On the Value of the Cosmical Constant.

By Sir Arthur Eddington, F.R.S.

(Received August 11, 1931.)

1. The cosmical constant $\lambda$ occurs in Einstein's law of gravitation $G_{\mu\nu} = \lambda g_{\mu\nu}$. In the resulting equations of motion the term containing $\lambda$ represents a scattering force which tends to make all very remote bodies recede from one another; this phenomenon is the basis of the theories of de Sitter and Lemaitre concerning the "expansion of the universe." If the observed recession of the spiral nebulae is a manifestation of this effect the value of $\lambda$ can be found from the astronomical observations.

In this paper I put forward a simple geometrical interpretation of the term in the wave equation which contains the mass $m$ of an electron; this interpretation provides an alternative expression for the term. The new expression involves $\lambda$, and by equating it to the ordinary expression we find a theoretical value of $\lambda$, viz., $9.8 \cdot 10^{-55}$ cm.$^{-2}$. This agrees satisfactorily with the value found from the observed recession of the spiral nebulae (§ 8).

The wave equation for an electron moving in the electrostatic field due to a fixed electron is:

$$\frac{i\hbar}{2\pi} \frac{\partial \psi}{\partial t} = c \left\{ m^2 c^2 + \left( \frac{i\hbar}{2\pi} \right)^2 \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right\}^{1/2} \psi + \frac{\alpha^2}{r} \psi. \quad (1)$$

Let

$$\alpha = \frac{\hbar c}{2\pi e^2}, \quad \gamma = \frac{2\pi m c \alpha}{\hbar}. \quad (2)$$

Then (1) can be written

$$\left( \alpha \frac{\partial}{c \partial t} + \frac{i}{r} \right) \psi + \left\{ \alpha^2 \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right\}^{1/2} \psi = 0. \quad (3)$$

Following Dirac, we can rationalise (3) by introducing four perpendicular matrices which in my usual notation would be denoted by $E_{14}$, $E_{24}$, $E_{34}$, $E_{4}$; but I will here write them $i_{14}$, $i_{24}$, $i_{34}$, $i_{4}$, as a reminder that they are all square roots of $-1$, though they must be distinguished from the ordinary $i$. The resulting linear equation is

$$\left\{ i \left( \alpha \frac{\partial}{c \partial t} + \frac{i}{r} \right) + i_{14} \alpha \frac{\partial}{\partial x} + i_{24} \alpha \frac{\partial}{\partial y} + i_{34} \alpha \frac{\partial}{\partial z} + i_{4} (\dot{\gamma}) \right\} \psi = 0. \quad (4)$$

The argument of this paper is that $\gamma$ has the interpretation

$$\gamma = \sqrt{N/R},$$

where $R$ is the radius of curvature of the universe in a static state or eigenstate, and $N$ the number of electrons in the universe. It will be seen that there is a suggestive resemblance in the structure of the terms

$$i\left(\frac{i}{r}\right) \quad \text{and} \quad i_4\left(\frac{i\sqrt{N}}{R}\right)$$

in (4); and my theory turns on this similarity. The suggestion is briefly that, if the energy of a specified electron due to a singularity at distance $r$ is $i/r$, the energy of $N$ indistinguishable electrons due to a singularity at distance $R$ (the centre of the spherical universe) will be $i\sqrt{N}/R$.

The suggestion is not based merely on a general analogy. I think I have been able to show that this simple relation of the two terms is definitely demanded by the geometry of the problem (see § 6).

2. In trying to understand the origin of mass we keep in mind the following principles. According to relativity theory, mass is associated with curvature of space-time. So long as we keep to the flat space-time ordinarily adopted in wave-mechanics we cannot expect to obtain light on the nature of the term involving the mass; it must remain an empirical term. But by contemplating curved space-time we may hope to assimilate the mass-term into the theory, and see that it arises in a similar manner to the terms involving other energies recognised in quantum theory.

The Machian view that the mass of a body owes its existence to the presence of other bodies was at one time strongly urged by Einstein, and has been more or less recognised in relativity literature. The conceptions of the quantum theory favour it—in particular, the recognition that an electron has no individuality and is not separable from all the other electrons in the universe in the way that the classical picture supposed. I take the view that the mass of an electron is an interchange energy with all the other charges in the universe suitably averaged.

