

# If the Metric Were Not Dynamical, Counterfactuals in General Relativity Would Be Easy<sup>†</sup>

Erik Curiel<sup>‡</sup>

7. Apr 2023

## Contents

<a href="#">1</a>	<a href="#">Précis</a>	<a href="#">2</a>
<a href="#">2</a>	<a href="#">Counterfactuals in Physical Theories</a>	<a href="#">3</a>
<a href="#">3</a>	<a href="#">The Problem in Detail</a>	<a href="#">5</a>
<a href="#">4</a>	<a href="#">Severity of the Problem</a>	<a href="#">8</a>
	<a href="#">References</a>	<a href="#">10</a>

## ABSTRACT

General relativity poses serious problems for counterfactual propositions peculiar to it as a physical theory. Because these problems arise solely from the dynamical nature of spacetime geometry, they are shared by all schools of thought on how counterfactuals should be interpreted and understood. Given the role of counterfactuals in the characterization of, *inter alia*, many accounts of scientific laws, theory confirmation and causation, general relativity once again presents us with idiosyncratic puzzles any attempt to analyze and understand the nature of scientific knowledge must face.

**Keywords:** general relativity; spacetime structure; counterfactuals

---

<sup>†</sup>Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) – Projekt Nummer 312032894, CU 338/1-2.  
(Funded by the German Research Foundation (DFG) – project number 312032894, CU 338/1-1.)

<sup>‡</sup>I thank David Malament and David Wallace (private correspondence) and Harvey Brown, Dennis Lehmkuhl and Oliver Pooley (conversation) for vigorously pressing me on the paper’s arguments and conclusions. I thank The Young Guns of the Spacetime Church of the Angle Brackets for suffering through a previous version of this paper and giving me, as always, insightful help with a smile. **Author’s affiliations:** Munich Center for Mathematical Philosophy, LMU Munich; Black Hole Initiative, Harvard University; **email:** [erik@strangebeautiful.com](mailto:erik@strangebeautiful.com)

# 1 Précis

In his elegant, magisterial exposition of the foundations of general relativity, Malament (2012, ch. 2, §1, pp. 120–121) provides three interpretive principles to endow the mathematical framework of Lorentzian geometry with physical content:<sup>1</sup>

For all smooth curves  $\gamma : I \rightarrow M$  [where  $I \subset \mathbb{R}$  is an open interval and  $M$  is a candidate spacetime manifold]:

(C1)  $\gamma$  is timelike iff  $\gamma[I]$  could be the worldline of a point particle with positive mass;

(C2)  $\gamma$  can be reparametrized so as to be a null geodesic iff  $\gamma[I]$  could be the trajectory of a light ray;

(P1)  $\gamma$  can be reparametrized so as to be a timelike geodesic iff  $\gamma[I]$  could be the worldline of a *free* point particle with positive mass.

(Emphases are Malament’s; ‘C’ indicates the proposition pertains to the interpretation of conformal structure, ‘P’ to projective structure; (C1) articulates the physical meaning of timelike curves, (C2) that of null geodesics, and (C2) that of timelike geodesics.) He immediately offers four comments and qualifications to address possible concerns one may have with these propositions as interpretative principles, touching on questions about the exclusion of tachyons, the restriction to smooth curves, the status of point particles in general relativity, and, of most interest for our purposes, the modal character of the propositions. He concludes (*ibid.*, p. 122),

Though these four concerns are important and raise interesting questions about the role of idealization and modality in the formulation of physical theory, they have little to do with relativity theory as such.

I agree with his conclusion in all parts, except for the concern about the role of modality. I think there are important problems with modality in general, and with the understanding of counterfactuals in particular, peculiar to general relativity as a physical theory, problems that have gone unremarked in the philosophy and the physics literature.<sup>2</sup> Malament’s formulation and discussion of the interpretative principles allows them to be drawn out with great clarity.<sup>3</sup>

About (C2) he says (*loc. cit.*), “We are considering what trajectories are available to light rays when no intervening material media are present—*i.e.*, when we are dealing with light rays in vacuo.” Now, surely we want to talk as well about the null cones even at those places where matter

1. I follow Malament (2012) in all relevant conventions (the signature of the spacetime metric, the definition of the Weyl tensor, *etc.*). The reader should consult that work or Wald (1984) for exposition of all concepts and results about general relativity I rely on in this paper, except where explicitly noted otherwise.

