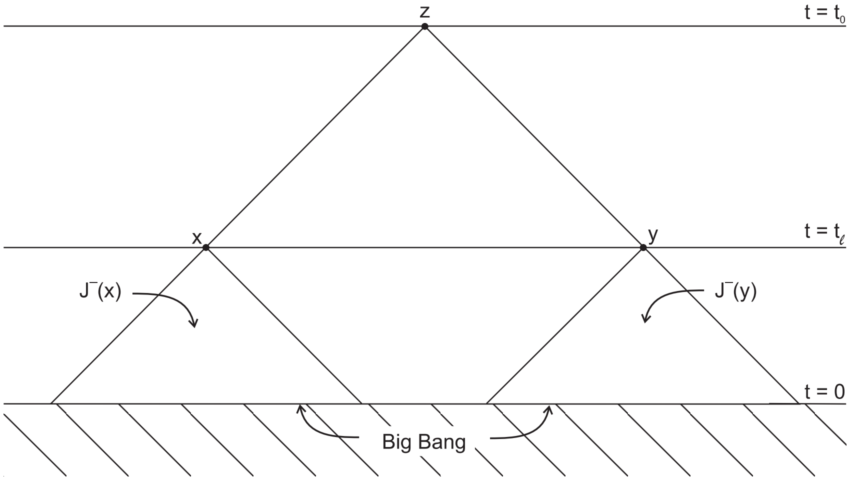

Fig. 4. Conformal diagram of a standard big bang model

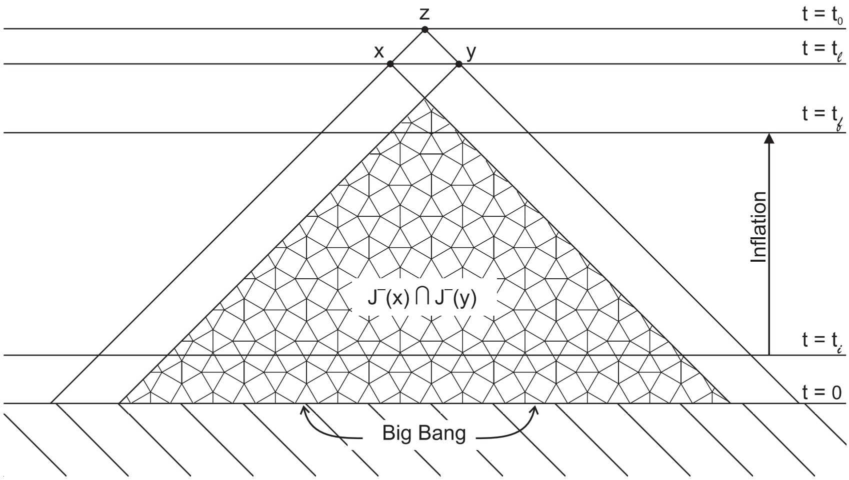

Fig. 5. Inflationary modification of the standard big bang Model

physics. The “inflaton field,” as it came to be called, is an innocent looking scalar field Φ. The stress-energy tensor for such a field has the form

$$T^{\Phi}_{\mu\nu} = \nabla_{\mu}\Phi\nabla_{\nu}\Phi - g_{\mu\nu}\left(\frac{1}{2}g^{\alpha\beta}\nabla_{\alpha}\nabla_{\beta}\Phi + V(\Phi)\right) \tag{24}$$

where $V(\Phi)$ is the potential of the field. When Φ is (approximately) constant, the stress-energy tensor reduces to

$$T^{\Phi}_{\mu\nu} \approx -g_{\mu\nu}V(\Phi) \tag{25}$$

which is the form of a cosmological constant term with $\Lambda_{\Phi} = \frac{V(\Phi)}{8\pi\kappa}$. If Λ_{Φ} is large enough, its inclusion into the Einstein field equations generates the kind of behavior pictured in Fig. 5. Guth's original inflationary scenario identified the inflaton field with the Higgs field responsible for symmetry breaking of the strong and electro-weak forces in a particular version of the grand unified theories (or GUTs) of elementary particles. This hypothetical mechanism for inflation had to be abandoned because the exit from inflation left the universe in too inhomogeneous a state to be compatible with observations. Much work has gone into finding another version of inflation that allows for a graceful exit and at the same time does not involve such an unpalatable "fine tuning" of the parameters of the model as to rekindle the original charge of *ad hoc* explanations. But this is not the place to recount such work.[34]

