Einstein’s mistake and the cosmological constant

Alex Harvey
Visiting Scholar, New York University, New York, New York 10003

Engelbert Schucking
Physics Department, New York University, 4 Washington Place, New York, New York 10003

(Received 5 August 1999; accepted 20 November 1999)

A brief history of the cosmological constant in the equations of general relativity is presented. Particular attention is paid to (a) a misunderstanding by Einstein of both its function as a repulsive force and new vacuum state rather than the relativistic analog of an exponential potential cutoff he thought he had introduced and to (b) a common misunderstanding of the function of the cosmological constant. © 2000 American Association of Physics Teachers.

I. INTRODUCTION

In late 1915 Albert Einstein submitted his completed theory of general relativity to the Prussian Academy of Sciences. Following this it was as natural for him as it had been for Newton to apply it to the structure of the cosmos in the large. In this endeavor he was guided by the efforts of Newton and his successors. Newton, having constructed a law of gravitation and a viable mechanics, had the tools he needed. His considerations were based on a uniform, static mass density distributed over an infinite Euclidean three-dimensional space. He soon realized that such an infinite uniform distribution of mass was unstable and would collapse. He never resolved the difficulty and it remained a troublesome problem.

It received continuing attention. Laplace suggested a remedy in the form of an exponential damping factor for the force law viz.,

$$F = -\frac{m_1 m_2 e^{-\mu r}}{r^3} \hat{r}.$$  \hfill (1)

This, unfortunately, does not integrate readily to yield a manageable expression for the potential.

This problem was neatly outflanked by the Königsberg theoretician Carl Neumann, who ignored the force law and applied an exponential cutoff directly to the gravitational potential:

$$\phi = \int_0^\infty \rho e^{-r\xi} d\nu.$$  \hfill (2)

This had a number of advantages. The kernel of this integral is a solution of the modified Laplace equation:

$$\nabla^2 \phi - \lambda \phi = 0.$$  \hfill (3)

This both anticipated and pointed in the direction of a supposed resolution to Einstein’s concern about the disastrous influence of distant stars on the local potential. (The problems of constructing viable cosmological models within the context of Newtonian mechanics have long since been successfully addressed. Milne and McCrea in 1934 constructed satisfactory dynamical cosmological models. Research in Newtonian cosmology is a continuing enterprise.)

II. EINSTEIN’S COSMOLOGY

Einstein addressed the problem head-on in his paper "Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie." In reviewing the difficulties besetting the Newtonian model he noted that the problem of obtaining a homogeneous static universe could be solved by replacing the Poisson equation

$$\nabla^2 \phi = 4\pi \kappa \rho$$  \hfill (4)

by the modified equation

$$\nabla^2 \phi - \lambda \phi = 4\pi \kappa \rho,$$  \hfill (5)

with the solution

$$\phi = -\frac{4\pi \kappa \rho}{\lambda}.$$  \hfill (6)

The fourth section of Einstein’s cosmology paper titled "On an Additional Term for the Field Equations of Gravitation" introduced the cosmological constant $\lambda$. Instead of his field equation (13) (see the Appendix for the conventions we use)

$$R_{\mu\nu} = -\kappa(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T)$$  \hfill (7)

(where $\kappa = 8\pi G/c^4$) he suggested now the modified equation (13a),

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

thus introducing the cosmological constant. In the second section of the paper he states that this modification corresponds perfectly to the transition from Poisson’s equation (1) to equation (2) of Sec. I.

When Einstein stated that adding the cosmological term corresponded perfectly to the transition from Poisson’s equation, Eq. (4), to Eq. (5) he was wrong. Nonetheless, generations of physicists have parroted this nonsense. Wolfgang Pauli, that most penetrating of critics, failed to see the error. Abraham Pais writes in his magisterial Einstein biography about the analogy between the $\lambda$ terms in Poisson’s and Einstein’s equations: "he (Einstein) performs the very same transition in general relativity."’ It seemed so deceptively obvious: In Newtonian approximation with $c = 1$ Eq. (7) yields

$$g_{00} = - (1 + 2\phi).$$  \hfill (8)

Thus, adding the term $-\lambda \phi$ to the Poisson equation to obtain Eq. (5) should correspond to adding the term $\lambda g_{\mu\nu}$ to the Einstein equations.

