Gravitational collapse and space-time singularities

Re-written from original hardcopy by K. Braun (not authorized by the author)

Roger Penrose
Department of Mathematics, Birkbeck College, London, England
(Received 18 December 1964)

Physical Review Letters, Vol. 14, No 3, 18 January 1965

The discovery of the quasistellar radio sources has stimulated renewed interest in the question of gravitational collapse. It has been suggested by some authors [1] that the enormous amounts of energy that these objects apparently emit may result from the collapse of a mass of the order of $(10^6 - 10^8)M_{\otimes}$ to the neighborhood of its Schwarzschild radius, accompanied by a violent release of energy, possible in the form of gravitational radiation. The detailed mathematical discussion of such situations is difficult since the full complexity of general relativity is required. Consequently, most <u>exact</u> calculations concerned with the implications of gravitational collapse have employed the simplifying assumption of spherical symmetry. Unfortunately, this precludes any detailed discussion of gravitational radiation – which requires at least a quadripole structure. The general situation with regard to a spherically symmetrical body is well known [2]. For a sufficiently great mass, there is no final equilibrium state. When sufficiently thermal energy has been radiated away, the body contracts and continues to contract until a <u>physical singularity</u> is encountered at $r = 0$. As measured by local comoving observers, the body passes <u>within</u> its Schwarzschild radius $r = 2m$. (The densities at which this happens need not be enormously high if the total mass is large enough). To an outside observer the contraction to $r = 2m$ appears to make infinite time. Nevertheless, the existence of a singularity presents a serious problem for any <u>interior</u> region.

The question has been raised as to whether this singularity is, in fact, simply a property of the high symmetry assumed. The matter collapses radially inwards to the <u>single point</u> at the center, so that a resulting space-time catastrophe there is perhaps not surprising. Could not the presence of perturbations which destroy the spherical symmetry alter the situation drastically? The recent rotating solution of Kerr [3] also possesses a physical singularity, but since a high degree of symmetry is still present (and the solution is algebraically special), it might again be argued that this is not representative of the general situation [4]. Collapse <u>without</u> assumptions of symmetry will be discussed here.

Consider the time development of a Cauchy hypersurface C^3 representing an initial matter distribution. We may assume Einstein's field equations and suitable equations of the state governing the matter. In fact, the <u>only</u> assumption made hereabout these equations of state will be the non-negative definiteness of Einstein's energy expression (with or without cosmological term). Suppose this matter distribution undergoes gravitational collapses in a way which, at first, qualitatively resembles the spherically symmetrical case. It will be shown that, after a certain critical condition has been fulfilled, <u>derivations from spherical symmetry cannot prevent space-time singularities from arising</u>.

If, as seems justifiable, actual physical singularities in space-time are not to be permitted to occur, the conclusion would appear inescapable that inside such a collapsing object at least one of the following holds:

(a) Negative local energy occurs. [6]

(b) Einstein's equations are violated.

(c) The space-time manifold is incomplete. [7]

(d) The concept of space-time loses its meaning at very high curvature – possible because of quantum phenomena. [2]

In fact (a), (b), (c), (d) are somewhat interrelated, the distinction being partly one of attitude of mind.

Before examining the asymmetrical case, consider a sherically symmetrical matter distribution of finite radius in C^3 which collapses symmetrically. The empty region surrounding the matter will, in this case, be a Schwarzschild field, and we can conveniently use the metric

$$ds^2 = -2drdv + dv^2(1 - 2m/r) - r^2(d\theta^2 + \sin^2\theta d\varphi^2),$$

with an advanced time parameter to describe it [8]. The situation is depicted in figure 1. Note that an exterior observer will always see matter outside $r = 2m$, the collapse through $r = 2m$ to the singularity at $r = 0$ being invisible to him.

