The Analysis of Singular Spacetimes

Author(s): Erik Curiel

Source: *Philosophy of Science*, Sep., 1999, Vol. 66, Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part I: Contributed Papers (Sep., 1999), pp. S119-S145

Published by: The University of Chicago Press on behalf of the Philosophy of Science Association

Stable URL: http://www.jstor.com/stable/188766

The Analysis of Singular Spacetimes

Erik Curiel†‡
Stanford University

Much controversy surrounds the question of what ought to be the proper definition of 'singularity' in general relativity, and the question of whether the prediction of such entities leads to a crisis for the theory. I argue that a definition in terms of curve incompleteness is adequate, and in particular that the idea that singularities correspond to 'missing points' has insurmountable problems. I conclude that singularities per se pose no serious problem for the theory, but their analysis does bring into focus several problems of interpretation at the foundation of the theory often ignored in the philosophical literature.

> The mind of man, by nature a monist, cannot accept *two* nothings; he knows there has been *one* nothing, his biological inexistence in the infinite past, for his memory is utterly blank, and *that* nothingness, being, as it were, past, is not too hard to endure. But a second nothingness—which perhaps might not be so hard to bear either—is logically unacceptable.
>
> V. Nabokov, *Ada*

†Philosophy Department, Stanford University, Palo Alto, CA 94305-2155; e-mail: encuriel@csli.stanford.edu.

‡This paper began life as a small criticism of a few points John Earman makes in Chapter 2 of his book *Bangs, Crunches, Whimpers and Shrieks*, and grew as I grew to realize more fully the complexity and subtlety of the issues involved. I shall not always point out where I am in agreement or disagreement with Earman, much less always discuss why this is so, though I shall try to on the most important points. The reader ought to keep in mind, though, that Earman's book is the constant foil lurking in the background. I thank R. Geroch and D. Malament for stimulating conversations on all these topics. I am also grateful to M. Dorato for writing a review of Earman 1995 that made me realize the need to reread it and think more about singular structure, and to the History and Philosophy of Science Department at Pittsburgh, where I presented an earlier, briefer, version of this paper in a colloquium, for stimulating questions.

Philosophy of Science, 66 (Proceedings) pp. S119–S145. 0031-8248/99/66supp-0010$0.00

1. Introduction. I suspect that, for many, talk of a singularity in the context of general relativity conjures up the image of something like a rent in the fabric of spacetime. This metaphor, evocative as it may be, perhaps misleads: a web of cloth exists in space and time, and one naturally would rely (implicitly, at least) upon this fact were one to define what one meant in saying the cloth had a hole—one would say that points were missing from the cloth, a notion made precise by the embedding of the cloth in physical space. When thinking of a singular spacetime,[1] though, one does not have the luxury of imagining it embedded in any physically meaningful way in a larger space with respect to which one can try to define what one means by saying there are points missing from spacetime.

On a manifold endowed with a positive-definite Riemannian metric, one can give a precise characterization of what it is for there to be missing points that accords quite well with our intuitions. The manifold has no missing points if and only if it is Cauchy complete as a metric space. On a manifold with a pseudo-Riemannian metric of Lorentz signature, such as a spacetime in general relativity, there is no natural way to construct a metric that measures the distance between points of the manifold, so one cannot employ this technique to test whether a spacetime has missing points. By the Hopf-Rinow theorem, the manifold in the Riemannian case is Cauchy complete with respect to the constructed metric if and only if it is geodesically complete with respect to the Riemannian metric.[2] This naturally suggests that we define a spacetime to have missing points if and only if it is geodesically incomplete with respect to the spacetime pseudo-Riemannian metric. Now one faces a severe problem, which lies at the heart of the difficulty in giving a precise and intuitively satisfying definition of a spacetime singularity as a point missing from spacetime: there is no natural way to take a Cauchy-like completion of the spacetime manifold in order to give substance to the idea that there really are points that in some sense 'ought to have been included in the spacetime' in the first place.

The usual tack taken at this point in the physics literature is simply to bracket the question of missing points and define a spacetime to be singular if and only if it contains incomplete, inextendible curves of a certain specified type, and the spacetime manifold itself satisfies a few collateral conditions. More precisely, the commonly accepted schema for fixing a rigorous definition of a singular spacetime is:

1. By 'spacetime', I shall always mean a smooth 4-dimensional connected paracompact manifold with a fixed smooth metric of Lorentz signature.

2. See Spivak 1979, Ch. 9, for a precise statement and proof of the theorem.

A spacetime (M, g_{ab}) satisfying ____ is *singular* if and only if there exists a curve γ incomplete in the sense that ____.[3]

Such a conception of singular structure actually has a lot to say for itself, as capturing the idea that singular structure is somehow physically *outré*, even if one is not able to hook it up cleanly to an idea of missing points: an observer travelling along such an incomplete curve would be able to experience only a finite total amount of proper time.

This paper has several concrete aims: to investigate particular ways that have been proposed to fill in the blanks of the schematic definition with an eye to determining whether they capture the spirit of the idea that an incomplete curve corresponds to singular structure; to argue that the idea of missing points ought not be central in thought about singular structure; and to argue that the reasons most often given for eschewing singular structure as unphysical do not withstand scrutiny. It also has one overarching, more inchoate aim: to try to give a sense of the marvellous philosophical riches still waiting to be mined from thorough investigation of the foundations of general relativity—which is to say, a sense of how little we still comprehend of and about this astounding theory, and how much we stand in need of that comprehension if we wish to understand the world.

2. Curve Incompleteness. The path-breaking work of the mid-1960s demonstrating the existence of singular structure in generic solutions to the Einstein field equations invoked timelike or null geodesic incompleteness as a sufficient condition for classifying a spacetime as singular, in so far as timelike and null geodesics represent possible worldlines of particles and observers and it appears prima facie physically suspect for an observer or a particle to be allowed to pop in or out of existence right in the middle of spacetime, so to speak.[4] There was, however, no consensus on what ought to count as a necessary condition.

Geroch (1968b) gave the first extended discussion of the difficulty of framing a satisfactory definition of a singular spacetime. He settled provisionally on simple geodesic incompleteness as the criterion for singular structure, conceding that the definition is perhaps overly inclusive, but better to brand 10 innocents than to leave one guilty man unmarked. The possible innocents include spacetimes that are timelike and null geodesically complete but possess incomplete spacelike geodesics (null and timelike incomplete and spacelike complete, for short).

3. See, for example, Clarke 1993, 10; Joshi 1993, 161–162; Wald 1984, 212–216; and Hawking and Ellis 1973, 256–261.

4. Cf. Geroch 1966, Hawking 1965, Hawking 1967, and Penrose 1965.

Spacelike incompletenesss (in the absence of the other two types of incompleteness) sets off no serious alarms, for an incomplete spacelike geodesic seems to represent structure of the spacetime that is not physically accessible to any observer (I will discuss this matter more thoroughly in §6). Moreover, not only does geodesic incompleteness lock up a few possible innocents but, as Geroch proceeds to show, it almost certainly fails to nab a few clever guilty parties, for a spacetime can be geodesically complete and yet possess an incomplete timelike curve of bounded total acceleration—that is to say, an inextendible curve traversable by a rocket with a finite amount of fuel, along which an observer could experience only a finite amount of proper time.

Because of these problems, null and timelike geodesic incompleteness continued to be used as a sufficient condition for anointing a spacetime singular, but was considered to be inadequate as a definition. To analyze the structure of non-geodetic curves in the search for a necessary condition, a method is required to characterize their completeness. Schmidt (1971) appears to have been the first to propose using so-called generalized affine parameters to define the completeness of general curves. Any curve of unbounded proper length automatically has an unbounded generalized affine parameter, but not vice versa—any inextendible timelike curve of unbounded total acceleration and finite total proper time in Minkowski space, for example, has an unbounded generalized affine parameter. A spacetime in which every inextendible curve has an unbounded generalized affine parameter will be referred to as *b-complete.*[5] Thus, one has what Earman (1995, 36) refers to as the "semioffical view": a spacetime is said to be singular if and only if it is b-incomplete.[6] This definition is more general than geodesic completeness, in that it implies, but is not implied by, the latter, as Geroch's example demonstrates.

