On the Hoyle–Narlikar theory of gravitation

By S. W. Hawking

Department of Applied Mathematics and Theoretical Physics, Cambridge

(*Communicated by H. Bondi, F.R.S.—Received* 8 *October* 1964—
Revised 11 *February* 1965)

It is shown that the direct-particle action-principle from which Hoyle & Narlikar derive their new theory of gravitation not only yields the Einstein field-equations in the 'smooth-fluid' approximation, but also implies that the 'm'-field be given by the sum of half the retarded field and half the advanced field calculated from the world-lines of the particles. This is in effect a boundary condition for the Einstein equations, and it appears that it is incompatible with an expanding universe since the advanced field would be infinite. A possible way of overcoming this difficulty would be to allow the existence of negative mass.

1. Introduction

The success of Maxwell's equations has led to electrodynamics being normally formulated in terms of fields that have degrees of freedom independent of the particles in them. However, Gauss suggested that an action-at-a-distance theory in which the action travelled at a finite velocity might be possible. This idea was developed by Wheeler & Feynman (1945, 1949) who derived their theory from an action-principle that involved only direct interactions between pairs of particles. A feature of this theory was that the 'pseudo'-fields introduced are the half-retarded plus half-advanced fields calculated from the world-lines of the particles. However, Wheeler & Feynman, and in a different way Hogarth (1962), were able to show that, provided certain cosmological conditions were satisfied, these fields could combine to give the observed field. Hoyle & Narlikar (1964*a*) extended the theory to general space-times and obtained similar theories for their 'C'-field (1964*b*) and for the gravitational field (1964*c*). It is with these theories that the present paper is concerned. It will be shown that in an expanding universe the advanced fields are infinite, and the retarded fields finite. This is because, unlike electric charges, all masses have the same sign.

2. Notation

Space-time is represented by a four-dimensional Riemannian space with metric tensor g_{ij} of signature -2. Covariant differentiation in this space is indicated by a semi-colon. Particles are labelled $a, b, \ldots$, and $\mathrm{d}a, \mathrm{d}b, \ldots$ represent the differential of proper time along the world-lines of $a, b, \ldots$ respectively. When there is doubt as to which point a covariant derivative is to be taken at, a suffix will be added to the appropriate indices. The suffix a will indicate covariant differentiation at a point on the world-line of particle a, and so on.

$$\delta^4(X, X') = \delta(X_1 - X_1')\,\delta(X_2 - X_2')\,\delta(X_3 - X_3')\,\delta(X_4 - X_4'),$$

where X_1, X_2, X_3, X_4 are the coordinates of the point X, and $\delta(Y)$ is the Dirac delta-function. The operator $\Box$ is defined by

$$\Box f = g^{ij} f_{;ij} \text{ for any function } f.$$

3. The boundary condition

Hoyle & Narlikar derive their theory from the action

$$A = \sum_{a \neq b}\sum \iint G(a, b)\, \mathrm{d}a\, \mathrm{d}b,$$

where the integration is over the world-lines of particles $a, b, \ldots$. In this expression, G is a Green function that satisfies the wave equation:

$$G(X, X')_{;ij}\, g^{ij} + \tfrac{1}{6} R\, G(X, X') = \frac{\delta^4(X, X')}{\sqrt{-g}},$$

where g is the determinant of g_{ij}. Since the double sum in the action A is symmetrical between all pairs of particles a, b, only that part of $G(a, b)$ that is symmetrical between a and b will contribute to the action A, i.e. the action can be written

$$A = \sum_{a \neq b}\sum \iint G^*(a, b)\, \mathrm{d}a, \mathrm{d}b,$$

where $G^*(a, b) = \frac{1}{2} G(a, b) + \frac{1}{2} G(b, a)$. Thus G^* must be the time-symmetric Green function, and can be written: $G^* = \frac{1}{2} G_{\text{ret.}} + \frac{1}{2} G_{\text{adv.}}$ where $G_{\text{ret.}}$ and $G_{\text{adv.}}$ are the retarded and advanced Green functions. By requiring that the action be stationary under variations of the g_{ij}, Hoyle & Narlikar obtain the field-equations:

