Nonsingular general-relativistic cosmologies*

Jacob D. Bekenstein

Center for Relativity Theory, Department of Physics, University of Texas at Austin, Austin, Texas 78712 and Department of Physics, Ben-Gurion University of the Negev, Beer-Sheva 84120, Israel[†]

(Received 21 August 1974)

According to the "singularity theorems" of Penrose, Hawking, and Geroch, all general-relativistic cosmological models must have a singularity. However, the energy condition assumed by the theorems is not satisfied by all known forms of matter. (A notable exception is the massive Klein-Gordon field.) Our object here is to construct exact isotropic cosmological models without singularities by exploiting a violation of the energy condition which arises naturally from the basic physics, rather than being introduced *ad hoc* via an equation of state. We accomplish this with models in which the matter, envisaged as dust, interacts with a conformal scalar field whose field equation and stress-energy tensor come from an action principle. All the equations that govern the evolution of the models are solved exactly. For all possible topologies of the universe, the singularity can be avoided for a certain range of the parameters of the model, but only for the closed universe is the required range physically appealing.

I. INTRODUCTION

The "singularity theorems" of Penrose, Hawking, and Geroch[1] show that any general-relativistic model universe (a) satisfying reasonable causality and generality conditions, (b) possessing a surface to the past (future) of which the light cones start converging, and (c) containing matter which satisfies the energy condition

$$(T_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}T)t^\mu t^\nu \geq 0 \tag{1}$$

for any unit timelike vector t^μ must contain a timelike curve with a past (future) end point at finite proper time (a singularity). In most known examples this end point is associated with an infinite curvature singularity, but an exception exists.[2] The actual universe is expected to satisfy the causality and generality requirements, conditions for the existence in it of a surface of convergence of the past light cones are satisfied,[3] and its matter content is generally regarded as satisfying the energy condition. Does this mean that the universe actually passed through a singularity in the past, either an infinite curvature singularity (big bang) or a mild singularity[4] associated with the origin of matter particles?

Some workers are resigned to the singularity[5] and try to make it palatable by conjecturing that it is of the mild variety,[4] or else by relegating the big-bang singularity to the "infinite past" according to a physically appealing time scale.[6] Others contend that the cosmological singularity is only a symptom of the incompleteness of the theory, and they propose that it is absent in a more fundamental theory in which Einstein's equations are appropriately modified by quantum effects,[7] phenomenological quadratic terms in the curvature,[8] or the effects of torsion.[9]

Before taking the serious steps of accepting the cosmological singularity as real, or modifying the gravitational field equations, it seems timely to rediscuss the fundamental assumption of the theorems that the energy condition (1) is satisfied by all physically reasonable forms of matter. It is known that once the energy condition is given up, singularity-free cosmologies become possible. One example has been provided by Murphy,[10] who postulates matter with sufficiently large second viscosity to violate condition (1) but does not try to justify this property with a microscopic model. Another example is the closed Friedmann model of Fulling and Parker[11] in which the singularity can be prevented because the energy condition is violated in a natural way be a massive scalar field in certain coherent quantum states. It is not yet clear whether the singularity is avoided in every cycle of this model universe, or whether the quantum states envisaged are realistic.

It may not even be necessary to resort to quantum effects to violate the energy condition and prevent the singularity. For example, the stress-energy tensor for a *classical* massive scalar field ϕ with Compton wavelength m^{-1} is

$$T_{\mu\nu}=\phi_{,\mu}\phi_{,\nu}-\tfrac{1}{2}g_{\mu\nu}(\phi_{,\alpha}\phi^{,\alpha}+m^2\phi^2)\,. \tag{2}$$

We have

$$(T_{\mu\nu}-\tfrac{1}{2}g_{\mu\nu}T)t^\mu t^\nu=(\phi_{,\mu}t^\mu)^2-\tfrac{1}{2}m^2\phi^2\,, \tag{3}$$

which may be negative in violation of the energy condition. This is very significant because the strong interactions in nuclear matter can be regarded as mediated by a classical massive scalar field, the non-second-quantized part of the pion field.[12] Because of (3) it is no longer clear that nuclear matter in the dense stages of our universe (where the strong interactions are important) always obeys the energy condition. Therefore, it is no longer possible to state categorically that the cosmological singularity is unavoidable; judgment must await concrete model calculations which allow in detail for the effects of the strong interactions.

