

relativistic units, are chosen, so that

$$c = G = 1,$$

that is to say, we translate our units according to

$$1 \text{ s} = 3 \cdot 10^{10} \text{ cm} = 4 \cdot 10^{36} \text{ g}.$$

When $r = 2m$, the metric form (1) breaks down. The radius $r = 2m$ is referred to as the *Schwarzschild radius* of the body.

Let us imagine a situation in which the collapse of a spherically symmetrical (nonrotating) star takes place and continues until the surface of the star approaches the Schwarzschild radius. So long as the star remains spherically symmetrical, its *external field* remains that given by the Schwarzschild metric (1). The situation is depicted in Fig. 1. Now the particles at the surface of the star must describe *timelike* lines. Thus, from the way that the «angle» of the light cones appears to be narrowing down near $r = 2m$, it would seem that the surface of the star can never cross to within the $r = 2m$ region. However, this is misleading. For suppose an observer were to follow the surface of the star in a rocket ship, down to $r = 2m$. He would find (assuming that the collapse does not differ significantly from free fall) that the total proper time that he would experience as elapsing, as he finds his way down to $r = 2m$, is in fact *finite*. This is despite the fact that the world line he follows has the *appearance* of an «infinite» line in Fig. 1. But what does the observer experience after this finite proper time has elapsed? Two possibilities which suggest themselves are: i) the observer encounters some form of space-time singularity—such as infinite tidal forces—which inevitably destroys him as he approaches $r = 2m$; ii) the observer enters some region of space-time not covered by the (t, r, θ, φ) co-ordinate system used in (1). (It would be unreasonable to suppose that the observer's experiences could simply *cease* after some finite time, without his encountering some form of violent agency.)

In the present situation, in fact, it is possibility ii) which occurs. The easiest way to see this is to replace the co-ordinate t by an advanced time parameter v given by

$$v = t + r + 2m \log(r - 2m),$$

whereby the metric (1) is transformed to the form (EDDINGTON [3], FINKELSTEIN [4])

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

Gravitational Collapse: the Role of General Relativity.

R. PENROSE

Department of Mathematics, Birkbeck College - London

Stars whose masses are of the same order as that of the sun ($M_\odot$) can find a final equilibrium state either as a white dwarf or, apparently, (after collapse and ejection of material) as a neutron star. These matters have been nicely discussed in the lectures of HEWISH and SALPETER. But, as they have pointed out, for larger masses no such equilibrium state appears to be possible. Indeed, many stars are observed to have masses which are much larger than $M_\odot$ —so large that it seems exceedingly unlikely that they can ever shed sufficient material so as to be able to fall below the limit required for a stable white dwarf ($\sim 1.3 M_\odot$; CHANDRASEKHAR [1]) or neutron star ($\sim 0.7 M_\odot$; OPPENHEIMER and VOLKOFF [2]) to develop. We are thus driven to consider the consequences of a situation in which a star collapses right down to a state in which the effects of general relativity become so important that they eventually dominate over all other forces.

I shall begin with what I think we may now call the «classical» collapse picture as presented by general relativity. Objections and modifications to this picture will be considered afterwards. The main discussion is based on Schwarzschild's solution of the Einstein vacuum equations. This solution represents the gravitational field exterior to a spherically symmetrical body. In the original Schwarzschild co-ordinates, the metric takes the familiar form

$$(1) \quad ds^2 = (1 - 2m/r) dt^2 - (1 - 2m/r)^{-1} dr^2 - r^2(d\theta^2 + \sin^2\theta d\varphi^2).$$

Here θ and φ are the usual spherical polar angular co-ordinates. The radial co-ordinate r has been chosen so that each sphere $r = \text{const.}$, $t = \text{const}$ has intrinsic surface area $4\pi r^2$. The choice of time co-ordinate t is such that the metric form is invariant under $t \rightarrow t + \text{const}$ and also under $t \rightarrow -t$. The static nature of the space-time is thus made manifest in the formal expression for the metric. The quantity m is the mass of the body, where «general-

This form of metric has the advantage that it does not become inapplicable at $r = 2m$. The whole range $0 < r < \infty$ is encompassed in a nonsingular fashion by (2). The part $r > 2m$ agrees with the part $r > 2m$ of the original expression (1). But now the region has been extended inwards in a perfectly regular way across $r = 2m$ and right down towards $r = 0$.

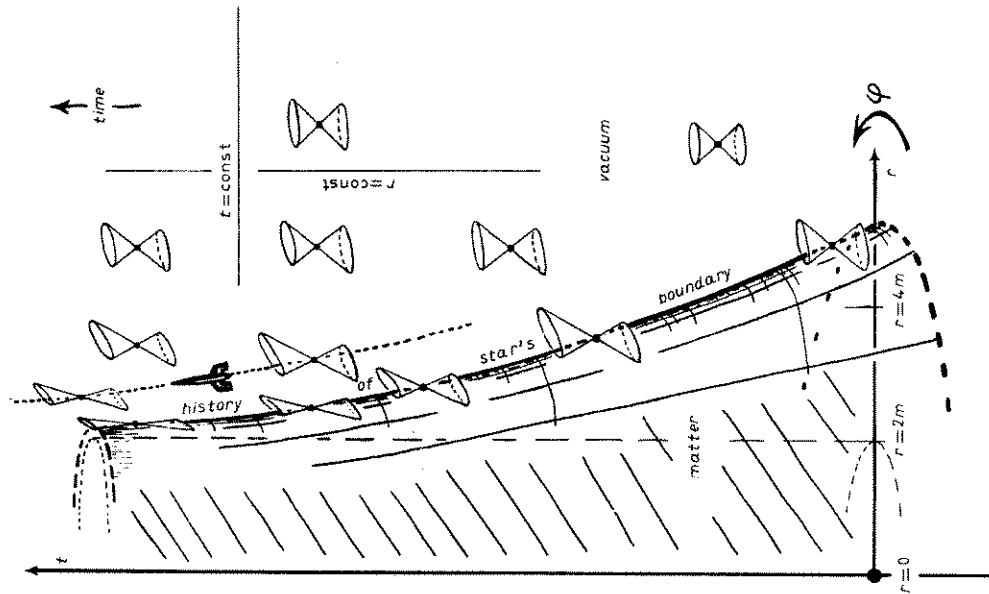


Fig. 1. - Spherically symmetrical collapse in the usual Schwarzschild co-ordinates.

The situation is as depicted in Fig. 2. The light cones tip over more and more as we approach the centre. In a sense we can say that the gravitational field has become so strong, within $r = 2m$, that even light cannot escape and is dragged inwards towards the centre. The observer on the rocket ship, whom we considered above, crosses freely from the $r > 2m$ region into the $0 < r < 2m$ region. He encounters $r = 2m$ at a perfectly finite time, according to his own local clock, and he experiences nothing special at that point. The spacetime there is locally Minkowskian, just as it is everywhere else ($r > 0$).

Let us consider another observer, however, who is situated far from the star. As we trace the light rays from his eye, back into the past towards the star, we find that they cannot cross into the $r < 2m$ region after the star has collapsed through. They can only intersect the star at a time *before* the star's surface crosses $r = 2m$. No matter how long the external observer waits, he can always (in principle) still see the surface of the star as it *was* just before it plunged through the Schwarzschild radius. In practice, however, he would soon see nothing of the star's surface—only a «black hole»—since the observed intensity would die off exponentially, owing to an infinite red shift.

Fig. 2. - Spherically symmetrical collapse in Edington-Finkelstein co-ordinates.

