Foundations of Physics, Vol. 26, No. 5, 1996

Tolerance for Spacetime Singularities

John Earman[1]

Received November 27, 1995

A common reaction to the Penrose–Hawking singularity theorems is that Einstein's general theory of relativity contains the seeds of its own destruction. This attitude is critically examined. A more tolerant attitude toward spacetime singularities is recommended.

"Progress in physics can proceed both from tolerance and intolerance."
C. W. Misner

1. INTRODUCTION

Charles Misner[1] distinguished three attitudes towards spacetime singularities in models of Einstein's general theory of relativity (GTR).[2] The first attitude ("Einstein avoids a singularity") holds that such singularities are merely artifacts of the unrealistic idealizations of the models (e.g., the perfect spherical symmetry and pressure-free dust matter of the Oppenheimer–Snyder model of gravitational collapse). This attitude was exemplified by the Russian school of Lifshitz, Khalatnikov, and co-workers, who claimed to have shown that a generic solution to Einstein's field equations (EFE) is singularity free.[2] They were eventually forced to recant[3] in the face of a series of theorems, due principally to Penrose and Hawking,[4] which were generally acknowledged as showing that singularities in solutions to EFE are to be expected in generic circumstances in both gravitational collapse and cosmology.[3]

[1] Department of History and Philosophy of Science, University of Pittsburgh, Pittsburgh, Pennsylvania 15260.

[2] Actually, Misner was concerned principally with the initial singularity in big bang cosmological models; but his remarks apply equally well to other types of spacetime singularities.

[3] A bit of caution is required here: the Penrose–Hawking theorems demonstrate the existence of singularities only in the sense of geodesic incompleteness; see Sect. 2.

0015-9018/96/0500-0623$09.50/0

The second attitude ("Nature avoids Einsteinian singularities") holds that, since GTR does entail singularities, the theory fails to accurately describe nature. This is no doubt the attitude that Einstein himself would have espoused had he lived to read the Penrose–Hawking theorems, at least if we are to believe Peter Bergmann's report of Einstein's intolerance of singularities:

> It seems that Einstein always was of the opinion that singularities in classical field theory are intolerable. They are intolerable from the point of view of classical field theory because a singular region represents a breakdown of the postulated laws of nature. I think one can turn this argument around and say that a theory that involves singularities and involves them unavoidably, moreover, carries within itself the seeds of its own destruction...[4]

The third, more tolerant and more optimistic, attitude ("Nature and Einstein are subtle but tolerant") was advocated by Misner himself. It views the existence of singularities in solutions to EFE "not as proof of our ignorance, but as a source from which we can derive much valuable understanding of cosmology."[5]

It is fair to say that advocates of this third attitude are few and far between. Not surprisingly, there have arisen two types of research programs to cope with the spacetime singularities of GTR. One seeks to modify classical GTR in such a way that singularities are avoided while retaining the verified predictions of GTR for weak gravitational fields.[(6, 7)] The other program does not attempt to tinker with classical GTR but seeks to show that quantizing GTR will smooth away spacetime singularities.[(8, 9)]

There is, perhaps, a parallel here with quantum mechanics (QM) in that a growing number of physicists and philosophers of science hold that the measurement problem shows that QM contains the seeds of its own destruction.[6] If these twin attitudes of intolerance are correct, we arrive at the stunning conclusion that, despite their many empirical and conceptual successes, the two main theories of twentieth century physics are self-refuting. This conclusion may well prove to be correct, but precisely because of its stunning quality, caution is in order. In the case of QM many attempts have been made to rebut the seeds-of-its-own-destruction charge. (In my opinion, the devices that have been invoked to preserve standard QM (many worlds, many minds, hidden variables,...[7]) are sufficiently

[4] Reference 5, p. 186.
[5] Reference 1, p. 1329.
[6] Bell and Nauenberg wrote: "It seems that the quantum mechanical description will be superseded. In this it is like all theories made by man. But to an unusual extent its ultimate fate is apparent in its internal structure. It carries in itself the seeds of its own destruction." (Ref. 10, p. 285).
[7] See Albert[(11)] for a review of various approaches to the measurement problem in QM.

implausible that if I had to make a bet, it would be on investigations that modify Schrödinger dynamics, perhaps along the lines suggested by Ghirardi, Rimini, and Weber.[12]) By contrast, very little has been done to investigate the viability of Misner's attitude of tolerance towards spacetime singularities. The present paper takes some tentative steps in that direction.

Section 2 defines the subject matter by giving a brief overview of attempts to pin down the elusive concept of spacetime singularity. Sections 3 and 4 critically discuss two versions of the seeds-of-its-own-destruction argument. The first version holds that, even taken in its own terms, classical GTR reveals itself to be false, or incomplete, or otherwise defective. The second version holds that the prediction of singularities by GTR is to be taken as an indication that the classical description breaks down near singularities because quantum effects become dominant and that singularities are absent in the correct quantum description. Section 5 presents some concluding remarks.