We intend nothing so complex here. Rather we study model isotropic universes whose matter contents mimic that of the early universe, but which can be solved exactly. They contain incoherent radiation, and pressureless matter coupled to a classical conformal massless scalar field after the manner of the standard pion-nucleon coupling. This field can also violate the energy condition; it is the closest analog to the pion field for which we have been able to find exact solutions. We find the solutions for all three possible topologies of the universe, and in every case the singularity can be prevented for a certain range of the parameters of the model. But only for the closed universe is the necessary range physically compelling.

II. THE COSMOLOGICAL EQUATIONS

Consider a point particle of rest mass μ coupled to a conformal scalar field ψ with coupling strength f. As the action we take

$$S=-\tfrac{1}{2}\int(\psi_{,\alpha}\psi^{,\alpha}+\tfrac{1}{6}R\psi^2)(-g)^{1/2}d^4x-\int(\mu+f\psi)d\tau\ , \tag{4}$$

where R is the curvature scalar and $d\tau$ is the element of proper time of the particle. The first part of S is the action for a free conformal scalar field,[13] the term proportional to μ is the familiar action for a free point particle, and the coupling term, which may be written

$$-f\int\psi d\tau = -f\int d^4x\left[(-g)^{1/2}\psi\int(-g)^{-1/2}\delta^4(x^\mu-x^\mu(\tau))d\tau\right]\ ,$$

is a classical analog of the standard pion-nucleon coupling. The integral over $d\tau$, a three-dimensional δ function, plays the role of the "density" $(\bar{\Psi}\Psi)$ of the nucleon idealized as a point.

Under a conformal mapping $g_{\mu\nu}\to g_{\mu\nu}\Omega^2$ and $\psi\to\psi\Omega^{-1}$ (arbitrary function Ω), both the free-field action[13] and the coupling action are left unchanged. Hence the scalar equation obtained by varying ψ in S,

$$\psi_{,\alpha}{}^{;\alpha}-\tfrac{1}{6}R\psi=f\int(-g)^{-1/2}\delta^4(x^\mu-x^\mu(\tau))d\tau\ , \tag{5}$$

is conformally invariant. Variation of S with respect to $g^{\mu\nu}$ gives the stress-energy tensor

$$\begin{aligned}T_{\mu\nu}&=\psi_{,\mu}\psi_{,\nu}-\tfrac{1}{2}g_{\mu\nu}\psi_{,\alpha}\psi^{,\alpha}-\tfrac{1}{6}\psi^2{}_{,\mu;\nu}+\tfrac{1}{6}\psi^2{}_{,\alpha}{}^{;\alpha}g_{\mu\nu}\\&\quad+\tfrac{1}{6}G_{\mu\nu}\psi^2\\&\quad+(\mu+f\psi)u_\mu u_\nu\int(-g)^{-1/2}\delta^4(x^\mu-x^\mu(\tau))d\tau\ ,\end{aligned} \tag{6}$$

where $u^\mu=dx^\mu/d\tau$ is the 4-velocity of the particle, and $G_{\mu\nu}$ is the Einstein tensor. We may add to the above stress-energy that of any radiation present.

Now consider a Robertson-Walker model universe containing a uniform distribution of identical particles of the kind considered above, plus isotropic radiation. The metric can be written as

$$ds^2=a(\eta)^2[-d\eta^2+(1+\tfrac{1}{4}kr^2)^{-2}(dr^2+r^2d\theta^2+r^2\sin^2\theta d\phi^2)] \tag{7}$$

for a closed $(k=+1)$, flat $(k=0)$, or open $(k=-1)$ universe. For this metric

$$R=6a^{-3}(a_{,\eta\eta}+ka)\ . \tag{8}$$

For simplicity we assume that the particles in question are each at fixed r, θ, and ϕ coordinates. In the continuum approximation we may then replace the integral over $d\tau$ in (5) and (6) by the number density of particles n. Assuming that the field ψ is uniform, and introducing the notations $F=\psi a$ and $N=na^3$, we may, with the help of (8), reduce the wave equation (5) to

$$F_{,\eta\eta}+kF=-fN\ . \tag{9}$$

We assume that the particle number is conserved; in this case N is a constant.

