preprint (25/1/2024), forthcoming in *The British Journal for the Philosophy of Science*,
please cite the published version if available

Counterparts, Determinism, and the Hole Argument

Franciszek Cudek

ABSTRACT

The hole argument concludes that substantivalism about spacetime entails the radical indeterminism of the general theory of relativity (GR). In this paper, I amend and defend a response to the hole argument first proposed by Butterfield ([1989]) that relies on the idea of counterpart substantivalism. My amendment clarifies and develops the metaphysical presuppositions of counterpart substantivalism and its relation to various definitions of determinism. My defence consists of two claims. First, *contra* Weatherall ([2018]) and others: the hole argument is not a blunder resulting from a mistaken view on how mathematical physics works, and requires a developed (meta)metaphysical response. Second, *contra* Melia ([1999]) and others: one can be content with a notion of determinism for GR which is not sensitive to merely haecceitistic differences.

1 Introduction

The hole argument concludes that substantivalism about spacetime entails the radical indeterminism of the general theory of relativity (GR). In this paper, I amend and defend a response to the hole argument first proposed by Butterfield ([1989]) that relies on the idea of counterpart substantivalism.

In section 2, I discuss substantivalism about spacetime and the doctrine of manifold substantivalism. In section 3, I present the hole argument of Earman and Norton ([1987]), and sketch some canonical responses, including counterpart substantivalism. I note the differences between my preferred version of counterpart substantivalism and Butterfield's ([1989]). In section 4, I show the place of the hole argument in the context of a more general problem about the (modal) metaphysics of spacetime, namely how to count possible spacetimes. In section 5, I argue against Weatherall's ([2018]) claim that the hole argument is based on a confusion about how mathematical physics works. In section 6, I consider the problem raised by Melia ([1999]) and others, namely that the definition of determinism endorsed by counterpart substantivalists classifies some intuitively indeterministic theories as deterministic.

2 Spacetime Substantivalism

There is much to be said about spacetime substantivalism, but I shall start by saying next to nothing: spacetime substantivalism is the view that spacetime exists, and is somewhat independent of its material content. On this thin reading, subtantivalism does not take a stance on whether spacetime is fundamental, or primitive, or what is its essence. It does not even take a stance on whether such notions are intelligible.[1] But for our purposes, that is to get the hole argument going in its usual formulation, I shall nevertheless elaborate on two issues.

First, I will take spacetime substantivalism as the view that spacetime points exist. Points are dimensionless simples that together make up the spacetime continuum. I admit it is perfectly coherent to be committed to the existence of spacetime, but not spacetime points. For one may claim that the nature of spacetime is 'gunky', that is, not made of simples.[2] This, however, does not make any difference to our purposes: the most ad-

[1] There are philosophers who argue that the substantivalism vs. relationalism debate should be concerned with these issues if it pretends to have any robust (meta)physical import. See, for example, Dasgupta ([2011], p. 116) and North ([2021], sec. 5.4). I do not dispute these claims here: indeed, the weaker the kind of substantivalism one assumes to be the target of the hole argument, the bigger the threat it presents.

[2] The term 'gunk' comes from Lewis ([1991], p. 20). See Arntzenius ([2008]) for an exposition and

vanced topologies of gunk reconstruct the notion of a point as a location in space(time),[3] so one could in principle rephrase all that follows in terms of locations, rather than points.[4]

Second, the kind of spacetime substantivalism that I focus on goes by the name of 'manifold substantivalism', and it involves some commitments about the individuation of spacetimes.[5] It is often characterised as the view that 'the [smooth] manifold of events represents spacetime',[6] or as 'classifying the metric as content rather than as an aspect of the spacetime container'.[7] Importantly, it is not a wildly implausible view that regards any two models of general relativity whose base manifolds are related by any diffeomorphism as physically equivalent.[8] (I briefly describe these technicalities in section 3.) Even though I don't disagree with the above glosses, I am not convinced that they suffice in explicating manifold substantivalism.