2. WHAT IS A SPACETIME SINGULARITY?

Talk of singularities is thing talk—it encourages the notion that singularities are localizable objects. But the official definition of a relativistic spacetime does not seem to accommodate such talk: a spacetime is a pair M, g_{ab}, where M is a differentiable manifold and g_{ab} is a Lorentz signature metric that is defined and $C^n(n \geq 2(?))$ for all of M; so there are no singular points or regions in the spacetime. This does not mean that object talk about singularities is beyond the pale, but it does mean that the objects have to be constructed. Three steps are involved in the construction: first, the singular spacetimes have to be distinguished from the non-singular ones; second, for a singular M, g_{ab} we can try to represent singularities as boundary points of M; and third, the set of boundary points needs to be equipped at least with a topological structure and, one hopes, differentiable and metric structures as well. There are a number of attempts to implement this scheme, but all of the extant ways of carrying out the second and third steps lead to counterintuitive results. Geroch, Liang, and Wald[13] argue that this is no accident. They show that any construction satisfying some seemingly natural conditions will have the consequence that the singular boundary points are not Hausdorff separated from points of M. It seems that we must be prepared to find that there is no satisfying way to talk about spacetime singularities as localizable objects.

We are left with the task of fashioning a definition that captures the set of singular spacetimes. My thesis here is that no simple definition will be forthcoming since to call a spacetime singular is to call attention to one or more members of a large family of conceptually distinct but interrelated pathologies that can infect relativistic spacetimes. I will mention four such pathologies, trying to give the main ideas while omitting technical details.[8]

The most obvious and intuitively appealing idea is that a singular spacetime is one in which one or more scalar curvature invariants "blow up." A little more precisely: a singular spacetime is one in which a curvature invariant becomes unbounded along a curve in the spacetime. But what kind of curve? If a curvature invariant becomes unbounded only as the curve trails off to spatial or temporal infinity, then no singularity in the spacetime itself seems indicated.[9] Thus, it is necessary to stipulate that, to use Einstein's[(15)] phrase, the singular behavior occurs "at a finite distance." There are various ways to make this stipulation precise, but in order to have a concrete proposal at hand, let us say that a spacetime is singular if and only if a curvature invariant becomes unbounded within a finite affine distance along a geodesic in the spacetime. This definition counts as singular familiar examples such as the Robertson–Walker big bang models and Kruskal–Schwarzschild spacetime. However, there are spacetimes where the definition fails but nevertheless the curvature is sufficiently ill-behaved that arguably these spacetimes should be counted as singular. For example, all curvature scalars may remain bounded at finite affine distances even though some of the physical components of the Riemann curvature tensor as measured in a parallely propagated (p.p.) orthonormal tetrad frame diverge at finite affine distances.[10] Furthermore, the blow-up of curvature is not the only way curvature can be ill-behaved. Suppose that a spacetime contains a half-geodesic (a geodesic which has an endpoint and which is extended as far as possible in some direction from that point) that is incomplete (i.e., has finite affine length). And suppose also that although the physical components of the Riemann tensor in any p.p. orthonormal frame along the half-geodesic remain bounded, some of the physical components oscillate wildly as an affine parameter approaches its limiting value. Does such behavior qualify the spacetime as singular? The original intuition with which we started has become fuzzy.

Another idea is that geodesic incompleteness by itself is indicative of a singular spacetime. Certainly the ghost of an observer whose world line is an incomplete timelike geodesic would have grounds for complaining that the spacetime is pathological. There is a link to the first idea since curvature blow-up at a finite distance induces geodesic incompleteness. But

[8] See Chap. 2, Reference 14, for more details and references to the literature.
[9] But see the discussion below of naked singularities.
[10] If e_i^a, $i = 1, 2, 3, 4$, is an orthogonormal tetrad field, the physical components of the Riemann curvature tensor R_{abcd} in this field are $R_{(i)(j)(k)(l)} \equiv R_{abcd} e_i^a e_j^b e_k^c e_l^d$.

the converse implication fails even in inextendible spacetimes. For example, in the Curzon spacetimes (which belong to the Weyl class of axisymmetric vacuum solutions to EFE), some geodesics approaching the axis of symmetry are incomplete even though the curvature is well behaved. Even if incompleteness is taken as the touchstone of singular spacetimes, it is not clear what the appropriate sense of incompleteness is. Geroch[(16)] produced an example of a spacetime that is geodesically complete but contains timelike half-curves of bounded acceleration and finite proper length. Such a curve could be instantiated by a rocketship using a finite amount of fuel. The astronaut who steers the rocket ship would find that even if he has drunk of the fountain of youth, the spacetime structure does not allow him to live beyond a certain finite number of years. It is tempting to say that the reason for this bizarre limitation is that the spacetime is singular. Still more inclusive notions of incompleteness can be formulated using the concept of generalized affine length.[11] Before moving on to another idea it should be noted that the singularity theorems of Penrose and Hawking and their generalizations all use geodesic incompleteness as the criterion of a singular spacetime. The proof techniques of these theorems do not lend themselves to demonstrating the existence of singularities in the other senses discussed here.