As a sidelight we mention an incident. Many years ago Otto Heckmann commented to one of us (ES): ‘‘Einstein’s Argument ist natürlich Quatsch.’’ And the late Hamburg cosmologist was right. If c is not set equal to 1 then $\phi$ should
be written as \( \phi/c^2 \) and can be neglected compared to 1 in first approximation. The Newtonian limit of Einstein’s equation (5) but

\[
\nabla^2 \phi + \lambda c^2 = 4 \pi \kappa \rho.
\]

(9)

With Eq. (7a) Einstein had not introduced an exponential cutoff for the range of gravitation but a new repulsive force \( (\lambda > 0) \), proportional to mass, that repelled every particle of mass \( m \) with a force

\[
\vec{F} = mc^2 \frac{\lambda}{3} \vec{\dot{x}}.
\]

(10)

It is this repulsive force which is the basis for the old saw that in the Einstein cosmos there was “matter without motion” whereas in the de Sitter cosmos there was “motion without matter.” In both cases the cosmological constant provides a repulsive force. In the Einstein cosmos this force balances the attractive force of the distribution of mass; in the de Sitter case the absence of this mass distribution allows a particle placed at any point away from the origin of coordinates to fly off. This was clearly stated by Arthur Eddington.\(^{11}\)

Instead of getting a shielded gravitational force one had now at large distances almost naked repulsion. This was quite different from the expected bargain. But, it did provide precisely the cosmological model Einstein desired. It provided \textit{inter alia} a static, closed universe.

One cannot know with certainty how Einstein arrived at the modified field equations (7a). Much of the history of its introduction has been treated in detail in the recent book by Kerszberg.\(^{12}\) What Kerszberg, in company with many physicists, does not realize is that Einstein made a remarkable mistake in \textit{interpretation} when he introduced the cosmological constant. Based on our experience it seems that a coherent, concise version of the background to its introduction by Einstein and its proper significance might be useful.

### III. THE EINSTEIN TENSOR

It is convenient to rewrite Eq. (7a) in the more common form

\[
R_{\mu \nu} - \frac{1}{2} g_{\mu \nu} R + \lambda g_{\mu \nu} = - \kappa T_{\mu \nu}.
\]

(11)

For obvious reasons the structure of the left-hand side came under close scrutiny by his colleagues. Einstein had spent almost a decade to find the first two terms. Now a third term was added. What should be the defining characteristics of this gravitational field tensor? Are there other terms which might be added or other modifications?

The left-hand side of Einstein’s field equations has the following properties:

- It is a second rank tensor constructed solely from the metric tensor and its first and second derivatives.
- It is linear in terms of the second differential order.
- It has a vanishing covariant derivative.

It was first discussed by Vermeil\(^ {13}\) in 1917 and somewhat later by Weyl\(^ {14}\) in 1922.

In his article surveying relativity theory Pauli\(^ {15}\) stated that the field equations for the metric tensor in the presence of matter must have the general form

\[
c_1 R_{\mu \nu} + c_2 R g_{\mu \nu} + c_3 g_{\mu \nu} = - \kappa T_{\mu \nu}
\]

(12)

and that, necessarily, \( c_1 = 1 \), \( c_2 = -1/2 \), and \( c_3 \) is just the cosmological constant.