I will treat manifold substantivalism as an interpretational postulate applicable to any particular spacetime theory T. I also prefer to explicate it in terms of constraints on the relation between the models of T and the possible physical situations such models attempt to represent (I will explain what I mean by a 'possible physical situation' later in this section). In particular, manifold substantivalism ascribes to possible (physical) spacetimes fairly minimal sufficient conditions for identity. Such conditions are generally captured by the identity of certain mathematical objects in T through the following schema, (S):

> (S). If two models M_1 and M_2 of T contain, or are identical to, the same mathematical object of kind K, then, within any reasonable representational convention, they attempt to represent possible physical situations involving the same spacetime.[9]

Manifold substantivalists endorse a version (S-Man) of (S), where the objects K are four-dimensional smooth manifolds (more on the technicalities in section 3):

> (S-Man). If two models M_1 and M_2 of T contain the same smooth manifold

moderate defence of the idea of gunky spacetime.

3 See, for example, Roeper ([1997], sec. 2) and Lando and Scott ([2019], sec. 7).

4 What a gunk enthusiast may reject are examples of transformations where spacetime points are permuted non-smoothly, or even non-continuously, as in (Melia [1999], p. 642; Maudlin [1988], pp. 84–5). But this does not bear on the hole argument which relies on a smooth transformation.

5 See, in particular, Earman and Norton ([1987], pp. 518–9) and Norton et al. ([2023], sec. 4).

6 Norton et al. ([2023], sec. 4), see also Earman and Norton ([1987], p. 518).

7 Pooley and Read ([forthcoming], p. 15).

8 This is a view assigned by Halvorson and Manchak ([forthcoming], sec. 4) to a fictional character named Carsten.

9 To emphasise: what is represented by a model of a spacetime theory T, that is what I call a 'possible physical situation', may describe more than just a spacetime. It may describe items which are not spacetime, for example, matter and radiation fields.

> M, then, within any reasonable representational convention, they attempt to represent possible physical situations involving the same spacetime.

Some philosophers maintain that (S-Man) is too weak, since spacetimes should be individuated by more structured mathematical objects than smooth manifolds. How these claims bear on the hole argument debate is also discussed in section 3.

But before I move on, let me address one more issue here (indeed a very important one): what's the motivation behind adopting any particular version of schema (S)? Why should any version of (S) guide the way in which we interpret our physical theories? These questions cannot be answered independently of other parts of the notoriously multifaceted topic of what physical theories tell us about the world (and how they do so). So what follows in this paper can be read as conditionally dependent on adopting some version of (S). But since this endeavour would be pointless if (S) were really indefensible, I'm going to say—tersely and, no doubt, without full justification—why I am sympathetic to it, and also what I mean by 'representational convention', why models only 'attempt' to represent possible physical situations, and what I take these possible situations to be.

I believe that models of spacetime theories both guide and constrain the truth-conditions of modal propositions made relative to any particular theory, and the truth (or falsity) of such propositions is what is fundamentally at stake both in the hole argument as well as in many other debates about the intepretation of physical theories. The orthodox way of thinking about the role of models is as follows: models of a theory T represent metaphysically possible worlds that conform to the laws of T, which in turn provide an extensional analysis of modal propositions made relative to T. The manner in which the orthodox path is usually followed carries along an optimistic assumption that there is no need to specify what possible worlds that conform to T in fact are. The justification of this optimistic assumption, I presume, is that each party in the debate aims to establish a conclusion that constitutes a necessary condition for any adequate metaphysical account of possibility.

I do not share that optimism, and I will not say that models represent possible worlds so understood. Metaphysically possible worlds are maximal, in that they settle the truth-value of every proposition, whereas what is represented by models of spacetime theories (such as GR), and what I shall henceforth call a 'possible (physical) situation (described by GR)' need not be maximal in this sense. In particular, possible situations described by GR should only settle propositions that comprise the subject matter of GR.[10] They need

