Mon. Not. R. astr. Soc. (1974) **168**, 399–415.

BLACK HOLES IN THE EARLY UNIVERSE

B. J. Carr and S. W. Hawking

(Received 1974 February 25)

SUMMARY

The existence of galaxies today implies that the early Universe must have been inhomogeneous. Some regions might have got so compressed that they underwent gravitational collapse to produce black holes. Once formed, black holes in the early Universe would grow by accreting nearby matter. A first estimate suggests that they might grow at the same rate as the Universe during the radiation era and be of the order of 10^{15} to 10^{17} solar masses now. The observational evidence however is against the existence of such giant black holes. This motivates a more detailed study of the rate of accretion which shows that black holes will not in fact substantially increase their original mass by accretion. There could thus be primordial black holes around now with masses from 10^{-5} g upwards.

1. INTRODUCTION

Black holes are normally thought of as being produced by the collapse of stars or possibly galactic nucleii. However, one would also expect there to be a certain number of black holes with masses from 10^{-5} g upwards which were formed in the early stages of the Universe (Hawking 1971). This is because the existence of galaxies implies that there must have been departures from homogeneity and isotropy at all times in the history of the Universe. These could have been very large in the early stages and even if they were small on average there would be occasional regions in which they were large. One would therefore expect at least a few regions to become sufficiently compressed for gravitational attraction to overcome pressure forces and the velocity of expansion and cause collapse to a black hole. We shall refer to such black holes as primordial.

A region in the early Universe of radius R has a potential energy of self-gravitation

$$\Omega \sim -\mu^2 R^5$$

and kinetic energy of expansion

$$T \sim \mu R^3 \dot{R}^2$$

where μ is the energy density and units are such that $G = c = 1$. In a $k = 0$ Friedmann universe the sum of these energies is zero. Thus

$$\left(\frac{\dot{R}}{R}\right)^2 \sim \mu.$$

In the radiation epoch, when most of the particles are relativistic, the pressure p is $\frac{1}{3}\mu$ and μ is proportional to R^{-4}. Thus

$$\mu \sim t^{-2}, \quad R \propto t^{1/2}.$$

In the very early stages ($t < 10^{-4}$ s) it is possible that the number of different species of particles present may increase very sharply (Hagedorn 1970). This could mean that most of the particles were non-relativistic and that p was of the order $\log \mu \ll \mu$. In this case

$$\mu \sim t^{-2}, \quad R \propto t^{2/3}.$$

If the density in a region is somewhat higher than average or the rate of expansion somewhat lower, the gravitational forces may be able to halt the expansion by overcoming both the kinetic energy of expansion and the pressure forces. To overcome the pressure forces requires that the gravitational energy, $-\Omega$, should be greater than the internal energy U. When $p = \frac{1}{3}\mu$, $U \sim \mu R^3$ and so a necessary condition for collapse is $\mu R^2 > \sim 1$. When $p \sim \mu_0 \log \mu/\mu_0$, $U \sim \mu_0 R^3 \log \mu/\mu_0$. Thus a necessary condition for collapse is

$$\mu R^2 > \sim \frac{\mu_0}{\mu} \log \frac{\mu}{\mu_0}.$$

(The constant μ_0 is about 10^{14} g cm^{-3}.) These inequalities place a lower limit on the size of a region that can undergo gravitational collapse. This limit is just the Jean's length for the epoch in question. A region that is about to collapse has also an upper limit to its size at the moment at which it begins to contract. To see how this arises, consider a spacelike hypersurface orthogonal to the matter flow which crosses the region at the moment when the rate of expansion is zero. The $R^{00} - \frac{1}{2}g^{00}R = 8\pi T^{00}$ constraint equation implies that the 3-geometry of this hypersurface has positive curvature of order μ in the region where the rate of expansion is zero. If this positive curvature extended over a sufficiently large region, the spacelike hypersurface would close up on itself to form a disconnected compact 3-space of radius about $\mu^{-1/2}$. In this case the region would form a separate closed universe which was completely disconnected from our Universe. Such a situation would not correspond to a black hole formed by collapse of matter in our Universe. This shows that for black hole formation $\mu R^2 < \sim 1$. Together with the previous conditions this implies that at the moment of recollapse $\mu R^2 \sim 1$ for the case $p = \frac{1}{3}\mu$ and $\mu_0/\mu \log \mu/\mu_0 < \mu R^2 < \sim 1$ for the case $p \sim \mu_0 \log \mu/\mu_0$. One can interpret these conditions in the following way. The energy density μ is of order t^{-2} in both cases. Thus $\mu R^2 \sim 1$ implies $R \propto t$. This shows that, in the $p = \frac{1}{3}\mu$ case, the size of the region at the moment of recollapse must be of order of that of the particle horizon, the distance light could have travelled since the beginning of the Universe. In the $p \ll \mu$ case, it could be much smaller.

The condition that the region should be within its Schwarzchild radius is that the mass ($\sim \mu R^3$) should be greater than the radius R, i.e. $\mu R^2 > 1$. Thus in the $p = \frac{1}{3}\mu$ case the region would be within a (future-directed) trapped surface and so would be a black hole at about the time when it began to recollapse. On the other hand in the $p \ll \mu$ case, a region smaller than the particle horizon which begun to recollapse would have to contract quite a lot before it became a black hole. In fact random turbulent motions might prevent it from ever collapsing sufficiently.

The earliest time at which one can hope to apply classical general relativity is the Planck time $\sqrt{G\hbar/c^5} \sim 10^{-43}$ s. A black hole formed at this time would have an initial mass of about 10^{-5} g and radius 10^{-33} cm. A black hole formed at the time of Helium formation when the temperature was 10^9 K would have a mass of about 10^7 solar masses.

After the formation of a black hole in the early Universe one would expect it to grow by accreting some of the nearby matter. The first estimate of the rate of accretion was made by Zeldovich & Novikov (1967). They considered the accretion as a quasi-stationary process. In this case the velocity of the matter crossing the horizon ($r_g = 2M$) will be of order the velocity of light (unity). Thus the rate of accretion

$$\frac{dM}{dt} \sim \mu r_g^2 \sim \mu M^2$$

where μ is the density of the background Friedmann universe. But $\mu \sim t^{-2}$. Thus