The most recent examination resides in a searching theorem constructed by Lovelock\(^ {16}\) which severely delimits the form of this tensor. He has shown that if the field equations are to be derived from a variational principle then: \textit{In a four-dimensional space the only type (2,0) tensor density whose components satisfy

\[
A^{ij} = A^{ij} (g_{ab} \partial_a \partial_b \epsilon_{cd}), \quad A^{ij} = 0
\]

is given by

\[
A^{ij} = \alpha \sqrt{g} (R^{ij} - \frac{1}{2} g^{ij} R) + \lambda \sqrt{g} g^{ij}.
\]

IV. THE COSMOLOGICAL TERM

It is clear that the most general form of the Einstein tensor subject to the various constraints is

\[
R_{\mu \nu} - \frac{1}{2} g_{\mu \nu} R + \lambda g_{\mu \nu}.
\]

The inclusion of the term \( \lambda g_{\mu \nu} \) is not at all arbitrary. It is not an adjustable parameter if one does not take offense at the fact that the constant \( \lambda \) is not dimensionless. Indeed, its omission requires justification. Its presence changes the \textit{character} of the field equations and delimits the kind of solutions available. One has for the field equations of the vacuum

\[
R_{\mu \nu} = \lambda g_{\mu \nu}.
\]

(13)

the “flattest” solution for which is a space of constant curvature. Given that no general criteria exist for deciding if a solution to the field equations has a physical relevance, the inclusion of \( \lambda \) should be resolved on the basis of observational data, a point ultimately advocated by Einstein himself.\(^ {17}\)

Adding the term \( \lambda g_{\mu \nu} \) to the left-hand side of the field equations for positive \( \lambda \) is equivalent to adding to the vacuum a positive constant pressure \( \lambda \) and a negative energy density \( \lambda \). This is made immediately manifest by the field equations with stress energy tensor for a perfect fluid with pressure,

\[
R_{\mu \nu} + \frac{1}{2} R g_{\mu \nu} + \lambda g_{\mu \nu} = - \kappa [(\rho + p)u_{\mu} u_{\nu} + pg_{\mu \nu}].
\]

(14)

V. GENERAL EFFECT OF THE COSMOLOGICAL CONSTANT

To see the effect of a cosmological term on the motion of particles, it is convenient to consider the equation of geodesic deviation. We give a brief derivation. Let \( u^\mu(x^\lambda) \) be the unit tangent vector of a congruence of timelike geodesics

\[
u^\mu u_{\mu} = -1, \quad \nabla_\nu u^\mu = 0.
\]

(15)

Further, let \( \eta^\mu(x^\lambda) \) be a vector field orthogonal to the congruence such that \( \epsilon \eta^\mu(x^\lambda) \) is an infinitesimal vector connecting a geodesic with a neighboring one. We have then

\[
u^\mu \eta_\mu = 0
\]

with vanishing Lie derivative of \( \nu^\mu \) with respect to \( \eta^\nu \),

\[
\mathcal{L}_\nu u^\mu = u^\mu \eta^\nu - \nu^\mu \eta_\nu = 0.
\]

(16)

We write \( \eta^\mu \) for the directional derivative of the vector \( \eta^\mu \) along a timelike geodesic with respect to arclength \( s \) and by virtue of Eq. (16) the result is
For the second derivative \( \ddot{\eta} \) we obtain

\[
\ddot{\eta} = (u^\mu ; : \eta) ; \eta = (u^\mu ; : \eta^\rho) u^\rho + u^\mu ; \eta^\rho , \eta^\rho ,
\]

the second term of which may be converted by virtue of Eq. (17) so that

\[
\ddot{\eta} = (u^\mu ; _\rho \eta^\rho) u^\rho + u^\mu ; u^\nu , \eta^\rho .
\]

Interchange the summation indices in the first term and note that

\[
u^\mu ; u^\nu = (u^\mu ; v) ; v - u^\mu ; \nu u^\nu .
\]

The first term on the right-hand side vanishes so that

\[
\ddot{\eta} = (u^\mu ; _\rho v) \eta^\rho u^\nu = R^\rho _{\kappa \nu \lambda} u^\lambda u^\nu \eta^\rho .
\]

Define

\[
E_{\mu \rho v} = R^\rho _{\kappa \nu \lambda} u^\lambda u^\nu = E_{\rho \mu},
\]

so that

\[
\ddot{\eta} = E_{\rho \mu} \eta^\rho .
\]

The tensor \( E_{\mu \rho v} \) is a symmetric, purely spatial tensor, i.e.,

\[
u^\mu E_{\mu \rho v} = 0,
\]

and we know that

\[
E_{\mu \rho v} = - R_{\kappa \nu \lambda} u^\nu u^\lambda .
\]

For reasons indicated below we shall refer to it as the Pirani tensor.