[10] For a philosophical analysis of the notion of subject matter, see Hamblin ([1958]), Lewis ([1998]), and Russell ([2015], sec. 5). I agree that it might be vague, or indeterminate, whether some proposition belongs to the subject matter of GR (or any other physical theory), but I don't think this is in any way problematic for the view I espouse here. It is merely a feature of physical theories understood, in

not settle propositions whose truth-value is independent of the claims that any particular model of a theory, jointly with interpretative principles agreed upon by the theory's practitioners, plausibly warrants. Such propositions might include propositions about phenomena that do not fall within the purview of the theory (such as those dependent on, say, thermodynamical or quantum-mechanical considerations), or propositions using concepts the theory explicitly dispenses of (such as absolute simultaneity, or absolute velocity). I believe that possible situations (described by a theory T) so understood, in conjunction with a counterpart-theoretic framework for assessing *de re* modal claims, would still offer an adequate extensional analysis of modal propositions relativised to a particular theory T.[11] But, to stay in line with the literature, I will continue to use the term 'possible world' as a synonym of a 'possible situation' in the sense just described, and reserve the term 'metaphysically possible world' for maximal worlds—I believe this will not lead us astray.

Moving on from purely modal concerns, I say that an assignment of a model to represent some possible situation sets a 'representational convention'.[12] What counts as a reasonable representational convention is no doubt interest-relative and theory-dependent, but I can give you an example of what would be unreasonable: interpreting each model of GR as representing, say, the actual world (insofar as GR is concerned). I restrict my

all their complexity, as bodies of knowledge, or ways of conceptualising a significant portion of our experience, that their boundaries might be vague or indeterminate.

11 Admittedly, combining possible situations (or worlds) with a counterpart-theoretic framework for *de re* modality leads to some troubles. Let me briefly discuss two most pressing ones. First, such a combination is sensitive to the distinction between possible worlds (or situations) and possibilities (understood as extensional truth-makers that comprise possible worlds (or situations) and counterpart relations), and—following the line of argument applied to Newtonian shifts by Skow ([2008])—only the latter are relevant to the kind of considerations raised by the hole argument. Second, Teitel ([2022], p. 251) argues that selection of a given counterpart relation between spacetime points to account for the notion of physical possibility is trumped by considerations regarding the nature of metaphysical necessity within the framework given by Lewis ([1986]). Let me summarise (at the unfortunate cost of not giving these worries their proper due) the response that I endorse: after some work, it should be possible to recover the talk about relevant possibilities——in an interesting, even if unorthodox, sense—just from the models of T jointly with a choice of a counterpart function, and a collection of interpretative facts that relate the abstract elements of models to physical objects and concepts. The importance of the difference between possible worlds (or situations) and possibilities would dissolve once one rejects same-worldly (or same-situational) counterparts. This move is justified, because the notion of possibility at play is not metaphysical compossibility with the laws of T, but a *sui generis* notion of possibility-according-to-T, and the usual intuition pumps for same-worldly counterparthood relation might no longer apply. This remark also mitigates Teitel's worries, for one could merely stipulate that the relevant counterparthood relation for GR-possibility is the one that takes into account geometric similarity between the manifold points representing the spacetime points in question. In general, such modality-according-to-T could be put into broader theoretical use as an example of an 'objective' modality, from which metaphysical modality could be recovered along the lines proposed by Williamson ([2016]). This, I admit, is a very broad sketch, but I hope to develop it in future work.

12 Pooley and Read ([forthcoming], p. 16) use 'representational context' for the same idea, whereas Gomes ([2021], sec. 1.2.1) uses 'representational convention', although his use is more technical.

interest to reasonable representational conventions where the models of GR attempt to represent possible situations concerning the whole physical universe, and not merely one of its subsystems (though this is usually the case in physicists' practice), since only these conventions are relevant to the hole argument.

Admittedly, I don't like the idea that, as a matter of principle, all isomorphic models represent the same possible situation within any single representational convention. But I would endorse a weaker principle: namely that each of a class of isomorphic models is equally fit to represent any of the possible situations that can be represented by these models. That is, representational conventions that differ only over which of two isomorphic models is assigned to a given possible situation are equally adequate.[13] (I am also sympathetic to the idea (but I will not assume it as a principle) that in many (most? all?) theories, including GR, there is just one possible situation to be represented by a class of isomorphic models, even though it may not be representable by all these models within a single representational convention.)