It is difficult to think of a more comprehensive criterion of completeness than *b*-completeness, and I suspect its popularity arises therefrom, but that it sits comfortably with some of the intuitions that drove the search for a definition of singular structure in the first place is not so clear on reflection. For the moment, I shall accept *b*-incompleteness as the definition of singular structure—when I refer to 'incomplete

5. '*b*' for 'bundle': with this construction one tacitly defines a natural (basis-dependent) Riemannian metric on the bundle of frames of the spacetime manifold to define curve completeness. See Schmidt 1971 for more details.

6. Strictly speaking, this is not the standardly accepted definition, since I have not mentioned anything about the maximality of the spacetime in question, whether, that is, it can be embedded in (thought of as merely a part of) a larger spacetime in such a way as to make previously incomplete, inextendible curves extendible. I shall take up this issue in §6.

curves', unless I explicitly state otherwise I shall mean *b*-incomplete, inextendible curves. I shall return to some of these questions in §6.

3. Curvature Blowup Along Incomplete Curves. While curve incompleteness seems to capture one aspect of the intuitive picture of singular structure, it completely ignores a different aspect, curvature pathology. One may measure the growth and diminution of spacetime curvature in various ways, but it turns out that the unbounded growth of curvature according to any of these measures is neither necessary nor sufficient for the existence of incomplete, inextendible curves. To get an idea of the independence of the existence of incomplete curves from the presence of curvature pathology, consider the striking ease with which examples of a spacetime with everywhere vanishing Riemann tensor and incomplete geodesics can be constructed: excise from 2-dimensional Minkowski space a closed set in the shape of an echidna. This example may strike one as cheating, since one has only to restore the excised set to restore geodesic completeness (or, in fancier terms, one has only to isometrically embed the mutilated spacetime by the natural inclusion map back into Minkowski spacetime to restore completeness). So a slightly more sophisticated example: for some $0 < \phi_0 < \frac{\pi}{2}$, excise from Minkowski space, represented in polar coordinates, the wedge consisting of all points with azimuthal coordinate $0 < \phi < \phi_0$; identify the corresponding points on the hyperplanes $\phi = 0$ and $\phi = \phi_0$. By a suitable redefinition of the coordinate neighborhoods of the points on $\phi = 0$, the resulting space can be given the manifold structure of $\mathfrak{R}^4$, and the Minkowski metric can be smoothly extended to the points at $\phi = 0$, $r > 0$. It cannot be smoothly extended to the points $r = 0$, however, and so these points must be excised from the spacetime. The Riemann tensor of this spacetime vanishes everywhere, but any geodesic that previously passed through the line $r = 0$ will now be incomplete; there is, moreover, no other spacetime into which this spacetime can be embedded and in which the metric can be smoothly extended.[7] This sort of structure is known as a 'conical singularity', since it corresponds to taking a wedge out of the 2-dimensional real plane and pasting the edges together to form a cone.

The two most commonly used methods of measuring the growth of curvature intensity are the behavior of scalar curvature invariants along some particular curve through the region of interest, and the behavior of the physical components of the Riemann tensor as mea-

7. This example is from Wald 1984, 214. See Ellis and Schmidt 1977, 921–923, for further discussion of this sort of singular structure.

sured by a frame parallel-propagated along some particular curve through the region of interest (if any of the physical components grow without bound in such a frame on a particular curve, then they will in all such frames on that curve).[8] In accordance with customary usage, the existence of an incomplete curve along which the physical components of the Riemann tensor in a parallel-propagated frame do not approach a definite finite limiting value will be referred to as *p.p.-singular structure*, and the same of some scalar curvature invariant along an incomplete curve will be referred to as *s.p.-singular structure* ('s.p.' for 'scalar polynomial'). The existence of an incomplete curve along which the physical components of the Riemann tensor in parallel-propagated frames and all its scalar invariants converge to definite, finite values will be called *quasi-regular singular structure.*[9] Note that curvature pathology on these definitions occurs not only if some feature of the curvature grows without bound along an incomplete curve, but also if it oscillates indefinitely (even if only within finite bounds), never settling down to a limiting value.

I believe there are two primary motivations for using a parallel-propagated frame in which to express the components of the Riemann tensor. First, one naturally expects the presence of curvature pathology to show itself, at the least, in misbehavior of the tidal forces an observer would experience.[10] The intensity of tidal force, as measured in any frame, is directly proportional to the components of the Riemann tensor in that frame, as one can see from the equation of geodesic deviation (Hawking and Ellis 1973, 80). In a back-of-the-envelope sort of way, the unbounded growth of the components of the Riemann tensor in a parallel-propagated frame would seem to indicate that an observer traversing that curve would experience unbounded tidal forces as well. Second, Clarke (1973) demonstrated that an incomplete curve in a singular spacetime has a local extension if and only if the relevant incom-

8. A *frame* is a pseudo-orthonormal complete set of basis vectors for the tangent plane over a point of a manifold. A *frame-field* is an assignment of frames to points in some specified region, e.g. along a curve or in an open set.

9. Quasi-regular singular structure is perhaps the most psychologically disturbing, since it can be absolutely inobservable until one runs into it, so to speak, creating a hair-raising hazard for spacetime navigation.

10. Tidal force is generated by the differential in intensity of the gravitational field, so to speak, at neighboring points of spacetime. For example, when I stand, my head is farther from the center of the Earth than my feet, so it feels a (practically negligible) smaller pull downward than my feet. For a graphic illustration of the effects of tidal forces on observers in strong gravitational fields, see the description in Misner, Thorne, and Wheeler 1973, §32.6, of what would happen to a person standing on the surface of a collapsing star—not for the faint of heart, or weak of stomach.

plete curve constitutes quasi-regular singular structure. A local extension is an isometric embedding of an open subset of the spacetime manifold containing the incomplete curve into another spacetime in which the curve can be extended. Local extensions can exist even when the singular spacetime as a whole is not embeddable as a proper open submanifold into a larger spacetime in which the incomplete curves can be extended (Ellis and Schmidt 1977, 928–929). Many take the existence of local extensions to indicate that nothing *local*, such as curvature pathology (narrowly construed), goes wrong in quasi-regular singular spacetime, but rather some global structure impedes the extension of spacetime. The motivation for using scalar curvature invariants to scout for curvature pathology is somewhat more straightforward. A scalar curvature invariant at a point does not depend on what curve through that point or what frame on a curve through that point one uses to probe the point: it is, as the name suggests, invariant. Unbounded growth of a scalar curvature invariant, moreover, is logically equivalent to the unbounded growth of the components of the Riemann tensor as measured in *every* frame-field along the curve, parallel-propagated or not.

S.p.-singular structure implies, but is not implied by, p.p.-singular structure. In fact, all scalar curvature invariants can be zero and yet the Riemann tensor not be equal to zero, as in plane gravitational wave spacetimes (Penrose 1960, 189). Colliding thick gravitational wave spacetimes provide examples of p.p.-singular structure in regions where all scalar curvature invariants are well-behaved (Konkowski and Helliwell 1992). More strikingly, colliding sandwich plane gravitational wave spacetimes can exhibit p.p.-singular structure and yet all scalar curvature invariants remain identically zero; finally, colliding plane gravitational wave spacetimes also provide less artificial examples than the conical singularity above of the existence of incomplete curves in regions of a spacetime in which the Riemann tensor itself vanishes, viz. quasi-regular singular structure (Konkowski and Helliwell 1992). Thus the existence of incomplete curves does not ipso facto necessitate any sort of curvature pathology as conventionally quantified. That the misbehavior of the physical components of the Riemann tensor in a parallel-propagated frame or of a scalar curvature invariant in the limit as one traverses a curve does not suffice to ensure that the curve be *b*-incomplete follows from examples of spacetimes produced by Sussmann (1988) in which scalar curvature invariants diverge asymptotically along complete timelike and null geodesics.

