LETTERS TO NATURE

PHYSICAL SCIENCES

The Universe as a Black Hole

SINCE Einstein applied his general theory of relativity to study the structure of the universe as a whole[1], cosmologists have wondered if the universe is geometrically closed or open. Neither theory nor observation has been able to settle this question unambiguously. Several authors have hoped that the universe may after all be a closed, yet unbounded, system. This would solve many problems regarding the nature and origin of the universe, and would fit many of the observations of distant sources made at radio, optical and other wavelengths[2]. Here I demonstrate that the universe may not only be a closed structure (as perceived by its inhabitants at the present epoch) but may also be a black hole, confined to a localized region of space which cannot expand without limit.

An object of mass M is a black hole if it is confined to a region of space bounded by a surface of area $4\pi R_S^2$, where R_S is the Schwarzschild radius of the object. By definition, the Schwarzschild radius is characterized by the vanishing of the element g_{44} of the metric tensor g_{ik}, and is (ordinarily) given by $2GM/c^2$. For a typical star, R_S is of the order of 10^5–10^6 cm; for the universe as a whole, R_S would be of the order of 10^{28} cm. Because the linear dimensions associated with the universe are also of the order of 10^{28} cm, the question arises: Is the universe itself a black hole? To investigate this question, the customary view of the universe, which is necessarily internal, is not sufficient; it has to be supplemented with an external view—I assume that there exists, outside our universe, an external world from which one may take a "detached" look at our universe. It turns out that these two views are not only mutually compatible but also show considerable similarities. Both views are presented here.

For studying the space-time structure of the universe from inside, we regard it as a smeared-out system with uniform mass density $\rho(t)$ and radius of curvature $R(t)$. The geometry of the space-time, in a comoving frame of reference, is governed by the Robertson-Walker metric[3,4]

$$ds^2 = \{R(t)\}^2 \frac{dx^\alpha dx^\alpha}{\left\{1 + \frac{\kappa}{4}(x^\alpha x^\alpha)\right\}^2} - c^2\,dt^2 \quad (\alpha = 1,2,3) \quad (1)$$

where κ ($= +1$, 0 or -1) is the index of curvature. $R(t)$ is a function of the pressure p and density ρ of the universe and is determined by Einstein's field equations. For $p \ll \rho c^2$

$$\left(\frac{dR}{dt}\right)^2 = c^2\left(\frac{1}{3}\Lambda R^2 + \frac{C}{R} - \kappa\right) \quad (2)$$

where Λ is the cosmological constant and C is the mass constant of the universe

$$C = 8\pi G(\rho R^3)/3c^2 \quad (3)$$

Equation (2) leads to a variety of cosmological models[4,5]. I wish to concentrate on the case $\kappa = +1$ which corresponds to a space of positive curvature—the three-dimensional analogue of the surface of a sphere. In this case

$$\left(\frac{dR}{dt}\right)^2 = c^2\left(\frac{1}{3}\Lambda R^2 + \frac{C}{R} - 1\right) \geq 0 \quad (4)$$

At the present epoch the universe is in a phase of expansion and the rate of expansion seems to be on the decrease; accordingly

$$H_0 = \left(\frac{\dot{R}}{R}\right)_{t=t_0} = \frac{c}{R_0}\left(\frac{1}{3}\Lambda R_0^2 + \frac{C}{R_0} - 1\right)^{1/2} > 0 \quad (5)$$

and

$$q_0 = \left(-\frac{\ddot{R}R}{\dot{R}^2}\right)_{t=t_0} = -\frac{\left(\frac{1}{3}\Lambda R_0^2 - \frac{C}{2R_0}\right)}{\left(\frac{1}{3}\Lambda R_0^2 + \frac{C}{R_0} - 1\right)} > 0 \quad (6)$$

Here, H_0 and q_0 denote the present-day values of the Hubble parameter H and the deceleration parameter q. Equations (4)–(6) imply that at a later epoch the function $R(t)$ may require a maximal value R_{max}, which will be determined by the conditions

$$\left(\frac{1}{3}\Lambda R_{max}^2 + \frac{C}{R_{max}} - 1\right) = 0 \quad (7)$$

and

$$\left(\frac{2}{3}\Lambda R_{max} - \frac{C}{R_{max}^2}\right) < 0 \quad (8)$$

Several possibilities arise here:

$\Lambda \leq 0$. In this case we have only one real solution

$$R_{max} = \frac{2}{|\Lambda|^{1/2}} \sinh\left\{\frac{1}{3}\sinh^{-1}\left(\frac{|\Lambda|}{\Lambda_c}\right)^{1/2}\right\} \quad \left(\Lambda_c = \frac{4}{9C^2}\right) \quad (9)$$

which decreases monotonically from the value $R_{max} = C$ to $R_{max} = 0$ as Λ decreases from 0 to $-\infty$.