Now, here is some motivation why I deny the stronger principle (taken as a matter of principle), but endorse the weaker one. If a theory describes some physical object as being such-and-such in some possible world, it does so through a model. In that case, some part of that model should represent the object—that's clear enough. But what if the theory attempts to describe, once again through its models, the ways in which a given object *could* be? In other words, what if a theory attempts to make some *de re* modal claims? That physical theories should attempt to do so is *prima facie* a reasonable expectation: 'this electron is spin-up, but could have been spin-down', 'this spacetime region is flat, but could have been curved', 'this particle decayed in two years, but could have decayed in four' are all intelligible scientific propositions. Naively, we would like to assess such claims by looking at what properties a given object has across different possible worlds. For this, we need a notion of 'trans-world identity'. Although I shall ultimately dispense with the notion of strict trans-world identity in favour of the counterparthood relation, I think that strict trans-world identity should be expressible in our framework for interpreting a physical theory, because I don't like to dismiss any metaphysical views on the grounds that they are 'inexpressible', especially if I believe I understand what they're trying to get at.

Here is how I accommodate these beliefs, and mould them into something like (S): if one sets a representational convention by representing some possible situation which

[13] Equivalent versions of this weaker principle are widely endorsed by parties disagreeing over the philosophical significance of the hole argument. See Weatherall ([2018], p. 330), Fletcher ([2020], p. 233) (his (REME)), Roberts ([2020], p. 3) (his (Weak Leibniz Equivalence)), and Pooley and Read ([forthcoming], p. 16). In section 5, I discuss why adopting this principle doesn't dissolve the hole argument.

has an object o with a model M that contains a mathematical object o_m which represents o, then another model M' which also contains o_m attempts to represent (within that representational convention) some possible situation where o exists. Here, in order not to commit anyone automatically to the doctrine of trans-world identity, I added the qualification that M' only attempts to represent some possible world. I may deny that it succeeds, because I don't believe in strict trans-world identity. I may also deny, for some other reasons, that there is more than one possible world that is to be represented by all models isomorphic to M. But, I'm not obliged to do it solely because of any 'illiberal' doctrines which impose severe restrictions on the relation between models and possible situations such as outright denial of trans-world identity.

Now, we can turn to GR and the hole argument itself.

3 General Relativity and the Hole Argument

First, let us try to present GR in a slightly more formal way.[14] We can think of GR, in its standard formalism, as comprising a collection of mathematical objects (called 'models') of the form $\langle M, g, T \rangle$, where: M is a four-dimensional smooth manifold, g is a metric tensor field, and T is a stress-energy tensor field.[15]

Three caveats. First, I shall be interested only in dynamically possible models, that is, models which not only match the structure of spacetime postulated by the theory, but also satisfy its dynamical laws. In the case of GR, the dynamical laws are given by the Einstein Field Equation (EFE) that relates g (the geometry of spacetime) to T (spacetime's mass-energy content) in a specific way. Second, in what follows, I suppress the stress-energy tensor field and work with models of the form $\langle M, g \rangle$ (Lorentzian manifolds), which purport to represent physical possibilities without matter and radiation fields. This is not an important change: my rationale is not only the fact that we can carry over the discussion to include these fields, but also the fact that Lorentzian manifolds have dominated the recent literature.[16] Let's also note that models of the form $\langle M, g \rangle$ are sometimes called 'relativistic spacetimes' in the literature.[17] Here, I stress that I wish to stick to calling these mathematical objects 'models', and I reserve the term 'spacetime' or 'physical spacetime' for a (possible) object that one might attempt to represent with (a part

[14] See Dewar ([2022], sec. 5.2) for a discussion of similarities and differences between physical and formal theories as regards satisfaction by models.

[15] See Wald ([1984]) for a modern presentation of general relativity.

[16] See, for example, Weatherall ([2018]), Roberts ([2020]), Halvorson and Manchak ([forthcoming]), and Pooley and Read ([forthcoming]).