$$\begin{aligned}
&\Big[\sum_{a \neq b}\sum \tfrac{1}{6} m^{(a)}(X)\, m^{(b)}(X)\Big]\, (R_{ik} - \tfrac{1}{2} g_{ik} R) \\
&\qquad = -T_{ik} + \sum_{a \neq b}\sum \tfrac{1}{3}\big[m^{(a)}(g_{ik}\, m^{(b)r}_{;r} - m^{(b)}_{;ik}) + 2(m^{(a)}_{;i}\, m^{(b)}_{;k} - \tfrac{1}{4} g_{ik}\, m^{(a);r} m^{(b)}_{;r})\big],
\end{aligned}$$

where $m^{(a)}(x) = \int G^*(x, a)\, \mathrm{d}a$. However, as a consequence of the particular choice of Green function, the contraction of the field-equations is satisfied identically. There are thus only 9 equations for the 10 components of g_{ij}, and the system is indeterminate.

Hoyle & Narlikar therefore impose $\Sigma m^{(a)} = m_0 = \text{constant}$, as the tenth equation. By then making the 'smooth-fluid' approximation, that is by putting $\sum_{a \neq b}\sum m^{(a)} m^{(b)} \simeq m_0^2$, they obtain the Einstein field-equations:

$$\tfrac{1}{6} m_0^2 (R_{ik} - \tfrac{1}{2} R g_{ik}) = -T_{ik}.$$

There is an important difference, however, between these field-equations in the direct-particle interaction theory and in the usual general theory of relativity. In the general theory of relativity, any metric that satisfies the field-equations is admissible, but in the direct-particle interaction theory only those solutions of the field-equations are admissible that satisfy the additional requirement:

$$\begin{aligned}
m_0(x) = \Sigma m^{(a)}(x) &= \Sigma \int G^*(x, a)\, \mathrm{d}a \\
&= \tfrac{1}{2}\, \Sigma \int G_{\text{ret.}}(x, a)\, \mathrm{d}a + \tfrac{1}{2}\, \Sigma \int G_{\text{adv.}}(x, a)\, \mathrm{d}a.
\end{aligned}$$

This requirement is highly restrictive; it will be shown that it is not satisfied for the cosmological solutions of the Einstein field-equations, and it appears that it cannot be satisfied for any models of the universe that either contain an infinite amount of matter or undergo infinite expansion.

The difficulty is similar to that occurring in Newtonian theory when it is recognized that the universe might be infinite.

The Newtonian potential ϕ obeys the equation:

$$\Box\phi = -\kappa\rho \quad (\rho > 0),$$

where ρ is the density.

In an infinite static universe, ϕ would be infinite, since the source always has the same sign. The difficulty was resolved when it was realized that the universe was expanding, since in an expanding universe the retarded solution of the above equation is finite by a sort of 'red shift' effect. The advanced solution will be infinite by a 'blue shift' effect. This is unimportant in Newtonian theory, since one is free to choose the solution of the equation and so may ignore the infinite advanced solution and take simply the finite retarded solution.

Similarly in the direct-particle interaction theory the m-field satisfies the equation:

$$\Box m + \tfrac{1}{6}Rm = N \quad (N > 0),$$

where N is the density of world-lines of particles. As in the Newtonian case, one may expect that the effect of the expansion of the universe will be to make the retarded solution finite and the advanced solution infinite. However, one is now not free to choose the finite retarded solution. For the equation is derived from a direct-particle interaction action-principle symmetric between pairs of particles, and one must choose for m half the sum of the retarded and advanced solutions. We would expect this to be infinite, and this is shown to be so in the next section.