What is relevant to present concerns, however, is the fact that the simple inflationary model pictured in Fig. 5 seems to require the services of Λ. The dimensionless density parameters associated with matter and Λ respectively are defined by

$$\Omega^M := \frac{8\pi\kappa\rho}{3H^2}, \quad \Omega^\Lambda := \frac{\Lambda}{3H^2}. \tag{26}$$

It is easy to show that for the simple version of inflation, the presence of enough inflation to solve the horizon problem forces the current value of the total density parameter $\Omega_o^{tot} := \Omega_o^M + \Omega_o^\Lambda$ to be very close to unity. But according to current dynamical estimates of matter, $\Omega_o^M \approx 0.3$. Thus, either Λ has to assume its by now familiar fixer role, or else more complicated models of inflation have to be invented that will solve the horizon problem while allowing $\Omega_o^M < 1$ without the assistance of Λ. Such models have been constructed [see, for example, Bucher et al. (1995)], but they seem somewhat contrived. And in any case, there is mounting evidence that Λ, or some close cousin of Λ, is in fact with us.

12. Lambda *redux*

In the 1990s glimmerings of Λ began to appear in cosmological observations. For example, models of structure formation from cold dark matter obtain a better fit to the observed power spectrum of galaxy clustering if $\Lambda > 0$ (see Efstathiou et al. (1990), Kofman et al. (1993)). But, of course, any doubts about the models are visited upon Λ. Another glimmering came from Yoshii and Peterson's (1995) analysis of galaxy number counts as a function of apparent magnitude, which favored a low density universe with a positive Λ. However, the case for Λ is rather delicate because it turns on the issue of how to handle the bias due to selection effects that are built into the methods used to detect faint galaxies. Real excitement about Λ began to build only recently with the use of supernovae to measure the deacceleration parameter.

The dimensionless deacceleration parameter q is defined by

$$q := \frac{-a\ddot{a}}{\dot{a}^2}. \tag{27}$$

[34] For a very readable account of the origins and development of inflationary cosmology, see Guth (1997). For a somewhat more skeptical overview, see Earman and Mosterin (1999).

Of course, $\dot{a}^2 > 0$ and $a > 0$. Thus, $\ddot{a} < 0 \Longleftrightarrow q > 0$. A positive q is then a measure of the slowing down of the rate of expansion of the universe, while a negative q is a measure of the speeding up of expansion.

For a Friedmann model with Λ included, Eq. $(8')$ gives

$$q = \frac{1}{2}\Omega^M \left(1 + \frac{3p}{\rho}\right) - \Omega^\Lambda \tag{28}$$

where the density parameters Ω^M and Ω^Λ for matter and Λ respectively were defined in the preceding section. When $\Lambda \leq 0$ and the matter is not strange $(\rho > 0,\ p \geq 0)$, $q > 0$ and the rate of expansion must slow down. To get $q < 0$, or speeding up of expansion, one must have either a big enough $\Lambda > 0$ or else a strange form of matter.

Let us for the moment ignore strange forms of matter. The present epoch is matter dominated, i.e. $(p/\rho) \ll 1$, allowing us to set

$$q_o \approx \frac{1}{2}\Omega_o^M - \Omega_o^\Lambda. \tag{29}$$

For sake of illustration, take $k = 0$, the case of flat space preferred by most inflationary models. Then

$$\Omega_o^{tot} := \Omega_o^M + \Omega_o^\Lambda \approx 1. \tag{30}$$

Combining (29) and (30) gives

$$q_o \approx \frac{1}{2} - \frac{3}{2}\Omega_o^\Lambda. \tag{31}$$

As already noted, current observational evidence favors $\Omega_o^M \approx 0.3$, which together with (30) implies that $\Omega_o^\Lambda \approx 0.7$. Inserting this value in (31) gives $q_o \approx -0.55$, so that the expansion rate of the universe should be speeding up.