Though there is no necessary connection of any sort between the existence of incomplete curves and curvature pathology as quantified in the standard ways, the *b*-completeness criterion does allow one to categorize singular spacetimes according to the behavior of the curvature along the incomplete curves as quantified in the standard ways sketched above. Earman (1995, 37, 43–44) goes so far as to proclaim one of the most seminal virtues of the *b*-completeness definition that it allows for a categorization of this sort. The categorization has a binary branching structure: first, an incomplete curve is said to constitute *essential singular structure* if there is no larger spacetime into which the singular spacetime can be embedded as a proper open submanifold such that the curve is extendible in the larger spacetime, and otherwise it is said to be *inessential*; essential singular structure is then subdivided into quasi-regular and p.p.-singular structure; finally, p.p.-singular structure is subdivided into s.p.-singular and non-s.p.-singular structure (Ellis and Schmidt 1977).

The thought behind the putative importance of the categorization scheme seems to be as follows. Very little is known about singular structure at the present time, in part due to the difficulty of the mathematics involved in analyzing singular structure rigorously and in part due to the vanishingly small amount of experimental access we can get to singular structure in the foreseeable future. Nevertheless, the singularity theorems indicate that the spacetime we actually inhabit is singular, so it behooves us to try to understand such structure as much as possible. Categorizing singular structure appears to be a way for us to organize and begin to get a grip on such a daunting task. To be appropriate for such a task, I submit, the mathematically different species of singular structure ought to exhibit prima facie different sorts of physical behavior, as near as one can judge that sort of thing with the crude tools at our disposal; otherwise it will be difficult to see the physical relevance of this so far purely mathematical categorization.

As already noted, in a spacetime with s.p.-singular structure, the Riemann tensor components will behave badly as expressed in *any* frame-field along the relevant incomplete curve, and, moreover, will do so in general along any curve close enough, as it were, to the incomplete curve.[11] The tidal forces a body will feel as it moves along a curve in spacetime are naturally measured in a spacelike 3-frame fixed rigidly in the body, orthogonal to the timelike unit vector tangent to the curve,

11. More precisely, in general there will exist an open neighborhood of the incomplete curve such that every curve completely contained in the open neighborhood has Riemann components that are as badly-behaved as one likes in all frames along the curve. The 'in general' hedges against the case where the scalar curvature invariant oscillates wildly along the incomplete curve; in this case, it may be possible for nearby curves to weave cleverly around the incomplete curve in such a way as to avoid the peaks of oscillation, and so have well-behaved Riemann tensor components. No hard results are known either way in such cases.

used to fill out the full 4-frame. Based on what has already been said, one might expect that the state of motion of the observer along the curve, whether the observer is slowing down and speeding up, or spinning on his or her axis, would have no effect on how the observer experiences the curvature pathology: when a scalar curvature invariant grows without bound along a curve, after all, the tidal forces as measured in *any* frame along the curve also will grow without bound. Interestingly enough, however, the state of motion of the observer as it traverses an incomplete curve, so-called inertial effects, can be important in determining the physical response of an object to the curvature pathology. Whether the object is spinning on its axis or not, for example, or accelerating slightly in the direction of motion, may determine whether the object gets crushed to zero volume along an s.p.-singular curve or whether it survives (roughly) intact all the way along the curve (Ellis and Schmidt 1977, 944–947).

The effect of the observer's state of motion on his or her experience of tidal forces can be even more pronounced in the case of p.p.-singular structure that is not s.p.-singular, which is precisely the existence of an incomplete curve along which there is a frame-field (necessarily not parallel-propagated) relative to which the components of the Riemann tensor approach definite, finite limiting values along the curve (Ellis and Schmidt 1977, 939). In such a case, the frame-field in which the physical components of the Riemann tensor stably approach a limit is related to any parallel-propagated frame-field by a Lorentz transformation that, in an appropriate sense, behaves pathologically in the limit along the curve. For a non-geodetic curve, the proper mode of transport along a curve of a frame rigidly fixed in the body of an object traversing that curve is not parallel-propagation but Fermi-transport (Hawking and Ellis 1973, 80–81). A Fermi-transported frame is related to a parallel-propagated frame by a continuously varying Lorentz transform. It can happen, therefore, that an observer cruising along a p.p.-singular curve that is not s.p.-singular would experience unbounded tidal forces and so be torn apart while another observer, in a certain technical sense approaching the same limiting point as the first observer, accelerating and decelerating in just the proper way, would experience perfectly well-behaved tidal force, though he would approach as near as one likes to the other poor fellow in the midst of being ripped to shreds. Again, certain gravitational plane wave spacetimes provide good examples of this phenomenon: an observer travelling along the incomplete timelike geodesic constituting the singular structure would experience unbounded tidal acceleration, whereas *any* observer travelling arbitrarily close by would not (Ellis and Schmidt 1977, 937).

Things can get stranger still. An incomplete geodesic contained entirely within a compact subset of a spacetime, with accumulation point p, that satisfies a certain genericity condition necessarily constitutes p.p.-singular structure, so that an observer freely falling along such a curve would be torn apart by unbounded tidal forces; it can easily be arranged in such circumstances, though, that a separate observer, who actually travels through p, will experience perfectly well-behaved tidal forces (Hawking and Ellis 1973, 290–292). Here we have an example of an observer being ripped apart by unbounded tidal forces right in the middle of spacetime, as it were, while other observers cruising peacefully by could reach out to touch him or her in solace during the final throes of agony.

This discussion points to a startling conclusion: curvature pathology, as standardly quantified, is not a well-defined property of a region of spacetime *simpliciter*, but may in fact sensitively depend on how one probes spacetime regions with various curves, and the state of motion of an observer or test particle along the curves! These matters are far more subtle and complicated than many, notably Earman (1995), would lead one to believe. I believe there is far more work to be done straightening out the physical consequences of the existence of singular structure.

Ellis and Schmidt say, vis-à-vis their classificatory scheme (the canonical one):

> It is not claimed here that the singularities discussed are *likely* to occur in physically realistic situations, but rather that only when we understand which singularities can occur (a) in general space-times, and (b) in space-times with the field equations satisfied for particular matter content, can we hope to discuss fruitfully their occurrence, equations of motion, and so on (1977, 918).

I do not mean to argue with the motivation for their classificatory scheme, but they beg a serious question with their 'which' in the phrase "when we understand which singularities can occur": clearly the correlative demonstratives of this relative interrogative refer to the different classes of their categorization, but why ought one think that their classification picks out physically relevant differences among all possible singular structures? This question becomes more poignant when one reflects on the fact that curvature pathologies provide the differentiae for their speciation, and I have attempted to show that curvature pathology as customarily quantified is not a straightforward concept with clear and unambiguous physical content. The mathematics has outrun the physics, but still masquerades as such.

Taub is the only person I know in print who shares my apprehension about the status of the canonical classification scheme:[12]

> I have difficulty understanding the usefulness of the classification scheme of singularities proposed . . . by Ellis and Schmidt. . . . I think that the important work on singularities now being done would become much more important if it turned toward learning how to deal with the physics associated with singularities. . . . (Taub 1979, 1009)

He appears to be saying that one ought to concentrate first on trying to work out the behavior associated with various singular structures we are more or less familiar with in a clear and unambiguous way, and only then should one feel confident enough to begin classifying singular structures, based on that clear physical knowledge, not on a purely mathematical scheme that becomes murky as soon as one tries to think about it in physical terms. I heartily concur.[13]

4. Missing Points. We now have a precise definition of a singular spacetime, and some ideas about what such structure implies and does not imply about the curvature of spacetime, but, as Earman notes, "it is not true to an idea that is arguably a touchstone of singularities in relativistic spacetimes: spacetime singularities correspond to missing points" (Earman 1995, 40). For those who would argue missing points ought to be such a touchstone, Earman sketches what seems to me the most (initially) promising position, that, though the idea of missing points and that of curve incompleteness lead to *prima facie* different concepts of singular structure, they are extensionally equivalent in all physically reasonable singular spacetimes, and so the two concepts are for all practical purposes in agreement (Earman 1995, 42). I shall argue with this: missing points ought not be a touchstone of discussion of singular structure in relativistic spacetimes.