4. The cosmological solutions

The Robertson–Walker cosmological metrics have the form

$$\mathrm{d}s^2 = \mathrm{d}t^2 - R^2(t)\left[\frac{\mathrm{d}r^2}{1-Kr^2} + r^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\right].$$

Since they are conformally flat, one can choose coordinates in which they become

$$\begin{aligned}\mathrm{d}s^2 &= \Omega^2[\mathrm{d}\tau^2 - \mathrm{d}\rho^2 + \rho^2\,\mathrm{d}\theta^2 + \rho^2\sin^2\theta\,\mathrm{d}\phi^2],\\ &= \Omega^2\eta_{ab}\,\mathrm{d}x^a\,\mathrm{d}x^b\end{aligned}$$

where η_{ab} is the flat-space metric tensor and

$$\Omega = \Omega(\tau,\rho) = \frac{R(t)}{\sqrt{\{[1+\frac{1}{4}K(\tau+\rho)^2][1+\frac{1}{4}K(\tau-\rho)^2]\}}}$$

(cf. Infeld & Schild 1945).

For example, for the Einstein–de Sitter universe

$$\begin{aligned}&K = 0, \quad R(t) = (t/T)^{\frac{2}{3}} \quad (0 < t < \infty),\\ &\Omega = R = (\tau/T)^2 \quad (0 < \tau < \infty),\\ &r = \rho \quad (\tau = T^{\frac{2}{3}}t^{\frac{1}{3}}).\end{aligned}$$

For the steady-state (de Sitter) universe

$$K = 0, \quad R(t) = \mathrm{e}\, t/T \quad (-\infty < t < \infty),$$
$$\Omega = R = -T/\tau \quad (-\infty < \tau < 0),$$
$$r = \rho \quad (\tau = -T\,\mathrm{e}^{-t/T}).$$

The Green function $G^*(a, b)$ obeys the equation

$$\Box G^*(a,b) + \tfrac{1}{6} R G^*(a,b) = \delta^4(a,b)/\sqrt{-g}.$$

From this it follows that

$$\frac{1}{\Omega^4}\frac{\partial}{\partial x^a}\left(\Omega^2 \eta^{ab}\frac{\partial}{\partial x^b} G^*\right) + \frac{\partial}{\partial x^a}\left(\eta^{ab}\frac{\partial}{\partial x^b}\Omega\right)\Omega^{-3} G^* = \Omega^{-4}\delta^4(a,b).$$

If we let $G^* = \Omega^{-1} S$, then

$$\Omega \frac{\partial}{\partial x^a}\left(\eta^{ab}\frac{\partial}{\partial x^b} S\right) = \delta^4(a,b).$$

This is simply the flat-space Green function equation, and hence

$$G^*(\tau_1, 0; \tau_2, \rho) = \frac{\Omega^{-1}(\tau_1)}{8\pi}\left[\frac{\delta(\rho - \tau_2 + \tau_1)}{\Omega(\tau_2)\,\rho} + \frac{\delta(\rho + \tau_2 - \tau_1)}{\Omega(\tau_2)\,\rho}\right],$$

The 'm'-field is given by

$$m(\tau_1) = \int G^* N \sqrt{-g}\,\mathrm{d}x^4 = \tfrac{1}{2}(m_{\text{ret.}} + m_{\text{adv.}}).$$

For universes without creation (e.g. the Einstein–de Sitter universe), $N = R^{-3} n$, $n = \text{const}$. For universes with creation (steady state) $N = n$, $n = \text{const.}$,

$$m_{\text{adv.}}(\tau_1) = \Omega^{-1}(\tau_1)\int \frac{N\Omega^3(\tau_2)}{4\pi r} 4\pi r^2\,\mathrm{d}r,$$

where the integration is over the future light cone. This will normally be infinite in an expanding universe, e.g. in the Einstein–de Sitter universe

$$m_{\text{adv.}}(\tau_1) = \left(\frac{\tau_1}{T}\right)^{-2}\int_{\tau_1}^{\infty} n(\tau_2 - \tau_1)\,\mathrm{d}\tau_2$$
$$= \infty.$$

In the steady-state universe

$$m_{\text{adv.}}(\tau_1) = \left(\frac{-T}{\tau_1}\right)^{-1}\int_{\tau_1}^{0} -n\left(\frac{T}{\tau_2}\right)^{3}(\tau_2 - \tau_1)\,\mathrm{d}\tau_2.$$
$$= \infty.$$