2. Strictly speaking, as the discussion in §3 will make clear, the problem is not restricted to general relativity, but rather infects all theories with spacetime structure that is dynamical by virtue of a particular kind of coupling with matter sources, such as  $f(R)$ -gravity. (See, *e.g.*, de Felice and Tsujikawa 2010 for an extended review.) I focus on general relativity as it our most strongly confirmed and most widely accepted theory of spacetime structure, and everything I say about it can be translated easily into the context of similar theories. The problem does not arise, however, simply by virtue of *any* non-trivial relation a theory may posit between geometry and matter: the problem does not arise, for instance, in classical Yang-Mills theory on Minkowski spacetime.

3. I emphasize that I am not criticizing Malament or trying to draw attention to weaknesses or errors in his exposition, quite the contrary. It is the exemplary (and characteristic) clarity, precision and thoroughness of his discussion that allows a previously unrecognized problem to be brought to light.

is present. In order to do so, and in order to formulate the analogue of (C2) for those spacetime regions (in order to give a physical interpretation to the null cones at those points), we must say something along the following lines: the null geodesics where matter is present are those paths light rays would follow if the matter there were removed. But on its face, that modal statement makes no sense in the context of general relativity, because however we make sense of the idea of “removing matter” from a spacetime region, the metric will *eo ipso* be different in that region from what it was, and it will generically be the case that the new metric in that region will not agree with the original metric on what it counts as null vectors, much less on what it counts as null geodesics, among many other differences.<sup>4</sup> If the paths light rays “could follow” after matter is removed are generically not the paths light rays “could follow” before the matter is removed, how can removing matter possibly address the question in a cogent way? The distribution of matter in a region of spacetime in large part informs the metrical structure there, so what sense can be made, in the context of the theory more generally, in asking what the metrical structure *would* be if the matter actually there *were not* there? And now we face the heart of the problem: the lack of a unique vacuum solution for the Einstein field equation forces an ineliminable ambiguity in the idea of what it may mean to “remove matter from a region of spacetime”, guaranteeing that we have no way to conclude on any principled basis “what the metric would then look like there”.

## 2 Counterfactuals in Physical Theories

[\*\*\* two kinds of counterfactuals:

1. theoretical/formal: basically, comparing solutions to equations of motion or field equations with differing initial or boundary conditions, or comparing theorems, general principles, and local or global structures under different local and global assumptions; they can tell you things about the nomic or conceptual structure of the theory, but not about “modality in the real world” (see classification of metaphysical, nomological, principle and physical possibility)
2. experimental: taking account of full complexity of modeling real systems (*e.g.*, schematized representations of instruments, *etc.*), comparing results of different calculations (“predictions” or “kinematic characterizations”), and so on

\*\*\*]

One of the simplest and *prima facie* most promising ways to begin to try to get a grip on the problem, and to look for its resolution, is to treat the proposed counterfactual changes as represented by a change in initial conditions, and so to use the machinery of general relativity’s initial-value formulation to fix the solution. Indeed, this is the natural and effective way of dealing with such counterfactuals in all other physical theories. When one wants to know how the state of a system would change if some of its properties or some properties of its environment were to change,

---

4. Indeed, in regions of spacetime filled, say, with a non-trivial dielectric, there simply are no null curves that are “the possible paths of light rays” because the full resources of electromagnetic theory tell us in this case that, under any reasonable construal of “propagate”, light does not propagate through a dielectric with the speed of light in vacuo.

The same problem arises for timelike curves in regions of spacetime already occupied by matter, *i.e.*, for (C1) and (P1), but I focus on the case of null rays to simplify the exposition.

one first writes down the initial-value formulation (or boundary-value problem, depending on the details of the theory and the problem at hand), plugs in the original values for all parameters, and calculates the solution; one then does the same thing using the new, counterfactually changed values for the parameters, and compares the new solution to the old. Easy as pie, and about as philosophically unproblematic as things can possibly get for a large class of fundamentally important counterfactuals.

What makes this procedure work in almost all physical theories is the context of a family of fixed background structures, often spatiotemporal ones, with respect to which one has natural ways of identifying and so comparing entities like “the same quantity of the same system at the same place and same time, under otherwise different conditions or in otherwise different states”. To see how this works in a simple case, one that has especial relevance to the problem in general relativity, consider the situation in Newtonian gravitational theory. It makes perfect sense in Newtonian theory to reason counterfactually about the behavior of a given kind of system in the presence or absence of any other kind of system, since that presence or absence won’t affect the kinematical structure of Newtonian spacetime: it is always  $\mathbb{R}^4$  with a fixed foliation of simultaneity slices, and a fixed, flat affine structure defining the possible trajectories of all freely falling bodies (Stein 1997).<sup>5</sup> There is no problem in principle in computing the counterfactual change in gravitational forces in a region induced by any counterfactual changes in the distribution of matter anywhere in the spacetime. For example, one may be interested in the question: what would happen to the orbits of the planets in the Solar System if the sun were to vanish? Nothing simpler. Plug in the new values for the sun’s size and mass (*viz.*, zero), compute the new orbits, and compare them to the original ones by using the background simultaneity and affine structures as referential framework.