But what will be the fate of our original observer on the rocket ship? After crossing the Schwarzschild radius, he finds that he is compelled to enter regions of smaller and smaller r . This is clear from the way the light cones tip over towards $r = 0$ in Fig. 2, since the observer's world line must always remain a timelike line. As r decreases, the space-time curvature mounts (in proportion to r^{-3}), becoming theoretically infinite at $r = 0$. The physical effect of space-time curvature is experienced as a *tidal force*: objects become squashed in one direction and stretched in another. As this tidal effect mounts to infinity, our observer must eventually (*) be torn to pieces—indeed, the very atoms of which he is composed must ultimately individually share this same fate!

Thus, the true *space-time singularity*, resulting from a spherically symmetrical collapse, is located not at $r = 2m$, but at $r = 0$. Although the hypersurface $r = 2m$ has, in the past, itself been frequently referred to as the « Schwarzschild singularity », this is really a misleading terminology since $r = 2m$ is a singularity merely of the t co-ordinate used in (1) and not of the space-time geometry. More appropriate is the term « event horizon », since $r = 2m$ represents the absolute boundary of the set of all events which can be observed in principle by an external inertial observer. The term « event horizon » is used also in cosmology for essentially the same concept (cf. RINDLER [5]). In the present case the horizon is less observer-dependent than in the cosmological situations, so I shall tend to refer to the hypersurface $r = 2m$ as the *absolute event horizon* (**) of the space-time (2).

This, then, is the standard spherically symmetrical collapse picture presented by general relativity. But do we have good reason to trust this picture? Need we believe that it necessarily accords, even in its essentials, with physical reality? Let me consider a number of possible objections:

- a) densities in excess of nuclear densities inside,
- b) exact vacuum assumed outside,
- c) zero net charge and zero magnetic field assumed,
- d) rotation excluded,
- e) asymmetries excluded,

(*) In fact, if m is of the order of a few solar masses, the tidal forces would already be easily large enough to kill a man in free fall, even at $r = 2m$. But for $m > 10^8 M_\odot$ the tidal effect at $r = 2m$ would be no greater than the tidal effect on a freely falling body near the Earth's surface.

(**) In a general space-time with a well-defined external future infinity, the absolute event horizon would be defined as the boundary of the union of all timelike curves which escape to this external future infinity. In the terminology of PENROSE [6], if $\mathcal{H}$ is a weakly asymptotically simple space-time, for example, then the absolute event horizon in $\mathcal{H}$ is $I_-[\mathcal{I}^+]$.

- f) possible λ -term not allowed for,
- g) quantum effects not considered,
- h) general relativity a largely untested theory,
- i) no apparent tie-up with observations.

As regards a), it is true that for a body whose mass is of the order of $M_\odot$, its surface would cross $r = 2m$ only after nuclear densities had been somewhat exceeded. It may be argued, then, that too little is understood about the nature of matter at such densities for us to be at all sure how the star would behave while still outside $r = 2m$. But this is not really a significant consideration for our general discussion. It could be of relevance only for the least massive collapsing bodies, if at all. For, the larger the mass involved, the smaller would be the density at which it would be expected to cross $r = 2m$. It could be that very large masses indeed may become involved in gravitational collapse. For $m > 10^{11} M_\odot$ (e.g. a good-sized galaxy), the averaged density at which $r = 2m$ is crossed would be less than that of air!

The objections b), c), d), e) and, to some extent, f) can all be partially handled if we extract, from Fig. 2, only that essential qualitative piece of information which characterizes the solution (2) as describing a collapse which has passed a « point of no return ». I shall consider this in more detail shortly. The upshot will be that if a collapse situation develops in which deviations from (2) near $r = 2m$ at one time are not too great, then two consequences are to be inferred as to the subsequent behaviour. In the first instance an absolute event horizon will arise. Anything which finds itself inside this event horizon will not be able to send signals to the outside worlds. Thus, in this respect at least, the qualitative nature of the « $r = 2m$ » hypersurface in (2) will remain. Similarly, an analogue of the physical singularity at $r = 0$ in (2) will still develop in these more general situations. That is to say, we know from rigorous theorems in general-relativity theory that there must be *some* space-time singularity resulting inside the collapse region. However, we do not know anything about the detailed *nature* of this singularity. There is no reason to believe that it resembles the $r = 0$ singularity of the Schwarzschild solution very closely.

In regard to c), d) and f) we can actually go further in that *exact* solutions are known which generalize the metric (2) to include angular momentum (KERR [7]) and, in addition, charge and magnetic moment (NEWMAN *et al.* [8]), where a cosmological constant may also be incorporated (CARTER [9]). These solutions appear to be somewhat special in that, for example, the gravitational quadrupole moment is fixed in terms of the angular momentum and the mass, while the magnetic-dipole moment is fixed in terms of the angular momentum, charge and mass. However, there are some reasons for believing that these solutions may actually represent the general exterior asymptotic limit resulting

from the type of collapse we are considering. Any extra gravitational multipole moments of quadrupole type, or higher, can be radiated away by gravitational radiation; similarly, extra electromagnetic multipole moments of dipole type, or higher, can be radiated away by electromagnetic radiation. (I shall discuss this a little more later.) If this supposition is correct, then e will to some extent also be covered by an analysis of these exact solutions. Furthermore, b would, in effect, be covered as well, provided we assume that all matter (with the exception of electromagnetic field—if we count that as «matter») in the neighbourhood of the «black hole», eventually falls into the hole. These exact solutions (for small enough angular momentum, charge and cosmological constant) have absolute event horizons similar to the $r = 2m$ horizon in (2). They also possess space-time curvature singularities, although of a rather different structure from $r = 0$ in (2). However, we would not expect the detailed structure of these singularities to have relevance for a generically perturbed solution in any case.

It should be emphasized that the above discussion is concerned only with collapse situations which do not differ too much initially from the spherically symmetrical case we originally considered. It is not known whether a gravitational collapse of a *qualitatively different character* might not be possible according to general relativity. Also, even if an absolute event horizon *does* arise, there is the question of the «stability» of the horizon. An «unstable» horizon might be envisaged which itself might develop into a curvature singularity. These, again, are questions I shall have to return to later.

As for the possible relevance of gravitational quantum effects, as suggested in g), this depends, as far as I can see, on the existence of regions of space-time where there are extraordinary local conditions. If we assume the existence of an absolute event horizon along which curvatures and densities remain small, then it is very hard to believe that a classical discussion of the situation is not amply adequate. It may well be that quantum phenomena have a dominating influence on the physics of the deep interior regions. But whatever effects this might have, they would surely not be observable from the outside. We see from Fig. 2 that such effects would have to propagate outwards in *spacelike* directions over «classical» regions of space-time. However, we must again bear in mind that these remarks might not apply in some qualitatively different type of collapse situation.

We now come to h), namely the question of the validity of general relativity in general, and its application to this type of problem in particular. The inadequacy of the observational data has long been a frustration to theorists, but it may be that the situation will change somewhat in the future. There are several very relevant experiments now being performed, or about to be performed. In addition, since it has become increasingly apparent that «strong» gravitational fields probably play an important role in some astrophysical phenomena, there appears to be a whole new potential testing-ground for the theory.

