Some "No Hole" Spacetime Properties Are Unstable*

JB Manchak†

[Forthcoming in *Foundations of Physics*.]

Abstract

We show a sense in which the spacetime property of effective completeness – a type of "local hole-freeness" or "local inextendibility" – is not stable.

1 Introduction

It has been argued that "it is a general feature of the description of physical systems by mathematics that only conclusions which are stable, in an appropriate sense, are of physical interest" (Geroch, 1971, 70). Whatever one thinks of such a position, an investigation concerning the (in)stability of spacetime properties has been a central area of research on global structure.[1] Here, we consider the spacetime property of effective completeness (Manchak 2014) which is a type of "local hole-freeness" (see Earman 1989) or "local inextendibility" (Hawking and Ellis 1973). We show a strong sense in which effective completeness is not stable; a spacetime free of holes in can be arbitrarily close (even in especially fine topologies) to spacetimes with them.

2 Preliminaries

We begin with a few preliminaries concerning the relevant background formalism of general relativity.[2] An n-dimensional, relativistic *spacetime* (for $n \geq 2$) is a pair of mathematical objects (M, g_{ab}). M is a connected n-dimensional manifold (without boundary) that is smooth (infinitely differentiable). Here, g_{ab} is a smooth, non-degenerate, pseudo-Riemannian metric of Lorentz signature $(-, +, ..., +)$ defined on M.

*Special thanks to Thomas Barrett, Jeremy Butterfield, Erik Curiel, Sam Fletcher, David Malament, Jim Weatherall, and an anonymous referee for helping to improve and earlier draft.

†Department of Logic and Philosophy of Science, University of California, Irvine, 3151 School of Social Sciences, Plaza A, Irvine, CA 92697, e-mail: jmanchak@uci.edu

[1] See Beem and Ehrlich (1996) and the references there.

[2] The reader is encouraged to consult Hawking and Ellis (1973), Wald (1984), and Malament (2012) for details. Less technical surveys of the global structure of spacetime are given by Geroch and Horowitz (1979) and Manchak (2013).

Note that M is assumed to be *Hausdorff*; for any distinct $p, q \in M$, one can find disjoint open sets O_p and O_q containing p and q respectively. We say two spacetimes (M, g_{ab}) and (M', g'_{ab}) are *isometric* if there is a diffeomorphism $\varphi : M \to M'$ such that $\varphi_*(g_{ab}) = g'_{ab}$. A spacetime (M, g_{ab}) is *extendible* if there exists a spacetime (M', g'_{ab}) and a (proper) isometric embedding $\varphi : M \to M'$ such that $\varphi(M) \subset M'$. Here, the spacetime (M', g'_{ab}) is an *extension* of (M, g_{ab}). A spacetime is *inextendible* if has no extension.

For each point $p \in M$, the metric assigns a cone structure to the tangent space M_p. Any tangent vector ξ^a in M_p will be *timelike* if $g_{ab}\xi^a\xi^b > 0$, *null* if $g_{ab}\xi^a\xi^b = 0$, or *spacelike* if $g_{ab}\xi^a\xi^b < 0$. Null vectors create the cone structure; timelike vectors are inside the cone while spacelike vectors are outside. A *time orientable* spacetime is one that has a continuous timelike vector field on M. A time orientable spacetime allows one to distinguish between the future and past lobes of the light cone. In what follows, it is assumed that spacetimes are time orientable and that an orientation has been chosen.

For some open (connected) interval $I \subseteq \mathbb{R}$, a smooth curve $\gamma : I \to M$ is *timelike* if the tangent vector ξ^a at each point in $\gamma[I]$ is timelike. Similarly, a curve is *null* (respectively, *spacelike*) if its tangent vector at each point is null (respectively, spacelike). A curve is *causal* if its tangent vector at each point is either null or timelike. A causal curve is *future-directed* if its tangent vector at each point falls in or on the future lobe of the light cone.