The equations of motion provided by the divergence of a stress-energy tensor for a perfect fluid with four-velocity \( u^\mu \), density \( \rho \), and pressure \( p \) [see Eq. (14)] are

\[
(\rho + p) u^\mu = - p, v \delta^\mu _{v} + u^\mu u^\nu .
\]

For vanishing spatial gradient the motion of the fluid becomes geodesic and the left-hand side vanishes. The field equations (14) give

\[
R_{\mu \nu v} = \kappa \left( \frac{(\rho - p) \delta_{\mu \nu} - (\rho + p) u^\mu u^\nu + \lambda g_{\mu \nu}}{c^2} \right).
\]

We thus have

\[
E_{\mu \nu v} = + \frac{\kappa}{2} (\rho + 3p) - \lambda = \frac{4 \pi G}{c^2} (\rho c^2 + 3p) - \lambda .
\]

The interpretation of these formulas and their relevance to a more precise understanding of the role of the cosmological constant should be placed in historical perspective. The formula for geodesic deviation was introduced by Carl Gustav Jacobi in his study of geodesics on ellipsoids. Given a geodesic \( \Gamma (s) \) with arclength \( s \) measured from an arbitrary point \( P \), consider now a neighboring geodesic on their common surface and let \( \eta (s) \) as defined above the (infinitesimal) distance of an immediately adjacent geodesic \( \Gamma' \) from \( \Gamma \), Jacobi found that

\[
\frac{d^2 \eta}{ds^2} + K(s) \eta = 0,
\]

where \( K(s) \) is the Gaussian curvature of the surface along the geodesic. Tullio Levi-Civita generalized this beautiful intrinsic formula to geodesics on an \( n \)-dimensional Riemannian manifold. Another derivation can be found in the Tensor Calculus of Synge and Schild. But, it was Pirani who recognized the fundamental importance of this tensor for understanding of Einstein’s theory of gravitation and he used it to derive Einstein’s field equations. ²¹

To establish now a connection with Newtonian gravity, we pick a time-like geodesic \( \Gamma_i \) and attach to it by parallel transport a set of four orthonormal vectors \( \{ \mathbf{e}_0, \mathbf{e}_j \} \) with

\[
\mathbf{e}_0 ^\mu = u_\mu (s)
\]

on \( \Gamma_i \). We put

\[
y^j = \eta^\mu \mathbf{e}_\mu ^j .
\]

Because

\[
\dot{\mathbf{e}}_0 ^\mu = 0
\]

along \( \Gamma_i \), the \( y^j \) are spatial coordinates in a freely falling local inertial system and Eq. (21) of geodesic deviation takes the form

\[
y^j = E_{\mu v} y^\mu , \quad E_{\mu v} = E_{\mu v} \mathbf{e}_\mu ^j / e_v ^j.
\]

In the origin of a freely falling system, i.e., an Einstein elevator, the gravitational potential and its gradient vanish and the equation of motion of a particle at the position \( y^j \) is, in first approximation,

\[
y^j = - \phi _{j,k} y^k
\]

at \( y^k \neq 0 \). Because \( E_{00} \) vanishes according to Eq. (22) we have

\[
E_{\mu \nu} = E_{\mu j} = + \nabla^2 \phi .
\]

The trace of Pirani’s tensor has to be identified with the negative Laplacian of the gravitational potential.