Missing points, could they be defined, would correspond to a boundary for a singular spacetime—actual points of an extended spacetime

12. Though R. Geroch has told me in conversation that he does not see the use of the classification scheme either, because he is not sure what physical content it has.

13. A physically unambiguous sense of curvature pathology occurs in, *e.g.*, the Friedmann-Roberston-Walker metrics, wherein physical quantities such as the mass-density of ponderable matter grow without bound along incomplete curves and thus scalar curvature invariants correlatively grow without bound as well. This sort of idea is developed nicely in a not very well known paper (to judge by its citation record) by Thorpe (1977). I think it would be of interest to see whether a categorization scheme based on some of Thorpe's ideas could be constructed and compared to the canonical categorization.

at which curves incomplete in the original spacetime would terminate.[14] My argument therefore will alternate between speaking of missing points and speaking of boundary points, with no difference of sense intended. Before I begin examining the primary attempts to define boundary points for singular spacetimes,[15] it is well to note an oddity of the situation: compact spacetimes can contain incomplete, inextendible geodesics, as shown by a simple example due to Misner (1963). In a sense that can be made precise, Hausdorff compact sets, from a topological point of view, 'contain every point they could possibly be expected to contain',[16] one manifestation of which is that a compact manifold cannot be embedded as an open submanifold of any other manifold, a necessary prerequisite for attaching a boundary to a singular spacetime—a manifold-with-boundary minus its boundary is embeddable by the identity map as an open submanifold into itself. This already suggests that, even were one able to come up with a satisfactory definition of missing points in the context of Lorentzian metrics, it may not be extensionally equivalent to the existence of incomplete curves, unless we are willing to swallow unpalatable topological structure.

Schmidt (1971) produced the most well-known boundary construction for singular spacetimes, the so-called *b*-boundary based on the *b*-completeness criterion. The relativity community at first embraced Schmidt's construction with enthusiasm, to judge by the remarks in Chapter 8 of Hawking and Ellis's canonical work *The Large Scale Structure of Space-Time.* Shortly thereafter, however, Bosshard and Johnson showed that the *b*-boundary had undesirable properties in the most physically relevant spacetimes known, the Friedmann-Robertson-Walker spacetimes, which to a quite high degree of approximation accurately model the large scale structure of the actual universe, and the Schwarzschild spacetimes, which represent the neighborhood of spherically symmetric isolated bodies, such as stars.[17] For closed Friedmann-Robertson-Walker spacetimes, the *b*-boundary consists of a single point (the same for the big bang as for the big crunch) that is not Hausdorff-separated from any point in the interior of the spacetime. Not only does one reach the same point, then, by travelling either for-

14. Strictly speaking, such a space would not be a manifold in the usual sense of the term, but a manifold with boundary. See Spivak 1979.

15. I shall not consider in this paper the 'ideal-point' boundary construction of Geroch, Kronheimer, and Penrose (1972), as it requires the singular spacetime to be past- and future-distinguishing, a fairly strong causality condition. I intend to sidestep all questions about the physical plausibility or necessity of such conditions.

16. See Geroch 1985, §30, for a discussion of this precise sense.

17. Cf. Bosshard 1976 and Johnson 1977.

ward or backward in time, but that point is, in a certain sense, arbitrarily near every single spacetime event! Similarly, the *b*-boundary of a Schwarzschild spacetime consists of a single point not Hausdorff-separated from any interior point of the spacetime. This certainly will not do for the advocates of missing points.

The reactions to these problems vary widely. Clarke (1993) still embraces the *b*-boundary construction, and defines a singularity to be a point on the *b*-boundary of a singular spacetime (§3.4). He barely mentions these problems, noting only in passing that the topological structure of the singular spacetime with boundary can be "very strange," (40) which I do not think an adequate address. Wald (1984), on the other hand, does not like the *b*-boundary construction precisely because of these problems (cf. 213–214), and Joshi (1993) does not even mention the possibility of attaching boundaries to singular spacetimes, speaking only of incomplete curves.

A second method of constructing a boundary for singular spacetimes due to Geroch (1968a) fares much better with physically relevant spacetimes. In this construction, the so-called *g-boundary*, geodesic incompleteness rather than *b*-incompleteness defines singular structure, and one defines a boundary point to be an equivalence class of incomplete geodesics under the equivalence relation 'approach arbitrarily close to each other' (in a certain technical sense). The set of boundary points can be given a topology and, in many cases of physical interest, can even be given a differentiable and metric structure, so that one can locally analyze the structure of spacetime at a 'singularity' rather than mess around with troublesome limits along incomplete curves.[18] The *g*-boundary construction, moreover, yields the boundaries one might have expected on physical grounds in spacetimes of particular physical interest: the *g*-boundary of a Schwarzschild spacetime is a spacelike 3-surface, topologically $S^2 \times \Re$, and that of a closed Friedmann-Robertson-Walker spacetime is the disjoint union of two spacelike S^3's. Pathological topology rears its head here as well, though, in the case of Taub-NUT spacetime: the *g*-boundary of this spacetime contains a point that again is not Hausdorff-separated from any point in the interior of the spacetime.[19]

The advocate of missing points may at this point retort that Taub-NUT spacetime hardly constitutes a physically relevant spacetime for other reasons, namely that it violates strong causality, which is to say

18. In certain *outré* examples, there is an ambiguity in choice of topology for the *g*-boundary, but I shall waive this concern for the sake of argument. There are bigger fish to fry.

19. Cf. Hawking and Ellis 1973, §5.3, for a thorough account of Taub-NUT spacetime.

that it contains causal curves that come arbitrarily close to intersecting themselves. While I do not think this reply carries much weight,[20] I have a better example at hand. Geroch, Can-bin, and Wald (1982) construct a geodesically incomplete spacetime with no causal pathology for which a very large class of boundary constructions, including the *b*- and the *g*-boundary, will yield pathological topology in the completed spacetime (the conditions that a boundary construction must satisfy to fall prey to this example are quite weak). The advocate of missing points may point out that the example appears artificial and contrived, with closed sets excised here and conformal factors plastered on there, and in short has no physical relevance. Ellis and Schmidt (1977, 932) exemplify this sort of simplicity chauvinism: "We know lots of examples of [flat singular spacetimes], all constructed by cutting and gluing together decent space-times; and because of this construction, we know that these examples are not physically relevant."

I would reply that this judgment has its roots in the schooling our intuitions have received in our contemplation of well-worked out examples of physical theories, which by and large tend to include mathematical structures that strike us as 'simple' and 'natural'. This ought not escape our notice: most such examples of physical theories are demonstrably false (Newtonian mechanics and classical Maxwell theory) or have at the moment insuperable problems of interpretation (quantum mechanics) or experimental accessibility (general relativity). We should beware of relying too much on intuitions trained in such schools—especially when one also recalls how much of our contemplation of those theories involves models of systems with physically unrealistic perfect symmetries and vaguely justified approximations and simplifications. It may turn out, for all we know, that spacetime instantiates just such topological structure as $\mathfrak{R}^4$ with certain closed sets excised. Perhaps the most important point to notice, though, is that "$\mathfrak{R}^4$ with certain closed sets excised" is a *misleading* description of such a manifold. It suggests that we built that manifold from a more fundamental one, viz. $\mathfrak{R}^4$. But that manifold *simply is a manifold* all on its own, with no intrinsic reference to $\mathfrak{R}^4$, or indeed any other manifold. Because of certain facts about how we practice mathematics, the most convenient presentation of that manifold happens to be "$\mathfrak{R}^4$ with certain closed sets excised." One could as legitimately present $\mathfrak{R}^4$ as that manifold glued together with certain other manifolds-with-boundary. There are no good grounds I can see for suspecting that the universe heeds our preferred methods for organizing mathematical structures.