Among the recently performed experiments, designed to test general relativity, one of the most noteworthy has been that of DICKE and GOLDENBERG [10], concerning the solar oblateness. Although the results have seemed to tell against the pure Einstein theory, the interpretations are not really clear-cut and the matter is still somewhat controversial. I do not wish to take sides on this issue. Probably one must wait for further observations before the matter can be settled. However, whatever the final outcome, the oblateness experiment had, for me, the importance of forcing me to examine, once more, the foundations of Einstein's theory, and to ask what parts of the theory are likely to be «here to stay» and what parts are most susceptible to possible modification. Since I feel that the «here to stay» parts include those which were most revolutionary when the theory was first put forward, I feel that it may be worth-while, in a moment, just to run over the reasoning as I see it. The parts of the theory I am referring to are, in fact, the geometrical interpretation of gravity, the curvature of space-time geometry and general-relativistic causality. These, rather than any particular field equations, are the aspects of the theory which give rise to what perhaps appears most immediately strange in the collapse phenomenon. They also provide the physical basis for the major part of the subsequent mathematical discussion.

To begin with, let us agree that it is legitimate to regard space-time as constituting a four-dimensional smooth manifold (or «continuum»). I do not propose to give a justification of this, because on an ordinary macroscopic level it is normally taken as «obvious». (On the other hand, I think that at a deeper submicroscopic level it is almost certainly «false», but this is not likely to affect the normal discussion of space-time structure—except perhaps at a space-time singularity!) Next, we must establish the existence of a *physically well-defined metric* ds which defines for our manifold a (pseudo-) Riemannian structure, with signature $(+ - - -)$. The meaning of ds is to be such that when integrated along the world line of any particle, it gives the lapse of proper times as experienced by that particle. Thus, the existence of ds depends on the existence of *accurate clocks* in nature. These clocks must behave locally according to the laws of special relativity. Also, for any two such clocks following the same world line, the time rates they register must agree with one another along the line and should not depend on, say, differing histories for the two clocks. That such clocks do seem to exist in nature, in effect, is a consequence of the fact that any mass m has associated with it a natural frequency mh^{-1} . Thus, the existence of accurate clocks comes down ultimately, via quantum mechanics, to the existence of well-defined masses in nature, whose relative values are in strict proportion throughout space-time. Of course, it might ultimately turn out that the mass ratios of particles are not constant throughout space-time. Then different particles might define slightly different (conformally related)

metrics for space-time. But the evidence at present is strongly against any appreciable difference existing.

If two neighbouring events in space-time have a separation such that $ds^2 > 0$, then according to special relativity, it is possible for one to have a causal influence on the other; if $ds^2 < 0$, then it is not. We expect this to persist also on a global scale. Thus, it is possible, of two events, for one to influence the other causally if and only if there is a timelike or null curve connecting them.

The existence of a physically well-defined metric and causal structure for space-time, then, seems to be fairly clearly established. It is not so clear, however, that this metric, as so defined, is going to be nonflat. However, we can take the experiment of POUND and REEKA [11] as almost a direct measurement establishing the nonflat nature of space-time. (For this, strictly speaking, the experiment would have to be repeated at various points on the Earth's surface.) The measured ds near the Earth's surface and the ds further from the Earth's surface cannot both be incorporated into the same Minkowskian framework because of the 'clock slowing' effect (cf. SCHILD [12]). Furthermore, owing to energy balance considerations it is clear that it is with *gravitational* fields that this 'clock slowing' effect occurs (owing to the fact that it is *energy*, i.e. *mass* which responds to a gravitational field). Thus gravitation must be directly related to space-time curvature.

Since we have a (pseudo-) Riemannian manifold, we can use the standard techniques of differential geometry to investigate it. In particular, we can construct a *physically meaningful* Riemann tensor $R_{\alpha\beta\gamma\delta}$ and thence its Einstein tensor

$$G_{\alpha\beta} = R_{\alpha\beta} - \frac{1}{2} R g_{\alpha\beta}.$$

Because of the contracted Bianchi identities we know that this satisfies the usual vanishing divergence law. But we also have a symmetric tensor $T_{\alpha\beta}$, namely the local energy-momentum tensor (composed of all fields but gravitation), which must satisfy a similar vanishing divergence law. It does not then necessarily follow that

$$(3) \quad G_{\alpha\beta} + \lambda g_{\alpha\beta} = -8\pi T_{\alpha\beta}$$

for some constant λ , but it is worth remarking that if we do *not* postulate this equation, then we have not just one, but *two* (linearly unrelated) conserved 'energylike' quantities, namely $G_{\alpha\beta}$ and $T_{\alpha\beta}$. In fact, this is just what happens in the theory of BRANS and DICKE [13]. (Such a motivation for the choice of Einstein's field equations (3), does not to my mind have quite the force of the earlier argument, so alternatives to (3) are certainly well worth considering.) Finally, the geodesic motion of monopole test particles may be taken as a consequence

of the vanishing divergence condition on $T_{\alpha\beta}$ (EINSTEIN and GROMMER [14]).

So I want to admit the possibility that Einstein's field equations may be wrong, but *not* (that is, in the macroscopic realm, and where curvatures or densities are not fantastically large) that the general pseudo-Riemannian geometric framework may be wrong. Then the mathematical discussion of the collapse phenomenon can at least be applied. It is interesting that the general mathematical discussion of collapse actually uses very little of the details of Einstein's equations. All that is needed is a certain inequality related to positiveness of energy. In fact, the adoption of the Brans-Dicke theory in place of Einstein's would make virtually no qualitative difference to the collapse discussion.

The final listed objection to the collapse picture is h , namely the apparent lack of any tie-up with observed astronomical phenomena. Of course it could be argued that the prediction of the 'black hole' picture is simply that we will not see anything—and this is precisely consistent with observations since no 'black holes' have been observed! But the real argument is really the other way around. Quasars *are* observed. And they apparently have such large masses and such small sizes that it would seem that gravitational collapse ought to have taken over. But quasars are also long-lived objects. The light they emit does not remotely resemble the exponential cut-off in intensity, with approach to infinite red shift, that might be inferred from the spherically symmetrical discussion. This has led a number of astrophysicists to question the validity of Einstein's theory, at least in its applicability to these situations.

My personal view is that while it is certainly possible (as I have mentioned earlier) that Einstein's equations may be wrong, I feel it would be very premature indeed to dismiss these equations just on the basis of the quasar observations. For, the *theoretical* analysis of collapse, according to Einstein's theory, is still more or less in its infancy. We just do not know, with much certainty, what the consequences of the theory really are. It would be a mistake to fasten attention just on those aspects of general-relativistic collapse which *are* known and to assume that this gives us essentially the complete picture. (It is perhaps noteworthy that many general-relativity theorists have a tendency, themselves, to be a bit on the sceptical side as regards the 'classical' collapse picture.) Since it seems to me that there are a number of intriguing largely unexplored possibilities, I feel it may be worth-while to present the 'generic' general-relativistic collapse picture as I see it, not only as regards the known theorems, but also in relation to some of the more speculative and conjectural aspects of the situation.

To begin with, let us consider what the general theorems *do* tell us. In order to characterize the situation of collapse 'past a point of no return', I shall first need the concept of a *trapped surface*. Let us return to Fig. 2. We ask what qualitative peculiarity of the region $r < 2m$ (after the star has collapsed

through) is present. Can such peculiarities be related to the fact that everything appears to be forced inwards in the direction of the centre? It should be stressed again that *apart* from $r=0$, the space-time at any individual point inside $r=2m$ is perfectly regular, being as a locally Minkowskian * as any other point (outside $r=0$). So the peculiarities of the $0 < r < 2m$ region must be of a partially * global * nature. Now consider any point T in the (θ, r) -plane of Fig. 2 ($r < 2m$). Such a point actually represents a *spherical 2-surface* in space-time, this being traced out as the θ, φ co-ordinates vary. The surface area of this sphere is $4\pi r^2$. We imagine a flash of light emitted simultaneously over this spherical surface T . For an ordinary spacelike 2-sphere in flat space-time, this would result in an ingoing flash imploding towards the centre (surface area decreasing) together with an outgoing flash exploding outwards (surface area increasing). However, with the surface T , while we still have an ingoing flash with decreasing surface area as before, the * outgoing * flash, on the other hand, is in effect also falling inwards (though not as rapidly) and its surface area also decreases. The surface T ($\theta = \text{const}$, $r = \text{const} < 2m$) of metric (2) serves as the prototype of a trapped surface. If we perturb the metric (2) slightly, in the neighbourhood of an initial hypersurface, then we would still expect to get a surface T with the following property:

T is a spacelike closed (*) 2-surface such that the null geodesics which meet it orthogonally all *converge* initially at T .