[\*\*\* Stein (1995, p. 22), on how to develop the idea of “lines of causal influence” in Newtonian gravitational theory, even though interactions are simultaneous and continual:

ask *if a certain change* were introduced—conceptually, as a thought-experiment—at a certain moment, how would the *new* entire process differ from the old?

\*\*\*]

The virtues of this method for the representation, interpretation and evaluation of counterfactuals in the context of physical theory are legion. For our purposes, perhaps the greatest virtue is this: pragmatics plays absolutely no role. It does not matter what the investigator’s purposes are, or the exact experimental techniques one envisions as being employed, or anything of the sort. There is no need for the *ad hoc* fixing of comparison classes of systems, or for the *ad hoc* fixing of methods for identifying “the same quantity of the same system at the same place and same time, under otherwise different conditions or in otherwise different states”. It’s all fixed naturally and canonically from the start.<sup>6</sup>

5. Nothing of importance for the example would change if we were to work in the more sophisticated context of geometrized Newtonian gravity (Malament 2012). I suspect, however, that in generalizations of geometrized Newtonian gravity, in which one does not restrict the topology of spacetime to  $\mathbb{R}^4$  and one does not (effectively) require that the simultaneity slices be spatially flat (Malament 2012), then similar problems would arise. They would, however, arise for the same reasons as they do in general relativity: the absence of fixed, absolute, background spatiotemporal structure because of dynamical coupling with matter (footnote 2).

6. This is not to say there no choices at all to be made in this context. One might decide in the Newtonian example, for instance, to consider only smooth solutions, or to allow continuous or even distributional solutions,

Famously in general relativity, there is no fixed, absolute, background structure one could use as such a referential framework for the comparison of properties of different solutions to the Einstein field equation. Thus, *prima facie*, there is no way to employ the standard machinery of the initial-value formulation (or the boundary-value problem) to represent, render meaning to and evaluate such counterfactuals in general relativity.

[\*\*\* Maxwell theory: “if one were to have hit a charged particle with a hammer, how would the world have changed?” Is there ambiguity here? Yes, but it’s subtle, and not the same kind as in general relativity. If one chooses any spacelike hypersurface appropriate for fixing initial data (say, a slice that is wholly spacelike), then, so long as that hypersurface contains the event at which one considers the possibility of having hit the particle with a hammer, so one can change the initial data there, the forward and backward evolution is independent of choice of hypersurface. If one chooses an achronal slice not wholly spacelike, but asymptotically null to the past, one has the ambiguity of possibly adding free incoming radiation. \*\*\*]

### 3 The Problem in Detail

To begin to come to grips with the problem in more detail, let us consider the problem of the sun’s vanishing translated from the Newtonian context into that of general relativity: what would Schwarzschild spacetime look like if its central mass were removed? Perhaps the immediate, intuitive (and naive) response is to say: obviously the result will be Minkowski spacetime. The plethora of available vacuum solutions, however, shows the choice is not so simple.

To see the issues involved more clearly, let us try to articulate the problem in a more explicit, precise and rigorous fashion, using the idea of taking the limit of a continuously varying family of spacetimes. We start with Schwarzschild spacetime of a given mass, and consider a family of Schwarzschild spacetimes parametrized by mass, starting from that initial one, continuously shrinking to zero. It may be initially surprising to learn that such a limiting family has no unique limit. (This follows immediately from the fact that any reasonable topology on a non-trivial family of spacetimes is non-Hausdorff; see, *e.g.*, Curiel 2015.) Working the example out in a little detail shows clearly what is going on.<sup>7</sup> In Schwarzschild coordinates, using the parameter  $\lambda \equiv M^{-1/3}$  (the inverse-third root of the Schwarzschild mass), the metric takes the form

$$\left(1 - \frac{2}{\lambda^3 r}\right) dt^2 - \left(1 - \frac{2}{\lambda^3 r}\right)^{-1} dr^2 - r^2(d\theta^2 + \sin^2 \theta d\phi^2) \quad (1)$$

This clearly has no well defined limit as  $\lambda \rightarrow 0$ . Now, apply the coordinate transformation

$$\tilde{r} \equiv \lambda r, \quad \tilde{t} \equiv \lambda^{-1} t, \quad \tilde{\rho} \equiv \lambda^{-1} \theta$$

In these coordinates, the metric takes the form

$$\left(\lambda^2 - \frac{2}{\tilde{r}}\right) d\tilde{t}^2 - \left(\lambda^2 - \frac{2}{\tilde{r}}\right)^{-1} d\tilde{r}^2 - \tilde{r}^2(d\tilde{\rho}^2 + \lambda^{-2} \sin^2(\lambda \tilde{\rho}) d\phi^2)$$

---

in the fixing of the comparison class. This choice, however, does not affect the naturalness and canonicity of the procedure.