In any problem relating to the actual world we may divide the particles into two groups, (a) certain particles specifically mentioned in the problem, and (b) the unspecified particles. The unspecified particles are replaced by an averaged distribution of their states which is called the field. If the field has directional characteristics this must be definitely stated in the enunciation of the problem, viz., that the specified particles are in a particular gravitational or electromagnetic field, and equation (1) must be modified accordingly.
Otherwise we are entitled to assume that the field is of the most featureless kind possible; it is then in one aspect a pure inertial field and in another aspect it is a uniform macroscopic space-time. If we do not take account of (b) in our problem there is no inertial field and the specified particles (a) are without inertia; they have then no energy other than that of their mutual interaction, and the mass-term will be missing from (1).

3. The physical conception of location in space is inseparable from the conception of interaction; for there is no meaning in saying that a body is located at A rather than at B unless it makes some difference to something that it is at A and not B. The significance of location is that the symbols specifying it (co-ordinates) enter as parameters into the functions describing interaction effects.

If we consider one specified particle in a uniform space so that there is nothing to distinguish one possible position of the particle from another, the interaction of the particle with the space (i.e., with the averaged state of the unspecified particles, which the uniform space or inertial field replaces) must be independent of its co-ordinates. At first sight we might think that the interaction must therefore be zero. But we are familiar in dynamics with cases in which the Hamiltonian, whilst not involving a co-ordinate explicitly, involves the corresponding velocity. Such co-ordinates are cyclic, and the energy associated with them must be inserted in the Hamiltonian as shown in the theory of ignorance of co-ordinates. It seems reasonable to think that all terms in (4) of potential energy type, i.e., not explicitly involving derivatives, represent energy corresponding to ignored cyclic co-ordinates.

The definition of a cyclic co-ordinate θ is that by continued application of the change dθ we reach a state which is counted as identical with the state started from. In flat space-time cyclic co-ordinates can only exist in connection with a singularity; and we see in (4) that corresponding to a specified singularity (electron or proton) a term (i/τ) appears in the Hamiltonian which presumably represents the energy of the cyclic co-ordinate admitted by it. In a spherical world continued progress in a uniform direction brings us to a state reckoned to be identical with the state started from; I consider that the term (iγ) in the Hamiltonian represents the corresponding cyclic momentum.

It is hardly necessary to emphasise that the importance of cyclic transformations is that they are a pitfall for statistical theory. States which are represented as distinct in our continuous analysis are counted as identical for statistical purposes. The primary effect of adopting a curved world in our theory instead of the flat space usually treated is that we introduce a new
connectivity which must be allowed for in the statistical theory. The interpretation of irregularities of curvature as gravitational fields is a side development which does not here concern us, and we confine attention to a uniform and therefore spherical space.

4. We now see how material particles "produce" curvature of space-time. As individual singularities they would have cyclic co-ordinates associated with them. In averaging, we may smooth out the singularities and abolish the individual cyclic co-ordinates provided that we furnish an equivalent cyclic co-ordinate in the space which we substitute; and, of course, the space cannot furnish such a cycle unless it is re-entrant. The substituted space is the macroscopic space of which we are aware, which is accessible to experiment and measurement; and we may assume that its dimensions are such as to provide a cyclic co-ordinate dynamically equivalent to the suppressed co-ordinates.

Although only one* particle is specified, the system which we are considering includes all the particles in the universe; for unless we include the unspecified particles the mass of the specified particle is non-existent and we shall not obtain equation (1). The only statical arrangement of N particles is in the form of an Einstein world and we shall deal with this case.

Our "expanding universe" is probably far from the static configuration, and the question arises whether the value of \( m \) which we shall deduce for an Einstein universe applies to it. This would be a difficulty if we were attempting to establish equation (1) by a new theoretical argument. That is not our intention. We accept equation (1) as the orthodox wave equation which is supposed to be true whatever the state of the universe,† and apply it to a simple tractable case from which we can obtain a useful result. According to quantum mechanics (1) is satisfied when the specified particle is in an eigenstate whatever the state of the N — 1 unspecified particles, provided that they do not produce an electric or gravitational field; it is therefore satisfied when the whole N particles are in an eigenstate and form an Einstein world, so that we are at liberty to apply it to this case.