This convergence is taken in the sense that the local surface area of cross-section *decreases*, in the neighbourhood of each point of T , as we proceed into the future. (These null geodesics generate, near T , the boundary of the set of points lying causally to the future of the set T .) Such a T is called a *trapped surface*.

We may ask whether any connection is to be expected between the existence of a trapped surface and the presence of a physical space-time singularity such as that occurring at $r=0$ in (2). The answer supplied by some general theorems (PENROSE [6, 15], HAWKING and PENROSE [16]) is, in effect, that the presence of a trapped surface always *does* imply the presence of some form of space-time singularity.

There are similar theorems that can also be applied in cosmological situations. For example (HAWKING [17-19], HAWKING and PENROSE [16]), if the universe is *spatially closed*, then (excluding exceptional limiting cases, and assuming $\lambda < 0$) the conclusion is that there must be a space-time singularity. This time we expect the singularity to reside in the past (the * big bang *). Other theorems (HAWKING [17, 19], HAWKING and PENROSE [16]) can be applied also to spatially open universes. For example, if there is any point

(e.g. the Earth at the present epoch) whose past light cone starts * converging again * somewhere in the past (i.e. objects of given size start to have *larger* apparent angular diameters again when their distance from us exceeds some critical value), then, as before, the presence of space-time singularities is implied ($\lambda < 0$). According to HAWKING and ELLIS [20], cf. also HAWKING and PENROSE [16] the presence and isotropy of the 3°K radiation strongly indicates that the above condition on our past light cone is actually satisfied. So the problem of space-time singularities does seem to be very relevant to our universe, also on a large scale.

The main significance of theorems such as the above, is that they show that the presence of space-time singularities in exact models is not just a feature of their high symmetry, but can be expected also in generically perturbed models. This is *not* to say that *all* general-relativistic curved space-times are singular—far from it. There are many exact models known which are complete and free from singularity. But those which resemble the standard Friedmann models or the Schwarzschild collapse model sufficiently closely must be expected to be singular ($\lambda < 0$). The hope had often been expressed (cf. LINDQVIST and WHEELER [21], LIFSHITZ and KHALATNIKOV [22]) that the actual space-time singularity occurring in a collapsing space-time model might have been a consequence more of the fact that the matter was all hurtling simultaneously towards one central point, than of some intrinsic feature of general-relativistic space-time models. When perturbations are introduced into the collapse, so the argument could go, the particles coming from different directions might * miss * each other, so that an effective * bounce * might ensue. Thus, for example, one might envisage an * oscillating * universe which on a *large* scale resembles the cycloidal singular behaviour of an * oscillating * spatially closed Friedmann model; but the *detailed* behaviour, although perhaps involving enormous densities while at maximal contraction, might, by virtue of complicated asymmetries, contrive to avoid actual space-time singularities. However, the theorems seem to have ruled out a singularity-free * bounce * of this kind. But the theorems do *not* say that the singularities need resemble those of the Friedmann or Schwarzschild solutions at all closely. There is some evidence (cf. MISNER [23], for example) that the * generic * singularities may be very elaborate and possess a qualitative structure very different from that of their smoothed-out counterparts. Very little is known about this, however.

It is worth mentioning the essential basic assumptions that enter into the theorems. In the first place we require an * energy condition * which, by virtue of Einstein's equations (3), may be stated as a negative-definiteness condition on the Ricci tensor:

$$R_{ab}t^at^b < 0, \quad (4)$$

$$t^at_a = 1 \quad \text{implies} \quad R_{ab}t^at^b < 0,$$

that is to say, the time-time component R_{00} of R_{ab} is nonpositive in any ortho-

(*) By a * closed * surface, hypersurface, or curve, I mean one that is * compact without boundary *.

normal frame. If we assume $\lambda = 0$ in Einstein's equations (3), then (4) becomes

$$t^*t_* = 1 \quad \text{implies} \quad T_{\alpha\beta}t^{\alpha}t^{\beta} \geq \frac{1}{2}T^c_c.$$

This, when referred to an eigenframe of $T_{\alpha\beta}$, can be stated as

$$(5) \quad E + p_{\alpha} > 0 \quad \text{and} \quad E + \sum p_{\alpha} > 0,$$

where $\alpha = 1, 2, 3$. Here E is the energy density (referred to this frame) and p_1, p_2, p_3 are the three principal pressures. If (3) holds with $\lambda < 0$, then it is still true that (4) is a consequence of (5). The significance of the energy condition (4) lies in the effect of RAYCHAUDHURI [24] which states that whenever a system of timelike geodesics normal to a spacelike hypersurface starts converging, then this convergence inevitably increases along the geodesics until finally the geodesics cross over one another (assuming the geodesics are complete).

There is a corresponding focussing effect in the case of null geodesics. This depends on the 'weak energy condition':

$$(6) \quad l^{\alpha}l_{\alpha} = 0 \quad \text{implies} \quad R_{\alpha\beta}l^{\alpha}l^{\beta} \leq 0.$$

This condition (6) is a consequence of (4) (as follows by a limiting argument) but not conversely. If we assume Einstein's equations (3) with, now, any value of λ , then (6) is equivalent to

$$(7) \quad E + p_{\alpha} \geq 0$$

for $\alpha = 1, 2, 3$. The conditions (7) are, in fact, a consequence of the nonnegative definiteness of the energy density:

$$t^*t_* = 1 \quad \text{implies} \quad T_{\alpha\beta}t^{\alpha}t^{\beta} \geq 0$$

(that is $T_{\alpha\beta} > 0$ in each orthonormal frame). Thus, there is a strong physical basis for (6). The physical basis for (4) is not quite so strong, but provided $\lambda < 0$, we would certainly expect (4) to hold for all normal matter. (Note that if $E > 0 \geq \lambda$, only large negative pressures could cause trouble with (4). Usually people only worry about large positive pressures!) It is the 'strong' condition (4) that is required for the proofs of most of the theorems, but much can be said, concerning the qualitative nature of a collapse situation, even on the basis of the 'weak' condition (6) alone (cf. PENROSE [15]).

A remark concerning the condition on the cosmological constant λ seems appropriate here. It is a weakness of the theorems that most of them do require $\lambda \leq 0$ for their strict applicability. However, it would appear that the condition $\lambda < 0$ is only really relevant to the initial setting of the global conditions on the

space-time which are required for applicability of the theorem. If curvatures are to become large near a singularity, then (from dimensional considerations alone) the λ -term will become more and more insignificant. So it seems unlikely that a λ -term will really make much difference to the singularity structure in a collapse. The relevance of λ is really only at the cosmological scale.