7. The following analysis is taken from Geroch (1969).

The limit  $\lambda \rightarrow 0$  exists and yields

$$-\frac{2}{\tilde{r}}d\tilde{t}^2 + \frac{\tilde{r}}{2}d\tilde{r}^2 - \tilde{r}^2(d\tilde{\rho}^2 + \tilde{\rho}^2 d\phi^2)$$

a flat solution discovered by Kasner (1921). If instead of that coordinate transformation we apply the following to the original Schwarzschild form (1),

$$x \equiv r + \lambda^{-4}, \quad \rho \equiv \lambda^{-4}\theta$$

then the resulting form also has a well defined limit, which is the Minkowski metric. The two limiting processes yield different spacetimes because it happens behind the scenes that “the same points of the underlying manifold get pushed around relative to each other in different ways”. Because the coordinate relations of initially nearby points differ in different coordinate systems, those differences get magnified in the limit, so that their final metrical relations differ. Thus, the limits in the different coordinates yield different metrics, with no natural or preferred way to say which is the “correct” limit—which, if either, is the correct spacetime that results from counterfactually removing the sun from the Solar System.

The root of the problem lies in the fact that metrical curvature is *only* in part informed by the distribution of matter: the Weyl curvature at a point, exactly that part of the curvature encoding conformal information, such as what counts as a null vector, is independent of the value of the stress-energy tensor at that point—the value of the Weyl tensor, point by point, is not constrained by the presence or absence of matter. In regions without matter, moreover, metrical curvature is governed entirely by the Weyl tensor. Still, the Weyl tensor  $C^a_{bcd}$  is subtly related to the distribution of matter at neighboring points, when there is such matter, in a way that can be made precise by using the Bianchi identity formulated using the so-called Lanczos tensor.<sup>8</sup> Thus, in “removing matter” from a spacetime region, there can be no principled way to determine what the “remaining curvature” will be. One may decide to keep the Weyl tensor the same. But precisely its relation to stress-energy by way of the Lanczos tensor means that this is not an unproblematic way to proceed, and is likely even incoherent or inconsistent.<sup>9</sup> One may sum all these issues up by adverting to the fact that there is not a unique vacuum solution to the Einstein field equation: the form of the Einstein field equation for a vacuum region remains the same no matter what the field equations are for matter immediately outside the vacuum region, but those exterior matter fields do not suffice to fix the solution for the vacuum region.<sup>10</sup>

8. The Lanczos tensor is defined as follows:

$$\begin{aligned} J_{abc} &:= \frac{1}{2}\nabla_{[a}R_{b]c} + \frac{1}{6}g_{c[b}\nabla_{a]}R \\ &= 4\pi\nabla_{[a}T_{b]c} - \frac{1}{12}g_{c[b}\nabla_{a]}T \end{aligned} \tag{2}$$

The Bianchi identity may then be rewritten

$$\nabla_n C^n_{abc} = J_{abc}$$

Thus the value of the Weyl tensor at a point does depend in an indirect way on the distribution of matter at nearby points.

9. It should therefore be clear that these sorts of problem arise not only for counterfactuals involving changes in the distribution of matter, but also for any involving changes in the curvature more generally. One may, for example, try to consider how the behavior of test-particles in a vacuum spacetime would change if one were to “alter a component of the ambient gravitational radiation”.

10. It was, as a matter of fact, exactly this issue that first led Einstein to formulate his infamous Hole Argument (Curiel 2018).

Now, the problem with this example is, in fact, even more acute than the discussion so far shows: there are an innumerable number of possible spacetimes, not all of them flat, with many different topologies possible, that one can derive by rigorous limiting processes from Schwarzschild spacetime as the central mass goes to zero.<sup>11</sup> The same problems already discussed arise for all of these, compounded by the great number and variety of possible topologies the limits may respectively take. Let us say, however, that we have, in some way or other, decided that we are interested in the limit that takes Schwarzschild spacetime to Minkowski spacetime, and in particular we are interested in the ways that geodesics near the event horizon will change over the course of the limit. The natural topology of the manifold of Schwarzschild spacetime is  $\mathbb{R}^2 \times \mathbb{S}^2$ . The natural topology of Minkowski spacetimes, however, is  $\mathbb{R}^4$ . Thus, in an intuitive sense spacetime points will “disappear” in the limiting process, because one must “de-compactify two topological dimensions” to derive  $\mathbb{R}^4$  from  $\mathbb{R}^2 \times \mathbb{S}^2$ . There are many ways to effect such a de-compactification; all the simplest, such as that based on Alexandrov compactification, work by the removal of a point or set of points from the topological manifold.<sup>12</sup> How is one to identify which spacetime points in the limit manifold  $\mathbb{R}^4$  are those that were “close to the event horizon” in the original Schwarzschild spacetime manifold  $\mathbb{R}^2 \times \mathbb{S}^2$ ?