The present radius of our expanding universe is greater than the radius \( R \) of the Einstein world; but, inasmuch as it does not correspond to an eigen-

* Or two, if we count the supernaturally fixed electron which is introduced to provide a convenient comparison term.
† The investigation in § 7 throws some light on the question why (1) continues to hold as the wave equation for a specified electron when the unspecified particles are not in an eigenstate.
Value of Cosmical Constant. 609

state, it is irrelevant to the theory of the wave equation. That is why the
Einstein radius and not the actual radius is given by (5).

5. We know in a general way that the term \((i/r)\) in (4) arises from the inter­
changeability of the two electrons. The line \(r\) joining them has a (concealed)
way of rotating so that its two ends become interchanged. If \(\theta\) is the angular
variable measuring this rotation, \(\psi\) must contain a factor \(e^{i\theta}\), since according
to the Fermi-Dirac principle the sign of \(\psi\) is reversed when the rotation is
through an angle \(\pi\) and the two electrons are interchanged. The term \((i/r)\) \(\psi\)
could be written

\[
\frac{\partial}{r \partial \theta} \psi,
\]

so that it is comparable with \(\partial/\partial x, \partial/\partial \epsilon t\), occurring elsewhere in (4). Or we
may consider the momentum \(p\) associated with a circuit of radius \(r\) by the
elementary quantum rule \(2\pi r \cdot p = \hbar\), so that

\[
p = \frac{\hbar}{2\pi} \cdot \frac{1}{r},
\]

the factor \(h/2\pi\) having been removed from the Hamiltonian in (4).*

Our spherical space provides a circuit of radius \(R\), and on the same principles
this should furnish a term \((i/R)\) \(\psi\). There will be \(N\) such terms, one for each of
the \(N\) electrons whose eigenstate is being considered; for, although they are
unspecified and their co-ordinates are " ignored " in the Hamiltonian, their
cyclic momenta must be included. The question remains whether the \(N\)
equal terms will compound additively to give a resultant \(iN/R\) or perpendicularly
to give a resultant \(i\sqrt{N/R}\). Both kinds of combination are exemplified in
wave mechanics. It seems clear that one or other of these resultants will
appear in the wave equation and the only possible term in (4) with which it
can be identified is the mass-term \(i\gamma\). The question is whether it is the term
\((i\gamma)\) in (4) or the term \((i\gamma)^2\) in (3) which is the simple sum of \(N\) elementary terms.
In the latter case we can still exhibit the \(N\) terms separately in (4) as

\[
i'_4 \cdot (i/R) + i''_4 \cdot (i/R) + i'''_4 \cdot (i/R) + \ldots,
\]

where \(i'_4, i''_4, i'''_4, \ldots\) are different anticommuting square roots of \(-1\).

In favour of \((iN/R)\) it may be urged that in the wave equation for two
particles the masses are simply added. But that is an altogether different

* It is not necessary to consider here why the potential energy escapes the factor
\(\alpha\) (= 137) which multiplies the kinetic energy. (I have written on this in other papers.)
We shall proceed by comparing potential energy with potential energy, or cyclic energy
with cyclic energy, so that the factor (whatever its nature) is eliminated.
application; the mass term is then doubled, not because there are more particles in the system (for the change is merely that an unspecified particle becomes specified), but because the number of dimensions of the configuration space adopted is doubled. A stronger argument is that when there is more than one fixed electron the corresponding terms \((i/r)\) are added. But these terms occur linearly both in (3) and (4), so that there is here no suggestion of perpendicular combination. The fact that in the known case of linear composition the equations indicate it unambiguously seems to make it all the more likely that their hint of quadratic composition of \((i\gamma)\) is to be accepted.*