20. See Earman 1995, Chs. 6–7, for a discussion of why a violation of strong causality *simpliciter* does not constitute an argument for the unphysicality of a spacetime.

I refer those unmoved by this sermon to a remark that Geroch, Can-bin, and Wald (1982, 435) make: "The purpose of [a boundary] construction, after all, is merely to clarify the discussion of various physical issues involving singular space-times: general relativity as it stands is fully viable with no precise notion of 'singular points.' " When we contemplate potential phenomena that we have little or no observational access to, I submit that the standards for what can count as a *physical* account of a situation ought to be priggishly severe, if we are not unwittingly to degenerate into pure mathematical discourse.[21] A construction that yields topological pathology, and contains no precise criteria for what ought to count as a 'physically relevant' spacetime, does nothing to clarify discussion of the physical issues involved in analyzing singular spacetimes.

The abstract boundary construction, or *a-boundary*, proposed by Scott and Szekeres (1994) appears at first glance to have the most promise for those wanting a natural, workable definition of missing points for singular spacetimes.[22] It also nicely exemplifies a feature of all missing point constructions I know of or can easily imagine, their dependence on a prior characterization of incomplete curves. For these two reasons, I shall consider it in a little more detail than the previous two. An *envelopment* of a manifold $\mathcal{M}$ is an ordered pair $(\mathcal{N}, \phi)$ consisting of a manifold $\mathcal{N}$ and an embedding ϕ into $\mathcal{N}$ of $\mathcal{M}$ as a proper open submanifold of the same dimension.[23] Scott and Szekeres propose that singular structure always arises by the deletion of points from an envelopment of a singular manifold. Given an envelopment $(\mathcal{N}, \phi)$ of $\mathcal{M}$, a subset of its topological boundary in $\mathcal{N}$ will be called a *boundary set*. Now, as it clearly is possible to envelop a given manifold in many ways (if the manifold has any envelopment at all), one does not want to consider merely boundary sets of manifolds under particular envelopments, but rather equivalence classes of boundary sets under some appropriate equivalence relation. To this end, Scott and Szekeres propose the following:

21. R. Geroch stressed this point to me in a conversation in which he also dismissed the adequacy of his own g-boundary construction *merely because* it gave unphysical results in the admittedly contrived example of Geroch, Can-bin, and Wald (1982). It gives very nice results in almost all other known types of examples.

22. Whether the a-boundary construction satisfies the conditions of Geroch, Can-bin, and Wald (1982), and so necessarily leads to pathological topology for certain space-times, is not clear, for as of yet Scott and Szekeres have not defined a topology for their construction at all. From the structure of the construction, I suspect that any topology one would define for it would satisfy Geroch, Can-bin, and Wald's conditions.

23. When it can cause no confusion, I shall often identify $\mathcal{M}$ with its image under the envelopment mapping.

Definition 4.1 *A boundary set B of $\mathcal{M}$ in an envelopment $(\mathcal{N}, \phi)$ is said to* cover *the boundary set B' of $\mathcal{M}$ in an envelopment $(\mathcal{N}', \phi')$ if for every open neighborhood U' in $\mathcal{N}'$ of B' there exists an open neighborhood U in $\mathcal{N}$ of B such that*

$$\phi \circ \phi'^{-1}[U' \cap \phi'[\mathcal{M}]] \subset U.$$

A boundary set B may cover another boundary set B' while B' does not cover B. One easily sees, however, that defining B and B' to be equivalent if they mutually cover each other does in fact yield an equivalence relation; the equivalence class of the boundary set B under this relation will be written '$[B]$' and called an *abstract boundary set.* An equivalence class that contains a singleton as a representative member will be called an *abstract boundary point.* The collection of all abstract boundary points is the *abstract* or *a-boundary*, written '$\mathcal{B}[\mathcal{M}]$'.

Although $\mathcal{B}[\mathcal{M}]$ by itself is defined without reference to any particular geometrical structure on $\mathcal{M}$, such as a pseudo-Riemannian metric or an affine connection, which Scott and Szekeres take to be one of its cardinal virtues, to define singular structure they must select a class of curves $\mathcal{C}$ on $\mathcal{M}$ satisfying what they call the bounded-parameter property: roughly speaking, the curves in $\mathcal{C}$ must cover the manifold and must be such that the parameter along any of the curves grows without bound if and only if it grows without bound along every nice reparametrization of the curve. The class of geodesics on a manifold with affine connection and the class of C^1 curves parametrized by generalized affine parameter on a manifold with affine connection provide two examples of classes of curves satisfying the bounded-parameter property. The idea is that curves in $\mathcal{C}$ will be used to probe the boundary to distinguish points 'at infinity' from points that can be reached in a finite parameter interval and hence are candidate singular points. The details of the construction and definitions hereon out become quite complicated, so I shall sketch only the most salient points. First, for a candidate singular spacetime $\mathcal{M}$, Scott and Szekeres wish to remove from consideration all abstract boundary points that have a representative singleton boundary point in some envelopment through which, in a certain technical sense, the spacetime metric can be smoothly extended. In this case, the thought is, the original spacetime simply had not been made as 'large' as it reasonably could have. Such points will be called *regular*, and need not apply as potential singular points. Next, one fixes the class of curves $\mathcal{C}$, and defines the *$\mathcal{C}$-boundary* to be the class of *a*-boundary points that have a singleton representative in some envelopment that is the limit point of a curve in $\mathcal{C}$; such points are referred to as *approachable.* All other *a*-boundary points are *unapproachable.* It is straightforward to show that the property of being approachable or unapproachable is invariant under the defining a-boundary equivalence relation, but one must keep in mind that it depends crucially on the class of curves $\mathcal{C}$ chosen. A non-regular point in an envelopment $\mathcal{N}$ on the boundary of $\mathcal{M}$ that is not the limit point of any curve of bounded parameter in $\mathcal{C}$ will be called a *point at infinity*; if, moreover, it cannot be covered by any regular boundary set of another envelopment, it will be called an *essential* point at infinity. This property is clearly invariant under the a-boundary equivalence relation, and so one speaks of a-boundary points at infinity. A non-regular boundary point p of $\mathcal{M}$ in the envelopment $\mathcal{N}$ that is the limit point of some curve in $\mathcal{C}$ of bounded parameter will be called a *singular point*. If there exists a non-singular boundary set of another envelopment that covers p, then it is said to be *removable*; otherwise it is *essential*. Again, this property is invariant under the a-boundary equivalence relation, so one says that $[p]$ is an essentially singular a-boundary point. These, finally, are the missing points Scott and Szekeres aimed to construct.

The most obvious problem facing the a-boundary approach is its physical significance. First off, a 'point' of the a-boundary is not clearly a point in any usual sense of the term: an individual boundary point of one envelopment of a manifold can always be made to cover an uncountable number of boundary points in another envelopment. It is the case that every representative boundary set of an a-boundary point must be compact, but it is not even true that every compact boundary set is a representative of some a-boundary point, nor does the a-boundary point equivalence relation preserve connectedness and simple-connectedness—ought one think of a candidate singularity as a single point or as a non-simply connected, non-connected compact set? Then there is the unapproachability of some a-boundary points: it can happen, for instance, that regular a-boundary points of a pseudo-Riemannian manifold are not approachable by any geodesic of the metric. The existence of such extraneous points makes one wonder about the physical relevance of those boundary points that are approachable by curves in the spacetime. It is also not clear what relevance the 'covering' relation they define has to anything physical: for a given $\mathcal{C}$, $\mathcal{C}$-boundary sets may cover unapproachable boundary sets; non-regular unapproachable boundary sets may cover approachable regular boundary sets; essential boundary points at infinity may cover anything except singular boundary sets and may be covered by anything except regular points; essential singular points may cover any kind of boundary set. Given the promiscuity of possible covering relations, I believe an argument is needed why this definition captures any physically relevant information, an argument they do not provide.