$$M \sim \frac{t}{1+\frac{t}{t_0}\left(\frac{t_0}{M_0}-1\right)}$$

where M_0 is the initial mass of the black hole and t_0 is the time of formation. If M_0 were small compared to t_0, i.e. if the black hole was small compared to the particle horizon at the time of formation, then $M-M_0$ remains small and so there would not be much accretion. On the other hand if M_0 were of the same order as t_0 (which, as was shown above, is likely to be the case in the $p = \frac{1}{3}\mu$ era), then the Zeldovich–Novikov argument indicates that $M \sim t$. In other words, the accretion would cause the black hole to grow at the same rate as the particle horizon. If this growth continued up to the present time, the black hole would be of the order of the Hubble radius and we would have fallen into it or be just about to. This would not be in accordance with observations which indicate that the Universe is homogeneous on a large scale. On the other hand, one might suppose that black holes would grow at the same rate as the particle horizon only during the $p = \frac{1}{3}\mu$ era when the radiation pressure would drive matter into the black hole. If this were the case, the black holes would grow to a mass of 10^{15} to 10^{17} solar masses, the mass within the particle horizon at the time when radiation ceased to dominate matter. The obvious place to look for such giant black holes would be in clusters of galaxies where they might provide the missing mass necessary to bind the cluster gravitationally. The observational evidence is that the missing mass of the Virgo Cluster is not in the form of black holes of masses greater than about 10^{10} solar masses (Van den Bergh 1969).* Also one would expect that the existence of such very large black holes at the time of decoupling would produce appreciable fluctuations in the microwave background on small angular scales whereas none have been observed (Boynton & Partridge 1973).

It seems therefore either that there were no black holes formed in the early Universe or that, contrary to what is indicated by the Zeldovich–Novikov argument, they did not grow at the same rate as the particle horizon. As mentioned before, it is difficult to believe with any theory of random perturbations that there would not have been at least a few black holes formed. It therefore seems worthwhile to examine more closely the Zeldovich–Novikov argument. The assumption of quasi-stationary accretion, on which this is based, is probably a reasonable approximation when the black hole is small compared to the particle horizon. Thus the conclusion

* Van den Bergh concluded from the lack of observed tidal distortion of galaxies that the missing mass could not be in the form of compact objects in the mass range 10^{10} to 10^{13} solar masses. However, one can remove the upper limit on this range because the lower number of higher mass objects would be balanced by the fact that each object could induce tidal distortion in a larger volume of space.

that such a black hole would not grow very much is probably valid. However, the assumption breaks down in the critical case of a black hole whose size is of the same order as the particle horizon because the expansion of the Universe has to be taken into account.

In this paper we shall investigate the crucial question of whether a black hole can grow at the same rate as the Universe. Such a situation would be described by a similarity solution, that is, one in which all lengths increase at the same rate. If the Zeldovich–Novikov argument were correct, one would expect a black hole formed in the early Universe to approach such a solution asymptotically as time proceeds. For simplicity we shall consider only the case of spherical symmetry. One would expect departures from spherical symmetry such as turbulence and vorticity to decrease the rate of accretion. If therefore this rate is insufficient in the spherical case to make a black hole grow at the same rate as the particle horizon, it would also be insufficient in the more realistic non-spherical case.

To determine the evolution of a black hole in a cosmological model, one has to specify the initial conditions on some spacelike hypersurface near the initial singularity. In the $p = 0$ case (which is a good approximation to the $p \ll \mu$ case) these initial conditions determine the evolution in a very direct manner: each spherical shell of matter moves on a geodesic in the gravitational field of the mass interior to it and its total energy E per unit mass is constant. If $E \geq 0$ the shell expands indefinitely and if $E < 0$ the shell will reach a maximum radius and collapse into the black hole. In the $k = 0$ Friedmann model $E = 0$ for all shells. If one introduced a black hole of mass Δm into such a model but left the density and the velocities of expansion unchanged on the spacelike hypersurface, E would be of the form $-\Delta m/R$ where R is the radial coordinate which measures the area of spherical shells. Such a choice of initial conditions would cause the black hole to grow indefinitely and each spherical shell of matter would eventually fall into it. Succeeding shells of matter would however take longer and longer to fall into the black hole which would not grow as fast as the particle horizon. The maximum radius obtained by these shells would be much larger than the size of the black hole. Small departures from spherical symmetry such as rotation would grow large during the collapse of the shell from its maximum radius and could prevent it from falling into the black hole. One might therefore expect that a black hole would not grow very much after it had been left some way behind by the particle horizon.

In Section 4 we shall show that in the $p = 0$ case the energy E has to be negative and constant in order to have a similarity solution, that is, in order to have the black hole growing at the same rate as the particle horizon. However, because of the existence of particle horizons, regions on the initial spacelike hypersurface at large spatial distances from the black hole will not have any causal communication with the region that collapsed to produce the hole. One would therefore expect that the existence of a black hole at the origin would affect the initial conditions in these regions only through the Coulomb field of the initial mass Δm. As mentioned above this gives $E = -\Delta m/R$ and not E constant as would be required for a similarity solution. It seems therefore that black holes formed by local inhomogeneity in the $p \ll \mu$ era will not grow very much by accretion.

One might expect black holes formed in the $p = \frac{1}{3}\mu$ era to grow rather faster because radiation pressure would drive matter into the holes. As in the $p = 0$ case one would not expect the existence of the hole to affect the initial conditions in regions to which no signal could yet have propagated from the region which underwent collapse. However, one could imagine a situation in which one had a decompression wave expanding into an exactly Friedmann universe. Matter behind the wavefront would be accelerated inwards by the pressure gradient and could fall into the black hole. The question is whether the pressure gradients in the decompression wave could cause a sufficient flow of matter into the black hole to make the black hole grow at the same rate as the particle horizon. We shall show in fact that they cannot: there are no similarity solutions containing a black hole which are exactly Friedmann outside some expanding wave-front. As in the $p = 0$ case, there are similarity solutions containing a black hole which are only asymptotically and not exactly Friedmann at large distances. Like the $p = 0$ case, these solutions require that the initial energy E of matter at large spacial distances from the black hole should be a negative constant and not merely $-\Delta m/R$ as would be expected if the initial conditions were altered only by the introduction of the Coulomb field of a black hole at the origin. Also in the $p = \frac{1}{3}\mu$ similarity solutions the pressure gradients are directed outwards and so are hindering rather than helping matter to fall into the hole. In other words in order to get the black hole to grow at the same rate as the particle horizon, the initial conditions have to be such that the matter is really thrown into the hole. One would not expect the existence of a black hole produced by local fluctuations to cause the initial conditions to be arranged in this way. Therefore as in the $p = 0$ case one would not expect a black hole to grow as fast as the particle horizon. After it has fallen some way behind the particle horizon, the Zeldovich–Novikov argument suggests that it will probably not grow very much further. The main conclusion of this paper is therefore that black holes formed in the early Universe would not grow very much by accretion and so could be around today with any mass from 10^{-5} g upwards.