If \(\psi\) is normalised so as to correspond to one electron in spherical space, \(\sqrt{N} \cdot \psi\) represents the same distribution with density \(N\) times greater, corresponding therefore to \(N\) electrons in spherical space. I think, therefore, we get a simple view of (4) if we write it as

\[
\left\{ i \left( \alpha \frac{\partial}{\partial t} + \frac{i}{r} \right) + i_{14} \alpha \frac{\partial}{\partial x} + i_{24} \alpha \frac{\partial}{\partial y} + i_{34} \alpha \frac{\partial}{\partial z} \right\} \psi_1 + \left\{ \frac{\dot{\psi}}{R} \right\} \psi_N = 0, \tag{6}
\]

where \(\psi_N = \sqrt{N} \cdot \psi_1\). Here the first part, consisting of terms referring to one specified electron, is multiplied by the wave function normalised to represent one electron; and the second part, containing the term common to all the electrons, is multiplied by the wave function normalised to represent \(N\) electrons.

Finally, we may perhaps even in the midst of a theoretical discussion not close our eyes altogether to experience and, inasmuch as the two alternative values of the term differ by a factor of about \(10^{39}\), there is no doubt that \(\sqrt{N}/R\) is the one which accords better with observation.

6. In this section I endeavour to show that the correspondence of \(i/r\) and \(i/R\) is not a vague analogy, but is definitely required by the conditions of the problem.

A difficulty in treating ideally simplified problems is that the equations, if they are to be of practical use, introduce quantities which have no meaning in the ideal conditions and depend on the introduction of some degree of complexity for their physical interpretation. For example, the equations for an unperturbed atom determine quantities which could only have a physical meaning on the assumption that the atom is interacting with its surroundings and therefore not unperturbed. In practice a “simple problem” often means

* The singularities are definite centres of rotation, whereas the unspecified electrons are indistinguishable, so that a system of two electrons with one singularity or "fixed" electron is not necessarily equivalent statistically to one electron and two singularities.
the limit of a more complex problem as the complexity tends to zero. When simplicity is used in this sense, the task of developing a theory by proceeding from the simple to the complex is like developing a theory of number by proceeding from the "simple" conceptions of 0 and \( \infty \).

In our problem of a fixed and a movable electron in a uniform spherical space, no meaning can be given to co-ordinates or distinctions of position other than those indicated by change of \( r \); the equation introduces them in anticipation of approximate application to more complex problems in which space is not so uniform and landmarks less infrequent. We have therefore to distinguish the elements of the problem which are logically simple (and therefore form a suitable starting point for the theory) from those which are simple only as a limiting case of complexity and logically can only be introduced at a later stage. We may regard the problem of which (1) is the solution as gradually taking shape when more and more complexity is introduced; and at each stage we cannot obtain more of the mathematical equation than applies to our partial picture. In the problem stated the first conception to emerge is that of two spherical or hyperspherical loci of radius \( r \) and \( R \) respectively, expressing that our movable particle is at a distance \( r \) from the fixed particle and at a distance \( R \) from the centre of the universe. This geometrical picture goes rather too far, for logically we cannot think of a spherical locus before thinking of positional space, and, as already stated, complete distinction of position does not appear until the next stage of complexity is reached; we should rather aim, as best we can, at a pre-geometrical picture of something ready to become a spherical locus when distinction of position is allowed. (In particular, the question of the number of dimensions of these loci would be premature; the conception of connectivity should, I think, precede that of dimensionality in this order of development.) The fact that \( R \) will ultimately be treated as a natural constant and \( r \) as a variable co-ordinate should be ignored; the difference is only one of degree, for theoretically any alteration of \( r \) would slightly alter the equilibrium of the whole universe and increase or decrease \( R \).

Clearly then we must look on equation (1) as developing out of a bipolar problem in which the two radii vectores are denoted by \( r \) and \( R \); so that a term \( i/r \) must have coupled with it a term \( i/R \). These two terms form the skeleton. We next turn attention to what is supposed to be occupying the two loci. The datum is that just one electron* occupies the sphere of radius \( r \), and \( N \) indistinguishable electrons occupy the sphere of radius \( R \) which

* For if there were another electron within range we should not apply equation (1).
constitutes the whole of space. Here the argument of § 5 finds place indicating that a factor $\sqrt{N}$ is introduced, and the coupled terms become $i/r$ and $i\sqrt{N/R}$.