A third conception derives from Misner's[(18)] idea that singular spacetimes are those that have resulted from cutting out singular (or even regular) points from a larger manifold. How to detect and characterize these "missing points" remains unsettled, but the tools developed by Scott and Szekeres[(19)] seem helpful. The details are too complicated to present here, and I will simply mention that the present conception diverges from the incompleteness and curvature blow-up conceptions of singularities. Missing regular points (as indicated by the fact that the spacetime is properly extendible) give rise to geodesic incompleteness. But geodesic incompleteness need not indicate missing points, regular or singular, as shown by Misner's[(18)] example of a compact and geodesically incomplete spacetime. The incompleteness in this example cannot be due to missing points since a compact manifold cannot be imbedded as a proper subset of another (Hausdorff) manifold of the same dimension. Curvature blow-up is one reason for missing singular points, but presumably missing points can also arise from other kinds of pathologies not associated with curvature singularities.

A fourth pathology corresponds to Penrose's idea of a naked singularity.[(20)] There are many ways to try to make this idea precise, but the strongest reasonable requirement for the *absence* of a naked singularity

[11] This concept is discussed in Reference 17, Sec. 8.1.

is the condition of global hyperbolicity.[12] A weaker requirement would tolerate breakdowns of global hyperbolicity as long as they are confined to the interiors of black holes. In one guise the cosmic censorship hypothesis is the conjecture that under physically reasonable conditions EFE do not permit naked singularities to develop from regular initial data. Much effort has gone into stating precise versions of this conjecture and either proving or refuting the conjectures. The fate of cosmic censorship remains unsettled.[13] What I want to emphasize is that the cosmic censorship idea represents a radical departure from the first two conceptions of singularities since nasty behavior which takes place only "at infinity" can lead to violations of cosmic censorship and lead one to classify a spacetime as nakedly singular. An artificial mathematical example suffices to illustrate the point. Remove from Minkowski spacetime $\mathbb{R}^4$, η_{ab} a compact ball B to the future of the time slice $t=0$. Define a scalar field Ω whose value goes to ∞ rapidly as the (missing) B is approached. The spacetime $\mathbb{R}^4-B$, $\Omega^2\eta_{ab}$ is geodesically complete, curvature does not blow up at a finite distance, etc. But it is not globally hyperbolic; the time slice $t=0$ is not a Cauchy surface for the new spacetime, and initial data on $t=0$ do not suffice to determine future development.

Other pathologies that might deservedly qualify a spacetime as singular can surely be formulated. But I trust that enough has been said to make it clear that there is no univocal concept with a simple and precise mathematical expression underlying our pre-analytic intuitions of a singular spacetime.

Before closing this section an embarrassing admission needs to be made. When we speak of singularities or singular spacetimes, we do not know what we are talking about. The point is not that we don't know how to choose among the family of pathologies listed above. Rather the point is that what we should be interested in are essential singularities, singularities that cannot be removed by extending the spacetime; but this is an ill-defined notion until the continuity/differentiability conditions on extensions are specified. The Penrose–Hawking theorems prove the existence of essential singularities (in the sense of geodesic incompleteness) under the assumption that an allowable extension is at least C^{2-}. Now any physically relevant extension must surely permit one to make sense of EFE at least in the language of distributions; but presumably this can be

[12] A spacetime M, g_{ab} is globally hyperbolic if and only if it is strongly causal (i.e., for every $p\in M$ any open neighborhood $N(p)$ of p there is a subneighborhood $N'(p)$ which no causal curve reenters once it leaves) and $J^-(p)\cap J^+(q)$ is compact for all $p, q\in M$. Global hyperbolicity is equivalent to the condition that there exist a Cauchy surface, i.e., a spacelike hypersurface that is intersected exactly once by every causal curve without endpoint.

[13] See Chap. 3 of Reference 14 for a review of the literature.

possible even if the metric is not C^{2-}. So what are the minimal continuity/differentiability conditions for this to be possible? And under these minimal conditions which of the spacetimes of general relativity that are commonly said to be singular are counted as essentially singular? These problems are unresolved, but some progress has been made.[(21)]

3. DO SINGULARITIES SHOW THAT CLASSICAL GTR CONTAINS THE SEEDS OF ITS OWN DESTRUCTION?

In this section I will critically assess the reasons that have been advanced for giving a positive answer to this question under the assumption that quantum effects are to ignored. Thus, a positive answer says that even taken in its own terms, classical GTR is, in an important sense, self-refuting.

The first motivation comes from Kip Thorne:

> From a purely philosophical standpoint it is difficult to believe that physical singularities are a fundamental and unavoidable feature of our universe. On the contrary, when faced with a theory which predicts the evolution of a singular state, one is inclined to discard or modify that theory rather than accept the suggestion that the singularity actually occurs in nature. Such was the case with Rutherford's theory of the atom...[14]

The analogy with the Rutherford atom is not apt. The combination of the Rutherford model of the atom and classical electrodynamics provides an illustration of the grossest form of the seeds-of-its-own-destruction argument at work: the combined theory sends the orbital electrons crashing into the nucleus, a prediction falsified by the fact that the material world exists. (Similarly, if the seeds-of-its-own-destruction interpretation of the measurement problem in QM is correct, then QM reveals itself to be empirically inadequate in the worst way since it cannot account for the fact that measurements have determinate outcomes.) By contrast, the archetypical GTR predictions of singularities do not involve any such blatant empirical inadequacies. Indeed, the prediction of black holes formed in gravitational collapse is meeting with increasing empirical success. The standard big bang model would be in trouble if some recent estimates of the Hubble constant stand up to scrutiny since, when plugged into the standard model, these estimates produce an age of the universe that is far too young.[(23)] But the difficulty here is not due to the initial big bang singularity per se since the model can accommodate the estimates by adding a positive cosmological constant.