[17] See, for example, Malament ([2012], p. 119)

of) such a model. Third, recall the discussion of manifold substantivalism from section 2, especially the principle (S-Man). It's now evident that for a manifold substantivalist models $\langle M, g_1 \rangle$ and $\langle M, g_2 \rangle$ purport to describe the ways the same physical spacetime can be, since the manifold M features in both.

I also mentioned that some substantivalists are discontented with manifold substantivalism, because they believe that spacetimes should not be individuated by smooth manifolds, but rather by Lorentzian manifolds.[18] In other words, they maintain that Lorentzian manifolds, rather than smooth manifolds, are appropriate objects to replace the placeholder K in the schema (S). A common response to these worries is that in GR the metric tensor field also represents the gravitational field, and so possesses stress-energy. This stress-energy meshes in a specific way with matter fields, so that for such a 'Lorentzian manifold substantivalist', two models would purport to describe (possibly different) ways the same spacetime could be, only if such spacetime was already tied to a specific material content. Since the leading idea behind spacetime substantivalism is the view that spacetime is somewhat independent of its material content, this is, all hands agree, bad news.[19] So, I retain manifold substantivalism as the default substantivalist position, and I turn to the hole argument itself.

A diffeomorphism is a bijection d from a manifold M onto a manifold M' (possibly $M = M'$), such that both d and its inverse are smooth. Given any model $\langle M, g \rangle$, and any diffeomorphism $d : M \to M'$, we can define a new model $\langle M', d^*g \rangle$ by, roughly speaking, defining d^*g (the pushforward of g under d) as the metric tensor field which, for any point p in M, ascribes to $d(p)$ the metric properties which g ascribes to p. Any such diffeomorphism induces a map between these two models (*qua* Lorentzian manifolds): it is, by construction, an isometry between the models, in that it carries the domain-model's metric g into coincidence with codomain-model's metric, namely d^*g. Also, the Einstein Field Equation is diffeomorphism-invariant in the sense that if some $\langle M, g \rangle$ satisfies it, then so does $\langle M', d^*g \rangle$, for any diffeomorphism $d : M \to M'$. Finally, let us employ all this machinery in order to put some pressure on manifold substantivalism.

The hole argument has its roots in some of Einstein's considerations from the 1910s,

18 See, for example, Hoefer ([1996]). To illustrate the contrast: manifold substantivalists who endorse (S-Man) would say that $\langle \mathbb{R}^4, g_{\text{Minkowski}} \rangle$ and $\langle \mathbb{R}^4, g_{\text{Gödel}} \rangle$, where $g_{\text{Minkowski}}$ is the Minkowski metric, $g_{\text{Gödel}}$ is the Gödel metric, and $\mathbb{R}^4$ is a numerically identical set-theoretic construction across these models, describe different possible situations describing the same spacetime, whereas 'Lorenztian manifold substantivalists' such as Hoefer would insist that these models represent different possible situations describing different spacetimes.

19 For an argument along these lines, see Earman and Norton ([1987], p. 519) and Norton et al. ([2023], sec. 4). I admit that this response isn't uncontroversial, because some philosophers maintain that gravitational field does not possess energy. However, I won't pursue this matter any further. For discussion, see Duerr ([2019]), Read ([2020]), and Gomes and Rovelli ([unpublished]).

but entered contemporary philosophical debate through Earman and Norton ([1987]).[20] Here is a one version of it:[21] suppose that $\langle M, g\rangle$ is a dynamically possible model of GR, and that there is some connected open region $O \subset M$ called a 'hole'. Suppose also that $\langle M, g\rangle$ admits time orientation and a Cauchy surface Σ such that O is in the future of Σ.[22] Then, consider a 'hole diffeomorphism' $h : M \to M$ such that for each point outside O, h is the identity map, and for each point in O, h is not the identity map. By the procedure described above, h defines a new model $\langle M, h^*g\rangle$, which, by diffeomorphism invariance, also satisfies the EFE.