Recent red shift vs. distance measurements on Type Ia supernovae indicate that q_o is indeed negative for these distant objects [see Riess et al. (1998), Schmidt et al. (1998), and Perlmutter et al. (1999)].These supernovae are thought to serve as “standard candles” in that their intrinsic brightness serves as a reliable indicator of their distance. Their apparent brightness is significantly less than would be expected if $q_o \geq 0$. The case for a positive Λ is subject to doubts due to selection effects, extinction effects (due to the presence of cosmic dust), and possible evolutionary effects of the Type Ia supernovae. However, the case is considerably strengthened when the supernovae measurements are combined with measurements of the power spectrum of the anisotropies in the cosmic background radiation. The two types of measurements show “cosmic complementarity” in that they break each other’s degeneracies [see Tegmark et al. (1998)]. The upshot favors the inflationary prediction (30), with a low mass density universe and a $\Lambda > 0$. The preliminary analysis of the recent Boomerang data is consistent with this situation [see Lange et al. (2000)]. The wiggle room here is constrained by gravitational lensing of distant QSOs: Moaz and Rix’s (1993) analysis gives the bound $\Omega_o^\Lambda \leq 0.7$ while Kochanek (1996) obtains $\Omega_o^\Lambda \leq 0.66$ (at the 95% confidence level for a flat model). In sum, when the evidence from the Type Ia supernovae is combined with that from the cosmic microwave background radiation, the dynamical estimates of matter, and the other measurements mentioned at the beginning of this section, the case for a positive Λ is quite impressive.

It would be premature, however, to anticipate a happy ending to the story of Λ's long struggle for respectability. Astrophysicists who are unable to swallow Λ have postulated a new hypothetical form of matter called "quintessence" that will mimic some of the key of the effects of Λ [see Turner and White (1997), Carroll (1998), Cornish and Starkman (1998), Huey et al. (1999), Spergal and Steinhardt (1999), Zlatev et al. (1999), Zlatev and Steinhardt (1999), Wang et al. (1999)]. Quintessence is supposed to be smoothly distributed up to scales of galaxy clustering so that the amount of this new form of matter is not constrained by the current dynamical estimates. Thus, on the quintessence scenario, the dynamical constraint should be written $\Omega_o^{M_C} \leq 0.3$, where M_C stands for ordinary gravitationally clustering matter, allowing for the possibility that quintessential matter M_Q makes up for all of the missing mass required by the standard inflationary scenario in that $\Omega_o^{tot} := \Omega_o^{M_C} + \Omega_o^{M_Q} \approx 1$ without any help from Λ. The equation of state for quintessence (construed as a perfect fluid) is $p_Q = w_Q \rho_Q$, where $-1 < w_Q < 0$. In the simple case where w_Q is constant, the energy density $\rho_Q(t)$ of quintessence scales as $a^{-3(1+w)}$, where $a(t)$ is the cosmic scale factor. The boundary case of $w = -1$ corresponds, of course, to a non-zero energy density of the vacuum due to a real cosmological constant. "Phantom matter" has an equation of state with $w_P < -1$. Its density ρ_P grows with time, and as a result a cosmology that uses phantom matter rather than Λ to make up the missing mass gives predictions for the age of the universe, horizon distance, and gravitational lensing of quasars that differ from those of the analogous Λ model [see Caldwell (1999)].

In addition to providing a solution to the missing mass problem without invoking Λ, quintessence also offers an alternative explanation of an accelerating expansion of the universe. Rewriting (28) to take into account the different forms of matter, we have

$$q = \frac{1}{2}\Omega^{M_C}(1 + 3w_C) + \frac{1}{2}\Omega^{M_Q}(1 + 3w_Q) - \Omega^{\Lambda} \tag{32}$$

where $w_C \geq 0$. In the present era where $w_C = (p_C/\rho_C) \ll 1$, we have as before that $\Omega_o^{M_C}(1 + 3w_C) \approx \Omega_o^{M_C}$. So if $\Lambda = 0$,

$$q_o \approx \frac{1}{2}\left[\Omega_o^{M_C} + \Omega_o^{M_Q}(1 + 3w_Q)\right]. \tag{33}$$

Thus, when $w_Q < -\frac{1}{3}$, q_o is negative if quintessence dominates ordinary gravitationally clustering matter, as it must in order to have $\Omega_o^{M_C} + \Omega_o^{M_Q} \approx 1$ and $\Omega_o^{M_C} \approx 0.3$.