[14] Reference 22, p. 415.

One might in fact seek to turn the present form of the seeds-of-its-own-destruction argument around and view singularities in GTR not as seeds of destruction but rather as seeds of definitive confirmation. Hawking has noted that most of the classical tests of GTR involve weak gravitational fields. "Thus observations to determine whether singularities actually occurred would provide a test of the [EF]equations for strong fields."[15] The rub is that determining whether singularities actually occur is not like determining whether some localizable object exists in spacetime; rather, it involves (as I hope Sec. 2 made clear) determining whether or not spacetime has some large-scale or even global properties. But though the determination is difficult, it is not impossible in principle.

Such optimism might be open to challenge. How, for example, could it be verified that the spacetime we inhabit is timelike geodesically incomplete? Even if some volunteer could be found to sacrifice himself for the sake of scientific knowledge, how could he tell the difference between a case where his geodesic (let us suppose) world line is in principle inextendible beyond a certain finite proper time because, say, it encounters unbounded curvature vs. a case where his world line is extendible in principle but not in practice because of curvature that is bounded but so strong as to terminate any physical measuring instrument? The short answer is that our self-sacrificing scientist can't tell the difference and, thus, cannot definitely verify the singular nature of his spacetime. But by the same token he cannot verify predictions about the temperature at the core of the sun. The interesting issue is not whether we can definitively verify predictions about the temperature of the core of the sun or the singularity structure of spacetime but whether observation and previously accepted theory can combine to give us reasonable beliefs about such matters. There are skeptics who will give a negative answer for both cases. I have no response to hardline skeptics. My claim is only that to the extent that it is reasonable to form beliefs about nonverifiable theoretical assertions in physics, then assertions about the singularity structure of spacetime will be among them.

To pursue the theme of confirmation, reflect on the truism that in testing a scientific theory it is fruitful to consider the theory not in isolation but in confrontation with rival theories—in the case in point, classical GTR would be compared with rival classical theories of gravitation that avoid spacetime singularities. Unless such rivals are of the ad hoc cut-and-paste variety, they will surely yield predictions that depart from those of GTR even in the regime where singularities are not involved. Consider, for example, Moffat's[7] nonsymmetric gravitational theory (NGT). Einstein and others had utilized nonsymmetric metrics and affine connections in an

[15] Reference 24, p. 520.

attempt to unify electromagnetism and gravitation. Moffat employs some of the same formalism not in the service of a unified field theory but as a means of describing pure gravitation. The resulting NGT promises to avoid singularities and black holes as well. Moffat[7] has detailed numerous subtle and not so subtle ways in which the predictions of NGT differ from those of GTR even in regimes not involving singularities. Now suppose that experimental investigations favor the predictions of GTR in regimes not involving singularities over NGT and other rival theories that eschew singularities—otherwise GTR would be rejected on straightforward grounds of empirical adequacy having nothing to do with singularities. In that case I submit that at least on the classical level, we would have good reason to take the singularity predictions of GTR seriously.

Scientific theories are judged along several dimensions in addition to empirical adequacy. One version of the seeds-of-its-own destruction argument is aimed at the dimension of completeness. Thus, Brandenberger *et al.* charge that "The presence of singularities is an indication that G[T]R is an incomplete theory."[16] An analogy with the special theory of relativity (STR) suggests why this charge might seem to have merit. Consider a set of putative laws in the form, say, of ordinary or partial differential equations for a field (scalar, vector, or tensor) on Minkowski spacetime. Suppose that it is discovered that the putative laws allow singularities to develop at points of Minkowski spacetime in the sense that the field strength becomes unbounded as these points are approached. Then the putative laws are incomplete in that they have nothing to say about what happens at these points.[17] Further, insofar as laws of nature must be universal—must apply to all of space and time—the putative laws do not count as genuine laws but, at best, as approximations to the real laws. While this attitude seems basically correct, the analogy with GTR is inapt because, in contrast to STR, GTR does not specify a fixed spacetime that can serve as a backdrop in which the singular behavior of physical fields can be measured. In GTR the metric field g_{ab} is now a physical field whose singularity behavior we have to judge. And the judgment has to start from the fact that a general relativistic spacetime M, g_{ab} is such that g_{ab} is defined and differentiable at every point of M—there are no singular points of spacetime where the laws of GTR fail to apply.

This definitional move may seem to facile. In the case of a closed Robertson–Walker spacetime, Brandenberger *et al.* see a kind of incompleteness resulting from the initial and final singularities. "[T]he [final] singularity implies that we cannot answer the question what will happen

[16] Reference 6, p. 1629.