We can now establish the Newtonian equivalent of the \( R_{00} \) component of Einstein’s field equations for a perfect fluid without a pressure gradient. We get from Eq. (23)

\[
\nabla^2 \phi = \lambda = 4 \pi G \left( \rho + 3p / c^2 \right).
\]

In contrast to Poisson’s equation there are two additional terms. If the \( \lambda \) term is brought to the right-hand side it appears for positive \( \lambda \) as a negative density of active mass of \( \lambda / 4 \pi G \) which should give rise to general repulsion. The other term is the surprising appearance of the additional contribution of \( 3p / c^2 \) to the active mass density. For a relativistic fluid this would mean that its active mass density would be larger than the inertial. This is known as Tolman’s paradox because for closed systems active and inertial masses are equal. The paradox was resolved by Misner and Putnam, who showed that the stresses in the container for the fluid cancel the mass contribution of the \( 3p / c^2 \) term.

The influence of the \( \lambda \) term on Newton’s local equation of motion can also be easily seen. Einstein’s field equations with cosmological constant and absence of matter are

\[
R_{\mu \nu} = \lambda g_{\mu \nu}.
\]

As noted earlier, the simplest solution for the vacuum is a space–time of constant curvature the Riemann tensor for which is just

\[
R_{\mu \nu \rho} = - \frac{\lambda}{3} (g_{\mu \rho} g_{\nu \lambda} - g_{\mu \lambda} g_{\nu \rho}) .
\]

The corresponding Pirani tensor is

\[
E_{\mu \rho} = \frac{\lambda}{3} (g_{\mu \rho} + u_\mu u_\rho) .
\]
This leads to the equation of motion for free particles

\[ y^j = \frac{\lambda}{3} y^j. \]  

(37)

For \( \lambda > 0 \) this is a uniformly repulsive force proportional to mass.

The intrinsic existence of such a force in the de Sitter vacuum was pointed out by Eddington\(^1\) and used to explain the redshifts of distant galaxies measured by Slipher.\(^2\)

Otto Heckmann\(^25\) was, as far as we know, the first to point out Einstein’s mistake. In a footnote on p. 15 of his Theorien der Kosmologie he remarks,

The equation \( \Delta + \lambda \phi = 4\pi G \rho \) which is different from \( \Delta + \lambda g(t) = 4\pi G \rho \) is used by Einstein in his paper S.-B. Preuss. Acad. Wiss. 1917, 142 to explain the introduction of the term \( \lambda g_{\mu\nu} \) into his field equations. This suggestion for a change of Newton’s law (C. Neumann: “About the Newtonian Principle of Action at a Distance,” p. 1 and 2, Leipzig 1896—see also Leipziger Ber. Math.-Phys. Kl. 1874, 149) does not result as an approximation of the field equations of relativity theory. Thus, the argument which Heckmann and Siedentopf\(^26\) gave for their Eq. (5.18) is void.

From the last sentence it is clear that Heckmann also had once been fooled by Einstein’s erroneous argument. Heckmann’s book published during WWII never did find a wide readership nor did its re-publication in 1968. That it was written in German didn’t help. In this later edition he added the remark:

The \( \Lambda \)-term is kept in the whole book. The disdain towards this term seen again and again is Einstein’s own fault. It neglects the only rigorous way that we know for the derivation which gives also \( \Lambda \). Who trusts the Einstein theory should also trust \( \Lambda \) and should not carelessly put it equal to zero. Comp. McVittie in “H. P. Robertson in Memoriam,” p. 18ff, Philadelphia 1963.

This was also Albert Einstein’s original feeling. On 13 April 1917 he wrote to Willem de Sitter,\(^17\)

In any case, one thing stands. The general theory of relativity allows the addition of the term \( \lambda g_{\mu\nu} \) in the field equations. One day, our actual knowledge of the composition of the fixed star sky, the apparent motions of fixed stars, and the position of spectral lines as a function of distance, will probably have come far enough for us to be able to decide empirically the question of whether or not \( \lambda \) vanishes. Conviction is a good motive, but a bad judge.