A number of microphysical models for quintessence have been constructed, most of them using a scalar field evolving in a shallow potential to generate the quintessential mass, but at the present stage of development, these models are purely postulational and have not been related to fundamental physics. Depending upon the shape of the potential, the equation of state for quintessence can give a w_Q that may either be a constant or time varying, and in the latter case it can display oscillatory behavior or simple monotonic change. The so-called "tracker models," in which the amount of quintessence tracks the amount of ordinary matter as the universe expands would help to explain why the current value of the part of the density parameter (apparently) due to the cosmological constant is of the same order of magnitude as the current value of the part of the density parameter due to ordinary matter [see Zatlev and Steinhardt (1999)]. Cosmic microwave background measurements that are within present and planned experimental capabilities can in principle distinguish between Λ models and a wide range of quintessence models. The differences in predictions trace to the facts that the energy density of quintessence varies with time and that quintessence can gravitationally clump at very large scales, while vacuum energy density displays neither of these features. Perlmutter, Turner, and White (1999) argue that present observational evidence favors Λ over quintessence. However, Huey et al. (1999) conclude that new observational techniques may be needed to rule definitively in favor of Λ or quintessence.

13. Conclusion

Λ has the most checkered history of any constant in physics. It has been alternately reviled and praised, and it has been counted out many times, only to stage one comeback after another. At one extreme there are those who, like Einstein, feel that a cosmological term in the gravitational field equations is an ugly thing, so ugly that they cannot believe such a thing would be realized in nature. About such feeling it is hard to argue. But one might argue, as did McCrea (1971), that more than esthetics is involved; for if GTR is to be considered a self-contained theory and if it admits Λ as an arbitrary constant, then its predictions become "too indeterminate." But both Newtonian gravitational theory and GTR already contain an arbitrary constant – the gravitational constant. This does not render their predictions indeterminate; indeed, once one solar system test fixes the value of this constant, the predictions for other tests become perfectly determinate. Similarly, once the value of Λ is fixed by one cosmological test, the predictions of GTR for other cosmological tests become perfectly determinate. At the other extreme is the view that, like it or not, Λ is inevitable because it is to be interpreted as the effect of the vacuum energy density of quantum fields [see Carroll, Press, and Turner (1992)]. While some version of this view may prove viable, the ones given to date are less than convincing, and some of the arguments for a non-zero vacuum energy density border on incoherency. Between these two extremes is the more pragmatic stance that Λ is to be taken off the shelf when and only when it is needed to fix a problem in cosmology. Much of Λ's career has been spent in this fixer role: first used by Einstein to make possible a static general relativistic cosmology, it was later invoked to solve the age problem, to explain structure formation, and to explain redshifts of QSO's. Still later inflationary cosmology invoked an effective cosmological constant to overcome the horizon problem and other perceived inadequacies of the standard big bang model. At the present time a positive Λ – or a strange form of matter that imitates some of the effects of a positive Λ – seems to be needed to explain an accelerating expansion of the universe and to make up for the missing mass required by the spatially flat universe favored by inflationary cosmology. Whether or not Λ will survive in this role remains to be seen. But if the past is any guide to the future, Λ will always return in some form or other.

Acknowledgment. I am grateful to George Ellis, Helge Kragh, John Norton, Chris Smeenk, and Henrik Zinkernagel for a number of helpful suggestions on an earlier draft of this paper.

References

[1] Bergmann, P.G. (1942) *Introduction to the Theory of Relativity*. Englewood Cliffs, N.J.: Prentice-Hall.

[2] Bjorken, J.D. and Drell, S.D. (1965) *Relativistic Quantum Fields*. New York: McGraw-Hill.

[3] Bondi, H. and Gold, T. (1948) "The Steady-State Theory of the Expanding Universe," *Monthly Notices of the Royal Astronomical Society* **108**: 252–270.

[4] Bonnor, W.B. (1957) "Jeans' Formula for Gravitational Instability," *Monthly Notices of the Royal Astronomical Society* **117**: 104–117.