We say a curve $\gamma : I \to M$ is not *maximal* if there is another curve $\gamma' : I' \to M$ such that I is a proper subset of I' and $\gamma(s) = \gamma'(s)$ for all $s \in I$. A curve $\gamma : I \to M$ in a spacetime (M, g_{ab}) a *geodesic* if $\xi^a \nabla_a \xi^b = \mathbf{0}$ where ξ^a is the tangent vector and ∇_a is the unique derivative operator compatible with g_{ab}. A spacetime (M, g_{ab}) is *geodesically complete* if every maximal geodesic $\gamma : I \to M$ is such that $I = \mathbb{R}$. A spacetime is *geodesically incomplete* if it is not geodesically complete.

For any two points $p, q \in M$, we write $p << q$ if there exists a future-directed timelike curve from p to q. We write $p < q$ if there exists a future-directed causal curve from p to q. These relations allow us to define the *timelike and causal pasts and futures* of a point p: $I^-(p) = \{q : q << p\}$, $I^+(p) = \{q : p << q\}$, $J^-(p) = \{q : q < p\}$, and $J^+(p) = \{q : p < q\}$. Naturally, for any set $S \subseteq M$, define $J^+[S]$ to be the set $\cup\{J^+(x) : x \in S\}$ and so on. A set $S \subset M$ is *achronal* if $S \cap I^-[S] = \varnothing$.

A point $p \in M$ is a *future endpoint* of a future-directed causal curve $\gamma : I \to M$ if, for every neighborhood O of p, there exists a point $t_0 \in I$ such that $\gamma(t) \in O$ for all $t > t_0$. A *past endpoint* is defined similarly. A causal curve is *future inextendible* (respectively, *past inextendible*) if it has no future (respectively, past) endpoint. If an incomplete geodesic is timelike or null, there is a useful distinction one can introduce. We say that a future-directed causal geodesic $\gamma : I \to M$ without future endpoint is *future incomplete* if there is an $r \in \mathbb{R}$ such that $s < r$ for all $s \in I$. A *past incomplete* causal geodesic is defined analogously.

For any set $S \subseteq M$, we define the *past domain of dependence of* S, written $D^-(S)$, to be the set of points $p \in M$ such that every causal curve with past endpoint p and no future endpoint intersects S. The *future domain of dependence of* S, written $D^+(S)$, is defined analogously. The entire *domain of dependence of* S, written $D(S)$, is just the set $D^-(S) \cup D^+(S)$. The *edge* of an achronal set $S \subset M$ is the collection of points $p \in S$

such that every open neighborhood O of p contains a point $q \in I^+(p)$, a point $r \in I^-(p)$, and a timelike curve from r to q which does not intersect S. A set $S \subset M$ is a *slice* if it is closed, achronal, and without edge. A spacetime (M, g_{ab}) which contains a slice S such that $D(S) = M$ is said to be *globally hyperbolic*.

Let (K, g_{ab}) be a globally hyperbolic spacetime. Let $\varphi : K \to K'$ be an isometric embedding into a spacetime (K', g'_{ab}). We say (K', g'_{ab}) is an *effective extension* of (K, g_{ab}) if, for some Cauchy surface S in (K, g_{ab}), $\varphi(K)$ is a proper subset of $int(D(\varphi(S)))$ and $\varphi(S)$ is achronal. Hole-freeness can then be defined as follows.[3] A spacetime (M, g_{ab}) is *hole-free* if, for every set $K \subseteq M$ such that (K, g_{ab}) is a globally hyperbolic spacetime with Cauchy surface S, if (K', g_{ab}) is not an effective extension of (K, g_{ab}) where $K' = int(D(S))$, then there is no effective extension of (K, g_{ab}).