The problems discussed here with the representation, interpretation and evaluation of counterfactuals in general relativity are not restricted to the case of removing or otherwise altering matter distributions. Similarly problematic modalities, requiring the fixation of a relevant comparison class of spacetimes and the identification and comparison of structure and properties across spacetime models, arise all over the place in general relativity. I offer a brief sample.

1. the proofs of all the Laws of black-hole mechanics (Wald 1994)
2. the proof of the No-Hair theorem for black holes (Heusler 1996)
3. proofs that various causality conditions imply each other, *e.g.*, that the absence of closed timelike curves implies the absence of closed causal curves (Hawking and Ellis 1973)
4. proofs of geodesic theorems (Geroch and Jang 1975; Ehlers and Geroch 2004)
5. well-posedness (in the sense of Hadamard) of the initial-value formulation (Wald 1984)
6. formulation of and arguments for the Cosmic Censorship Hypothesis (Joshi 2003)
7. formulation of and arguments for Penrose’s Conformal Curvature Hypothesis (Penrose 1979)
8. proof of the Topological Censorship Theorem (Friedman et al. 1993)
9. the possible extendibility of spacetimes past singularities, and the possible hole-freeness of a spacetime in general (Manchak 2014)
10. taking “small” perturbations off a fixed spacetime, *e.g.*, in considering slight inhomogeneities in the cosmological FLRW models (Szekeres 1975a), or in treating slightly aspherical gravitational collapse (Szekeres 1975b) or more generally distorted black holes (Geroch and Hartle 1982)
11. the genericity or scarcity (or stability or rigidity) of a given spacetime property, *e.g.*, the existence of singularities or of large-scale cosmological structure or of a given range of values for a constant of nature (Curiel 2015)

11. See, *e.g.*, Paiva et al. (1993) for explicit construction of a few different ones.

12. See, *e.g.*, Kelley (1955) for an account of methods of compactification, including the Alexandrov type.



Even if one accepts the idea of massive test-particles without internal structure or stress-energy as a way to make sense of certain kinematical structures in counterfactuals about motion in general relativity, that strategy will not work for other forms of matter. Massive test particles work because I know how to get a grip on the idea of “small enough” for a particle whose coupling to the ambient geometry is fixed by a single scalar (its mass). One can control the size of the perturbation because there is a natural norm on scalars, and linearizing the perturbation of the metric gives one a physically significant norm on the metric perturbations (the linear perturbations for a single particle are diagonal, so the norm is the trace of the metric). I don’t know how to do this for Maxwell fields. There are no natural, or otherwise physically significant, norms on the family of Faraday tensors at a point, nor on the family of stress-energy tensors. I don’t know what “small enough to to treat as negligible” means in this case. Think, however, of counterfactually “turning on an ordinary flashlight”—I know a circle of illumination would appear on the wall were I to turn it on. I think that is because I am using the fixed background of Minkowski spacetime, and I have an intuitive sense that the flashlight’s beam *is* “weak enough” not to sensibly perturb that metric. But how do I know this? Can I make the sense precise, and so generalize it? What about using  $F_{mn}F^{mn} + \epsilon_{mnrs}F^{mn}F^{rs}$  as norm (*viz.*,  $E^2 + B^2$  plus the norm of the Poynting vector)? The smallness of that quantity, however, does not straightforwardly translate into “smallness of metric perturbation”.

The fact that field equations and stress-energy tensors, even for “test matter”, depend on the metric only compound all these difficulties. Even when the metric does not appear explicitly in a set of field equations, the derivative operator always does.

It is also the case that energy conditions can be satisfied by matter fields “off-shell”, *i.e.*, even when all field equations (those of the fields by themselves and the Einstein field equation itself) may be violated. Kontou and Sanders (2020, §2.2), for instance, discuss several results of interest that relate the validity of energy conditions to the dynamics of the fields, including when validity on-shell implies validity off-shell. The case when energy conditions must be satisfied on-shell but may be satisfied off-shell (*e.g.*, when the fields are “weak” and so considered “test matter”) raises the fascinating possibility of pathological behavior for a given kind of field *only* in the “weak-field” regime.