[5] Brecher, K. and Silk, J. (1969) "Lemaître Universe, Galaxy Formation and Observations," *Astrophysical Journal* **158**: 91–102.

[6] Bucher, M., Goldhaber, A. and Turok, N. (1995) "An Open Universe from Inflation," *Physical Review D* **52**: 3314–3347.

[7] Caldwell, R.R. (1999) "A Phantom Menace?" astro-ph/9908168.

[8] Cao, Y.T. (ed.) (1999) *Conceptual Foundations of Quantum Field Theory*. Cambridge: Cambridge University Press.

[9] Carroll, S.M. (1998) "Quintessence and the Rest of the World," *Physical Review Letters* **81**: 3067–3070.

[10] Carroll, S.M., Press, W.H. and Turner, E.L. (1992) "The Cosmological Constant," *Annual Review of Astronomy and Astrophysics* **30**: 499–542.

[11] Cohn, J.D. (1998) "Living with Lambda," *Astrophysics and Space Sciences* **259**: 213–234.

[12] Cornish, N.J. and Starkman, G. (1998) "Of Matter Less Repulsive than a Cosmological Constant," astro-ph/9805027.

[13] De Sitter, W. (1917) "On the Curvature of Space," *Proceedings of the Section of Sciences, Koninklijke Akademie van Wetenschappen te Amsterdam* **20**: 229–242.

[14] De Sitter, W. (1931) "The Expanding Universe," *Scientia* **49**: 1–10.

[15] De Sitter, W. (1932) *Kosmos*. Cambridge: Cambridge University Press.

[16] De Sitter, W. (1933) "On the Expanding Universe and the Time Scale," *Monthly Notices of the Royal Astronomical Society* **93**: 628–634.

[17] Earman, J. (1999) "The Penrose-Hawking Singularity Theorems: History and Implications," in Goenner, H., Renn, J., Ritter, J. and Sauer, T. (eds.), *The Expanding Worlds of General Relativity. Einstein Studies Vol. 7*, 235–267. Boston: Birkhaüser.

[18] Earman, J. and Eisenstaedt, J. (1999) "Einstein and Singularities," *Studies in the History and Philosophy of Modern Physics* **30**: 185–235.

[19] Earman, J. and Mosterin, J. (1999) "A Critical Look at Inflationary Cosmology," *Philosophy of Science* **66**: 1–49.

[20] Eddington, A.S. (1923) *The Mathematical Theory of Relativity*. Cambridge: Cambridge University Press.

[21] Eddington, A.S. (1929) *The Nature of the Physical World*. New York: Macmillan.

[22] Eddington, A.S. (1930) "On the Instability of Einstein's Spherical World," *Monthly Notices of the Royal Astronomical Society* **90**: 668–678.

[23] Eddington, A.S. (1933) *The Expanding Universe*. New York: Macmillan.

[24] Eddington, A.S. (1939) "The Cosmological Controversy," *Science Progress* **34**: 225–236.

[25] Efstathiou, G. (1995) "An Anthropic Argument for a Cosmological Constant," *Monthly Notices of the Royal Astronomical Society* **274**: L73–L76.

[26] Efstathiou, G., Maddox, S. and Sutherland, W. (1990) "The Cosmological Constant and Cold Dark Matter," *Nature* **348**: 705–707.

[27] Einstein, A. (1917a) "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie," *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*. 142–152. [English translation as "Cosmological Considerations on the General Theory of Relativity," in Perrett and Jeffrey (eds.), *The Principle of Relativity*, 177–188. New York: Dover Books, 1952.]