Most of the theorems (but not all, cf. HAWKING [19]) require, as an additional assumption, the nonexistence of closed timelike curves. This is a very reasonable requirement, since a space-time which possesses closed timelike curves would allow an observer to travel into his own past. This would lead to very serious interpretative difficulties! Even if it could be argued, say, that the accelerations involved might be such as to make the trip impossible in 'practice' (cf. GÖDEL [25]), equally serious difficulties would arise for the observer if he merely reflected some light signals into his own past! In addition closed timelike curves can lead to unreasonable consistency conditions on the solutions of hyperbolic differential equations. In any case, it seems unlikely that closed timelike curves can substitute for a space-time singularity, except in special unstable models.

Some of the theorems require an additional 'generality' condition, to the effect that every timelike or null geodesic enters some region in which the curvature is not everywhere lined up in a particular way with the geodesic. (More precisely, $t_{\alpha}R_{\beta\gamma\alpha\delta}t^{\alpha}t^{\beta}t^{\gamma}t^{\delta} \neq 0$ somewhere along the geodesic, t^{α} being its tangent vector.) This condition plays a role in the mathematics, but from the physical point of view it is really no condition at all. We would always expect a little bit of matter or randomly oriented curvature along any geodesic in a physically realistic solution. It is only in very special limiting cases that we would expect the condition to be violated. (Curiously enough, however, practically every explicitly known solution does violate the condition!)

Finally, it should be remarked that none of the theorems directly establishes the existence of regions of approaching infinite curvature. Instead, all one obtains is that the space-time is not geodesically complete (in timelike or null directions) and, furthermore, cannot be extended to a geodesically complete space-time. ('Geodesically complete' means that geodesics can be extended indefinitely to arbitrarily large values of their length or affine parameter—so that inertially moving particles or photons do not just 'fall off the edge' of the space-time.) The most 'reasonable' explanation for why the space-time is not inextendible to a complete space-time seems to be (and I would myself believe this to be the most likely, in general) that the space-time is confronted with, in some sense, infinite curvature at its boundary. But the theorems do not quite say this. Other types of space-time singularity are possible, and theorems of a somewhat different nature would be required to decide which is the most likely type of singularity to occur.

We must now ask the question whether the theorems are actually likely

to be relevant in the case of a collapsing star or superstar. Do we, in fact, have any reason to believe that trapped surfaces can ever arise in gravitational collapse? I think a very strong case can be made that at least *sometimes* a trapped surface must arise. I would not expect trapped surfaces necessarily always to arise in a collapse. It might depend on the details of the situation. But if we can establish that there can be nothing *in principle* against a trapped surface arising—even if in some very contrived and outlandish situation—then we must surely accept that trapped surfaces must at least occasionally arise in real collapse.

Rather than use the trapped-surface condition, however, it will actually be somewhat easier to use the alternative condition of the existence of a point whose light cone starts «converging again». From the point of view of the general theorems, it really makes no essential difference which of the two conditions is used. Space-time singularities are to be expected in either case. Since we are here interested in a collapse situation rather than in the «big bang», we shall be concerned with the *future* light cone C of some point p . What we have to show is that it is possible in principle for enough matter to cross to within C , so that the divergence of the null geodesics which generate C changes sign somewhere to the future of p . Once these null geodesics start to converge, then «weak energy condition» (6) will take over, with the implication that an absolute event horizon must develop (outside C). As a consequence of the stronger «energy condition» (4) it will also follow that space-time singularities will occur.

Since we ask only that it be possible *in principle* to reconverge the null rays generating C , we can resort to an (admittedly far-fetched) «gedanken experiment». Consider an elliptical galaxy containing, say, 10^{11} stars. Suppose, then, that we contrive to alter the motion of the stars slightly by eliminating the transverse component of their velocities. The stars will then fall inwards towards the centre. We may arrange to steer them, if we like, so as to ensure that they all reach the vicinity of the centre at about the same time without colliding with other stars. We only need to get them into a volume of diameter about fifty times that of the solar system, which gives us plenty of room for all the stars. The point p is now taken near the centre at about the time the stars enter this volume (Fig. 3). It is easily seen from the orders of magnitude involved, that the relativistic light deflection (an *observed* effect of general relativity) will be sufficient to cause the null rays in C to reconverge, thus achieving our purpose.

Let us take it, then, that absolute event horizons can sometimes occur in a gravitational collapse. Can we say anything more detailed about the nature of the resulting situation? Hopeless as this problem may appear at first sight, I think there is actually a reasonable chance that it may find a large measure of solution in the not-too-distant future. This would depend on the validity

of a certain result which has been independently conjectured by a number of people. I shall refer to this as the *generalized* (*) *Israel conjecture* (abbreviated GIC). Essentially GIC would state: if an absolute event horizon develops in an asymptotically flat space-time, then the solution exterior to this horizon approaches a Kerr-Newman solution asymptotically with time.

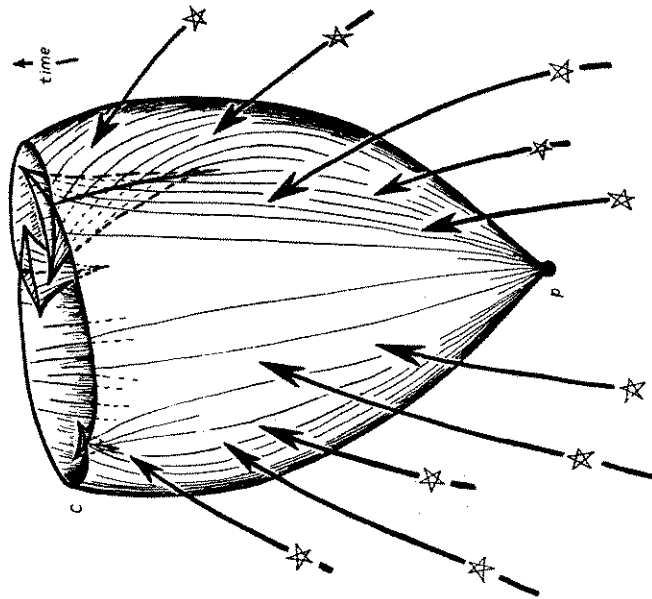


Fig. 3. — The future light cone of p is caused to reconverge by the falling stars.

The Kerr-Newman solutions (KERR [7], NEWMAN *et al.* [8]) are explicit asymptotically flat stationary solutions of the Einstein-Maxwell equation ($\lambda = 0$) involving just *three* free parameters m , a and e . As with the metric (1), the *mass*, as measured asymptotically, is the parameter m (in gravitational units). The solution also possesses angular momentum, of magnitude am . Finally, the total charge is given by e . When $a = e = 0$ we get the Schwarzschild solu-

(*) ISRAEL conjectured this result only in the stationary case, hence the qualification «generalized». In fact, ISRAEL has expressed sentiments opposed to GIC. However, Israel's theorem [26, 27] represents an important step towards establishing of GIC, if the conjecture turns out to be true.

tion. Provided that

$$m^2 > a^2 + e^2$$

the solution has an absolute event horizon. CARTER [9] has shown how to obtain all the geodesics and charged orbits for this solution, reducing the problem to a single quadrature. Thus, if GIC is true, then we shall have remarkably complete information as to the asymptotic state of affairs resulting from a gravitational collapse.