[17] Unless the laws can be recast in a distributional form that applies to singular points.

after the 'big crunch' or (in the case of an expanding universe) what was before the 'big bang'."[18] In response I would pose a dilemma. Either the initial and final singularities are essential—in the sense that it is impossible to extend through them in a way that EFE make sense even in the language of distributions (see Sec. 2)—or not. If so, then by the lights of GTR, talk about "before" the big bang and "after" the big crunch is physically meaningless. If not, then GTR can say something about the before and after. In either case GTR does not stand convicted out of its own mouth of raising meaningful questions it cannot answer.

There remains one respect in which the charge of incompleteness may be justified. If the singularities are of the naked variety there is a breakdown of determinism. So if determinism is a necessary condition for a classical relativistic theory to be complete, GTR is incomplete if the cosmic censorship hypothesis is false. Note that it is not just any spacetime singularities that sustain the charge of incompleteness; in particular, the singularities of the Robertson–Walker models do not fall under this heading since the initial and final singularities are not counted as naked (since the spacetimes of these models are globally hyperbolic). Note also what has to be done to defend GTR against this form of the incompleteness accusation: either cosmic censorship theorems have to be proved, or else it has to be argued that a globally well-posed initial value problem is not essential to completeness and that a locally well-posed problem—which in fact obtains for EFE—is sufficient. The latter move is not very appealing. Presumably evolution does not just cease at the Cauchy horizons;[19] something happens beyond the horizons, and if GTR does not determine what it is, then it is incomplete. Thus, the fate of the incompleteness objection rests on the fate of cosmic censorship.

Closely related to the incompleteness objection is the complaint that singularities in a field show that the field is not fundamental. Once again the motivating analogy relies on STR. In STR the Maxwell electromagnetic field is regarded as fundamental since, among other things, solutions to the Maxwell equations exhibit the global existence property. By contrast, a perfect fluid is not regarded as fundamental since the development of singularities can wreck global existence.[(25)] The blame for the wreck lies not with nature but in the oversimplified idealization of the perfect fluid description. But yet again the analogy with GTR is faulty. In the GTR the field at issue is the metric field, and in GTR we confront the unprecedented

[18] Reference 6, p. 1629.

[19] If a spacelike hypersurface S is not a Cauchy surface it will have nontrivial Cauchy horizons. The future Cauchy horizon $H^+(S)$ of S is defined as the future boundary of the future domain of dependence $D^+(S)$ of S. $D^+(S)$ consists of all those spacetime points p such that every causal curve which passes though p and which has no past endpoint meets S.

situation where the spacetime metric is not part of a fixed background against which it plays itself out but is itself a dynamical element that takes part in the drama. Is there reason to think that spacetime singularities indicate that GTR gives an oversimplified description of the dynamics of the metric field? A negative answer is indicated if, as supposed above, GTR continues to give better experimental predictions than rival classical theories of gravitation that eschew singularities. There is, however, a grain of truth to the current objection, but it leads back to the problem of cosmic censorship. Consider the initial value problem for the vacuum EFE. Choose a complete Riemannian three-manifold Σ. Given the first and second fundamental forms h_{ab} and k_{ab} of Σ, there is a unique (up to diffeomorphism) maximal development M, g_{ab} of (Σ, h_{ab}, k_{ab}) which is a solution of the vacuum EFE and for which Σ is a Cauchy surface. One can then ask whether under generic conditions this solution is a global one in that M, g_{ab} does not admit of proper extensions. If not, GTR is open to the charge that it does not provide a fundamental description of the metric field since such a description should fix global solutions. Note that it is not just any singularity but only naked singularities that give rise to a negative answer. Indeed, one way to formulate the cosmic censorship hypothesis is to say that a positive answer is a consequence of EFE.

In addition to empirical adequacy, completeness, and fundamentalness of description, physical theories are also judged in terms of their explanatory power. The standard big bang model has been found wanting on the last grounds. The complaint is that the explanation the model furnishes for the uniformity of the observed cosmic microwave background radiation lacks robustness since it must postulate special conditions near the big bang. Some skepticism about the robustness requirement is in order.[20] But in any case the source of the robustness complaint is not the big bang singularity itself but the presence of particle horizons which seem to block a satisfying causal account of how generic inhomogeneous and anisotropic initial conditions can smooth themselves enough to accord with present observations. And the currently most popular solution of the horizon problem, inflationary cosmology, proceeds not by eliminating the initial singularity but by allowing even the locations on the surface of last scattering that are separated by large angular distances to have a common causal past.

Finally, I want to examine Cornish and Moffat's[26] charge that the black hole singularities of GTR generate two paradoxes. The first paradox is formulated as follows:

[20] See Reference 14, Chap. 5.