By contrast, on the other hand, we have

$$m_{\text{ret.}}(\tau_1) = \Omega^{-1}(\tau_1)\int \frac{N\Omega^3}{4\pi r} 4\pi r^2\,\mathrm{d}r,$$

where the integration is over the past light cone. This will normally be finite, e.g. in the Einstein–de Sitter universe

$$m_{\text{ret.}}(\tau_1) = \left(\frac{\tau_1}{T}\right)^{-2}\int_{0}^{\tau_1} -n(\tau_2 - \tau_1)\,\mathrm{d}\tau_2 = \tfrac{1}{2} n T^2,$$

while in the steady-state universe

$$m_{\text{ret.}}(\tau_1) = -\left(\frac{-T}{\tau_1}\right)^{-1}\int_{-\infty}^{\tau_1} n\left(\frac{T}{\tau_2}\right)^3(\tau_2-\tau_1)\,\mathrm{d}\tau_2 = \tfrac{1}{2}nT^2.$$

Thus it can be seen that the solution m = const. of the equation

$$\Box m + \tfrac{1}{6}Rm = N$$

is not, in a cosmological metric, the half-advanced plus half-retarded solution since this would be infinite. In fact, in the case of the Einstein-de Sitter and steady-state metrics, it is the pure retarded solution.

5. Conclusion

It is one of the weaknesses of the Einstein theory of relativity that although it furnishes field-equations it does not provide boundary conditions for them. Thus it does not give a unique model for the universe but allows a whole series of models. Clearly a theory that provided boundary conditions and thus restricted the possible solutions would be very attractive. The Hoyle–Narlikar theory does just that (the requirement that $m = \frac{1}{2}m_{\text{ret.}} + \frac{1}{2}m_{\text{adv.}}$ is equivalent to a boundary condition). Unfortunately, as we have seen above, this condition excludes those models that seem to correspond to the actual universe, namely the Robertson–Walker models.

The calculations given above have considered the universe as being filled with a uniform distribution of matter. This is legitimate if we are able to make the 'smooth-fluid' approximation to obtain the Einstein equations. Alternatively if this approximation is invalid, it cannot be said that the theory yields the Einstein equations.

It might possibly be that local irregularities could make $m_{\text{adv.}}$ finite, but this has certainly not been demonstrated and seems unlikely in view of the fact that, in the Hoyle–Narlikar direct-particle interaction theory of their 'C'-field, which is derived from a very similar action-principle, it can be shown without assuming a smooth distribution that the advanced 'C' field will be infinite in an expanding universe with creation (see Appendix).

The reason that it is possible to formulate a direct-particle interaction theory of electrodynamics that does not encounter this difficulty of having the advanced solution infinite is that in electrodynamics there are equal numbers of sources of positive and negative sign. Their fields can cancel each other out and the total field can be zero apart from local irregularities. This suggests that a possible way to save the Hoyle–Narlikar theory would be to allow masses of both positive and negative sign. The action would be

$$A = \sum_{a\neq b}\sum q_a q_b \iint G^*(a,b)\,\mathrm{d}a\,\mathrm{d}b \quad (q_a, q_b = \pm 1),$$

where q_a, q_b are gravitational charges analogous to electric charges. Particles of positive q in a positive 'm'-field and particles of negative q in a negative 'm'-field would have the normal gravitational properties, that is, they would have positive gravitational and inertial masses. A particle of negative q in a positive 'm'-field would still follow a geodesic. Therefore it would be attracted by a particle of positive q. Its own gravitational effect however would be to repel all other particles. Thus it would have the properties of the negative mass described by Bondi (1957); that is, negative gravitational mass and negative inertial mass.

Since there does not seem to be any matter having these properties in our region of space (where $m \doteqdot$ const. > 0), there must clearly be separation on a very large scale. It would not be possible to identify particles of negative q with antimatter, since it is known that antimatter has positive inertial mass. However, the introduction of negative masses would probably raise more difficulties than it would solve.

The author would like to thank Professor F. Hoyle, F.R.S. and Dr J. V. Narlikar for making available the manuscripts of their papers and for discussions on them, and also to thank Dr D. W. Sciama for his help in preparing this paper.