[28] Einstein, A. (1917b) *Über die spezielle und die allgemeine Relativitätstheorie*. Braunschweig: Vieweg. English translation as *Relativity: The Special and the General Theory.* New York: Crown Publishers.
[29] Einstein, A. (1918) "Kritisches zu einer von Herrn De Sitter gegebenen Lösung der Gravitationsgleichungen," *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*: 270–272.
[30] Einstein, A. (1919) "Spielen Gravitationsfelder im Aufber der materiellen Elementarteilchen eine wesentliche Rolle?" *Königlich Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*: 349–356. [English translation as "Do Gravitational Fields Play an Essential Part in the Structure of Elementary Particles?" in Perrett and Jeffrey (eds.), *The Principle of Relativity*, 191–198. New York: Dover Books, 1952.]
[31] Einstein, A. (1922) "Bemerkung zu der Arbeit von A. Friedmann: Über die Krümmung des Raumes," *Zeitschrift für Physik* **11**: 326.
[32] Einstein, A. (1923) "Notiz zu der von A. Friedmann: Über die Krümmung des Raumes," *Zeitschrift für Physik* **16**: 228.
[33] Einstein, A. (1931) "Zum kosmologischen Problem der allgemeinen Relativitätstheorie,"*Preussische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*. 235–237.
[34] Einstein, A. (1946) *The Meaning of Relativity,* 2nd ed. Princeton, N.J.: Princeton University Press.
[35] Enz, C.P. (1974) "Is the Zero-Point Energy Real?" in Enz, C.P. and Mehra, J. (eds.), *Physical Reality and Mathematical Description*, 124–132. Dordrecht: D. Reidel.
[36] Friedmann, A. (1922) "Über die Krümmung des Raumes," *Zeitschrift für Physik* **10**: 377–386.
[37] Gamow, G. (1958) "The Evolutionary Universe," in *The Universe. A Scientific American Book*, 59- . London: G. Bell and Sons.
[38] Gamow, G. (1970) *My World Line*. New York: Viking Press.
[39] Guth, A. (1981) "Inflationary Universe: A Possible Solution for the Horizon and Flatness Problems," *Physical Review D* **23**: 347–356.
[40] Guth, A. (1997) *The Inflationary Universe*. Reading, Mass.: Addison-Wesley.
[41] Hoyle, F. (1948) "A New Model for the Expanding Universe," *Monthly Notices of the Royal Astronomical Society* **108**: 372–382.
[42] Hoyle, F. (1948) "On the Cosmological Problem," *Monthly Notices of the Royal Astronomical Society* **109**: 365–371.
[43] Hoyle, F. and Narliker, J.V. (1964) "A New Theory of Gravitation." *Proceedings of the Royal Society (London)* **278**: 191–207.
[44] Hubble, E. (1929) "A Relation between Distance and Radial Velocity Among Extra-Galactic Nebulae," *Proceedings of the National Academy of Sciences* **15**: 168–173.
[45] Huey, G., Wang, L., Dave, R., Caldwell, R.R. and Steinhardt, P.J. (1999) "Resolving the Cosmological Missing Energy Problem," *Physical Review D* **59**: 063005, 1–6.
[46] Kardashev, N. (1968) "Lemaître Universe and Observations," *Astrophysical Journal Letters* **150**: L135–L139.
[47] Kochanek, C.S. (1996) "Is There a Cosmological Constant?" *Astrophysical Journal* **466**: 638–659.
[48] Kofman, L., Gnedin, N. and Bachall, N. (1993) "Cosmological Constant, *COBE* Cosmic Microwave Background Anisotropy, and Large-Scale Clustering," *Astrophysical Journal* **413**: 1–9.
[49] Kragh, H. (1996) *Cosmology and Controversy*. Princeton, N.J.: Princeton University Press.
[50] Lange, A.E. et al. (2000) "First Estimates of Cosmological Parameters from Boomerang," astro-ph/0005004.
[51] Lemaître, G. (1925) "A Note on De Sitter's Universe," *Journal of Mathematical Physics* **4**: 188–192.

[52] Lemaître, G. (1927) “Un universe homogène de masse constante et de rayon croissant rendant compte de la vitesse radiale des nébuleuses extra-glactiques,” *Société Scientifique de Bruxelles. Annales A* **47**: 49–59.

[53] Lemaître, G. (1931a) “A Homogeneous Universe of Constant Mass and Increasing Radius Accounting for the Radial Velocity of Extra-galactic Nebulae,” *Monthly Notices of the Royal Astronomical Society* **91**: 483–489. [English translation of Lemaître (1927).]

[54] Lemaître, G. (1931b) “The Expanding Universe,” *Monthly Notices of the Royal Astronomical Society* **91**: 490–500.

[55] Lemaître, G. (1933a) “Condensations sphèriques dans l’universe en expansion,” *Comptes Rendus* **196**: 903–904.