VI. FINAL COMMENTS

The cosmological constant has been a fruitful source of controversy ever since Einstein added it to his original field equations. Now it returns once again from limbo meriting serious consideration. Perhaps it should be called the Phoenix Constant.

In addition to being controversial among specialists in cosmology it has attracted the obsessive attention of popular science writers who uniformly object to its usage. Their most common characterization is “fudge factor.” This species of scientific illiteracy is well-exemplified by a recent popular article by the well-known science writer Nicholas Wade titled Star Spangled Scandal\(^27\) in which it is suggested that usage of the cosmological constant by astrophysicists to resolve a problem concerning the age of the universe was “…a scandal of the intellectual kind. You know like iatrogenic disease or the Constitution’s provisions for slavery.”

A similar presentation of nonsense was given by Donald Goldsmith.\(^28\) In the first paragraph on p. 7 he states, “In 1917 … Einstein realized to his dismay that the equations of his theory of general relativity implied either an expanding or contracting universe.” This is simply not true. That the equations contained such solutions was not made known until the work of Friedmann\(^29\) in 1922. What neither Wade nor Goldsmith nor others of their ilk understand is that the cosmological constant is not a mere adjustable parameter and far worse lack any understanding of the structure of Einstein’s argument for its inclusion; its presence or absence changes the character of the field equations.

Though Einstein ultimately rejected the inclusion of the cosmological constant in his equations, it was not on the basis that it was a pressure rather than an exponential cutoff, but rather that he deemed it unnecessary. It is doubtful that Einstein ever said it was the “biggest blunder” of his life. The source of this is George Gamow,\(^30\) who had a well-established reputation as a joker and given to hyperbole. The comment doesn’t appear in Einstein’s writings,\(^31\) It is possible that Einstein made some mild comment in German to Gamow at his having persuaded himself that the term was necessary.

With keen hindsight we observe now that the introduction of the cosmological constant amounted to a redefinition of the vacuum state for the universe—the replacement of Minkowski space–time by de Sitter space–time.

APPENDIX: CONVENTIONS AND UNITS

Einstein and Eddington both used \( G_{\mu\nu} \) to denote the Ricci tensor. We conform to current practice and use \( R_{\mu\nu} \). We (usually) follow the source in presenting the various equations and expressions. The velocity of light is variously taken to be 1 or \( c \) depending on the context. Greek and Roman summation indices have, respectively, the ranges \([0, 1, 2, 3]\) and \([1, 2, 3]\).  

Footnotes:
\footnote{Professor Emeritus, City University of New York; electronic mail: harvey@scires.acf.nyu.edu}
\footnote{Electronic mail: els@is8.nyu.edu}
\footnote{A. Einstein, The Principle of Relativity (Dover, New York, 1923), pp. 107–164.}
\footnote{S. P. Laplace, Mécanique Céléste, Livre xvi, Chap. 4.}
\footnote{C. Neumann, Allgemeine Untersuchungen über das Newtonsche Prinzip der Fernwirkungen (Teubner, Leipzig, 1896), pp. 1, 2.}
\footnote{See H. Bondi, Cosmology (Cambridge U.P., Cambridge, 1961), pp. 75–89.}
\footnote{See Ref. 1, pp. 177–188.}
\footnote{W. Pauli, Theory of Relativity (Pergamon, London, 1958), pp. 159–161.}
\footnote{A. Pais, Subtle is the Lord (Oxford U.P., Oxford, 1982), p. 286.}
\footnote{W. Weinberg, Gravitation and Cosmology (Wiley, New York, 1972), pp. 77–78.}
\footnote{Einstein’s argument is of course baloney.}
If science ceases to be a rebellion against authority, then it does not deserve the talents of our brightest children. I was lucky to be introduced to science at school as a subversive activity of the younger boys. We organized a Science Society as an act of rebellion against compulsory Latin and compulsory football. We should try to introduce our children to science today as a rebellion against poverty and ugliness and militarism and economic injustice.