> An observer at spatial infinity would see falling matter "freeze" before it can form an event horizon, whereas a freely falling observer can fall through the event horizon without difficulty, but be unable to communicate this fact to the observer at infinity. This means that the spacetime is separated into two disconnected parts by a null surface and there exists no communication between the two spacetimes. A paradox arises, for these two observers completely disagree about what they see in a spacetime containing a black hole.[21]

Note that the paradox here relates not to singularities per se but to event horizons. Similar paradoxes can arise even in the context of STR. Consider, for example, an observer γ in Minkowski spacetime who undergoes Born hyperbolic acceleration (see Fig. 1). Such an observer will have an even horizon $E(\gamma)$. An unaccelerated observer γ' can pass though $E(\gamma)$ without difficulty but be unable to communicate this fact to γ—γ will experience the lapse of an infinite amount of proper time before γ' crosses $E(\gamma)$. I submit that this paradox and the Cornish–Moffat paradox alike are not genuine paradoxes but—like the so-called twin paradox—are simply illustrations of basic features of relativistic spacetimes.[22]

The second paradox–Hawking's so-called information loss paradox—is genuinely disturbing. Hawking[(27)] established that black holes can evaporate by emitting radiation with a thermal spectrum. He later argued that the most plausible upshot of this evaporation is for the black hole to disappear completely without giving away any information about the black hole state.[(28, 29)] In quantum mechanical terms, an external observer would see an initially pure state converted into a mixture, with a resultant loss of information about phase relations. Such an information leak would be effectively plugged if black hole singularities were banished, as in Moffat's NGT. But other approaches are possible. For example, Banks and O'Laughlin[(30)] have suggested that the information is preserved in tiny black hole remnants, which if true would have important cosmological implications. Stephens *et al.*[(31)] have offered an alternative approach to black hole quantization in which quantum coherence is not lost. If these approaches to stemming information loss should prove wanting, it remains open to explore the consequences of a new element of unpredictability that results from combining GTR and QM. Whatever verdict the future will render on this matter, we have a good illustration of how Misner's attitude of tolerance can be fruitful in producing interesting new physics.

[21] Reference 26, p. 6628.

[22] Of course, in the case of a black hole the event horizon has an absolute or observer independent character since it is defined as the boundary of the region that can be seen from future null infinity. But this fact does not invalidate the point about the "paradox" of event horizons.

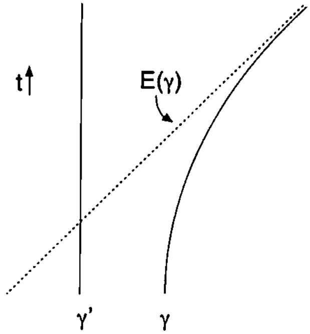

Fig. 1. Event horizons in Minkowski spacetime.

In introducing quantum considerations I have already strayed outside the first version of the seeds-of-its-own-destruction argument and into the second. The following section takes up the second version explicitly.

4. QUANTUM EFFECTS AND SINGULARITIES

The second version of the seeds-of-its-own-destruction argument does not claim that classical GTR, taken in its own terms, reveals itself to be empirically inadequate, incomplete, or otherwise defective. Rather, the second version makes a weaker charge that is contained in two claims: first, that quantum effects will come into play in regimes where GTR predicts singularities; and second, that these quantum effects will invalidate the classical description and will somehow lead to an avoidance of singularities. Ultimately these claims must be evaluated in the light of an adequate quantum theory of gravity, which we are very far from having. Absent such a theory, we are reduced to speculation. Still, the right sort of speculation can at least locate the relevant issues and suggest avenues of investigation.

The first point to be kept in mind in assessing the plausibility of the claims at issue is that the question has to be divided among the various types of spacetime singularities discussed in Sec. 2. Take first curvature singularities. It is plausible that quantum effects become important on scales corresponding to the Planck length $l_{\rm pl} \sim 10^{-33}$ cm and, thus, will dominate in cases where classical GTR says that, for example, the Riemannian curvature scalar approaches $l_{\rm pl}^{-2}$. As a first indication of why quantum effects could lead to an avoidance of curvature singularities, one could point to the amazing ability of ordinary QM to smooth away the singularities of classical mechanics. For instance, the motion of point mass particles under their mutual Newtonian $1/r^2$ gravitational forces is known to generate both collision and noncollision singularities (in the latter case the solution ceases after a finite time because all of the particles have disappeared to spatial infinity). By contrast, the quantum version of this problem is completely regular. The Hamiltonian is essentially self-adjoint so that the time evolution operator is defined for all times. This line of thought has been extended by Horowitz and Marolf,[(32)] who studied the motion of quantum test particles in singular spacetimes. They found that in some static spacetimes with timelike curvature singularities, the motion of the quantum test particles is well behaved (unitary time evolution).

The next step toward a quantum theory of gravity is the semiclassical approximation in which the quantum expectation value of the (renormalized) stress-energy tensor is inserted in EFE in order to calculate the backreaction of quantum fields on the spacetime metric. In some model calculations it is found that there are states that involve negative pressures large enough to violate the weak and strong energy conditions[23] used in the Penrose–Hawking singularity theorems, and as a result the model can bounce before reaching a curvature singularity.[(8)] Of course, there is no guarantee that such semiclassical calculations are reliable indicators of what will happen when the gravitational field itself is quantized, but in the meantime such results are encouraging for those who want to banish singularities.

The picture is not so rosy when we turn to other types of singularities. In the case where geodesic incompleteness is not accompanied by curvature blow-up, there is no a priori reason to think that quantum effects will dominate. Thus, the best hope for the anti-singularity league is to argue at the classical level that this type of singularity will not arise in physically reasonable circumstances. A result of Clarke[(33)] can be utilized to this end, at least if a strong version of cosmic censorship is true. He showed that in globally hyperbolic spacetimes that are not too specialized in a technical sense, the incompleteness of timelike geodesics is due to a curvature singularity.