3 Effective Completeness

Take any spacetime (M, g_{ab}) and remove a point p from M. The "hole" in the spacetime $(M - \{p\}, g_{ab})$ presumably renders the spacetime "physically unreasonable." Accordingly, one seeks a condition to rule out these and other "holes" in spacetime. (The condition ought to be necessary for "physical reasonableness" but need not be sufficient.) But "although one perhaps has a good intuitive idea of what it is that one wants to avoid, it seems difficult to formulate a precise condition to rule out such examples" (Geroch and Howowitz 1979, 275).

Because the singularity theorems of Hawking and Penrose (1970) show a sense in which some (if not all) "physically reasonable" spacetimes are geodesically incomplete, one cannot use geodesic completeness to rule out spacetimes with "holes." Instead, two "no hole" conditions – inextendibility and hole-freeness – are often used. But, while the satisfaction of each of these conditions has been taken to be necessary for a spacetime to be "physically reasonable," each condition is weak in the sense that it fails to rule out all of the unwanted "holes" (see Earman 1989). Naturally, one seeks the strongest (or, at least, a stronger) "no hole" condition which is necessary for "physical reasonableness." Along these lines, the condition of "effective completeness" has recently been proposed which seems promising (Manchak 2014). Consider the following.

Definition. A spacetime (M, g_{ab}) is *effectively complete* if, for every future or past incomplete timelike geodesic $\gamma : I \to M$, and every open set O containing $\gamma[I]$, there is no isometric embedding $\varphi : O \to M'$ into some other spacetime (M', g'_{ab}) such that $\varphi \circ \gamma$ has future and past endpoints.

The physical significance of the condition is this: if a spacetime fails to be effectively complete, then there is a freely falling observer who never records some particular watch reading but who "could have" in the sense that nothing in her vicinity precludes it. A violation of effective completeness turns out to be (modulo minor details) what Earman (1989) calls a "local hole" in spacetime. One can get a rough idea

[3] See Geroch (1977) for an earlier definition and Manchak (2009) for a discussion of why a revision was needed.

of the strength of such a "no holes" condition by noting the following implication relations (Manchak 2014).

$$\text{geodesic completeness} \quad \Longrightarrow \quad \text{effective completeness} \quad \begin{array}{l} \Nearrow \quad \text{inextendibility} \\ \\ \Searrow \quad \text{hole-freeness} \end{array}$$

4 Instability

Let us now construct the so called "C^k-open" topologies on a fixed manifold M (Hawking and Ellis 1973, 198). We begin with the following definition (cf. Fletcher 2016).

Definition. Let M be a manifold which admits a Lorentzian metric and let h^{ab} be a positive definite metric on M. At each point in M, the *distance function* $d(g_{ab}, g'_{ab}, h^{ab}, k)$ between the kth partial derivatives of any two Lorentzian metrics g_{ab} and g'_{ab} on M relative to h^{ab} is given by the following (here ∇_a is the unique derivative operator compatible with h^{ab}):

$$[h^{ac}h^{bd}(g_{ab} - g'_{ab})(g_{cd} - g'_{cd})]^{1/2} \qquad \text{for } k = 0$$
$$[h^{ac}h^{bd}h^{r_1 s_1}...h^{r_k s_k}(\nabla_{r_1}...\nabla_{r_k}(g_{ab} - g'_{ab}))(\nabla_{s_1}...\nabla_{s_k}(g_{cd} - g'_{cd}))]^{1/2} \qquad \text{for } k > 0$$

One can now use the distance function given above to define the C^k-open neighborhoods of a spacetime (M, g_{ab}). We have the following.

Definition. An C^k-open *neighborhood* of a spacetime (M, g_{ab}) is any collection of spacetimes which includes all spacetimes (M, g'_{ab}) such that $\text{Sup}_M[d(g_{ab}, g'_{ab}, h^{ab}, j)] < \epsilon$ for $j = 0, ..., k$ where h^{ab} is a positive definite metric on M and ϵ is a positive number.