After the matter has contracted within $r = 2m$, a spacelike sphere S^2 ($t = const$ $2m > r = const$) can be found in the empty region surrounding the matter. This sphere is an example of what will be called here a trapped surface – defined generally as a closed, spacelike, two-surface T^2 with the property that the two systems of null geodesics which meet T^2 orthogonally converge locally in future directions at T^2. clearly trapped surfaces will still exist if the matter region has no sharp boundary or if spherical symmetry is dropped, provided that the deviations fron the above situation are not too great.

Indeed, the Kerr solutions with $m > a$ (angular momentum ma) all possess trapped surfaces, whereas those for which $m \leq a$ do not [9]. The argument will be to show that the existence of a trapped surface implies –irrespective of symmetry – that singularities necessarily develop.

The existence of a singularity can never be interfered, however, without an assumption such as completeness for the manifold under consideration. It will be necessary, here, to suppose that the manifold M_+^4, which is the future time development of an initial Cauchy hypersurface C^3 (past boundary of the M_+^4 region), is in fact <u>null complete</u> into the future.

The various assumptions are, more precisely, as follows:

(i) M^4_+ is a nonsingular $(+---)$ Riemannian manifold for which the null half-cones form two separate systems ("past" and "future").

(ii) Every null geodesic in M^4_+ can be extended into the future to arbitrarily large affine parameter values (null completeness).

(iii) Every timelike or null geodesic in M^4_+ can be extended into the past until it meets C^3 (Cauchy hypersurface condition).

(iv) At every point of M^4_+ all timelike vectors t^μ satisfy

$$(-R_{\mu\nu} + \frac{1}{2} R g_{\mu\nu} - \lambda g_{\mu\nu}) t^\mu t^\nu$$

(non-negativeness of local energy).

(v) There exists a trapped surface T^2 in M^4_+.

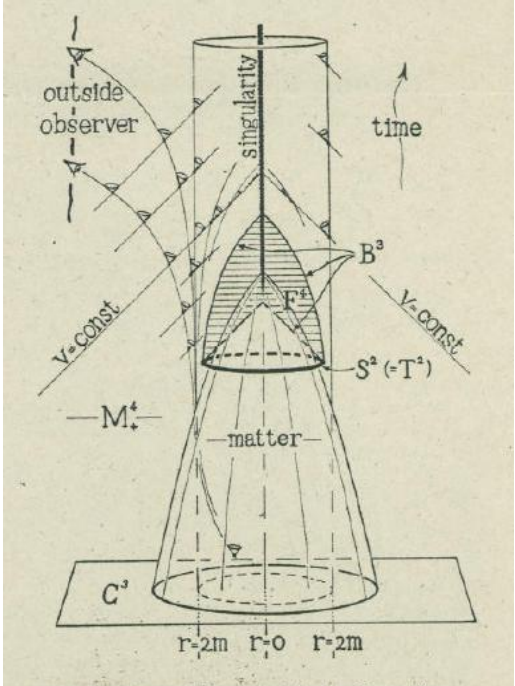

FIG. 1. Spherically symmetrical collapse (one space dimension surpressed). The diagram essentially also serves for the discussion of the asymmetrical case.

It will be shown here, in outline, that (i) – (v) are together inconsistent.

Let F^4 be the set of points in M^4_+ which can be connected to T^2 by a smooth timelike curve leading into the future from T^2. Let B^3 be the boundary of F^4. Local considerations show that B^3 is <u>null</u> where it is nonsingular, being generated by the null geodesic segments which meet T^2 orthogonally at the past endpoint and have a future endpoint if this is a singularity (on a caustic or crossing region) of B^3.