This brings us to naked singularities in the sense of features of spacetimes that would violate cosmic censorship. If, as I argued in Sec. 2, it is not singularities in general but only naked singularities that raise concerns about the adequacy of GTR, then, as I will now argue, one had

[23] The stress-energy tensor T_{ab} satisfies the strong energy condition if and only for any unit timelike V^a, $T_{ab}V^aV^b \geqslant -(1/2)\,\mathrm{Tr}(T_{ab})$. T_{ab} satisfies the weak energy condition just in case for any timelike V^a, $T_{ab}V^aV^b \geqslant 0$.

better hope that quantum gravity does not banish curvature singularities. The maximal extension of the Reissner–Nordström spacetime contains a curvature singularity that is naked in any reasonable sense of that term. This example, can be brushed aside as a counterexample to cosmic censorship on the grounds that it is highly nongeneric in the space of all solutions to EFE. Evidence for this comes from the fact that the Cauchy horizons in this solution are unstable: small perturbations on an initial value hypersurface are amplified by a blue shift to produce an infinite effect on the Cauchy horizon. However, when the black hole portion of the Reissner–Nordström solution is imbedded in de Sitter spacetime instead of an asymptotically flat spacetime to form Reissner–Nordström–de Sitter spacetime, the Cauchy horizons are found to be stable under a wide range of parameter values.[34] But quantum gravity may come to the rescue of cosmic censorship. Marković and Poisson[35] have argued that the Cauchy horizons are quantum mechanically unstable in the sense that the expectation value of the renormalized stress energy tensor diverges as the horizon is approached. The backreaction on the metric presumably cuts off future development. Observers are shielded from a naked singularity by having them terminated by a non-naked curvature singularity. This may seem a rather draconian way to protect cosmic censorship. But if, as I have argued, the real fear of singularities has to do with their nakedness, the watchword should be: better dead than uncensored.

Of course, one could hope that the need for such a draconian rescue is not necessary since in the fully quantized version of GTR the naked curvature singularity would not arise in the first place. But such a hope is inoperative for another way in which cosmic censorship can fail. The strong form of cosmic censorship can be violated not only by curvature singularities that are not hidden inside of black holes but also by the development of acausal features such as closed timelike curves (CTCs). An example is given by Misner's two-dimensional spacetime (Fig. 2) which serves to illustrate some of the features of Taub–NUT spacetime, which is a vacuum solution of EFE. The Taub region is as causally nice as one could desire, but the NUT region contains CTCs. Such features can be just as disruptive to predictability and determinism as naked curvature singularities. In the present case, the time slice S fails to be a Cauchy surface because of the development of the Cauchy horizon $H^+(S)$ which separates the Taub and NUT regions. There are in fact different (i.e., nondiffeomorphic) extensions of the Taub region, and the initial data on S fail to determine which will be actualized. In this example the chronology horizon is classically unstable. But this is not the case in general, and at present there are no indications of an effective mechanism in classical GTR that would censor the development of CTCs. Hawking[36] proved a chronology protection

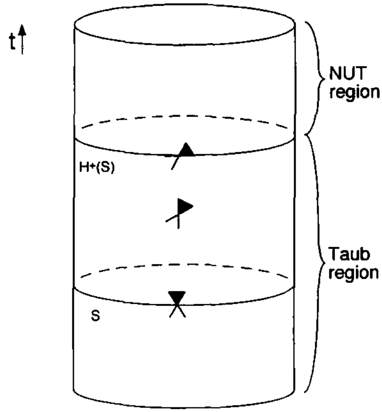

Fig. 2. Two-dimensional Misner spacetime illustrating some features of Taub–NUT spacetime.

theorem which was supposed to show that the combination of EFE and energy conditions do not permit the operation of a time machine which would manufacture CTCs. Although technically correct, the mathematical result does support this gloss.(37) At present, it seems that the only hope for censoring CTCs comes from the calculations of semi-classical quantum gravity that indicate the quantum instability of the Cauchy horizons that would result from the manufacture of CTCs.(38) Once again, curvature singularities are being pressed into the service of cosmic censorship.

5. CONCLUSION

Part of my message has been that a discussion of the implications of spacetime singularities in general and of the seeds-of-its-own-destruction argument against classical GTR in particular needs to be more sensitive to the fact that there are many different kinds of spacetime singularities (or better, many different ways in which a general relativistic spacetime can be singular). The other part of the message is that a generalized *horror singulariti* is not warranted and that not all forms of singularities call into question the soundness and completeness of GTR. If GTR allowed naked singularities to form in circumstances that are regarded as both physically reasonable and generic, then the completeness of the theory would be called into question. Thus, cosmic censorship continues to be perhaps the most important unresolved problem in classical general relativity. Whether or not a quantum theory of gravity will banish singularities remains to be seen. And it also remains to be seen whether we should want quantum gravity to rid us of singularities; for at present the only known mechanism enforcing cosmic censorship in certain circumstances is the divergence of the quantum expectation value of the stress-energy tensor on the Cauchy horizon and the curvature singularity resulting from the backreaction on the metric.