One can verify that even the C^0-open topology obtained from the above definition is quite fine. In particular, if (M, g_{ab}) is a spacetime and M is noncompact, then the one-parameter family of spacetimes $\{(M, \lambda g_{ab}) : \lambda \in (0, \infty)\}$ does not represent a continuous curve in the C^0-open topology; moreover, the induced topology on this one-parameter family is discrete (Geroch 1971). It seems, then, that the C^0-open topology is too fine to adequately capture, once and for all, what it means to say that one spacetime is "close" to another (cf. Fletcher 2016). And the same goes for the even finer C^k-open topologies for all $k > 0$. But the fact that there are too many open sets in the C^k-open topologies makes them ideal for proving instability results. Consider the following definition (cf. Beem and Ehrlich 1996).

Definition. A spacetime property $\mathscr{P}$ is *C^k-open stable* if, for each spacetime (M, g_{ab}) with $\mathscr{P}$, there is a C^k-open neighborhood of (M, g_{ab}) such that every spacetime in the neighborhood also has $\mathscr{P}$. A spacetime property $\mathscr{P}$ is C^k-*open unstable* if it is not C^k-open stable.

Note that if a spacetime property fails to be C^k-open stable for some $k \geq 0$, then it will fail to be stable relative to any other topology coarser than the C^k-open topology. Which properties are C^k-open unstable? Consider the following (Beem and Ehrlich 1996, 244-247).

Proposition. Geodesic completeness and geodesic incompleteness are C^k-open unstable for all $k \geq 0$.

Given the singularity theorems, the C^k-open instability of geodesic completeness is not too troubling. On the other hand, the C^k-open instability of geodesic incompleteness has been taken to be quite significant. Indeed, a great deal of work has been done to show that, if attention is appropriately limited to certain types of Robertson-Walker spacetimes, the C^0-open stability of geodesic incompleteness can be saved (see Beem and Ehrlich 1996, 250-257). One naturally wonders if any of the "no hole" conditions mentioned above can also be shown to be C^k-open unstable. Consider the following.

Proposition. Effective completeness is C^k-open unstable for all $k \geq 0$.

Proof. The proof proceeds in two dimensions for reasons of simplicity but can be easily generalized. Consider the spacetime (M, g_{ab}) where $g_{ab} = -2\nabla_{(a}t\nabla_{b)}\varphi$ and M is the quotient $\mathbb{R}^2/\sim$ under the equivalence relation given by $(t, \varphi) \sim (t, \varphi + 2\pi)$. When no confusion arises, we will speak of the point $(t, \varphi) \in M$ rather than the equivalence class of points $[(t, \varphi)] \in M$. One can show that (M, g_{ab}) is just Minkowski spacetime "rolled up" along one null direction and is therefore geodesically complete (see Beem and Ehrlich 1996, 245). It follows that the spacetime effectively complete (Manchak 2014). Let U be the region $\{(t, \varphi) \in M : -2 < t < 2\}$ and let V be the region $\{(t, \varphi) \in M : -3 \leq t \leq 3\}$. Let $f : M \to \mathbb{R}$ be any smooth function such that both $f_{|U} = t$ and $f_{|M-V} = 0$. For each $n \in \mathbb{Z}^+$, let $(M, g_{ab}(n))$ be the spacetime where $g_{ab}(n) = -2\nabla_{(a}t\nabla_{b)}\varphi - (f/n)\nabla_a\varphi\nabla_b\varphi$. Note that the spacetime $(U, g_{ab}(n))$ is isometric to a portion of Misner spacetime. (See equation (1) in Levanony and Ori (2011) and let $\psi = \varphi/n$ and $T = tn$.)