## 4 Severity of the Problem

There is no single algorithm or reasoning procedure to employ for all such possible cases of counterfactual (and more general modal) reasoning in general relativity, and certainly no natural or canonical one. One must handle such situations on a case-by-case basis, coming up with a method to fix the comparison class and similarity structure in the proposed counterfactual situation in some way that respects the particular constraints of the project the counterfactual reasoning is to play a part in, while at the same time making possible the identification and comparison of structure and properties across spacetimes in the comparison class: the intervention of pragmatics is inescapable, and will perforce depend on *ad hoc* constructions and arguments. Contrary to popular folk wisdom in the philosophy community, *ad hoc* does *not* always mean poorly justified, or arbitrary, or just having pulled it out of one’s ass—not to put too fine a point on it. Indeed, according to the *Oxford English Dictionary* (‘OED’), it *never* means that; philosophy has indulged



in an abusive and shallow malapropism, not for the first time. Rather, it is (still) continuous with the original Latin, which literally means “for this [purpose]”; the OED gives its definition as:

Created or done for a particular purpose; that answers a specific need or demand, rather than in accordance with a general policy, rule, etc.

And this is how one must handle counterfactuals in GR.

The astute reader, however, may now immediately want to respond that in many interesting accounts of counterfactuals in the literature, there are formal devices whose sole purpose is to manage the pragmatics of their interpretation and evaluation. For example, this is exactly what the freedom in the selection of “similarity” measures in a Lewisian or Stalnakerian account of counterfactuals is supposed to accommodate. But the problem is deeper here in general relativity. It is not simply that the choice of a similarity measure is a matter of pragmatics. It is rather that “the choice” in general is not a well defined set of alternatives *at all*. Even once one has decided in some way or other what the appropriate comparison class of spacetimes ought to be, and what is to count as the relevant measure of similarity among the spacetimes, still one simply does not know how, in a general, abstract way, to make sense of the idea “different values or forms of the same structure in different spacetime models” (or, in Lewisian terms, in different possible worlds in a family of mutually accessible ones)—to hearken back to the question that started the paper, there is no way to relate the null geodesics in the original spacetime with those in a metrically modified spacetime. Even putting aside the problems posed by differing topologies in spacetimes one may want otherwise to judge similar in some important fashion (as discussed in §3), the diffeomorphic freedom inherent in general relativity makes it in general impossible to say “what is the same point in different spacetimes”. There is no principled reason to use  $\mathbb{R}^2 \times \mathbb{S}^2$ , say, as the diffeomorphic presentation of Schwarzschild spacetime rather than  $\mathbb{R}^4$  with a line removed (which is diffeomorphic to  $\mathbb{R}^2 \times \mathbb{S}^2$ , but not canonically so)—but how is one to identify in a principled way a point in  $\mathbb{R}^2 \times \mathbb{S}^2$  with one in  $\mathbb{R}^4$  with a line removed? Thus, not only does general relativity block the standard methods of managing the pragmatics of counterfactual interpretation and evaluation, it blocks the prior, even more fundamental step: the representation of the counterfactual in the given language (in this case, the mathematical theory of Lorentz manifolds).

Still, one may want to ask, if this is such a problem, how do physicists handle it in practice? For they surely do—otherwise no progress would ever have been made in any of the problems listed at the end of §3. The answer is that, when engaging in arguments requiring such counterfactual and more generally modal reasoning, physicists do use *ad hoc* methods for the fixation of an appropriate comparison class of spacetimes and similarity measure on it. More importantly, they use *ad hoc* methods for the identification and comparison of structure and properties across the spacetimes in the fixed class, *e.g.*, by requiring that the spacetimes in the family have a fixed topology, or have a fixed diffeomorphic presentation, or satisfy certain symmetries, and then fixing a coordinate system for the identification of points across spacetimes. Though those methods are *ad hoc* in the strong (and etymologically literal) sense in that they are applicable only to the specific context in which they are used, and cannot be generalized in any way to a method applicable in all situations across the theory (much less generalized in a canonical or natural way), there are often strong physical justifications for their use.<sup>13</sup> This surely suffices for all the purposes the practical physicist has in

13. Even in many of these cases, however, the method used is (figuratively) *ad hoc* in the stronger sense that

his or her theoretical work, and only a fool or a philosopher would cavil at it. For we philosophers, however, who are interested in the understanding and comprehension of the foundations of the theory as a whole, the problem of the interpretation and evaluation of counterfactuals in the context of the theory cannot be considered settled, or even as having a promising attack made upon it, until we have a more detailed and deeper understanding of the epistemic principles that allow the pragmatics in any situation to support the principled formulation and evaluation of a counterfactual.