It is unnecessary for our purpose to proceed further. The next stage would be to introduce space and time co-ordinates (by treating a rather more complicated problem) and to explain, if we can, why the terms involving these are linked to the skeleton terms in the particular way shown in the equation. I do not profess to have solved this. The point which should be realised is that our view that the terms $i/r$ and $i\sqrt{N/R}$ arise in an identical manner is not shaken by the fact that they appear to be inserted rather unsymmetrically in the wave equation. In a logical development they are not inserted; the wave equation is built up round them. As to why the space-time part of the equation is linked to them unsymmetrically—it could scarcely be otherwise if one locus is to be given the current interpretation as the orbit of an electron and the other as all space.

7. When $\psi$ is a simple harmonic function of $t$ so that $\psi$ satisfies the same equation as $\psi$, another form of (1), omitting the fixed electron, is

$$\left(\frac{i\hbar}{2\pi}\right)^2 \frac{\partial^2 \psi}{\partial t^2} = c^2 \left(m^2 c^2 + \frac{(i\hbar)^2}{2\pi} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\right) \psi. \quad (7)$$

Reducing as before, this gives

$$\left\{\alpha^2 \left(\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2}\right) + \gamma^2\right\} \psi = 0. \quad (8)$$

According to our identification

$$\gamma^2 = N/R^2 = N\lambda \quad (9)$$

the cosmical constant $\lambda$ being the inverse square of the radius of the Einstein configuration. Thus (8) becomes

$$(\alpha^2 \Box + N\lambda) \psi = 0. \quad (10)$$

This may be compared with the equation of propagation of electromagnetic potential $\kappa_\mu$ in curved space satisfying $G_{\mu\nu} = \lambda g_{\mu\nu}$, viz.,

$$(\Box + \lambda) \kappa_\mu = 0. \quad (11)$$

This suggests that there is another equivalent treatment of our problem, depending on the construction of a theoretical proof of (10) analogous to the theory of (11). In one respect this would appear simpler, since the differential

properties of curvature are more amenable to analysis than the integral property of connectivity. Moreover (10) and (11) have no special reference to an Einstein world and apply immediately to any state of the universe. But to follow this out we should have to enter into the theory of the factor $\alpha$; the method we have used depends on a comparison in which $\alpha$ is not involved.

Equation (10) tempts me to a digression. My theoretical investigations of $\alpha$ lead me to think that the circuit corresponding to $R$ needs to be diminished in the ratio of the “packing fraction” $\frac{136}{137}$ in order to make it strictly comparable with $\eta$. I have regarded this correction (if real) as negligible for my purposes, and it is not taken account of elsewhere in the paper. But it has the effect that $\alpha$ in (10) should be taken as 136 (not 137). If protons and electrons are alike except in mass, the equation for the proton will be (10) with $\alpha^2 = 10$ very approximately; that is to say, the equations for free electrons and protons will be respectively

$$\begin{align*}
(136^2 \Box + N\lambda) \psi &= 0 \\
(10 \Box + N\lambda) \psi &= 0
\end{align*}$$

(12)

the numbers $136^2$ and 10 arising more or less in the way that I have suggested elsewhere.*

8. By (2) and (5) we have

$$\frac{2\pi m c \alpha}{h} = \frac{\sqrt{N}}{R}.$$  

(12)

The theory of the Einstein universe gives

$$GM_0/e^2 = \frac{1}{2}\pi R,$$

(13)

where $M_0$ is the total mass of the universe† and $G$ the constant of gravitation. Assuming that the number of protons is equal to the number of electrons, we have approximately

$$M_0 = Nm$$

where $m$ is the mass of a proton. Hence

$$N = \frac{1}{2}\pi e^2 R / GM$$

(14)

so that by (12)

$$\left(\frac{2\pi mc \alpha}{h}\right)^2 = \frac{1}{2}\pi e^2 \frac{1}{GM} \cdot \frac{1}{R}.$$  

(15)