If the pioneers of general relativistic physics had lost their collective nerve when they found that the theory yielded unexpected and seemingly bizarre consequences, our current understanding of the universe would be much impoverished. I suggest that a loss of nerve over the prediction of spacetime singularities could produce another kind of impoverishment. Trust the theory, even when it predicts singularities, and try to learn from these predictions.

REFERENCES

1. C. W. Misner, *Phys. Rev.* **186**, 1328 (1969).
2. E. M. Lifshitz, and I. M. Khalatnikov, *Adv. Phys.* **12**, 185 (1963).
3. I. M. Khalatnikov, E. M. Lifshitz, *Phys. Rev. Lett.* **24**, 76 (1970).
4. S. W. Hawking, and R. Penrose, *Proc. R. Soc. London A* **314**, 529 (1970).
5. P. Bergmann, *Some Strangeness in the Proportion*, H. Woolff, ed. (Addison-Wesley, Reading, Massachusetts, 1980), p. 156.
6. R. Branderberger, V. Mukhanov, and A. Sornborgern *Phys. Rev. D* **48**, 1629 (1993).
7. J. W. Moffat, in *Gravitation: A Banff Summer Institute*, R. Mann and P. Wesson, eds. (World Scientific, Singapore, 1991), pp. 523–597.
8. L. Parker, and S. A. Fulling, *Phys. Rev. D* **7**, 2357 (1973).
9. P. C. W. Davies, *Phys. Lett. B* **68**, 402 (1977).
10. J. S. Bell, and M. Nauenberg, in *Preludes in Theoretical Physics*, A. De Shalit, ed. (North-Holland, Amsterdam, 1966), pp. 279–286.
11. D. Albert, *Quantum Mechanics and Experience* (Harvard University Press, Cambridge, 1992).
12. G. C. Ghirardi, A. Rimini, and T. Weber, *Phys. Rev. D* **34**, 470 (1986).
13. R. P. Geroch, C. Liang, and R. M. Wald, *J. Math. Phys.* **23**, 432 (1982).
14. J. Earman, *Bangs, Crunches, Whimper, and Shrieks: Singularities and Acausalities in Relativistic Spacetimes* (Oxford University Press, Oxford, 1995).
15. A. Einstein, *König. Preuss. Akad. Wissen. (Berlin). Sitz.*, 270 (1918).
16. R. P. Geroch, *Ann. Phys.* **48**, 526 (1968).
17. S. W. Hawking, and G. F. R. Ellis, *The Large-Scale Structure of Space-Time* (Cambridge University Press, Cambridge, 1973).
18. C. W. Misner, *J. Math. Phys.* **4**, 924 (1963).
19. S. M. Scott, and P. Szekeres, *J. Geom. Phys.* **13**, 223 (1994).
20. R. Penrose, *Ann. N.Y. Acad. Sci.* **224**, 125 (1973).
21. C. J. S. Clarke, in *The Renaissance of General Relativity and Cosmology*, G. Ellis, A. Lanza, and J. Miller, eds. (Cambridge University Press, Cambridge), pp. 86–99.

22. K. Thorne, in *High-Energy Astrophysics*, C. DeWitt, E. Schatzman, and P. Véron, eds. (Gordon & Breach, New York, 1967), Vol. 3, pp. 259–441.
23. M. Rowan-Robinson, *Phys. World* **7**, 23 (Nov. 1994).
24. S. W. Hawking, *Proc. R. Soc. London A* **294**, 511 (1966).
25. V. Moncief, in *Lecture Notes in Physics*, Vol. 202 (Springer, 1984), pp. 185–196.
26. N. J. Cornish, and J. W. Moffat, *J. Math. Phys.* **35**, 6628 (1994).
27. S. W. Hawking, *Commun. Math. Phys.* **43**, 199 (1975).
28. S. W. Hawking, *Phys. Rev. D* **14**, 2460 (1976).
29. S. W. Hawking, *Commun. Math. Phys.* **87**, 395 (1982).
30. T. Banks, and M. O'Laughlin, *Phys. Rev. D* **47**, 4476 (1993).
31. C. R. Stephens, G. t'Hooft, and B. F. Whiting, *Class Quatum Gravit.* **11**, 621 (1994).
32. G. T. Horowitz, and D. Marolf, "Quantum Probes of Spacetime Singularities," UCSBTH-95-5, preprint.
33. C. J. S. Clarke, *Commun. Math. Phys.* **41**, 65 (1975).
34. F. Mellor, and I. Moss, *Class. Quantum Gravit.* **9**, L43 (1992).
35. D. Marković, and E. Poisson, *Phys. Rev. Lett.* **74**, 1280 (1995).
36. S. W. Hawking, *Phys. Rev. D* **46**, 603 (1992).
37. J. Earman, *Erkenntnis* **42**, 125 (1995).
38. K. Thorne, in *General Relativity and Gravitation* 1992, R. J. Gleiser, C. N. Kozameh, and O. M. Moreschi, eds. (Institute of Physics, New York, 1993).