But what reason is there for believing that GIC has any chance of being true? One indication comes from a perturbation analysis of the Schwarzschild solution (REGGE and WHEELER [28], DOROSHEVICH *et al.* [29]) which seems to indicate that all perturbations except rotation have a tendency to be damped out. Another indication is the theorem of ISRAEL [26] which states, in effect, that the Schwarzschild solution is the only static asymptotically flat vacuum solution with an absolute event horizon (although there is a nontrivial side-condition to the theorem; cf. also THORNE [30] for the axially symmetric case). ISRAEL [27] has also generalized his result to the Einstein-Maxwell theory, finding the spherically symmetric Reissner-Nordstrom solution to be the only asymptotically flat static solution with an absolute event horizon. CARTER [31] has made some progress, in the vacuum rotating case, towards the objective of establishing the Kerr solution ($e = 0$) as the general asymptotically flat stationary solution with an absolute event horizon. In addition, there are solutions of the vacuum equations known (ROBINSON and TRAUTMAN [32]), which are suitably asymptotically flat and nonrotating, which apparently possess absolute event horizons, but are nonstatic. As time progresses they become more and more symmetrical, approaching the Schwarzschild solution asymptotically with time [33]. In the process, the higher multipole moments are radiated away as gravitational radiation.

The following picture then suggests itself. A body, or collection of bodies, collapses down to a size comparable to its Schwarzschild radius, after which a trapped surface can be found in the region surrounding the matter. Some way outside the trapped surface region is a surface which will ultimately be the absolute event horizon. But at present, this surface is still expanding somewhat. Its exact location is a complicated affair and it depends on how much more matter (or radiation) ultimately falls in. We assume only a finite amount falls in and that GIC is true. Then the expansion of the absolute event horizon gradually slows down to stationarity. Ultimately the field settles down to becoming a Kerr solution (in the vacuum case) or a Kerr-Newman solution (if a nonzero net charge is trapped in the « black hole »).

Doubts have frequently been expressed concerning GIC, since it is felt that a body would be unlikely to throw off all its excess multipole moments

just as it crosses the Schwarzschild radius. But with the picture presented above this is not necessary. I would certainly not expect the body itself to throw off its multipole moments. On the other hand, the gravitational field *itself* has a lot of settling-down to do after the body has fallen into the « hole ». The asymptotic measurement of the multipole moments need have very little to do with the detailed structure of the body itself; the *field* can contribute very significantly. In the process of settling down, the field radiates gravitationally—and electromagnetically too, if electromagnetic field is present. Only the mass, angular momentum and charge need survive as ultimate independent parameters. (Presumably the charge parameter e would be likely to be very small by comparison with a and m .)

But suppose GIC is not true, what then? Of course, it may be that there are just a lot more possible limiting solutions than that of Kerr-Newman. This would mean that much more work would have to be done to obtain the detailed picture, but it would not imply any qualitative change in the set-up. On the other hand there is the more alarming possibility that the absolute event horizon may be *unstable*! By this I mean that instead of settling down to become a nice smooth solution, the space-time might gradually develop larger and larger curvatures in the neighbourhood of the absolute event horizon, ultimately to become effectively singular there. My personal opinion is that GIC is more likely than this, but various authors have expressed the contrary view (*).

If such instabilities are present then this would certainly have astrophysical implications. But even if GIC is true, the resulting « black hole » may by no means be so « dead » as has often been suggested. Let us examine the Kerr-Newman solutions, in the case $m^2 > a^2 + e^2$ in a little more detail. But before doing so let us refer back to the Schwarzschild solution (2). In Fig. 4, I have drawn what is, in effect, a cross-section of the space-time, given by $r - r = \text{const}$. The circles represent the location of a flash of light which had been emitted at the nearby point a moment earlier. Thus, they indicate the orientation of the light cones in the space-time. We note that for large r the point lies inside the circle, which is consistent with the static nature of the space-time (*i.e.* one can « stay in the same place » while retaining a timelike world line). On the other hand for $r < 2m$ the point lies outside the circle, indicating that all matter must be dragged inwards if it is to remain moving in a timelike direction (so, to « stay in the same place » one would have to exceed the local speed of light). Let us now consider the corresponding picture for the Kerr-Newman solutions with $m^2 > a^2 + e^2$ (Fig. 5). I shall not be concerned, here, with the curious nature of the solution inside the absolute

(*) Some recent work of NEWMAN [34] on the charged Robinson-Trautman solutions suggests that new features indicating instabilities may arise when an electromagnetic field is present.

mass $2^s m(1 - K)^s$. Now, the point is that *however small* K may in fact be, we can always choose N large enough so that $(1 - K)^s$ is as small as we please. Thus, in principle, we can extract an arbitrarily large fraction of the mass-energy content of Misner's galaxy.

But anyone at all familiar with the problems of detecting gravitational

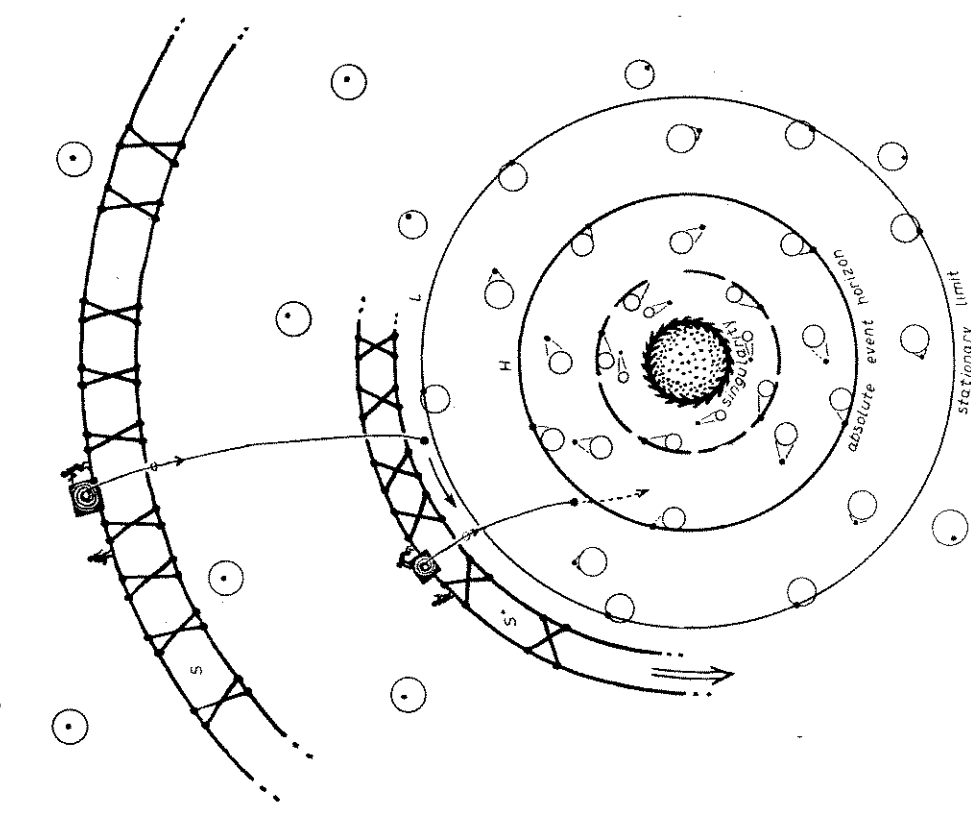


Fig. 5. - Rotating «black hole» (Kerr-Newman solution with $m^2 > a^2 + e^2$). The inhabitants of the structures S and S^* are extracting rotational energy from the «black hole».

event horizon H , since this may not be relevant to GIC. The horizon H itself is represented as a surface which is tangential to the light cones at each of its points. Some distance outside H is the «stationary limit» L , at which

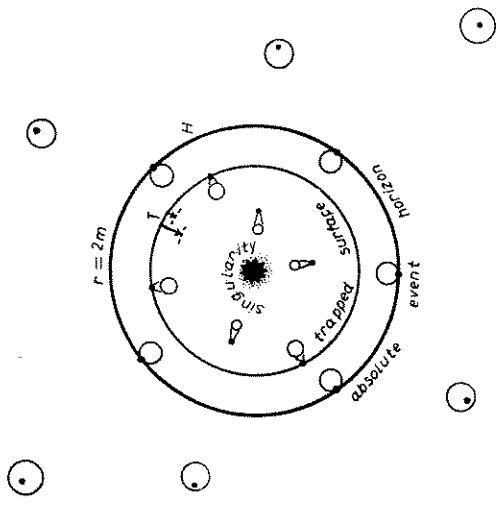


Fig. 4. - Spatial view of spherical «black hole» (Schwarzschild solution).

one must travel with the local light velocity in order to «stay in the same place».