From the measurement of the deceleration parameter (Sandage 1961), one can place an upper limit of about 10^{-28} g cm^{-3} on the present average mass density in black holes. On the other hand the present average density of observed luminous matter is only about 10^{-31} g cm^{-3}. Thus it is possible that most of the mass of the Universe at the present time is in the form of black holes. However, as one goes to earlier times, the mass density of black holes will increase as $(1+Z)^3$ where Z is the redshift, while the density of the microwave background will increase as $(1+Z)^4$. This means that at early times the mass in black holes was only a small fraction of the total mass of the Universe. In other words only a small fraction of the Universe can have undergone gravitational collapse at early times. This argument, which was first given by Zeldovich & Novikov (1967), places an upper limit on the degree of turbulence and inhomogeneity in the early Universe. Another upper limit on the inhomogeneity of the early Universe may be placed by an argument of Zeldovich & Sunyaev (1969) that the dissipation of a significant amount of turbulent energy at times later than about $10^9\,\Omega^8$ s (where Ω is the present average density of matter not in black holes in units of 10^{-29} g cm^{-3}) would cause distortions in the blackbody spectrum of the microwave background contrary to observation. This indicates that most of any primordial turbulence must have been damped out by this time and so one would not expect any primordial black holes bigger than about $10^{14}\,\Omega^8$ solar masses. Further information about the very early Universe would be known if it were possible experimentally either to detect the black holes of very low mass or to place better upper limits on their number density. This might be possible if such very small black holes carry an electric charge (Hawking & Gibbons 1974, to be published).

The plan of this paper is as follows. In Section 2 we extend the definition of black holes from asymptotically flat space-times to cosmological models, in particular to asymptotically Friedmann models. The form of the metric and field equations for spherically symmetric similarity solutions are given in Section 3. In Section 4 we discuss black holes in asymptotically Friedmann solutions with $p = 0$. We show that in order to cause a black hole to expand at the same rate as the Universe, the initial conditions have to be altered by more than would be accounted for simply by the Coulomb field of the black hole. Similar results are obtained for the $p = \frac{1}{3}\mu$ case in Section 5.

2. BLACK HOLES IN COSMOLOGY

In studying black holes one normally neglects the curvature of the Universe and treats the black hole as being in an asymptotically flat space-time. The black hole can then be defined as the region of space-time from which it is not possible to escape to future null infinity $\mathscr{I}^+$ (Hawking 1973). The boundary of the black hole, the event horizon, is formed by the wave-front which just fails to reach infinity. Assuming that there are no naked singularities, i.e. that the space is strongly future asymptotically predictable, one can show that the area of a two-dimensional section of the event horizon never decreases with time. Given a time coordinate t which defines a suitable family of spacelike surfaces at constant time, one can define the apparent horizon as the outer boundary of the region containing trapped surfaces lying in a surface of constant time. One can show that the event horizon always lies outside or coincides with the apparent horizon.

Asymptotically flat space-time is a very good approximation when dealing with black holes formed from stellar collapse since these will be very small compared to the radius of the Universe. However, it is clearly not appropriate when considering black holes in the early Universe since these may be of the order of the particle horizon at their time of formation. One would therefore like to define black holes in spaces which are not asymptotically flat but which asymptotically approach some cosmological model such as a Friedmann solution. This can be done provided that there is some infinity so that one can define the region of space-time from which one cannot escape to infinity. In order to define infinity we shall use the future causal boundary of space-time defined by Geroch–Kronheimer–Penrose (1972). Points of this boundary are represented by future-directed time-like curves which have no endpoint in the space-time manifold, two such curves defining the same boundary point if they have the same past. This boundary includes both points at infinity and points at finite distance which are singularities. We shall say that a subset $\mathscr{I}$ of the boundary is at infinity if every null geodesic reaching it (i.e. having the same past as a point of $\mathscr{I}$) has infinite affine length. One can then define the event horizon as the boundary of the past of a suitably chosen $\mathscr{I}$. As in the asymptotically flat case one wants to exclude the possibility of singularities which are naked, i.e. visible from $\mathscr{I}$. This can be done by extending the definition of strong asymptotic predictability to the cosmological case: one requires that there exist a spacelike surface S from which one can predict events at points near $\mathscr{I}$ and on the event horizon to the future of S. With this assumption one can prove that, as in the asymptotically flat case, the event horizon must lie outside the apparent horizon and its area cannot decrease with time.

In the $k = 0$ and -1 Friedmann models there is a future null infinity which is similar in many ways to that of asymptotically flat space (Penrose 1964; Hawking & Ellis 1973). One can therefore define black holes in spaces which are asymptotic to $k = 0$ or -1 Friedmann models. In the $k = +1$ model on the other hand there is no infinity as the space is closed and the model recollapses to a singularity within a finite time. One cannot, strictly speaking, define a black hole in such as space: in a certain sense every point is inside a black hole since nothing can escape and everything is doomed to annihilation in the singularity. However, there is a big difference between possible annihilation in about 2×10^{10} yr and annihilation in 1 s or less which is likely to occur if one fell into a black hole in the early Universe. Thus to a very good approximation in discussing such a situation one can neglect the eventual collapse of the Universe and treat it was being asymptotically a $k = 0$ Friedmann model: the difference between a $k = 0$ and a $k = \pm 1$ model would be less than one part in 10^{15} 1 s after the big bang.

3. SPHERICALLY SYMMETRIC SIMILARITY SOLUTIONS

In the early Universe the mean free path of particles such as photons will in most epochs be considerably smaller than the particle horizon. Thus the energy-momentum tensor will have the form of a perfect fluid:

$$T_{\mu\nu} = (\mu + p)\, u_\mu u_\nu - p g_{\mu\nu} \tag{3.1}$$

where u^μ is the unit flow vector of the matter. We shall consider two cases: $p = \frac{1}{3}\mu$ and $p = 0$ (which will be a good approximation to the very early epoch when $p \ll \mu$). In a spherically symmetric solution one can introduce a time coordinate t such that the surfaces of constant t are orthogonal to the flow lines and comoving coordinates (r, θ, ϕ) which are constant along each flow line. In these coordinates the metric takes the form

$$ds^2 = e^{2\nu}\, dt^2 - e^{2\lambda}\, dr^2 - R^2(d\theta^2 + \sin^2\theta\, d\phi^2). \tag{3.2}$$