We take into account the radiation by adding to $T_{\mu\nu}$ in (6) the term

$$Ba^{-4}(\tfrac{4}{3}u_\mu u_\nu+\tfrac{1}{3}g_{\mu\nu})\ ,$$

where B, a measure of the amount of radiation, is taken as a constant on the assumption of an adiabatic expansion (or contraction). It is understood that $u^\theta=u^\phi=u^r=0$ and $u^\eta=a^{-1}$ both for the radiation term and for the particle term. The trace of the *total* stress-energy tensor is

$$T_\alpha{}^\alpha=\psi\psi_{,\alpha}{}^{;\alpha}-\tfrac{1}{6}R\psi^2-(\mu+f\psi)n\ , \tag{10}$$

which reduces to $T_\alpha{}^\alpha=-\mu n$ by virtue of the wave equation (5). But the trace of Einstein's equations gives $R=-8\pi T_\alpha{}^\alpha$, so that (8) can now be rewritten as

$$a_{,\eta\eta}+ka=\frac{4\pi}{3}N\mu\ . \tag{11}$$

Finally, for the metric (7),

$$G_\eta{}^\eta = -\frac{3}{8\pi}\,(a_{,\eta}{}^2 + ka^2)\,. \tag{12}$$

The corresponding Einstein equation is

$$a_{,\eta}{}^2 + ka^2 = \frac{4\pi}{3}(F_{,\eta}{}^2 + kF^2 + 2Na\mu + 2NfF + 2B)\,. \tag{13}$$

Equations (9), (11) and (13) are the complete set of equations for the problem.

III. SOLUTIONS

The equations are solved as follows: First find the solution of (9) and (11) in terms of four constants of integration. In some cases one constant can be eliminated by an appropriate choice of the zero of η without loss of generality. Substitution of $a(\eta)$ and $F(\eta)$ into (13) then fixes another of the constants. Two or three are left free; denoting these by J, K, and L we have the following:

$k=+1$.

$$\psi = (-fN + J\sin\eta + K\cos\eta)a^{-1}\,, \tag{14}$$

$$a = \frac{4\pi}{3}N\mu\left\{1 + \left[1 + \frac{3}{4\pi}\,(N\mu)^{-2}(J^2+K^2+2B-N^2f^2)\right]^{1/2} \times\sin\eta\right\}\,;$$

$k=0$.

$$\psi = (-\tfrac{1}{2}fN\eta^2 + J\eta + K)a^{-1}\,,$$

$$a = \frac{2\pi}{3}\,N\mu\eta^2 + \left[\frac{8\pi}{3}\,(B + NfK + N\mu L + \tfrac{1}{2}J^2)\right]^{1/2}\eta + L\,; \tag{15}$$

$k=-1$ (type a).

$$\psi = (fN + J\sinh\eta + K\cosh\eta)a^{-1}\,,$$

$$a = \frac{4\pi}{3}N\mu\left\{-1 + \left[1 + \frac{3}{4\pi}(N\mu)^{-2}(K^2 - J^2 - N^2f^2 - 2B)\right]^{1/2} \times\cosh\eta\right\}\,; \tag{16}$$

$k=-1$ (type b).

$$\psi = (fN + J\sinh\eta + K\cosh\eta)a^{-1}\,,$$

$$a = \frac{4\pi}{3}N\mu\left\{-1 + \left[1 + \frac{3}{4\pi}(N\mu)^{-2}(J^2 - K^2 + N^2f^2 + 2B)\right]^{1/2} \times\sinh\eta\right\}\,; \tag{17}$$

$k=-1$ (type c).

$$\psi = \left[fN + Je^{\eta} + \frac{1}{4J}\left(2B + f^2N^2 - \frac{4\pi}{3}\,N^2\mu^2\right)e^{-\eta}\right]\,,$$

$$a = \frac{4\pi}{3}\,N\mu(-1 + Le^{\pm\eta})\,. \tag{18}$$

These solutions generalize those we found earlier[14] for the special models with no particles ($N \to 0$). In contrast with that case, in this the universe can always be made to "bounce" [$a(\eta)$ having a positive minimum] for all topologies. The conditions for a bounce are

$$\begin{aligned} &k=+1, && N^2f^2 > 2B + J^2 + K^2,\\ &k=0, && KNf < -B - \frac{1}{2}J^2,\\ &k=-1 \text{ (type } a), && K^2 > J^2 + N^2f^2 + 2B\,. \end{aligned} \tag{19}$$