$0 \leq \Lambda \leq \Lambda_c$. In this case all three solutions are real

$$R_{max} = \frac{2}{\Lambda^{1/2}} \cos\left\{\left(\psi + \frac{\pi}{6}\right), \left(\psi + \frac{5\pi}{6}\right), \left(\psi + \frac{9\pi}{6}\right)\right\} \quad (10a,b,c)$$

where

$$\psi = \frac{1}{3}\sin^{-1}(\Lambda/\Lambda_c)^{1/2} \quad (0 \leq \psi \leq \pi/6) \quad (11)$$

Solution (10*b*) is negative, and may be rejected at once. Solution (10*a*) corresponds to a minimum rather than a maximum; so it is of no interest at present. The only relevant solution in this case is (10*c*), which may also be written as

$$R_{max} = \frac{2}{\Lambda^{1/2}} \sin\left\{\frac{1}{3}\sin^{-1}\left(\frac{\Lambda}{\Lambda_c}\right)^{1/2}\right\} \quad (12)$$

It increases monotonically from the value $R_{max} = C$ to $R_{max} = \frac{3}{2}C$ as Λ increases from 0 to Λ_c.

$\Lambda > \Lambda_c$. In this case again there is only one real solution, but it is always negative. Accordingly, $R(t)$ never acquires a maximal value.

I now consider an external view of the universe. Because the expansion of the universe is isotropic, space-time in its exterior will be governed by a static, Schwarzschild metric (modified by the presence of the cosmological term[5,7])

$$ds^2 = \frac{dr^2}{1 - \frac{2\mu}{r} - \frac{1}{3}\Lambda r^2} + r^2(d\theta^2 + \sin^2\theta\,d\varphi^2) - \left(1 - \frac{2\mu}{r} - \frac{1}{3}\Lambda r^2\right)c^2\,dt^2 \quad (13)$$

where

$$\mu = \frac{4\pi G}{c^2}\int_0^R \rho(r)\, r^2\, \mathrm{d}r = \frac{4\pi G}{3c^2}(\rho R^3) \qquad (14)$$

The element g_{44} of the metric (13) is given by the expression

$$\left(-1 + \frac{C}{r} + \frac{1}{3}\Lambda r^2\right)$$

where C is the same constant as in equation (3). The Schwarzschild radius of the universe is therefore determined by the conditions

$$g_{44}\left(r \gtrless R_S\right) \lessgtr 0 \qquad (15)$$

which imply that

$$\left(-1 + \frac{C}{R_S} + \frac{1}{3}\Lambda R_S^2\right) = 0 \qquad (16)$$

and

$$\left(\frac{\partial g_{44}}{\partial r}\right)_{r=R_S} = \left(-\frac{C}{R_S^2} + \frac{2}{3}\Lambda R_S\right) < 0 \qquad (17)$$

Conditions (16) and (17) for R_S are identical with conditions (7) and (8) for R_{max}. This is true only for $\kappa = +1$, which therefore seems the natural choice. Solutions for R_S are therefore identically the same as for R_{max}—irrespective of the values of the parameters C and Λ, except that for $\Lambda > \Lambda_c$ neither R_S nor R_{max} exists. Now, the Schwarzschild radius of the universe is obtained from the static metric in the exterior of the universe while the maximal value of the function $R(t)$ was obtained from the nonstatic metric in the interior of the universe. The fact that the two, whenever they exist, are identically equal can hardly be a coincidence. It is therefore tempting to suggest that the identity $R_S = R_{max}$ is fundamental to the structure of the universe; accordingly, we must have

$$\kappa = +1 \text{ and } \Lambda \leq \Lambda_c \qquad (18)$$

It follows that at any epoch

$$R(t) \leq R_S \qquad (19)$$

which means that the universe is indeed in a black hole.

At present, the largest sphere that can be drawn in the universe has a surface area $4\pi R_0^2$. As expansion goes on, this may approach the maximal value $4\pi R_{max}^2$ $(\equiv 4\pi R_S^2)$. Being inside a black hole, we cannot hope to "shoot through" the Schwarzschild surface; we may approach it in infinite time $(\Lambda = \Lambda_c)$ or in a finite time $(\Lambda < \Lambda_c)$. In the latter case, the universe must retrace its steps and proceed along a phase of contraction, eventually producing densities where present understanding of physics breaks down (with, perhaps, a further expansion from the "primaeval" matter and so on, in an endless cycle of pulsations).