The conservation equations $T^{\mu\nu}{}_{;\nu} = 0$ give two first integrals

$$e^\nu = f(t), \qquad \mu\, e^\lambda R^2 = h(r) \quad \text{for} \quad p = 0 \tag{3.3}$$

$$\mu^{1/4} e^\nu = f(t), \quad \mu^{3/4} e^\lambda R^2 = h(r) \quad \text{for} \quad p = \tfrac{1}{3}\mu \tag{3.4}$$

where $f(t)$ and $h(r)$ are arbitrary functions of integration which can be given any value by appropriate coordinate transformations. The Einstein equations

$$R^{\mu\nu} - \tfrac{1}{2} R g^{\mu\nu} = 8\pi T^{\mu\nu} \tag{3.5}$$

have one other first integral:

$$m = \tfrac{1}{2} R(1 + e^{-2\nu} R_t{}^2 - e^{-2\lambda} R_r{}^2) \tag{3.6}$$

where

$$m_r = 4\pi\mu R^2 R_r \tag{3.7}$$

and

$$m_t = -4\pi p R^2 R_t. \tag{3.8}$$

From the equation (3.7)

$$m = 4\pi \int \mu R^2\, dR. \tag{3.9}$$

Thus one can interpret $m(r, t)$ as the mass within the comoving radius r at the time t. From equation (3.8) it follows that this is independent of time in the case $p = 0$

and it decreases during expansion because of the work done by the pressure in the case $p = \frac{1}{3}\mu$. The quantity $m(r, t)$ also governs the existence of trapped surfaces. Suppose that a flash of light is emitted from a 2-surface of constant r and t. As time proceeds there will be an ingoing and an outgoing wavefront from this 2-surface which will be formed by future-directed ingoing and outgoing null geodesics orthogonal to the 2-surface. The convergences, ρ_i and ρ_0, of these families of null geodesics will be

$$\begin{aligned} \rho_i &= -R^{-1}(e^{-\nu} R_t - e^{-\lambda} R_r) \\ \rho_0 &= -R^{-1}(e^{-\nu} R_t + e^{-\lambda} R_r). \end{aligned} \tag{3.10}$$

Thus

$$\frac{2m}{R} = 1 + R^2 \rho_i \rho_0. \tag{3.11}$$

Trapped surfaces occur when the convergence of the outgoing family of null geodesics is positive. Thus at the apparent horizon of the black hole, the outer boundary of the trapped surfaces,

$$\frac{2m}{R} = 1. \tag{3.12}$$

Condition (3.12) would also hold when $\rho_i = 0$. We shall call this the cosmological apparent horizon. It is related to the existence of the big bang singularity in the past (Hawking & Ellis 1973).

A solution of the Einstein equations is said to be a similarity solution if it admits a homothetic Killing vector, i.e. a vector field $\boldsymbol{\xi}$ such that

$$\xi_{\mu;\nu} + \xi_{\nu;\mu} = 2g_{\mu\nu}. \tag{3.13}$$

In other words, as one moves along the orbits of the vector field $\boldsymbol{\xi}$, all lengths increase at the same rate. Such a solution corresponds to the classical notion of a similarity solution and it is what one requires if one wants a black hole expanding at the same rate as the Universe. Cahill & Taub (1971) have investigated spherically symmetric similarity solutions in which the homothetic Killing vector is not parallel to the fluid flow vector. They show that by suitable coordinate transformations such solutions can be put in a form in which all dimensionless quantities such as ν, λ, R/r, μR^2 are functions only of the dimensionless variable $z \equiv r/t$. The Einstein field equations in the case $p = \frac{1}{3}\mu$ will be satisfied if

$$\ddot{S} + \dot{S} + \left(\frac{\dot{S}}{2S} - \frac{\dot{x}}{x}\right)(3S + 4\dot{S}) = 0 \tag{3.14}$$

$$\tfrac{1}{2}S^4 + 2x^4 z \frac{\dot{S}}{S} - \frac{\dot{x}}{x}(3V^2 - 1)\, S^4 = 4x^2 \tag{3.15}$$

where a dot indicates differentiation with respect to $\Omega \equiv \log z$, $S \equiv R/r$ and $x^4 \equiv 3/\mu r^2$. In terms of these quantities

$$e^{\nu} = \sqrt{z}x \tag{3.16}$$

$$e^{-\lambda} = S^2 x^{-3} \tag{3.17}$$

and

$$V \equiv e^{\lambda-\nu} z = \sqrt{z} x^2 S^{-2}. \tag{3.18}$$

The quantities S, x and V are functions only of z. V represents the velocity, relative to the flow lines of the matter, of the surfaces of constant z. These surfaces, which have the equation $r = z_0 t$ (z_0 = a constant), represent a family of spheres expanding through the matter. When $V < 1$, the surfaces are timelike, when $V = 1$, they are null and when $V > 1$, they are space like.

The quantity $M \equiv m/R$ is also a function only of z. One can derive two expressions for it in terms of x, z, and S. The equality of these two expressions is a first integral of equations (3.14) and (3.15):

$$2M = \frac{2S^2}{x^4}\left(3+\frac{4\dot{S}}{S}\right) \tag{3.19}$$

$$2M = 1+\frac{z\dot{S}^2}{x^2}-\frac{S^4}{x^6}(S+\dot{S})^2. \tag{3.20}$$

Similarity solutions with $p = \frac{1}{3}\mu$ will be considered further in Section 5. Solutions with $p = 0$ will be considered in the next section.

4. THE $p = 0$ CASE

In a spherically symmetric solution with $p = 0$ the quantity $m(r, t)$, which represents the mass within comoving radius r, is independent of t. The absence of pressure gradients means that each spherical shell of matter follows a geodesic path in the gravitational field of the mass $m(r)$ interior to it. The Einstein equations therefore admit an additional integral:

$$E(r) = \tfrac{1}{2}\,e^{-2\nu}\left(\frac{\partial R}{\partial t}\right)^2 - \frac{m(r)}{R}. \tag{4.1}$$

The first term on the right can be interpreted as the kinetic energy per unit mass of the shell of matter of comoving radius r, and the second term as the potential energy per unit mass. Thus $E(r)$ represents the total energy per unit mass of the shell. If $E \geq 0$, the shell will start from a point of zero area ($R = 0$) and expand indefinitely ($R \to \infty$). If $E < 0$, the shell will start from a point ($R = 0$), expand to some maximum value of R and recollapse to a point. Explicitly,

$$\left\{\begin{array}{lll} t-t_0(r) = \dfrac{\sqrt{2ER^2+2mR}}{2E} - \dfrac{2m}{(2E)^{3/2}}\sinh^{-1}\sqrt{\dfrac{ER}{m}} & & (E > 0) \\ R = \left(\tfrac{3}{2}\right)^{2/3}(2m)^{1/3}(t-t_0(r))^{2/3} & & (E = 0) \\ t-t_0(r) = \dfrac{\sqrt{2ER^2+2mR}}{2E} + \dfrac{2m}{(-2E)^{3/2}}\sin^{-1}\sqrt{\dfrac{-ER}{m}} & & (E < 0). \end{array}\right. \tag{4.2}$$

Equation (4.2) completely determines the evolution of a solution in terms of the functions $m(r)$, $E(r)$ and $t_0(r)$. From equations (3.7) and (3.6)

$$4\pi\mu = \frac{\partial m}{\partial r}\left(R^2\frac{\partial R}{\partial r}\right)^{-1} \tag{4.3}$$

$$e^{\lambda} = \left(\frac{\partial R}{\partial r}\right)(1+2E(r))^{-1/2}. \tag{4.4}$$

The $k = 0$ Friedmann solution is obtained with $t_0(r) = 0$, $E(r) = 0$ and $m(r)$ any positive monotonically increasing function of r which tends to infinity with r.