I want to consider the question of whether it is possible to extract energy out of a «black hole». One might imagine that, since the matter which has fallen through has been lost for ever, so also is its energy content irrevocably trapped. However, it is not totally clear to me that this need be the case. There are at least two methods (neither of which is very practical) which might be construed as mechanisms for extracting energy from a «black hole». The first is due to MISNER [35]. This requires, in fact, a whole galaxy of 2^s «black holes», each of mass m . We first bring them together in pairs and allow them to spiral around one another, ultimately to swallow each other up. During the spiraling, a certain fraction K of their mass-energy content is radiated away as gravitational energy, so the mass of the resulting «black hole» is $2m(1 - K)$. The energy of the gravitational waves is collected and the process is repeated. Owing to the scale invariance of the gravitational vacuum equations, the same fraction of the mass-energy is collected in the form of gravitational waves at each stage. Finally we end up with a single «black hole» of

radiation will be aware of certain difficulties! Let me suggest another method which actually tries to do something a little different, namely extract the « rotational energy » of a « rotating black hole » (Kerr solution). Consider Fig. 5 again. We imagine a civilization which has built some form of stabilized structure S surrounding the « black hole ». If they lower a mass slowly on a (light, inextensible, unbreakable) rope until it reaches L , they will be able to recover, at S , the entire energy content of the mass. If the mass is released as it reaches L then they will simply have bartered the mass for its energy content. (This is the highest-grade energy, however, namely wound-up springs!) But they can do better than this! They also build another structure S^* , which rotates, to some extent, with the « black hole ». The lowering process is continued, using S^* , to beyond L . Finally the mass is dropped through H , but in such a way that its energy content, as measured from S , is *negative*! Thus, the inhabitants of S are able, in effect, to lower masses into the « black hole » in such a way that they obtain *more* than the energy content of the mass. Thus they extract some of the energy content of the « black hole » itself in the process. If we examine this in detail, however, we find that the angular momentum of the « black hole » is also reduced.

Thus, in a sense, we have found a way of extracting *rotational energy* from the « black hole ». Of course, this is hardly a practical method! Certain improvements may be possible, *e.g.*, using a ballistic method (*). But the real significance is to find out what can and what cannot be done *in principle* since this may have some indirect relevance to astrophysical situations.

Let me conclude by making a few highly speculative remarks. In the first place, suppose we take what might be referred to, now, as the most « conservative » point of view available to us, namely that GIC is not only true, but it also represents the *only* type of situation that can result from a gravitational collapse. Does it follow, then, that nothing of very great astrophysical interest is likely to arise out of collapse? Do we merely deduce the existence of a few additional dark « objects » which do little else but contribute, slightly, to the overall mass density of the universe? Or might it be that such « objects », while themselves hidden from direct observation, could play some sort of catalytic role in producing observable effects on a much larger scale. The « seeding » of galaxies is one possibility which springs to mind. And if « black holes » are born of violent events, might they not occasionally be ejected with high velocities when such events occur! (The one thing we can be sure about is that they *would* hold together!) I do not really want to make any very specific

(*) Calculations show that this can indeed be done. A particle p_0 is thrown from S into the region between L and H , at which point the particle splits into two particles p_1 and p_2 . The particle p_2 crosses H , but p_1 escapes back to S possessing *more* mass-energy content than p_0 !

suggestions here. I only wish to make a plea for « black holes » to be taken seriously and their consequences to be explored in full detail. For who is to say, without careful study, that they cannot play some important part in the shaping of observed phenomena?

But need we be so cautious as this? Even if GIC, or something like it, is true, have we any right to suggest that the *only* type of collapse which can occur is one in which the space-time singularities lie hidden, deep inside the protective shielding of an absolute event horizon? In this connection it is

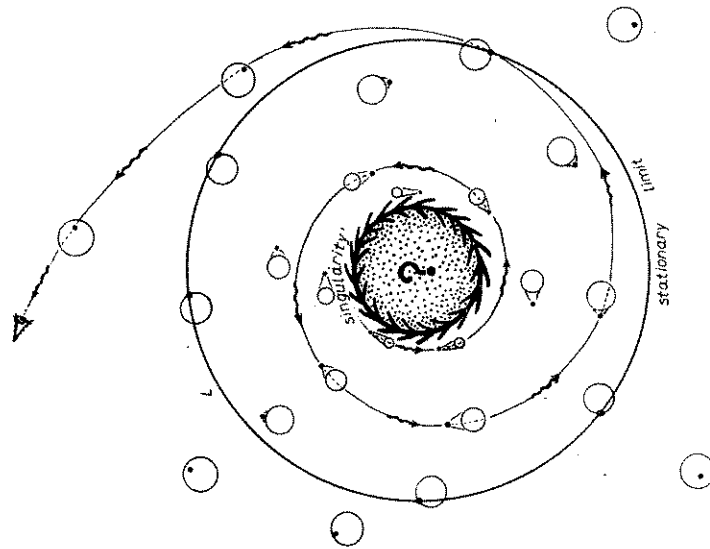


Fig. 6. — A « naked singularity » (Kerr-Newman solution with $m^2 < a^2 + e^2$).

worth examining the Kerr-Newman solutions for which $m^2 < a^2 + e^2$. The situation is depicted in Fig. 6. The absolute event horizon has now completely disappeared! A region of space-time singularity still exists in the vicinity of the centre, but now it is possible for information to escape from the singularity

to the outside world, provided it spirals around sufficiently. In short, the singularity is *visible*, in all its nakedness, to the outside world!

However, there is an essential difference between the logical status of the singularity marked at the centre of Fig. 6 and that marked at the centres of Fig. 4 and 5. In the cases of Fig. 4 and 5 there are trapped surfaces present, so we have a *theorem* which tells us that even with generic perturbation a singularity will still exist. In the situation of Fig. 6, however, we have no trapped surfaces, no known theorem guaranteeing singularities and certainly no analogue of GIC. So it is really an open question whether a situation remotely resembling Fig. 6 is ever likely to arise.

We are thus presented with what is perhaps the most fundamental unanswered question of general-relativistic collapse theory, namely: does there exist a « cosmic censor » who forbids the appearance of naked singularities, clothing each one in an absolute event horizon? In one sense, a « cosmic censor » can be shown *not* to exist. For it follows from a theorem of HAWKING [19] that the « big bang » singularity is, in principle, observable. But it is not known whether singularities observable from outside will ever arise in a generic *collapse*, which starts off from a perfectly reasonable nonsingular initial state.