Appendix. The 'C'-field

Hoyle & Narlikar derive their direct-particle interaction theory of the 'C'-field from the action

$$A = \sum_{a \neq b}\sum \iint \hat{G}(a,b)_{;i_a k_b}\, \mathrm{d}a^i\, \mathrm{d}b^k,$$

where the suffixes a, b refer to differentiation of $\hat{G}(a,b)$ on the world-lines of a, b respectively. $\hat{G}$ is a Green function obeying the equation

$$\Box \hat{G}(X, X') = \delta^4(X, X')/\sqrt{-g}.$$

We define the 'C'-field by

$$C(x) = \Sigma \int \hat{G}(x,a)_{;i_a}\, \mathrm{d}a^i,$$

and the matter-current J^k by

$$J^k(y) = \Sigma \int \delta^4(y,b)\, \mathrm{d}b^k.$$

Then

$$C(x) = \int \hat{G}(x,y)\, J^k(y)_{;k} \sqrt{-g}\, \mathrm{d}x^4,$$

$$\Box C = J^k{}_{;k}.$$

We thus see that the sources of the 'C'-field are the places where matter is created, or destroyed.

As in the case of the 'm'-field, the Green function G must be time-symmetric, that is

$$\hat{G}(a,b) = \tfrac{1}{2}\hat{G}_{\text{ret.}}(a,b) + \tfrac{1}{2}\hat{G}_{\text{adv.}}(a,b).$$

Hoyle & Narlikar claim that if the action of the 'C'-field is included along with the action of the 'm'-field, a universe will be obtained that approximates to the steady-state universe on a large scale although there may be local irregularities. In this universe, the value of C will be finite and its gradient time-like and of unit magnitude.

Given this universe, we may check it for consistency by calculating the advanced and retarded 'C'-fields and finding if their sum is finite. We shall not do this directly but will show that the advanced field is infinite while the retarded field is finite.

Consider a region in space-time bounded by a three-dimensional space-like hypersurface D at the present time, and the past light cone Σ of some point P to the future of D.

By Gauss's theorem

$$\int_V \Box C\sqrt{-g}\,\mathrm{d}x^4 = \int_{\Sigma+D}\frac{\partial C}{\partial n}\,\mathrm{d}S = \int J^k_{;k}\sqrt{-g}\,\mathrm{d}x^4.$$

Let the advanced field produced by sources within V be C'. Then C' and $\partial C'/\partial n$ will be zero on Σ, and hence

$$\int_V J^k_{;k}\sqrt{-g}\,\mathrm{d}x^4 = \int_D \frac{\partial C'}{\partial n}\,\mathrm{d}S.$$

But $J^k_{;k}$ is the rate of creation of matter $= n$ (const.) in the steady-state universe, and hence

$$\int_D \frac{\partial C'}{\partial n}\,\mathrm{d}S = nV.$$

As the point P is taken further into the future, the volume of the region V tends to infinity. However, the area of the hypersurface D tends to a finite limit owing to horizon effects. Therefore the gradient $\partial C'/\partial n$ must be infinite. A similar calculation shows the gradient of the retarded field to be finite. Their sums cannot therefore give the field of unit gradient required by the Hoyle–Narlikar theory.

It is worth noting that this result was obtained without assumptions of a smooth distribution of matter or of conformal flatness.

References

Bondi, H. 1957 *Rev. Mod. Phys.* **29**, 423.
Hogarth, J. E. 1962 *Proc. Roy. Soc.* A, **267**, 365.
Hoyle, F. & Narlikar, J. V. 1964*a* *Proc. Roy. Soc.* A, **277**, 1.
Hoyle, F. & Narlikar, J. V. 1964*b* *Proc. Roy. Soc.* A, **282**, 178.
Hoyle, F. & Narlikar, J. V. 1964*c* *Proc. Roy. Soc.* A, **282**, 191.
Infeld, L. & Schild, A. 1945 *Phys. Rev.* **68**, 250.
Wheeler, J. A. & Feynman, R. P. 1945 *Rev. Mod. Phys.* **17**, 157.
Wheeler, J. A. & Feynman, R. P. 1949 *Rev. Mod. Phys.* **21**, 425.