Neither do Scott and Szekeres broach a technical point that raises a serious difficulty for their approach at the very initial stages: some spacetimes, such as Taub spacetime, have two incomplete curves such that the spacetime can be extended so as to make either one or the other curve extendible, but no extension of the spacetime exists that makes both curves simultaneously extendible.[24] On Scott and Szekeres's account, both of these curves run into regular boundary points, and so neither will be counted as possible singularities, even though there is no actual envelopment of the spacetime in which both curves are extendible.

Finally, on this view, incomplete curves wholly contained in compact regions of spacetime cannot count as singular structure, trivially so since compact manifolds cannot be embedded as proper open submanifolds of another manifold. Scott and Szekeres not only gamely swallow this consequence, but actually claim that it is a "*sine qua non* of any successful theory of singularities" (Scott and Szekeres 1994, 34), and cite Shepley and Ryan 1978 as evidence for this claim. This is not only a contentious view, at best, which they do not bother to argue for, and not only seems to run counter to the spirit of most considerations forwarded in discussions of singular structure, which revolve around incomplete curves, but seems seriously to conflict with their own stated criterion for selecting those points of the a-boundary that will be singular points, viz. limit points of curves of bounded parameter, i.e., incomplete curves.

This last point brings out my final consideration against the idea of missing points as touchstones in the investigation of singular spacetimes, which is a simple one: the definition of singular spacetimes by incomplete curves is logically prior to the construction of missing points for singular spacetimes. All the missing point constructions I know of, and all the ways I can more or less easily imagine trying to concoct a new one, rely on probing the spacetime with curves of some sort or other to discover where points may be thought of as missing, just as in the Riemannian case one cannot complete a manifold until one knows which Cauchy sequences do not have a limit point, or equivalently which geodesics are incomplete. One, however, does not need any conception of a missing point, much less a definition of such, to define and investigate the existence of incomplete curves on a manifold. I therefore disagree with the gist of much of the discussion of Earman 1995, Ch. 2, wherein he suggests that unclarity plagues the semi-official definition of a singular spacetime, in terms of b-incompleteness, insofar as, on the face of it, one does not know how it relates to the idea of missing points. Incomplete curves seem to me a fine definition of singular structure on their own.

24. See, e.g., Ellis and Schmidt 1977, 920, and Hawking and Ellis 1973, §5.8.

5. Global vs. Local Properties of a Manifold. There is at least one prima facie good reason why it would be useful to have a precise characterization of points missing from singular spacetimes: one would then be able to analyze the structure of the spacetime 'locally at the singularity', instead of taking troublesome, perhaps ill-defined limits along incomplete curves. The power and elegance of Penrose's conformal construction of infinity for asymptotically flat spacetimes lie precisely in the ability one gains to perform such analysis locally at infinity, without relying on limits.[25] The example of Geroch, Can-bin, and Wald (1982) already discussed makes the prospects for a reasonable boundary construction for singular spacetimes poor. I believe this should not have been very surprising.

In desiring a boundary so as to have a place to 'analyze structure locally', one ought to be clear on what one means by 'locally'. One sometimes hears talk of a global, as opposed to a local, feature of a spacetime, but I know of no precise characterization of the difference. I believe this distinction plays a crucial role in a proper understanding of the standardly proposed definitions of a singular spacetime in terms of incomplete curves. I therefore offer the following precise definition of this distinction. I formulate it initially for topological properties both for the sake of generality and because I think it easier to get a feel for the definition in the sparser arena of topological structure than in the more cluttered arena of differentiable manifolds with an affine structure.

Consider the class $\mathfrak{T}$ of all topological spaces. A *topological property* $\mathfrak{P}$ is a subclass of this class. A topological space $\mathcal{S}$ has the property $\mathfrak{P}$ if $\mathcal{S} \in \mathfrak{P}$.

> **Definition 5.1** *A topological property* $\mathfrak{P}$ *is* local *if it has the following feature: a given topological space* $\mathcal{S}$ *has the property* $\mathfrak{P}$ *if and only if* $\mathcal{S}$ *is such that every neighborhood of every point has a subneighborhood that, considered as a topological space in its own right, with the restriction topology, has the property* $\mathfrak{P}$.[26]

Roughly speaking, a local property must hold in arbitrarily small neighborhoods of every point of a topological space, but not necessarily in every neighborhood of every point of the space; and conversely, if the property holds in arbitrarily small neighborhoods of every point of a space, it must hold for the entire space for it to be local.

25. See Wald 1984, §11.1, for an account of Penrose's construction.

26. This sense of 'local' has nothing to do with that often bandied about in discussions of the foundations of quantum mechanics.

Definition 5.2 *A topological property is* global *if and only if it is not local.*[27]

One could be sure of ascertaining for a given topological space whether the local property $\mathfrak{B}$ held or not by checking for $\mathfrak{B}$ at individual points of the space (quite a few points, to be sure), whereas a global property cannot be checked by examing the structure of the space at any collection of points. As one should expect, local compactness, local connectedness, and local simple connectedness for example all come out to be local on this definition, whereas compactness, paracompactness, connectedness, and simple connectedness come out to be global.[28]

In an analogous manner, one can now straightforwardly characterize properties of differentiable manifolds and of differentiable manifolds with an affine connection as either local or global. Non-trivial examples of local properties for a manifold include any structure residing entirely on the tangent planes over every point. For a manifold with affine connection, both the property of geodesic completeness and of geodesic incompleteness come out to be global properties, again as one should expect. One might initially have thought that geodesic incompleteness, at least, ought to have been a local property—if a geodesic came to an end abruptly, as it were, surely one ought to be able to pinpoint where this happens. If one could do this, however, then it also would seem that one could continue the geodesic. If there were a point on the manifold whereat the incomplete geodesic terminated, however, one could take a chart around that point diffeomorphic to some open set of $\mathfrak{R}^n$, push the geodesic and the connection down to $\mathfrak{R}^n$, where the geodesic obviously would be extendible, and pull the extended version back to the manifold, contradicting the hypothesis that the geodesic could not be continued.

A point of spacetime, in the usual way of thinking of these matters, represents an *event*, a highly localized occurrence in spacetime such as a snapping of fingers or the collision of two billiard balls. It represents an instant of some ponderable object, the specious 'now' of some sentient being. When thinking on cosmic scales, the sun, at a certain in-

27. By this 'not', I do not mean the logical negation of the definition of 'local' but rather the class complement of the class of local properties in the class of all topological properties—interestingly enough, these do not come to the same thing. Were the logical negation of the definition of 'local' used to define 'global', this would entail that a space with the global property $\mathfrak{B}$ would have a point and a neighborhood of that point such that *every* subneighborhood of that neighborhood did *not* have $\mathfrak{B}$. Compactness is clearly not a local property, and yet does not satisfy the negation of the definition of 'local'.

28. Cf. Hocking and Young 1988 for definitions of these topological properties.

stant, can profitably be thought of as occupying a single point of spacetime. In short, spacetime points pertain to discrete objects, very broadly construed, that can be localized in an intuitive sense. There is no a priori reason to suspect that the existence of an incomplete curve, a global phenomenon, could be tied in any natural or reasonable way to the existence of a particular point in an extended manifold. Incomplete curves are not discrete, localizable objects in the appropriate sense.