By making a coordinate transformation of the form $r' = f(r)$ one can arrange that $m(r) = Cr^3$ where C is a constant. However, for later convenience, in this paper we shall choose the radial coordinate r such that $m(r) = Cr$. One obtains a solution representing a black hole in an asymptotically $k = 0$ Friedmann universe if $t_0(r) = 0$ and E is negative for small r and is bounded as $r \to \infty$. If $E(r)$ is negative only for r less than some value r_0, then only those shells with $r < r_0$ will fall into the black hole and so the black hole will only grow to the mass $m(r_0)$. On the other hand if $E(r)$ is negative for all r, all the matter in the Universe will eventually fall into the black hole which will grow to an infinite size. Thus the amount of accretion by the black hole is completely determined by the function $E(r)$ which has to be specified right at the beginning of the Universe, before the hole has even formed. There seems no reason why the initial conditions should be arranged in such a way as to make an infinite amount of matter fall into the hole.

For the purpose of comparing with the $p = \frac{1}{3}\mu$ solutions considered in the next section, it is useful to study the $p = 0$ similarity solutions containing a black hole in an asymptotically Friedmann universe. As was stated in the previous section, in such a solution all dimensionless quantities are functions only of the dimensionless variable $z = r/t$. The quantity $E(r)$ is dimensionless and so is the ratio $m(r)/r$. Since both of these are functions only of r they must be constant in a similarity solution. By a coordinate transformation $m(r)/r$ can be put equal to 1 and from (4.1) one has the equation

$$\begin{aligned} E &= \tfrac{1}{2}r^2\left(\frac{\partial S}{\partial t}\right)^2 - \frac{1}{S} \\ &= \tfrac{1}{2}z^4\left(\frac{\partial S}{\partial z}\right)^2 - \frac{1}{S} \end{aligned} \tag{4.5}$$

where $S(z) = R/r$. This can be integrated to give (with integration constant D)

$$D \pm \frac{1}{z} = \begin{cases} \dfrac{\sqrt{ES^2+S\sqrt{2E+1}}}{\sqrt{2E}} - \dfrac{2\sqrt{1+2E}}{(2E)^{3/2}}\sinh^{-1}\sqrt{\dfrac{ES}{\sqrt{1+2E}}} & (E > 0) \\ \sqrt{\tfrac{2}{3}}\, S^{3/2} & (E = 0) \\ \dfrac{\sqrt{ES^2+S\sqrt{2E+1}}}{\sqrt{2E}} + \dfrac{2\sqrt{1+2E}}{(-2E)^{3/2}}\sin^{-1}\sqrt{\dfrac{-ES}{\sqrt{1+2E}}} & (E < 0). \end{cases} \tag{4.6}$$

The $k = 0$ Friedmann solution is obtained with $D = E = 0$. This gives

$$dS^2 = dt^2 - \left(\frac{9t^2}{2}\right)^{2/3}\left(\tfrac{1}{9}r^{-4/3}\,dr^2 + r^{2/3}(d\theta^2 + \sin^2\theta\, d\phi^2)\right). \tag{4.7}$$

The coordinate transformation

$$r'(r) = \left(\frac{9r}{2}\right)^{1/3}$$

puts this in the more usual form

$$dS^2 = dt^2 - t^{4/3}((dr')^2 + r'^2(d\theta^2 + \sin^2\theta\, d\phi^2)).$$

The solutions (4.6) with $D = 0$ but $E \neq 0$ are asymptotically Friedmann. Those with $E < 0$ contain a black hole. To analyse the behaviour of these solutions it is helpful to consider the quantities $V \equiv z\, e^{\lambda-\nu}$ and $M \equiv m/R$ which are functions only of z. In the Friedmann solution

$$V = (\tfrac{1}{6})^{1/3} z^{1/3}$$

$$M = (\tfrac{2}{9})^{1/3} z^{2/3}.$$

The surface $z = 6$ on which $V = 1$ is a null surface expanding from the origin $r = 0$ at the time $t = 0$. It can be thought of as the 'creation light cone' of 'particle horizon' of an observer at the origin. The surface $z = \frac{3}{4}$ on which $M = \frac{1}{2}$ is time-like and represents the locus of points at which the past light cones of an observer at the origin are focused and made to start to reconverge by the matter in the Universe. We shall call this the 'cosmological apparent horizon'.

In the solutions with $D = 0$ and E small and negative, V and M have the same asymptotic behaviour as in the Friedmann solution (Fig. 1). However, at small values of z they pass through a minimum and then increase again becoming infinite

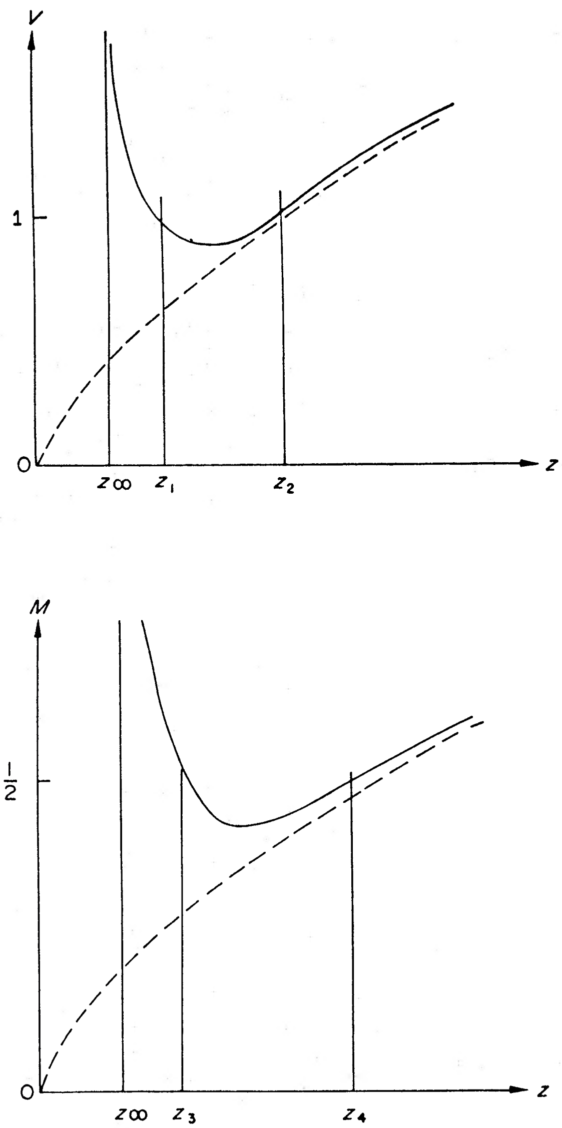

FIG. 1. *Graphs of V and M against z for p = 0 similarity solutions showing the positions of the particle and the event horizons and the cosmological and black hole apparent horizons. The dotted curves indicate the Friedmann values.*