Some of the immediate consequences of this picture may be expressed in terms of inequalities which must be satisfied by the various parameters characterizing the universe. Arising from conditions (18), these inequalities provide lower and upper bounds for the parameters Λ and ρ_0, and a lower bound for the parameter R_0, in terms of the observable quantities H_0 and q_0; so these parameters can be estimated from the kinematics of the universe alone. The numerical values resulting from these inequalities may not be very accurate because of the errors in the observed values of H_0 and q_0 (which, at the present time, may be as large as 20–40% (ref. 6)). But using $H_0 = 75$ km s^{-1} (Mpc^{-1}) and $q_0 = 1$ as "representative values"

$$-6.7 \times 10^{-57}\ \mathrm{cm}^{-2} < \Lambda \leq \Lambda_c \leq 1.0 \times 10^{-57}\ \mathrm{cm}^{-2} \qquad (20)$$

$$1.5 \times 10^{-29}\ \mathrm{g\ cm}^{-3} < \rho_0 \leq 2.3 \times 10^{-29}\ \mathrm{g\ cm}^{-3} \qquad (21)$$

$$R_0 \geq 1.1 \times 10^{28}\ \mathrm{cm} \qquad (22)$$

The ranges in which Λ and ρ_0 may lie are narrow.

Apart from these immediate consequences, there are deeper implications as well. For instance, we are now faced with several questions: How did the universe come to be a black hole—through a gravitational collapse, followed by a phase of expansion? In the cosmos, which includes the exterior as well as the interior of the universe, can our universe be unique? If not, what would its status be *vis-a-vis* other such structures in the cosmos? Investigation of these and other related questions, including the possible existence of an hierarchy of black holes, is clearly a matter of some importance.

As for our own universe, the concept of the Schwarzschild radius itself seems to be of considerable significance. Its relevance to other realms of physical phenomena will be discussed in a subsequent communication.

I thank Professor P. L. Bhatnagar, Dr J. Vankerkooy and Dr G. S. Kular for discussions, and Professor P. T. Landsberg for hospitality. This work was partly supported by the National Research Council of Canada.

R. K. PATHRIA

Department of Physics, University of Waterloo, Waterloo, Ontario

Received June 19; revised September 20, 1972.

1 Einstein, A., *Berlin Sitzungsberichte*, 142 (1917). English translation in *The Principle of Relativity* (Methuen, London, 1923).
2 *Nature* (*News and Views*), **232**, 440 (1971).
3 Robertson, H. P., and Walker, A. G., *Pub. Astron. Soc. Pacific*, **67**, 82 (1955).
4 Pathria, R. K., *The Theory of Relativity* (Hindustan Publishing Corporation, Delhi, 1963).
5 Robertson, H. P., and Noonan, T. W., *Relativity and Cosmology* (W. B. Saunders, Philadelphia, 1968).
6 Sandage, A. R., *Observatory*, **88**, 91 (1968).
7 Rindler, W., *Essential Relativity* (Van Nostrand Reinhold Co., New York, 1969).

Orbital Eccentricity of Mercury and the Origin of the Moon

A NUMBER of mechanisms for the formation of the Moon have been suggested; fission of the Earth, precipitation in a hot gaseous silicate atmosphere, independent formation in orbit about the Earth, and independent formation elsewhere in the solar system followed by capture by the Earth[1]. Although the last of these mechanisms has been admitted to be improbable by its proponents, they have shown that it is by no means impossible dynamically[2]. The principal objection to this mechanism is the strange composition of the Moon. It has been recognized for many years that the low mean density of the Moon implies that it is highly deficient in metallic iron. The lunar exploration programme has also shown that the Moon is much more deficient than the Earth in the more volatile of the condensable elements. Because of the apparent difficulty of satisfying these composition constraints in a theory in which the Moon is formed elsewhere in the solar system, I have tended to favour the other mechanisms mentioned above[3,4].

The situation now appears to be changed as a result of recent work by D. L. Anderson[5,6] (presented in most fully developed form at the IAGC). He justifies the postulate that, in addition to the above chemical abnormalities, the deep interior of the Moon is probably also very deficient in magnesium silicates, and that it is remarkably free of iron oxide in the silicate materials. Such a Moon would be close to the melting point in much of the deeper interior, yet its electrical conductivity would not exceed the bounds placed by analysis of lunar magnetometer measurements[7]. Anderson postulates that the Moon has the same basic composition as the calcium and aluminium-rich silicates in the inclusions in the Allende carbonaceous chondrite, which are believed to be very high temperature condensates within a cooling primitive solar nebula[8].

As a hot gas of solar composition cools, at a pressure of