A detractor will likely balk at this line of thought, pointing to the case of Riemannian manifolds, wherein incomplete curves can be naturally associated with points of an extended manifold. I would reply that it is merely a happy accident in the Riemannian case that one can arrange this. One has no grounds for suspecting that one will be able to do this in the general case, and in fact, as I endeavored to show, one has reasons to suspect that in general one will not be able to do this, since curve incompleteness is global and a missing point is, well, a point, and so prima facie 'local'. Of course, even for Lorentzian manifolds, in certain cases, one will be able to associate to an incomplete curve a missing point with ease—e.g. all the geodesics aimed at the origin in Minkowski spacetime (in some global coordinate system) with the origin removed—in general, though, one ought not expect the two to have anything to do with each other.

The demand that singular structure be localized at a *place* bespeaks an old Aristotelian substantivalism that invokes the maxim, "To exist is to exist in space and time."[29] When I speak of 'Aristotelian substantivalism' here, I refer to the fact that Aristotle thought that everything that exists is a substance and that all substances can be qualified by the Aristotelian categories, two of which are location in time and location in space. In particular, not only substantivalists but also relationalists in debates about the nature of spacetime points could (and often do, I think) consistently fall prey to this particular brand of substantivalism. By focusing attention on the way that spacetimes can have actual features that do not rely on the existence or absence of any particular point, and are not instantiated at any particular point, I think that this distinction between global and local properties of spacetime could have a salutary effect on the moribund debate between substantivalists and relationalists. I hope to work on this matter in the future.

I believe Geroch, Can-bin, and Wald (1982, 435) deserve the last word on this subject: "Perhaps the localization of singular behavior will go the way of 'simultaneity' and 'gravitational force.' "

6. The Finitude of Existence. I turn now to examine whether singular

29. This formulation of the maxim is due to Earman 1995, 28.

spacetimes as characterized are objectionable on physical or interpretive grounds, and whether one is forced to or ought to take them as indicating the 'breakdown' of classical general relativity, as some would have it. In the process, I shall examine whether *b*-completeness is wholly consistent with some of the explicit sentiments behind using curve incompleteness as a criterion for singular structure.

Two types of worries, one psychological, the other physical, give rise to the dissatisfaction with the existence of incomplete curves in relativistic spacetimes. Trying to imagine the experience of an observer traversing one of the incomplete curves provokes the psychological anxiety, for that observer would, of necessity, be able to experience only a finite amount of proper time's worth of observation, even were he, in Earman's evocative conceit, to have drunk from the fountain of youth. The physical worry arises from the idea that particles could pop in and out of existence right in the middle of a singular spacetime, and spacetime itself could simply come to an end, though no fundamental physical mechanism or process is known that could produce such effects. These two types of worries are not always clearly distinguished from each other in discussions of singular structure, but I think it important to keep in mind that in fact there are two distinct types of problems envisaged for incomplete curves, requiring to some degree two separate sorts of responses.

The existence of incomplete spacelike curves is often felt not to be so objectionable as that of incomplete timelike or null curves, on the grounds that it represents structure beyond the experience of any observer.[30] I submit that, on this criterion, neither ought one be so bothered by the existence of incomplete timelike or null curves, *for an observer travelling along such a curve will never experience the fact that he has only a finite amount of proper time to exist*—there is no spacetime point, no event in spacetime, that corresponds to the observer's ceasing to exist. This is not to say that the person traversing this worldline cannot surmise the fact that he has only a finite amount of time to exist, rather that there will never be an instant when the observer experiences himself dissipating, popping out of existence as it were.

These considerations also suggest a tension between the definition of singular structure by *b*-incompleteness and the intuitions that drove some to look to incomplete curves as marks of singular structure in the first place. Only the finitude of proper time matters so far as the experience of a possible observer goes—a generalized affine parameter has no clear physical significance—but, while a curve's being *b*-incomplete implies that the curve is of finite total proper time, the converse is not

30. See, e.g., Hawking and Ellis 1973, §8.1.

true: timeline curves of unbounded total acceleration in Minkowski space can be of finite total proper time and yet be *b*-complete. I would even say that such a curve should be more disturbing on reflection to those with such intuitions than an incomplete null geodesic, for the concept of 'proper time' does not apply to null curves at all, even though they are the possible paths of massless particles.

I speculate, with no hard evidence, that people have not wanted to count such curves as constituting singular structure because of vague worries about energy conservation. In general relativity, however, there is no 'energy conservation'—there is not even a general, rigorous, invariant definition of 'energy'![31] There is thus no a priori reason to suspect that anything in the structure of general relativity excludes a particle's getting shot out asymptotically 'to infinity' in finite total proper time, having started from perfectly regular (in whatever sense of that term one likes) initial data. An example of a spacetime that was *b*-complete for all timelike curves of bounded total acceleration but not for timelike curves of unbounded total acceleration would clarify some of these issues, and I conjecture that examples of such spacetimes exist. Those who would not want to count such a spacetime as singular would be forced to give up *b*-incompleteness as the criterion for singular structure—which, given the lack of a clear physical interpretation of *b*-incompleteness in general, as opposed to incompleteness with respect to total proper time, I would not mind. Of course, if incomplete timelike curves of unbounded total acceleration constituted singular structure, then every solution to Einstein's field equations would be singular. Many would reject this conclusion out of hand, but it does not seem so unbearable to me. Singular structure would simply be one more type of global structure that all spacetimes necessarily had, along with, e.g., paracompactness. Once so much was settled, then one could further classify spacetimes, according to the needs of the project at hand, by satisfaction of various more restrictive types of completeness.

On physical grounds, curve incompleteness has been objected to because it seems to imply that particles could be 'annihilated' or 'created' right in the middle of spacetime, with no known physical force or mechanism capable of pulling off such a feat.[32] The demand that a spacetime be maximal, i.e., have no proper extension, often rests on similar considerations: Clarke (1975, 65—66) and Ellis and Schmidt (1977, 920) conjecture that maximality is required by the lack of a physical process that could cause spacetime to draw up short, as it were, and not continue on as it could have, were it to have an extension. This sort of

31. See Curiel 1998.

32. Cf., e.g., Hawking 1967, 189.

argument, though, relies (implicitly) on a certain picture of physics that does not sit so comfortably in general relativity: that of the dynamical evolution of a system. From a certain quite natural point of view in general relativity, spacetime does not evolve at all. It just sits there, sufficient unto itself, very like the Parmenidean One. From this point of view, the question of a physical mechanism capable of causing the spacetime manifold not to have all the points it could have had, as it were, becomes less poignant, perhaps even misleading. An opponent of this point of view could argue that such a move could foreclose the possibility of deterministic physics, to which I would whole-heartedly agree, for we already know that general relativity does not guarantee deterministic physics: there may be no Cauchy surface in our spacetime, or there may even be so-called naked singularities.[33]

Perhaps a more serious worry is that such a viewpoint would seem to deny that certain types of potentially observable physical phenomena require explanation, when on their face they would look puzzling, to say the least. Were we to witness particles popping in and out of existence, the mettle of physics would surely demand an explanation. I would contend in such a case, however, that a perfectly adequate explanation was at hand: we would be observing singular structure. If there were no curvature pathology around, such a response might appear to be ducking the real issue, viz., why is there this anomalous singular structure when all our strongest intuitions and metaphysical principles tell us it should be impossible? Far from ducking the issue, the viewpoint I advocate is the only one I know of that gives a toehold on looking for precise answers to such questions—or, more precisely, on making such questions precise in the first place.

To make the point more clear, consider the big bang singular structure. Those who balk at this viewpoint ought be as equally troubled by the big bang singular structure as they are by the example under discussion, for it just as surely 'lacks an explanation'. From the viewpoint I advocate, questions about what happened 'before' the big bang, or why the universe 'came into being', can come from their former nebulosity into sharper definition, for they become questions about the presence of certain global structure in the spacetime manifold, in principle no different from paracompactness, connectedness, or the existence of an affine connection, and one can at least envisage possible forms of an answer to the question, 'Are there any factors that necessitated spacetime's having such and such global structure?'. And were we actually to observe particles popping in and out of existence, we could formulate and begin trying to answer the analogous questions.