at some value $z = z_\infty > 0$. There will be two values z_1 and $z_2 (z_2 > z_1)$ at which $V(z) = 1$. The outer surface $z = z_2$ can be regarded as the particle horizon. For $z_1 < z < z_2$ the surfaces of constant z are time-like. It would therefore be possible for an observer in a rocket to remain in the region $z > z_1$ by accelerating outwards at a suitable rate. However, if he should cross the null surface $z = z_1$, the surfaces of constant z would become spacelike and so he would inevitably hit the singularity which occurs at $z = z_\infty$. This shows that the surface $z = z_1$ is the event horizon of the black hole. In a similar manner there will be two values z_3 and $z_4 (z_3 < z_4)$ at which $M(z) = \frac{1}{2}$. The outer surface $z = z_4$ can be regarded as the cosmological apparent horizon. The inner surface $z = z_3$ will be the apparent horizon of the black hole. It will lie inside the event horizon.

For E small and negative the size of the event horizon of the black hole is small compared to the particle horizon. As E is made more negative the minimum of V moves upwards and the value of z_∞ moves to the right. At some critical value $E = E_0$ the minimum value of V is 1 and the event and particle horizons coincide. For $E < E_0$ the minimum of V is above 1 and so there will be no particle or event horizons. It can be shown that the minimum of M is always below $\frac{1}{2}$ so there will still be cosmological and black hole apparent horizons. Because the surfaces of constant z are always spacelike, any observer will eventually fall into the black hole's apparent horizon and hit the singularity. One can regard such a solution as representing a black hole which is expanding so fast that it envelopes the whole Universe.

5. THE $p = 1/3\mu$ CASE

The Einstein equations for a spherically symmetric solution with $p = \frac{1}{3}\mu$ form a hyperbolic system in two variables, r and t. The characteristic surfaces in this system move with the speed of sound, $1/\sqrt{3}$; there are no gravitational waves because of the spherical symmetry. In general this system of equations can be solved only on a computer. However, in the case of a similarity solution, these partial differential equations reduce to the set (3.14) and (3.15) of ordinary differential equations in the variable $\Omega = \log(r/t)$. A particular solution of these equations is

$$\begin{cases} x = 2^{3/2} z^{-1/2} \\ S = 2^{11/6} z^{-1/2}. \end{cases} \tag{5.1}$$

This gives the metric

$$ds^2 = 8\, dt^2 - 2^{5/3} z^{-1}\, dr^2 - 2^{11/3} rt (d\theta^2 + \sin^2\theta\, d\phi^2). \tag{5.2}$$

One can put this in a more familiar form by making the coordinate transformation

$$\hat{t} = 2^{3/2} t, \quad \hat{r} = 2^{13/12} r^{1/2}$$

which gives

$$ds^2 = d\hat{t}^2 - \hat{t}(d\hat{r}^2 + \hat{r}^2(d\theta^2 + \sin^2\theta\, d\phi^2)).$$

This is the $k = 0$ Friedmann solution with $p = \frac{1}{3}\mu$.

We shall be interested in similarity solutions which are asymptotically Friedmann, i.e. in which x and S approach the form (5.1) asymptotically at large z. It is therefore convenient to introduce new variables A and B defined by

$$\begin{cases} x = 2^{3/2} z^{-1/2}\, e^A \\ S = 2^{11/6} z^{-1/2}\, e^B. \end{cases} \tag{5.3}$$

In these variables equations (3.14), (3.15), (3.18), (3.19), (3.20) become

$$\ddot{B} = \dot{A}+4\dot{A}\dot{B}-\tfrac{3}{2}\dot{B}-3\dot{B}^2 \tag{5.4}$$

$$\dot{A}\left(3-\frac{1}{V^2}\right) = 2\dot{B}-\tfrac{1}{2}(e^{-2A}-1) \tag{5.5}$$

$$V = 4^{-1/3}\sqrt{z}\, e^{2A-2B} \tag{5.6}$$

$$2M = 2^{-4/3}z(1+4\dot{B})\, e^{2B-4A} \tag{5.7}$$

$$2M = 1+4\, e^{6B-6A}[V^2(\dot{B}-\tfrac{1}{2})^2-(\dot{B}+\tfrac{1}{2})^2]. \tag{5.8}$$

Given the values of A, B and $\dot{B}$, equations (5.7) and (5.8) determine the value of z, equation (5.6) determines V, equation (5.4) determines $\ddot{B}$ and, provided $V \neq 1/\sqrt{3}$, equation (5.5) determines $\dot{A}$. This means that at each point of the three-dimensional $(A, B, \dot{B})$ space there is a little arrow pointing to the value of A, B and $\dot{B}$ at $\Omega+\delta\Omega$. In other words equations (5.4)–(5.8) define a vector field on $(A, B, \dot{B})$ space. Integral curves of this vector field are parametrized by Ω (or equivalently z) and represent solutions to (5.4)–(5.8). There is thus a two-parameter family of spherically symmetric similarity solutions with $p = \frac{1}{3}\mu$. The Friedmann $k = 0$ solution is represented by the degenerate integral curve consisting of the origin alone.

In $(A, B, \dot{B})$ space a special role is played by the surface $V = 1/\sqrt{3}$. This is given by

$$\dot{B} = -1-\tfrac{1}{4}\, e^{-2A}+\tfrac{1}{4}\sqrt{12+6\, e^{-2A}+e^{-4A}+6\, e^{6A-6B}}. \tag{5.9}$$

On this surface equation (5.5) does not determine $\dot{A}$. This means that there can be a number of different integral curves representing a number of different solutions passing through the same point. However, it is easy to see that these integral curves cannot be extended beyond the point where they intersect $V = 1/\sqrt{3}$ unless $g = 0$ where g is defined to be the right-hand side of (5.5):

$$g \equiv 2\dot{B}-\tfrac{1}{2}(e^{-2A}-1). \tag{5.10}$$

The two-surface $g = 0$ in $(A, B, \dot{B})$ space will intersect the two-surface $V = 1/\sqrt{3}$ in a line Q. Any integral curve which intersects $V = 1/\sqrt{3}$ other than at the line Q cannot be continued further. From each point q of Q there will in general be two one-parameter families of integral curves, one with decreasing and one with increasing z. One can join any member of the first to any member of the second to obtain a similarity solution. This arbitrariness arises from the fact that the surface of constant z in the solution on which $V = 1/\sqrt{3}$ represents a sphere expanding from the origin at the speed of sound. At this wavefront new information can be fed into the solution.