If in fact naked singularities do arise, then there is a whole new realm opened up for wild speculations! Let me just make a few remarks. If we envisage an isolated naked singularity as a source of new matter in the universe, then we do not *quite* have unlimited freedom in this! For although in the neighbourhood of the singularity we have no equations, we still have normal physics holding in the space-time *surrounding* the singularity. From the mass-energy flux theorem of BONDI *et al.* [36] and SACHS [37], it follows that it is *not* possible for *more* mass to be ejected from a singularity than the original total mass of the system, *unless* we are allowed to be left with a singularity of *negative* total mass. (Such a singularity would *repel* all other bodies, but would still be attracted by them!)

While in the realm of speculation concerning matter production at singularities, perhaps one further speculative remark would not be entirely out of place. This is with respect to the manifest large-scale time asymmetry in the behaviour of matter in the universe (and also the apparent large-scale asymmetry between matter and antimatter). It is often argued that small observed violations of *T* (and *C*) invariance in fundamental interactions can have no bearing on the cosmological asymmetry problem. But it is not at all clear to me that this is necessarily so. It is a space-time singularity (*i.e.* presumably the « big bang ») which appears to govern the production of matter in the universe. When curvatures are fantastically large—as they surely are at a singularity—the local physics will be drastically altered. Can one be sure that the asymmetries of local interactions will not have the effect of being as drastically magnified?

When so little is known about the geometrical nature of space-time singularities and even less about the nature of the physics which takes place there, it is perhaps futile to speculate in this way about them. However, ultimately a theory will have to be found to cope with the situation. The question of the quantization of general relativity is often brought up in this connection. My own feeling is that the purpose of correctly combining quantum theory with general relativity is really somewhat different. It is simply a step in the direction of discovering how nature fits together as a whole. When eventually we have a better theory of nature, then perhaps we can try our hands, again, at understanding the extraordinary physics which must take place at a space-time singularity.

REFERENCES

- [1] S. CHANDRASEKHAR: *Roy. Astr. Soc. Month. Not.*, **95**, 207 (1935).
- [2] J. R. OPPENHEIMER and G. VOLKOFF: *Phys. Rev.*, **55**, 374 (1939).
- [3] A. S. EDDINGTON: *Nature*, **113**, 192 (1924).
- [4] D. FINKELSTEIN: *Phys. Rev.*, **110**, 965 (1958).
- [5] W. RINDLER: *Roy. Astr. Soc. Month. Not.*, **116**, 6 (1956).
- [6] R. PENROSE: in *Battelle Rencontres* (ed. C. M. DE WITT and J. A. WHEELER) (New York, 1968).
- [7] R. P. KERR: *Phys. Rev. Lett.*, **11**, 237 (1963).
- [8] E. T. NEWMAN, E. COUCH, K. CHINNAPARED, A. EXTON, A. PRAKASH and R. TORRENCE: *Journ. Math. Phys.*, **6**, 918 (1965).
- [9] B. CARTER: *Phys. Rev.*, **174**, 1559 (1968).
- [10] R. H. DICKE and H. M. GOLDENBERG: *Phys. Rev. Lett.*, **18**, 313 (1967).
- [11] R. V. POUND and G. A. REEBA: *Phys. Rev. Lett.*, **4**, 337 (1960).
- [12] A. SCHILD: in *Relativity Theory and Astrophysics*, Vol. 1: *Relativity and Cosmology* (ed. J. EHLERS) (Providence, R. I., 1967).
- [13] C. BRANS and R. H. DICKE: *Phys. Rev.*, **124**, 925 (1961).
- [14] A. EINSTEIN and J. GROMMER: *S. B. Preuss. Akad. Wiss.*, **1**, 2 (1927).
- [15] R. PENROSE: *Phys. Rev. Lett.*, **14**, 57 (1965).
- [16] S. W. HAWKING and R. PENROSE: *Proc. Roy. Soc., A* (in press) (1969).
- [17] S. W. HAWKING: *Proc. Roy. Soc., A* **294**, 511 (1966).
- [18] S. W. HAWKING: *Proc. Roy. Soc., A* **295**, 490 (1966).
- [19] S. W. HAWKING: *Proc. Roy. Soc., A* **300**, 187 (1967).
- [20] S. W. HAWKING and G. F. R. ELLIS: *Astrophys. Journ.*, **152**, 25 (1968).
- [21] R. W. LANDQUIST and J. A. WHEELER: *Rev. Mod. Phys.*, **29**, 432 (1957).
- [22] E. M. LIFSHITZ and I. M. KHALATNIKOV: *Advances in Phys.*, **12**, 185 (1963).
- [23] C. W. MISNER: *Phys. Rev. Lett.*, **22**, 1071 (1969).
- [24] A. K. RAYCHAUDHURI: *Phys. Rev.*, **98**, 1123 (1955).
- [25] K. GÖDEL: in *Albert Einstein Philosopher Scientist*, edited by P. A. SCHLIPF (New York, 1959), p. 557.
- [26] W. ISRAEL: *Phys. Rev.*, **164**, 1776 (1967).
- [27] W. ISRAEL: *Commun. Math. Phys.*, **8**, 245 (1968).

- [28] T. REGGE and J. A. WHEELER: *Phys. Rev.*, **108**, 1063 (1957).
 [29] A. G. DOROSHEVICH, YA. B. ZEL'DOVICH and I. D. NOVIKOV: *Zhurn. Éksp. Teor. Fiz.*, **49**, 170 (1965); English trans., *Sov. Phys. JETP*, **22**, 122 (1966).
 [30] K. S. THORNE: Ph. D. thesis, Princeton University Princeton, N. J. (1965).
 [31] B. CARTER: personal communication (1969).
 [32] I. ROBINSON and A. TRAUTMAN: *Proc. Roy. Soc.*, **A 265**, 463 (1962).
 [33] J. FOSTER and E. T. NEWMAN: *Journ. Math. Phys.*, **8**, 189 (1967).
 [34] E. T. NEWMAN: personal communication (1969).
 [35] C. W. MISNER: personal communication (1968).
 [36] H. BONDÍ, M. G. J. VAN DER BURG and A. W. K. METZNER: *Proc. Roy. Soc.*, **A 269**, 21 (1962).
 [37] R. K. SACHS: *Proc. Roy. Soc.*, **A 270**, 103 (1962).

Photon Statistics.

E. R. PIKE
Royal Radar Establishment - Malvern, Worcs.

Introduction.	277	4. Photon statistics of laser fields.	295
1. The description of an optical-radiation field: photons.	279	5. Photon statistics of mixed fields.	301
2. The detection of an optical-radiation field: photon counting.	282	6. Photon statistics of scattered laser light.	304
3. Photon statistics of thermal fields.	289	7. Application: intensity-fluctuation and Doppler spectroscopy.	305

Introduction.

The rapid growth in recent years of theoretical and experimental studies of the statistical properties of optical radiation or « photon statistics » can be attributed to a number of factors. Firstly, the experiments some fifteen years ago of FORRESTER, GUDMUNSON and JOHNSON [1] on the temporal correlation and of HANBURY BROWN and TWISS [2] on the spatial correlation of fluctuations of intensity of thermal radiation pioneered a new approach to optical spectroscopy and interferometry based on the known statistical behaviour of thermal light fields. This approach has proved to be very useful and practical. These were difficult experiments at the time, however, and a second important factor in the growth of statistical studies of « thermally » generated optical fields has been the steady development of the photomultiplier tube as a detector and of fast transistor electronic circuitry for data processing which makes such experiments very much easier today. Many laboratories are now using these techniques. A third factor which has even more transformed this type of work has been the introduction of the laser into optics. Laser light which is scattered by optically active elementary excitations of a medium which are in thermal equilibrium, such as for instance thermal longitudinal acoustic phonons in a liquid, has the same statistical character as thermally