In what follows, we restrict attention to the $k = 1$ case but the argument can be generalized to all $k \geq 0$ in the natural way. Let h^{ab} be any positive definite metric on M, and let ϵ be any positive number. If we let the smooth scalar fields $\alpha_0, \alpha_1 : M \to \mathbb{R}$ be given by $\alpha_0 = fh^{ab}\nabla_a\varphi\nabla_b\varphi$ and $\alpha_1 = [h^{ac}h^{bd}h^{rs}(\nabla_r(f\nabla_a\varphi\nabla_b\varphi))(\nabla_s(f\nabla_c\varphi\nabla_d\varphi))]^{1/2}$, then we see that for each $n \in \mathbb{Z}^+$, we have have the following.

$$d(g_{ab}, g_{ab}(n), h^{ab}, 0) = [h^{ac}h^{bd}(g_{ab} - g_{ab}(n))(g_{cd} - g_{cd}(n))]^{1/2} = \alpha_0/n$$
$$d(g_{ab}, g_{ab}(n), h^{ab}, 1) = [h^{ac}h^{bd}h^{rs}(\nabla_r(g_{ab} - g_{ab}(n)))(\nabla_s(g_{cd} - g_{cd}(n)))]^{1/2} = \alpha_1/n$$

By construction, $f_{|M-V} = 0$ and so this implies that $\text{Sup}_{M-V}[d(g_{ab}, g_{ab}(n), h^{ab}, k)] = 0$ for $k = 0, 1$. Now consider V. Because this region is compact, we know that there is an $m \in \mathbb{R}$ such that $\alpha_0(p) < m$ and $\alpha_1(p) < m$ for all $p \in V$. So for each n, we know $\text{Sup}_V[d(g_{ab}, g_{ab}(n), h^{ab}, k)] < m/n$ for $k = 0, 1$. But $m/n < \epsilon$ for large enough n. It follows that for $k = 0, 1$ we have $\text{Sup}_M[d(g_{ab}, g_{ab}(n), h^{ab}, k)] < \epsilon$ for large enough n. Therefore, in any C^1-open neighborhood of (M, g_{ab}), there is an $n \in \mathbb{Z}^+$ such that $(M, g_{ab}(n))$ is in the neighborhood. It remains to show that $(M, g_{ab}(n))$ is not effectively complete.

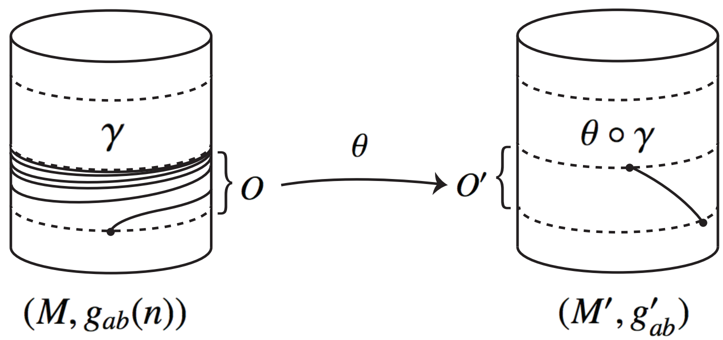

Figure 1: The isometry θ "unwinds" the geodesic γ.

Fix any $n \in \mathbb{Z}^+$ and consider the spacetime $(M, g_{ab}(n))$. Let O be the set $\{(t, \varphi) \in M : -2/n < t < 0\}$. One can show that the set $\{(t, \varphi) \in O : nt = -e^{-\varphi/n} - e^{-\varphi/2n}\}$ is the image of a timelike geodesic $\gamma : I \to O$. (See equation (10) in Levanony and Ori (2011) with $x_0 = -1$. Let $\psi = \varphi/n$ and $T = tn$.) One finds that $\gamma : I \to O$ approaches but never reaches the $t = 0$ line and is (depending on the temporal orientation) either future or past incomplete (see Hawking and Ellis 1973, 171-174). We are done once we are able to show that O can be isometrically embedded into another spacetime such that the image of γ under the embedding has both past and future endpoints.