The $k = 0$ Friedmann solution is given by $A = B = 0$ for all z. The asymptotic form of similarity solutions which approach the Friedmann solution for large z can be found by linearizing equations (5.4) and (5.5) in $\dot{A}$, A and B and neglecting the $(1/V)^2$ term:

$$\ddot{B} = -\tfrac{3}{2}\dot{B}+\dot{A} \tag{5.11}$$

$$3\dot{A} = 2\dot{B}+A. \tag{5.12}$$

These equations have the solution

$$A = -kz^{-1}+2k'z^{1/2} \tag{5.13}$$

$$B = -\beta-2kz^{-1}+k'z^{1/2} \tag{5.14}$$

where k, k' and β are constants. The relative spatial density gradient $\mu^{-1}/(\partial\mu/\partial r)$ equals $-4\dot{A}/r$. For an asymptotically Friedmann solution this goes to zero at large z so k' must be zero. Using equations (5.7) and (5.8) one can relate k and β by

$$k = \frac{4^{2/3}(e^{2\beta}-e^{-4\beta})}{18}. \tag{5.15}$$

The behaviour at large z of asymptotically Friedmann similarity solutions is thus determined by one parameter, k. The fact that B tends to a non-zero value $(-\beta)$ does not invalidate the linearization of equations (5.4) and (5.5) since B enters into these equations only through the factor $(1/V^2)$ which can be neglected to first order.

In the $k = 0$ Friedmann solution $V = 4^{-1/3}z^{1/2}$. The surface $z = 4^{2/3}$ on which $V = 1$ can be regarded as the particle horizon as in the $p = 0$ case. For a similarity solution to represent a black hole in an asymptotically Friedmann universe one requires that V has a similar $z^{1/2}$ behaviour at large z but that at smaller values of z it should have a minimum below 1. In such a solution $V = 1$ at two values of z, z_1 and z_2 $(z_1 < z_2)$, which represent respectively the event horizon of the black hole and the particle horizon of the Universe. We shall now show that A and B have to be negative and $\dot{A}$ and $\dot{B}$ have to be positive for all z. This will imply that k must be positive and that the solution cannot be exactly Friedmann at large z.

At z_1,

$$\dot{V} = \tfrac{1}{2}+2\dot{A}-2\dot{B}<0.$$

It follows from (5.5) that $A(z_1)<0$. Suppose that $\dot{A}(z_1)<0$. Then since A is continuous and zero at $z = \infty$ it must have a local minimum in (z_1, ∞). Let z_3 by the smallest value of z greater than z_1 at which A has a local minimum. At z_3 either $\dot{A} = 0$ or $V = 1/\sqrt{3}$. In either case

$$\dot{B}(z_3) = \tfrac{1}{4}(e^{-2A(z_3)}-1)>0$$

because $A(z_3)<A(z_1)<0$. This shows also that $\dot{B}(z_3)>\dot{B}(z_1)$. Hence $\ddot{B}>0$ somewhere in (z_1, z_3). But, since $\dot{A}\leqslant 0$ in (z_1, z_3),

$$\ddot{B} = \dot{A}+4\dot{A}\dot{B}-\tfrac{3}{2}\dot{B}-3\dot{B}^2>0 \quad \text{only if} \quad \dot{B}<0.$$

Since $\dot{B}(z_3)>0$, there would then exist a point z_4 in (z_1, z_3) where $\dot{B} = 0$ and $\ddot{B}>0$. But at z_4, $\ddot{B} = \dot{A}<0$. This establishes a contradiction which shows that the original assumption, $\dot{A}(z_1)<0$, was false. Thus $\dot{A}(z_1)\geqslant 0$. A similar argument shows that $\dot{A}\geqslant 0$ for $z\geqslant z_1$ provided that A remains less than or equal to 0. Suppose now that $A>0$ for some $z>z_1$ and let z_5 be the value of z at which A has its first positive local maximum (z_5 must exist since $A\to 0$ as $z\to\infty$). Then $\dot{A}\geqslant 0$ in $[z_1, z_5]$. At z_5,

$$\dot{B} = \tfrac{1}{4}(e^{-2A}-1)<0.$$

However, $\dot{B}(z_1)>\tfrac{1}{4}+\dot{A}(z_1)>0$. Therefore B has a maximum at some z_6 $(z_1<z_6<z_5)$. Since $\dot{B}$ is continuous there will be some $\epsilon>0$ such that $|\dot{B}|<\tfrac{1}{8}$ in $[z_6-\epsilon, z_6+\epsilon]$. Then from equation (5.4)

$$\ddot{B}>\tfrac{1}{2}\dot{A}-\tfrac{3}{2}\dot{B}-\tfrac{3}{8}|\dot{B}|. \tag{5.16}$$

One can integrate this to show that

$$\dot{B}(z_6)-\dot{B}(z_6-\epsilon)\geqslant C(A(z_6)-A(z_6-\epsilon)) \tag{5.17}$$

for some $C > 0$. This shows that $\dot{B}(z_6 - \epsilon) \leqslant 0$. However, $\dot{B}(z_6 - \epsilon) \geqslant 0$ since z_6 is the first local maximum of B for $z > z_1$. It follows from this that $\dot{B}$ would have to be 0 in $[z_6 - \epsilon, z_6]$. By repeating the above steps one could show that $\dot{B}$ was 0 for all z in the interval $[z_1, z_6]$. This establishes a contradiction which shows that the assumption that A went positive was false. Thus $A \leqslant 0$ and $\dot{A} \geqslant 0$ for all $z \geqslant z_1$.

The above proof shows that the constant k which governs the asymptotic behaviour of A and B as z goes to infinity must be greater than or equal to 0. Suppose that k were 0, i.e. that for large values of z the solution was exactly the Friedmann solution. By equations (5.4) and (5.5) the solutions for A and B are analytic functions of z except possibly where $V = 1/\sqrt{3}$. Thus A and B would be zero for $z \geqslant 4^{2/3}/3$, the value at which $V = 1/\sqrt{3}$ in the Friedmann solution. When $V = 1/\sqrt{3}$ equation (5.5) does not determine $\dot{A}$ and so it is possible for A and B to be non-zero for $z < 4^{2/3}/3$. Such a solution would correspond to a sound-wave expanding out from the origin into an exactly Friedmann universe. If $\dot{A}$ were negative for z just below $4/3^{2/3}$, the sound-wave would be a decompression wave and the matter would be accelerated inwards by the pressure gradient. However, the argument above shows that no black hole solution can have $\dot{A}$ negative at any point. This shows that there can be no similarity solution containing a black hole which is exactly Friedmann at large values of z.