33. See Earman 1995, Ch. 3 for a discussion of these phenomena.

The most serious problem I can imagine for the viewpoint I advocate is that of representing our subjective experience, experience that seems inextricably tied up with ideas of evolution and change. I suggest that this problem is not an idiosyncracy of the viewpoint I advocate, but in fact arises from the character of general relativity itself: 'dynamical evolution' and 'time' are subtle and problematic concepts in the theory no matter what viewpoint one takes, as attested by the most notorious and seemingly intractable problem in the drive to 'quantize' gravity, the so-called problem of time.[34] My viewpoint has the virtue of calling attention to this very fact, that, to judge by the preponderant mass of literature in both physics and philosophy, is easily overlooked: general relativity demands of us a profound rethinking of several dearly held, deeply related concepts and the relations among them.

It has become fashionable of late to say that such problems point to the need to find an 'interpretation' of general relativity in the same sense in which the measurement problem in quantum mechanics demands that that theory be 'interpreted'. Belot (1996), for instance, reaches this conclusion from speculating on the problems encountered in trying to develop a quantum theory of gravity. I think this is a serious misunderstanding. Quantum mechanics demands an interpretation because it is not clear how to model physical phenomena, how to model the outcomes of experiments *simpliciter*: the predictions of standard quantum theory are in some sense in contradiction with the outcomes of experiments, but not in such a way as to invalidate the theory—an extraordinary state of affairs. There is no analogous problem in general relativity. In a paper on the foundations of quantum mechanics, discussing the lack in general relativity of an explicit representation of our experience of a privileged instant in our history, the 'now', Stein (1984, 645) makes a most *à propos* remark: "although relativity does not give us a *representation* of that experience, there is no *incompatibility* between the experience and the theory: a gap is not a contradiction." There is a gap between the raw materials the theory provides us and the rich content of our experience to be explained—but it is no incompleteness of general relativity, no lack of an 'interpretation', that it does not illuminate the experience it predicts for an observer, no more than Newtonian mechanics fell short in so far as it did not show why I understand by certain irritations of my eardrum from perturbations in the ambient air pressure the import of the spoken word 'gap'.

Even though general relativity is not obscure in the sense that quantum mechanics is, it would be rash to assume that we have isolated and digested all the ways in which a proper understanding of general rel-

34. See Kuchař 1992 for a thorough discussion of this problem.

ativity requires us to modify in fundamental ways root concepts and the relations among them, as special relativity required us, for example, to modify our concept of 'mass' to attain to its proper understanding. We surely have much still to learn from it, much to *un*learn of old patterns of thought, if we are to understand more properly how to try to understand the physical world.

REFERENCES

Belot, G. (1996), "Why General Relativity Does Need an Interpretation", *Philosophy of Science* 63 (Proceedings): S80–S88.

Bosshard, B. (1976), "On the *b*-boundary of the closed Friedmann model," *Communications in Mathematical Physics* 46: 263–268.

Clarke, C. (1973), "Local Extensions in Singular Space-Times", *Communications in Mathematical Physics* 32: 205–214.

———. (1975), "Singularities in Globally Hyperbolic Space-Times", *Communications in Mathematical Physics* 41: 65–78.

———. (1993), *The Analysis of Space-Time Singularities.* Number 1 in Cambridge Lecture Notes in Physics. Cambridge: Cambridge University Press.

Curiel, E. (1998), "Is Energy a Well-Defined Quantity in the Theory of General Relativity?" (unpublished).

Earman, J. (1995), *Bangs, Crunches, Whimpers and Shrieks: Singularities and Acausalities in Relativistic Spacetimes.* Oxford: Oxford University Press.

Ellis, G. and B. Schmidt (1977), "Singular Space-times." *General Relativity and Gravitation* 8: 915–953.

Geroch, R. (1966), "Singularities in Closed Universes," *Physical Review Letters* 17: 445–447.

———. (1968a), "Local Characterization of Singularities in General Relativity", *Journal of Mathematical Physics* 9: 450–465.

———. (1968b), "What is a Singularity in General Relativity?", *Annals of Physics* 48: 526–540.

———. (1985), *Mathematical Physics.* Chicago: University of Chicago Press.

Geroch, R., L. Can-bin, and R. Wald (1982), "Singular Boundaries of Space-Times", *Journal of Mathematical Physics* 23: 432–435.

Geroch, R., E. Kronheimer, and R. Penrose (1972) "Ideal Points in Space-Time", *Philosophical Transactions of the Royal Society (London)* A327: 545–567.

Hawking, S. (1965), "Occurrence of Singularities in Open Universes," *Physical Review Letters* 15: 689–690.

———. (1967), "The Occurrence of Singularities in Cosmology, III", *Philosophical Transactions of the Royal Society (London)* A300: 187–210.

Hawking, S. and G. Ellis (1973), *The Large Scale Structure of Space-Time.* Cambridge: Cambridge University Press.

Hocking, J. and G. Young (1988), *Topology.* New York: Dover Publications, Inc. Originally published in 1961 by Addison-Wesley Publishing Co.

Johnson, R. (1977), "The Bundle Boundary in Some Special Cases", *Journal of Mathematical Physics* 18: 898–902.

Joshi, P. (1993), *Global Aspects in Gravitation and Cosmology.* Number 87 in International Series of Monographs on Physics. Oxford: Oxford University Press.

Konkowski, D. and T. Helliwell (1992), "Singularities in Colliding Plane-Wave Spacetimes", in G. Kunstatter (ed.), *General Relativity and Relativistic Astrophysics.* Singapore: World Scientific Publishing Co., 115–119.

Kuchař, K. (1992), "Time and Interpretations of Quantum Gravity", in G. Kunstatter et al. (eds.), *General Relativity and Relativistic Astrophysics*. Singapore: World Scientific Publishing Co., 211–314.

Misner, C. (1963), "The Flatter Regions of Newman, Unti, and Tamburino's Generalized Schwarzschild Space", *Journal of Mathematical Physics* 4: 924–937.

Misner, C., K. Thorne, and J. Wheeler (1973), *Gravitation.* San Francisco: Freeman Press.
Penrose, R. (1960), "A Spinor Approach to General Relativity." *Annals of Physics* 10: 171–201.
———. (1965), "Gravitational Collapse and Space-Time Singularities", *Physical Review Letters* 14: 57–59.
Schmidt, B. (1971), "A New Definition of Singular Points in General Relativity", *General Relativity and Gravitation* 1: 269–280.
Scott, S. and P. Szekeres (1994), "The Abstract Boundary—a New Approach to Singularities of Manifolds", *Journal of Geometry and Physics* 13: 223–253.
Shepley, L. and G. Ryan (1978), *Homogeneous Cosmological Models.* Princeton: Princeton University Press.
Spivak, M. (1979). *A Comprehensive Introduction to Differential Geometry*, 2nd ed. vol. 1. Houston: Publish or Perish Press. First edition published in 1970.
Stein, H. (1984), "The Everett Interpretation of Quantum Mechanics: Many Worlds or None?", *Noûs* 18: 635–652.
Sussmann, R. (1988), "On Spherically Symmetric Shear-Free Perfect Fluid Configurations (Neutral and Charged), II. Equation of State and Singularities", *Journal of Mathematical Physics* 29:945–970.
Taub, A. (1979), "Remarks on the Symposium on Singularities", *General Relativity and Gravitation* 10: 1009.
Thorpe, J. (1977), "Curvature Invariants and Space-Time Singularities", *Journal of Mathematical Physics* 18: 960–964.
Wald, R. (1984), *General Relativity.* University of Chicago Press.