[56] Lemaître, G. (1933b) “La formation des nébuleuses dans l’universe en expansion,” *Comptes Rendus* **196**: 1085–1087.

[57] Lemaître, G. (1934) “Evolution of the Expanding Universe,” *Proceeding of the National Academy of Sciences* **20**: 12–17.

[58] Lemaître, G. (1949) “The Cosmological Constant,” in Schilpp, P.A. (ed.), *Albert Einstein: Philosopher-Scientist*, Vol. 2, 439–456. New York: Harper and Row.

[59] Martel, H., Shapiro, P. and Weinberg, S. (1998) “Likely Values of the Cosmological Constant,” *Astrophysical Journal* **492**: 29–40.

[60] McCrea, W.H. (1951) “Relativity Theory and the Creation of Matter,” *Proceedings of the Royal Society (London)* **206**: 562–575.

[61] McCrea, W.H. (1971) “The Cosmical Constant,” *Quarterly Journal of the Royal Astronomical Society* **12**: 140–153.

[62] McCrea, W.H. and McVittie, G.C. (1931) “The Expanding Universe,” *Monthly Notices of the Royal Astronomical Society* **92**: 7–12.

[63] McMullin, E. (1993) “Indifference Principle and Anthropic Principle in Cosmology,” *Studies in History and Philosophy of Science* **24**: 359–389.

[64] Milonni, P.W. (1994) *The Quantum Vacuum: An Introduction to Quantum Electrodynamics*. Boston: Academic Press.

[65] Moaz, D. and Rix, H.-W. (1993) “Early Type Galaxies, Dark Halos, and Gravitational Lensing Statistics,” *Astrophysical Journal* **416**: 425–443.

[66] Neumann, C. (1896) *Allgemeine Untersuchungen über das Newton’sche Princip der Fernwirkungen mit besonderer Rückscicht auf die elektrische Wirkungen*. Leibzig: Teubner.

[67] North, J. (1965) *The Measure of the Universe*. Oxford: Clarendon Press.

[68] Norton, J.D. (1999) “The Cosmological Woes of Newtonian Gravitation Theory,” in Goenner, H., Renn, J., Ritter, J., and Sauer, T. (eds.), *The Expanding Worlds of General Relativity. Einstein Studies*, Vol. 7, 271–323. Boston: Birkhaüser.

[69] Norton, J.D. (2000) “Nature is the Realisation of the Simplest Conceivable Mathematical Ideas’: Einstein and the Canon of Mathematical Simplicity,” *Studies in the History and Philosophy of Modern Physics*, in press.

[70] Pais, B. (1982) *Subtle Is the Lord: The Science and Life of Albert Einstein*. Oxford: Oxford University Press.

[71] Pauli, W. (1921) “Relativitätstheorie,” *Encyklopädie der mathematischen Wissenschaften* **V19**. Leipzig: B.G. Teubner.

[72] Pauli, W. (1958) *Theory of Relativity.* New York: Pergamon Press. Revised version of Pauli (1921).

[73] Perlmutter, S., Turner, M.S. and White, M. (1999) “Constraining Dark Energy with SNeIa and Large-Scale Structure,” *Physical Review Letters* **83**: 347–356.

[74] Perlmutter, S. et al. (1999) “Measurements of Ω and Λ from 42 High Redshift Supernovae,” *Astrophysical Journal* **517**: 565–586.