Let l^μ (subject to $l^\mu{}_{;\nu} l^\nu = 0$), ρ ($= -\frac{1}{2} l^\mu{}_{;\mu}$), and $|\sigma|$ ($= \left[\frac{1}{2} l_{(\mu;\nu)} l^{\mu;\nu} - \frac{1}{4} (l^\mu{}_{;\mu})^2\right]^{1/2}$)

be, respectively, a future-pointing tangent vector, the convergence, and the shear for these null geodesics, [10] and let A a be a corresponding infinitesimal area of cross section of B^3. Then

$$\left[(A^{1/2})_{;\mu} l^{\mu;}\right]_{,\nu} l^\nu = -(A^{1/2}\rho)_{;\mu} l^{\mu;} = -(A^{1/2}(|\sigma|^2 + \Phi) \leq 0$$

where

$$\Phi = -\frac{1}{2} R_{\mu\nu} l^\mu l^\nu \geq 0 \quad \text{(by (iv)).}$$

Since T^2 is trapped, $\rho > 0$ at A, whence A becomes zero at a finite affine distance to the future of on each null geodesic. Each geodesic thus encounters a caustic. Hence B^3 is compact (closed), being generated by a compact system of finite segments. We may approximate B^3 arbitrarily closely by a smooth, closed, spacelike hypersurface B^{3*} . Let K^4 denote the set of pairs (P,s) with $P \in B^{3*}$ and $0 \le s \le 1$. Define a continuous map $\mu : K^4 \to M_+^4$ where, for fixed P , $\mu\{(P,s)\}$ is the past geodesic segment normal to B^{3*} at $P = \mu\{(P,1)\}$ and meeting C^3 (as it must, by (iii)) in the point $\mu\{(P,0)\}$. At each point Q of $\mu\{K^4\}$we can define the degree $d(Q)$ of μ to be the number of points of K^4 which map to Q (correctly counted). Over any region not containing the image of a boundary point of K^4, $d(Q)$ will be constant. Near B^{3*}, μ is 1-1, so $d(Q) = 1$. It follows that $d(Q) = 1$ near C^3 also, whence the degree of the map $B^{3*} \to C^3$ induced by μ when $s = 0$ must also be unity. The impossibility of this follows from the non-compactness of C^3.

Full details of this and other related results will be given elsewhere.

[1] F. Hoyle and W.A. Fowler, Monthly Notice Roy. Astron. Soc. 125, 169 (1963); E.M. Burbidge, Astrophys. J. 139, 909 (1964); W. A. Fowler, Rev. Mod. Phys. 36, 545 (1964); Ya.B. Zel'dovich and I.D. Novikov, Dokl. Akad. Nauk SSSR 155, 1033 (1964); I.S.Shklovskii and N.S. Kardashev Dokl. Akad. Nauk SSSR 156, 57 (1964), also various articles in the Proceedings of the 1963 Dallas Conference on Gravitational Collapse (University of Chicago, Illinois, 1964)

[2] J.R. Oppenheimer and H. Snyder, Phys. Rev., 56, 455 (1939). See also J.A. Wheeler, in Relativity, Groups and Topology, edited by C. deWitt and B. deWitt (Gordon and Breach Publishers., Inc.), New York, 1964, and reference 1

[3] R.P. Kerr, Phys. Rev. Letters 11, 237 (1963)

[4] see also E.M. Lifshitz and I.M. Khalatnikov, Advan. Phys. 12, 185 (1963)

[5] see also P.G. Bergmann, Phys. Rev. Letters, 12, 139 (1964)

[6] The negative energy of a "C field" may be invoked to avoid singularities: F. Hoyle and J.V. Narlikar, Proc. Roy. Soc (London) A278, 465 (1964). However, it is difficult to see how even the presence of negative energy could lead to an effective "bounce" if local causality is to be maintained.

[7] The "I'm all right, Jack" philosophy with regard to singularities would be included under this heading!

[8] D. Finkenstein, Phys. Rev. 110, 965 (1959).

[9] The case m<a is interesting in that here a singularity is "visible" to an outside observer. Whether or not "visible" singularities inevitable arise under appropriate circumstances is an intriguing question not covered by present discussion.

[10] For the notation, etc., see E. Newman and r. Penrose, J. Math. Phys. 3, 566 (1962)