Consider the spacetime (M', g'_{ab}) where $g'_{ab} = 2\nabla_{(a}t'\nabla_{b)}\varphi' - t'\nabla_a\varphi'\nabla_b\varphi'$ and M' is the quotient $\mathbb{R}^2/\sim$ under the equivalence relation given by $(t', \varphi') \sim (t', \varphi' + 2\pi)$. As before, we will speak of the point $(t', \varphi') \in M'$ rather than the equivalence class of points $[(t', \varphi')] \in M'$. Let O' be the set $\{(t', \varphi') \in M' : -2 < t' < 0\}$. Let $\theta : O \to O'$ be the isometry defined by $\theta((t, \varphi)) = (nt, \varphi/n + 2\ln(-nt))$. This isometry "unwinds" the geodesic γ (see Hawking and Ellis 1973, 171). One can verify that $\theta \circ \gamma[I]$ is the set $\{(t', \varphi') \in O' : t' = -e^{\varphi'} + e^{\varphi'/2})\}$. So, the curve $\theta \circ \gamma$ has two (one past and one future) endpoints at $(-2, 2\ln(2))$ and $(0, 0)$ (see Figure 1). □

5 Conclusion

The above result should give us pause concerning the physical significance of at least some "no hole" conditions.[4] Two lines of future work may help to clarify matters. First, can one show the C^k-open instability of inextendibility or hole-freeness for some $k \geq 0$?

[4]For more skepticism concerning "no hole" conditions, see Manchak (2009, 2011, 2016).

If so, the case for the physical insignificance of "no hole" conditions strengthens. If not, this fact may count as a strike against the condition of effective completeness. Second, can one show that the conjunction of effective completeness and some other condition necessary for "physical reasonableness" (e.g. a weak causality condition) is C^k-open stable for some $k \geq 0$? If so, perhaps there is a sense in which the stability of effective completeness can be saved for some spacetimes of interest (cf. Beem and Ehrlich 1996, 244-247).

References

[1] Beem, J. and P. Ehrlich (1996), *Global Lorentzian Geometry (Second Edition)*. New York: Marcel Dekker.

[2] Earman, J. (1989), *World Enough and Spacetime*. Boston: MIT Press.

[3] Fletcher, S. (2016), "Similarity, Topology, and Physical Significance in Relativity Theory," *The British Journal for the Philosophy of Science*, 67: 365-389.

[4] Geroch, R. (1971), "General Relativity in the Large," *General Relativity and Gravitation*, 2: 61-74.

[5] Geroch, R. (1977), "Prediction in General Relativity," *Foundations of Space-Time Theories, Minnesota Studies in the Philosophy of Science* Vol. VIII, J. Earman, C. Glymour, and J. Stachel (eds.), 81-93. Minneapolis: University of Minnesota Press.

[6] Geroch, R. and G. Horowitz (1979), "Global Structure of Spacetimes," *General Relativity: An Einstein Centenary Survey*, S. Hawking and W. Isreal (eds.), 212-293. Cambridge: Cambridge University Press.

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

[8] Hawking, S. and R. Penrose (1970). "The Singularities of Gravitational Collapse and Cosmology," *Proceedings of the Royal Society A*, 314:529-548.

[9] Levanony, D. and A. Ori (2011), "Extended Time-Travelling Objects in Misner Space," *Physical Review D*, 83: 044043.

[10] Malament, D. (2012), *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory*. Chicago: University of Chicago Press.

[11] Manchak, J. (2009), "Is Spacetime Hole-Free?" *General Relativity and Gravitation*, 41: 1639-1643.

[12] Manchak, J. (2011), "What is a Physically Reasonable Space-Time?" *Philosophy of Science*, 78: 410-420.

[13] Manchak, J. (2013), "Global Spacetime Structure," in R. Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press, pp. 587-606.

[14] Manchak, J. (2014), "On Spacetime Singularities, Holes, and Extensions," 81: 1066-1076.

[15] Manchak, J. (2016), "Epistemic "Holes" in Space-Time," *Philosophy of Science*, 83: 265-276.

[16] Wald, R. (1984), *General Relativity*. Chicago: University of Chicago Press.