[75] Petrosian, V. (1974) "Confrontation of Lemaître Models and the Cosmological Constant with Observations," in Longair, M.S. (ed.), *Confrontation of Cosmological Theories with Observational Data*, 31–46. Dordrecht: D. Reidel.
[76] Petrosian, V., Salpeter, E. and Sekeres, P. (1967) "Quasi-Stellar Objects in the Universe with Non-Zero Cosmological Constant," *Astrophysical Journal* **147**: 1222–1226.
[77] Petrosian, V. and Salpeter, E. (1970) "Lemaître Models and the Cosmological Constant," *Comments on Astrophysics and Space Physics* **2**: 109–115.
[78] Pirani, F.A.E. (1955) "On the Energy-Momentum Tensor and the Creation of Matter in Relativistic Cosmology," *Proceedings of the Royal Society* (London) **228**: 455–462.
[79] Rawson-Harris, D. (1969) "Formation of Galaxies in Lemaître's Universe," *Monthly Notices of the Royal Astronomical Society* **143**: 49–53.
[80] Riess, A.G. et al. (1998) "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant," *Astronomical Journal* **116**: 1009–1038.
[81] Rindler, W. (1969) *Essential Relativity*. New York: Van Nostrand.
[82] Robertson, H.P. (1928) "Relativistic Cosmology," *Philosophical Magazine* **5**: 835–848.
[83] Rugh, S.E., Zinkernagel, H. and Cao, T.Y. (1999) "The Casimir Effect and the Interpretation of the Vacuum." *Studies in the History and Philosophy of Modern Physics* **30**: 111–139.
[84] Rugh, S.E. and Zinkernagel, H. (2000) "On the Cosmological Constant Problem and the Reality of Vacuum Fluctuations," pre-print.
[85] Sahni, V. (1999) "The Case for a Positive Cosmological Λ-Term," astro-ph/9904398.
[86] Schmidt, B.G. et al. (1998) "The High-Z Supernovae Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type Ia Supernovae," *Astrophysical Journal* **507**: 46–63.
[87] Schulmann, R., Kox, A., Janssen, M. and Illy, J. (eds.) (1998) *The Collected Papers of Albert Einstein. The Berlin Years: Correspondence, 1914–1918*. Princeton, N.J.: Princeton University Press.
[88] von Seeliger, H. (1895) "Über das Newton'sche Gravitationsgesetz," *Astronomische Nachrichten* **137**: 129–136.
[89] von Seeliger, H. (1896) "Über das Newton'sche Gravitationsgesetz," *Bayerische Academie der Wissenschaften. Mathematische-Naturwissenschaftliche Klasse* **126**: 373–400.
[90] Spencer Jones, H. (1934) *General Astronomy* 2nd ed. New York: Longmans, Green and Co.
[91] Spergel, D.N. and Steinhardt, P.J. (1999) "Observational Evidence for Self-Interacting Cold Dark Matter," astro-ph/9909386.
[92] Tegmark, M., Eisenstein, D.J., Hu, W. and Kron, R.G. (1998) "Cosmic Complementarity: Probing the Acceleration of the Universe," astro-ph/9805117.
[93] Tolman, R.C. (1934) *Relativity, Thermodynamics, and Cosmology*. Oxford: Clarendon Press.
[94] Trautman, A. (1965) "Foundations and Current Problems of General Relativity," in Deser, S. and Ford, K.W. (eds.), *Lectures on General Relativity*, Vol. 1, 1–248. New York: Prentice-Hall.
[95] Turner, M. and White, M. (1997) "CDM Models with a Smooth Component," *Physical Review D* **56**: 4439–4443.
[96] Wald, R.M. (1994) *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics.* Chicago: University of Chicago Press.
[97] Wang, L., Caldwell, R.R., Ostriker, J.P. and Steinhardt, P.J. (1999) "Cosmic Concordance and Quintessence," astro-ph/9901388.
[98] Weinberg, S. (1989) "The Cosmological Constant Problem," *Reviews of Modern Physics* **61**: 1–23.
[99] Yoshii, Y. and Peterson, B.A. (1995) "Interpretation of Faint Galaxy Number Counts in the K Band," *Astrophysical Journal* **444**: 14–20.

[100] Zlatev, I. and Steinhardt, P.J. (1999) “A Tracker Solution to the Cold Dark Matter Cosmic Coincidence Problem,” *Physics Letters B* **459**: 570–574.

[101] Zlatev, I., Wang, L. and Steinhardt, P.J. (1999) “Quintessence, Cosmic Coincidence, and the Cosmological Constant,” *Physical Review Letters* **82**: 896–899.

[102] Zeldovich, Ya.B. (1967) “Cosmological Constant and Elementary Particles” *Journal of Experimental and Theoretical Physics. Letters* **61**: 316–317.

[103] Zeldovich, Ya.B. (1968) “The Cosmological Constant and the Theory of Elementary Particles,” *Soviet Physics Uspekhi* **11**: 381–393.

Department of History and Philosophy of Science
University of Pittsburgh
Pittsburgh, PA 15260, USA
jearman@pitt.edu

(*Received September 6, 2000*)