The problem I expose in this paper is severe: many influential philosophical approaches to many fundamental problems and issues in the philosophy of science—the meaning of fundamental interpretive principles, the nature of scientific laws, the stability of theory-confirmation, the nature of causation, *et al.*—rely in ineliminable ways on subjunctive conditionals for their formulation, analysis and application. Physicists certainly rely on such propositions in theoretical and experimental practice all the time, *e.g.*, to propose and perform tests of general relativity. What reason do we have to believe that we understand what is happening in such cases in the context of general relativity, much less to have confidence in any conclusions drawn? Because the problem arises solely from the dynamical nature of spacetime geometry in general relativity, moreover, what I say here is independent of one’s favorite account of counterfactuals—it depends only on the theoretical resources general relativity provides to model such situations and pose such propositions, no matter what ancillary tools or frameworks one uses to interpret and understand them.

I wanted in this paper only to draw attention to this serious problem, not to propose possible solutions. I think any decent attempt to do the latter will require a great deal of involved, technical work, including detailed examination of many non-trivial examples. I sincerely hope someone takes up the challenge. I am not entirely without hope for progress to be made here. Indeed, it seems likely that progress on this problem could suggest new avenues of attack on the traditional problem of understanding counterfactuals in general, and could suggest elaborations of and improvements to already existing accounts. In any event, beyond the possible benefits that progress on this problem may have for the general problem of counterfactuals, the mettle of philosophy and the needs of physics demand we understand what is going on here.

## References

- Curiel, Erik. 2015. “Measure, Topology and Probabilistic Reasoning in Cosmology.” [arXiv:1509.01878 \[gr-qc\]](https://arxiv.org/abs/1509.01878). Most recent draft available at <<http://strangebeautiful.com/phil-physics.html>>.
- . 2018. “On the Existence of Spacetime Structure.” *The British Journal for the Philosophy of Science* 69 (2, June): 447–483. First published online in 2016. Preprint: [arXiv:1503.03413 \[physics.hist-ph\]](https://arxiv.org/abs/1503.03413). A manuscript containing technical appendices working out details of some of

---

there is no known physically significant justification for it. This happens most often in those problems in which the analysis requires mathematical machinery of an even more advanced and heavy-duty sort than one needs in workaday general relativity. For instance, I know of no convincing physical interpretation for the Sobolev norms required to prove the well-posedness of general relativity’s initial-value formulation (Wald 1984). Physicists use them because they work mathematically, no other reason. Indeed, I have never even seen a physicist pay lip-service to the idea that the Sobolev norms have real physical significance.