Also in the Einstein world $\lambda = 1/R^2$, so that

$$\lambda = \left(\frac{2GM}{\pi}\right)^2 \left(\frac{2\pi mc \alpha}{h}\right)^4 = 9.79 \cdot 10^{-55}.$$  

(16)


† I adopt spherical (not elliptical) space since the treatment of the semi-circuits in elliptical space is not obvious. I think that it must be impossible to decide experimentally between the two forms of space so that both should give the same result.
This gives

\[ R = 1.01 \times 10^{27} \text{ cm.} = 328 \text{ million parsecs.} \]

\[ = 1070 \text{ million light years.} \]

The speed of recession of distant bodies per unit distance is given by \( c/R \sqrt{3} \).

The result for the above value of \( R \) is

\[ 528 \text{ km. per sec. per megaparsec.} \]

This represents the full effect of cosmic repulsion and would be diminished by any countervailing gravitational attraction between the galaxies. So far as can be judged from estimates of the average density of matter in the universe the reduction, though not entirely negligible, is not likely to be very large.

The observational value according to Hubble is 465 km. per sec. per mp. De Sitter finds nearly the same value; but, owing to the uncertainties of the distance-scale and other causes, the astronomical determination could perhaps be in error by as much as 20 per cent.

9. We have also by (12) and (14)

\[ \sqrt{N} = \frac{2\pi mc^2}{\hbar} \cdot \frac{2GMN}{\pi c^2}, \]

or, since \( c = \hbar/2\pi e^2 \),

\[ \frac{c^2}{GMm} = \frac{2}{\pi} \sqrt{N}. \]

(17)

The left side is the ratio of the electrostatic to the gravitational force between a proton and electron. Denoting this ratio by \( F \), we have

\[ N = \left(\frac{1}{2} \pi F\right)^2 = 1.29 \times 10^{79}. \]

(18)

This is the type of relation which I anticipated in an earlier paper, but it is reached mainly in a different way. The present investigation is independent of the theory of the proton.

Another formula equivalent to (16) is

\[ \lambda = (2mc^2/\pi e^2F)^2. \]

(19)

I have considered the alternative that the value of \( \gamma \) might be \( \sqrt{2N/R} \), \( 2N \) being the number of electrons and protons in the universe. I do not think that this would be in keeping with the general conceptions of the theory. We do not mind whether curved macroscopic space replaces \( N \) or \( 2N \) singularities; whatever the number and whatever their properties, the observed dimensions of space will be such as to compensate for their abolition. But we are

concerned with the fact that in order to make use of this replacement we must consider a system in which our electron is one of \( N \) indistinguishable electrons.

On the Value of the Cosmical Constant.

(Abstract.)

It is argued that, owing to the curvature of space, a term \( i\sqrt{N/R} \) should appear in the wave equation for an electron, \( N \) being the number of electrons in the universe and \( R \) the radius of the universe in an eigenstate, and having therefore the dimensions calculated for an Einstein world. The units are here supposed to be such that the electrostatic energy of two electrons at a distance \( r \) is simultaneously represented by a term \( i/r \). The theoretical term is identified with the term ordinarily attributed to the proper-mass of the electron. From this identification we find the cosmical constant \( \lambda = 9.79 \times 10^{-55} \), which gives a speed of recession of the spiral nebulae 528 km. per sec. per megaparsec. The observed speed according to Hubble is 465 km. per sec. per megaparsec.

The Diffraction of Electrons in Gases.

By F. L. Arnot, Ph.D., Lecturer in Natural Philosophy, The University, St. Andrews.

(Communicated by J. Chadwick, F.R.S.—Received August 10, 1931.)

Introduction.

The new wave mechanics has been eminently successful in correlating and accounting for the large amount of experimental data on the periodic properties of the atom. The applications of the new theory to aperiodic phenomena, though not nearly so numerous, have been attended with no less success; indeed, it is in these experiments on free electrons, in which diffraction patterns similar to those produced by beams of X-rays and light are obtained, that the wave nature of electrons is so clearly and objectively demonstrated.

Such diffraction effects have been obtained by a number of investigators* by scattering beams of homogeneous electrons in crystals, thin films and by a