On physical grounds we would like the scalar field to be generated entirely by the particles themselves (mathematically, the scalar field should be the particular solution of the inhomogeneous wave equation with no admixture of solutions to the homogeneous one). We achieve this by setting $J=K=L=0$; we find that now only the $k=+1$ universe can bounce provided

$$N^2f^2 > 2B\,. \tag{20}$$

Only in this case is the removal of the singularity physically appealing.

IV. DISCUSSION

In field theory the conformal scalar field is generally regarded as physically reasonable (if not necessarily existing in nature). The coupling with matter we have assumed is the most natural one. Furthermore, even though the scalar field can have negative energy density in some circumstances,[14] in the case of interest here [$k=+1$ with (20) satisfied] the total energy density $a^{-2}T_{\eta\eta}$ is positive. Thus we conclude that there exist physically reasonable classical cosmologies which are singularity-free as a result of violation of the energy condition assumed in the singularity theorems. Of course, since no massless scalar field is known to exist, the above has no immediate bearing on the actual universe. But our results do underline a point of principle, namely, that singularities in cosmology are not compulsory if certain subtle features, such as interactions, are taken into account.[15] It would appear, then, that only model calculations taking into account in detail the relevant physics of the matter in our universe will be able to clear up the question of whether there really was a big-bang singularity or not.

There are, however, qualifications to be made about our singularity-free model universe. The attractive scalar force between particles must, in any realistic situation, lead to the development of inhomogeneities (clumping of the particles). It is not clear whether the bounce is stable with respect to such inhomogeneous perturbations of our homogeneous model. In addition, in any realistic model entropy will be generated (we have ignored this effect) and, in the final analysis, it will manifest itself in an increase in the amount of radiation, i.e., in an increase in B. It is not clear whether under these circumstances the condition for a bounce, (20), will continue to hold indefinitely, or whether a collapse to a singularity will ensue after a finite number of cycles of the universe.

---

*Work partially supported by the N.S.F. under Grant No. GP 32039 to the University of Texas at Austin.

†Permanent address.

[1]R. Penrose and S. W. Hawking, Proc. R. Soc. A314, 529 (1970); R. P. Geroch, Phys. Rev. Lett. 17, 445 (1966).

[2]See review in M. Ryan and L. Shepley, *Cosmology* (Princeton Univ. Press, Princeton, N. J., 1974).

[3]S. W. Hawking and G. F. R. Ellis, Astrophys. J. 152, 25 (1968).

[4]G. F. R. Ellis and A. R. King, Commun. Math. Phys. 38, 119 (1974).

[5]For example, I. D. Novikov and Ya. B. Zel'dovich, Annu. Rev. Astron. Astrophys. 11, 387 (1974).

[6]C. W. Misner, Phys. Rev. 186, 1328 (1969).

[7]J. A. Wheeler in *Relativity, Groups and Topology*, edited by B. S. De Witt and C. M. De Witt (Gordon and Breach, New York, 1964); V. L. Ginzburg, Comments Astrophys. Space Phys. 3, 7 (1971).

[8]H. Nariai and K. Tomita, Prog. Theor. Phys. 46, 776 (1971).

[9]A. Trautman, Nat. Phys. Sci. 242, 7 (1973); J. Taffel, Phys. Lett. 45A, 341 (1973). However, see J. Stewart and P. Hájiček, Nat. Phys. Sci. 244, 96 (1973).

[10]G. L. Murphy, Phys. Rev. D 8, 4231 (1973).

[11]S. Fulling and L. Parker, Phys. Rev. D 7, 2357 (1973).

[12]Ya. B. Zel'dovich and I. D. Novikov, *Relativistic Astrophysics* (Univ. of Chicago Press, Chicago, 1971), Vol. I, p. 78.

[13]L. Parker, Phys. Rev. D 7, 976 (1973).

[14]J. D. Bekenstein, Ann. Phys. (N.Y.) 82, 535 (1974).

[15]L. Halpern (private communication) has stressed the same point in connection with the interaction of neutrino fields with matter.