- the constructions and arguments, and containing further discussion of the possible observability of different kinds of spacetime structure, is available at <<http://strangebeautiful.com/phil-physics.html>>, doi:[10.1093/bjps/axw014](https://doi.org/10.1093/bjps/axw014).
- De Felice, Anotonio De, and Shinji Tsujikawa. 2010. “ $f(R)$  Theories.” *Living Reviews in Relativity* 13:3. [arXiv:1002.4928 \[gr-qc\]](https://arxiv.org/abs/1002.4928), doi:[10.12942/lrr-2010-3](https://doi.org/10.12942/lrr-2010-3).
- Ehlers, Jürgen, and Robert Geroch. 2004. “Equation of Motion of Small Bodies in Relativity.” *Annals of Physics* 309 (1): 232–236. [arXiv:gr-qc/0309074](https://arxiv.org/abs/gr-qc/0309074), doi:[10.1016/j.aop.2003.08.020](https://doi.org/10.1016/j.aop.2003.08.020).
- Friedman, John L., Kristin Schleich, and Donald M. Witt. 1993. “Topological Censorship.” *Physical Review Letters* 71 (10, 6 September): 1486–1489. [arXiv:gr-qc/9305017](https://arxiv.org/abs/gr-qc/9305017). Erratum: *Physics Review Letters*, 75(1995):1872, doi:[10.1103/PhysRevLett.75.1872](https://doi.org/10.1103/PhysRevLett.75.1872)., doi:[10.1103/PhysRevLett.71.1486](https://doi.org/10.1103/PhysRevLett.71.1486).
- Geroch, Robert. 1969. “Limits of Spacetimes.” *Communications in Mathematical Physics* 13 (3): 180–193. Open access at <http://projecteuclid.org/euclid.cmp/1103841574>, doi:[10.1007/BF01645486](https://doi.org/10.1007/BF01645486).
- Geroch, Robert, and James Hartle. 1982. “Distorted Black Holes.” *Journal of Mathematical Physics* 23 (4, April): 680–692. doi:[10.1063/1.525384](https://doi.org/10.1063/1.525384).
- Geroch, Robert, and Pong Soo Jang. 1975. “The Motion of a Body in General Relativity.” *Journal of Mathematical Physics* 16 (1): 65–67. doi:[10.1063/1.522416](https://doi.org/10.1063/1.522416).
- Hawking, S. W., and G. F. R. Ellis. 1973. *The Large Scale Structure of Space-Time*. Cambridge: Cambridge University Press. doi:[10.1017/CBO9780511524646](https://doi.org/10.1017/CBO9780511524646).
- Heusler, Markus. 1996. *Black Hole Uniqueness Theorems*. Cambridge Lecture Notes in Physics 6. Cambridge: Cambridge University Press. doi:[10.1017/CBO9780511661396](https://doi.org/10.1017/CBO9780511661396).
- Joshi, Pankaj S. 2003. “Cosmic Censorship: A Current Perspective.” *Modern Physics Letters A* 17 (15): 1067–1079. [arXiv:gr-qc/0206087](https://arxiv.org/abs/gr-qc/0206087), doi:[10.1142/S0217732302007570](https://doi.org/10.1142/S0217732302007570).
- Kasner, Edward. 1921. “Geometrical Theorems on Einstein’s Cosmological Equations.” *American Journal of Mathematics* 43 (4, October): 217–221. doi:[10.2307/2370192](https://doi.org/10.2307/2370192).
- Kelley, John L. 1955. *General Topology*. The University Series in Higher Mathematics. Princeton: D. Van Nostrand Company, Inc.
- Kontou, Eleni-Alexandra, and Ko Sanders. 2020. “Energy Conditions in General Relativity and Quantum Field Theory.” *Classical and Quantum Gravity* 37 (19): 193001. [arXiv:2003.01815 \[gr-qc\]](https://arxiv.org/abs/2003.01815), doi:[10.1088/1361-6382/ab8fcf](https://doi.org/10.1088/1361-6382/ab8fcf).
- Malament, David B. 2012. *Topics in the Foundations of General Relativity and Newtonian Gravitational Theory*. Chicago: University of Chicago Press. Uncorrected final proofs for the book are available at <<http://strangebeautiful.com/other-texts/malament-founds-gr-ngt.pdf>>.
- Manchak, John Byron. 2014. “On Space-Time Singularities, Holes, and Extensions.” *Philosophy of Science* 81 (5, December): 1066–1076. doi:[10.1086/677696](https://doi.org/10.1086/677696).

- Paiva, F. M., M. J. Rebouças, and M. A. H. MacCallum. 1993. "On Limits of Spacetimes—A Coordinate-Free Approach." *Classical and Quantum Gravity* 10 (6): 1165–1178. [arXiv:gr-qc/9302005](https://arxiv.org/abs/gr-qc/9302005), doi:[10.1088/0264-9381/10/6/013](https://doi.org/10.1088/0264-9381/10/6/013).
- Penrose, Roger. 1979. "Singularities and Time-Asymmetry." In *General Relativity: An Einstein Centenary Survey*, edited by Stephen W. Hawking and Werner Israel, 581–638. Cambridge: Cambridge University Press.
- Stein, Howard. 1967. "Newtonian Space-Time." *Texas Quarterly* 10:174–200.
- . 1995. "How Does Physics Bear Upon Metaphysics; and Why Did Plato Hold that Philosophy Cannot Be Written Down?" *Studies in History and Philosophy of Modern Physics* 72:152–161. Published in 2020. The paper was delivered by Stein as a talk at a faculty colloquium (an informal affair) of the Department of Philosophy at the University of Chicago in November 1995. A scan of Stein's original typed manuscript can be found at <<http://strangebeautiful.com/other-texts/stein-physics-and-metaphysics-original.pdf>>, doi:[10.1016/j.shpsb.2020.06.004](https://doi.org/10.1016/j.shpsb.2020.06.004).
- Szekeres, P. 1975a. "A Class of Inhomogeneous Cosmological Models." *Communications in Mathematical Physics* 41 (1): 55–64. Open-access text available at: <<http://projecteuclid.org/euclid.cmp/1103860587>>., doi:[10.1007/BF01608547](https://doi.org/10.1007/BF01608547).
- Szekeres, Peter. 1975b. "Quasi-Spherical Gravitational Collapse." *Physical Review D* 12 (10, November): 2941–2948. doi:[10.1103/PhysRevD.12.2941](https://doi.org/10.1103/PhysRevD.12.2941).
- Wald, Robert M. 1984. *General Relativity*. Chicago: University of Chicago Press. doi:[10.7208/chicago/9780226870373.001.0001](https://doi.org/10.7208/chicago/9780226870373.001.0001).
- . 1994. *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. Chicago: University of Chicago Press.