The conclusion is therefore that in an asymptotically Friedmann solution containing a black hole the constant k must be positive. One can then use a similar argument as above to show that $\dot{A}$ and $\dot{B}$ must be positive everywhere and that A and B must be negative everywhere. The fact that $\dot{A}$ is positive implies that the pressure gradient is always directed outwards which means that pressure is hindering the accretion of matter into the black hole rather than helping it.

One can evaluate the quantity E which represents the total energy per unit mass of matter in these models. Unlike the $p = 0$ case, this will not be an exact constant of the motion. However, it is zero in the Friedmann model and is

$$E(z) = -18.2^{-7/3} k \, e^{-2\beta} + O(z^{-1})$$

in the asymptotically Friedmann similarity solutions. Thus in order to make the matter fall into the black hole fast enough to make the black hole grow at the same rate as the particle horizon one has to reduce its energy E at large spacelike separations on an initial surface by more than the amount $-\Delta m/R$ which would arise from the Coulomb field of the initial mass Δm of the black hole at the origin. If the black hole is formed by local fluctuations, there is no reason to believe that the initial energy of matter around it would be reduced in this way. We therefore conclude that the black hole will not grow as fast as the particle horizon. Once the particle horizon is bigger than the black hole, we would not expect the black hole to grow much more by the Zeldovich–Novikov argument.

Even though they may not occur in nature, it is of some interest to examine the asymptotically Friedmann similarity solutions which do contain black holes. We have integrated numerically equations (5.4) and (5.5) using the boundary conditions

$$A = -kz^{-1} \tag{5.17}$$

$$B = -\beta - 2kz^{-1} \tag{5.18}$$

appropriate to an asymptotically Friedmann universe at large z. For k large and positive the quantity V has $z^{1/2}$ behaviour as $z \to \infty$ as in the Friedmann solution.

At smaller values of z, V has a minimum above 1 and tends to infinity for some positive value of z (curve (a) in Fig. 2). In these solutions there are no particle or event horizons though there are black hole and cosmological apparent horizons. Any observer would inevitably fall through the apparent horizon and hit the singularity. These solutions, in which the whole Universe is inside the black hole, correspond to the $p = 0$ solutions with fairly negative E. As one makes k smaller there is a range of values for which the minimum of V occurs between 1 and $1/\sqrt{3}$ (curve (b) in Fig. 2). These solutions have particle and event horizons and represent

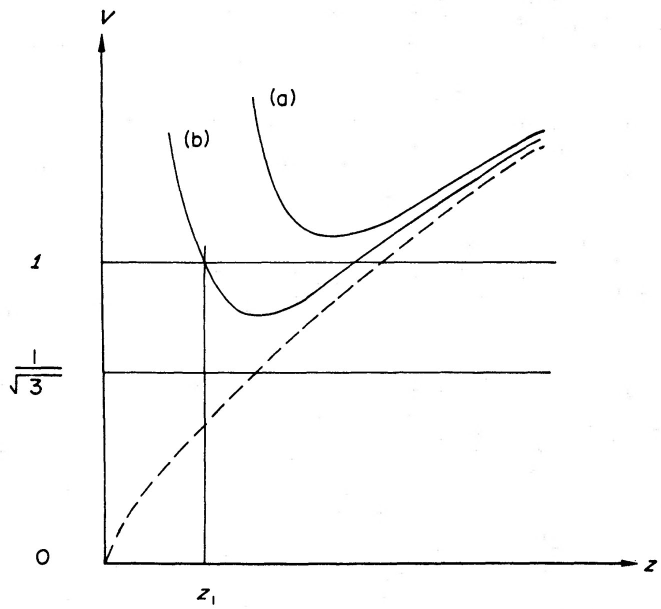

FIG. 2. *Graph of V against z for $p = \frac{1}{3}\mu$ similarity solutions. Curve (a) has a minimum above 1. There are no particle or event horizons and the whole Universe is inside the black hole. Curve (b) has a minimum between 1 and $1/\sqrt{3}$. Particle and event horizons exist in this solution which represents a black hole expanding at the same rate as the Universe.*

black holes which grow at the same rate as the Universe. Because V is always greater than $1/\sqrt{3}$ these solutions are acausal in that no sound-wave can propagate outwards. As in the $p = 0$ case the black hole's apparent horizon lies inside the event horizon. As k is further reduced, the minimum of V approaches $1/\sqrt{3}$. We can show that there is no solution with a minimum actually at $1/\sqrt{3}$. If the curve $V(z)$ intersects $1/\sqrt{3}$, it seems probable that it cannot rise above $1/\sqrt{3}$ again and reach 1. If this is the case then all black hole similarity solutions are acausal in the sense mentioned above and unlike the $p = 0$ case, there is a minimum ratio between the size of the black hole and the size of the particle horizon.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to J. E. Pringle and M. J. Rees for helpful discussions.

Institute of Astronomy and Department of Applied Mathematics and Theoretical Physics, University of Cambridge

REFERENCES

Boynton, P. E. & Partridge, R. B., 1973. *Astrophys. J.*, **181**, 243.
Cahill, A. H. & Taub, M. E., 1971. *Commun. Math. Phys.*, **21**, 1.
Geroch, R. P., Kronheimer, E. H., Penrose, R., 1972. *Proc. R. Soc. Lond.*, **A327**, 545.
Hagedorn, R., 1970. *Astr. Astrophys.*, **5**, 184.
Hawking, S. W., 1971. *Mon. Not. R. astr. Soc.*, **152**, 75.
Hawking, S. W., 1973. In *Black holes*, eds C. M. DeWitt and B. S. DeWitt, Gordon and Breach, New York, London, Paris.
Hawking, S. W. & Ellis, G. F. R., 1973. *The large scale structure of spacetime*, Cambridge University Press, London.
Penrose, R., 1964. *Relativity, groups and topology*, eds C. M. DeWitt and B. S. DeWitt, Gordon and Breach, New York.
Sandage, A., 1961. *Astrophys. J.*, **133**, 355.
Van den Bergh, S., 1969. *Nature*, **224**, 891.
Zeldovich, Y. B. & Novikov, I. D., 1967. *Sov. Astr. AJ.*, **10**, 602.
Zeldovich, Y. B. & Sunyaev, R. A., 1969. *Astrophys. Space Sci.*, **4**, 301.