Suppose $\langle M, g\rangle$ represents some possible world W. Keeping this representational convention fixed, one might wonder: what does $\langle M, h^*g\rangle$ represent? The manifold substantivalist regards M as representing some spacetime S. So, according to (S-Man) and the representational convention we've adopted, $\langle M, h^*g\rangle$ cannot represent W on pain of assigning different, contrary metric properties to the same spacetime point in S (represented by some point p in O) at once. But being a model of our theory—the thought goes—$\langle M, h^*g\rangle$ represents some possible world W^*. Besides, W and W^* are identical up to the time-slice represented by Σ, but they differ over the distribution of metric properties over the same spacetime points thereafter. (Since this difference is non-qualitative, we say in that case that they differ 'merely haecceitistically'.) So, by the lights of the manifold substantivalist, W and W^* are genuinely distinct possible situations.

Yet given all the data about our model up to Σ, the laws of GR do not determine whether we will 'end up' with $\langle M, g\rangle$, or with $\langle M, h^*g\rangle$. That is to say, the laws of GR do not determine whether the state of the world up to some time will evolve into W, or into

20 For a philosophico-historical account of Einstein's hole argument, see Stachel ([2014]). For a philosophico-historical account of Earman and Norton's version, see Weatherall ([2020]).

21 Not much hinges on this particular formulation apart from the convenient fact (which will be relevant in section 5) that it doesn't use isometry in order to compare the models in any way.

22 Some technical comments are in order. These two assumptions restrict our attention to the globally hyperbolic sector of GR, which makes the threat of indeterminism more intelligible. Cauchy surfaces and time orientation can give us a grip on what can count as an 'initial segment' of the possible situation and are generally assumed in the recent literature (see, for example, Roberts ([2020], p. 9) and Pooley and Read ([forthcoming], p. 1)). I also mentioned that I'm only interested in models which represent the whole universe. So, I follow Hawking and Ellis ([1973], p. 58) and Earman ([1995], p. 32) in considering only the models of GR which are inextendible, and which have a maximal development with respect to a given Cauchy surface (see Landsman ([2021], p. 166) for details; I'm grateful to Klaas Landsman and Henrique Gomes for bringing my attention to these technicalities). I admit that most of deep and interesting questions about determinism in GR concerns models which do not fulfill some of these conditions: see, for example, Doboszewski ([2017], [2019]). But I think that in the context of the hole argument debate, these conditions can be assumed even if they make determinism of GR almost a matter of mathematical proof (see footnote 46). Indeed, the reason why the hole argument is supposed to make the substantivalist uncomfortable is that this mathematically well-established sense in which GR is deterministic is apparently overridden due to philosophical reasons about the identity of spacetime points, that the hole argument will confront us with.

W^*. Since there are (generically) uncountably many different ways to construct a hole diffeomorphism on a given open subset O of M, one can arrive at many, many possible situations which, like W^*, describe apparently different futures than the one given by W.[23] So it appears that GR is radically indeterministic.

The reason why Earman and Norton take this result to pose a problem for substantivalism, rather than to be a discovery about GR, is that indeterminism would not result if $\langle M, h^*g\rangle$ represented W just as $\langle M, g\rangle$ does. And all that seems to preclude this move is the substantivalist doctrine.[24]

Early responses to the hole argument, such as Maudlin ([1988]) and Butterfield ([1989]), note that indeterminism can also be blocked if one denied that in the case of $\langle M, g\rangle$ representing some possible world, $\langle M, h^*g\rangle$ also represents a possible world. In other words, they use the fact that (S-Man) makes it possible for a model to fail to represent a possible world within a single representational convention. So what a manifold substantivalist needs is a plausible metaphysical picture which would both motivate this move and remain committed to (S-Man).

Maudlin's ([1988]) metric essentialism states that every spacetime point bears its metric properties necessarily, and thus if $\langle M, g\rangle$ represents a possible situation, then $\langle M, h^*g\rangle$ cannot. Metric essentialism faces several challenges, and I shall not discuss them here.[25] Butterfield's ([1989]) counterpart substantivalism, on the other hand, responds with a combination of (i) denial of strict trans-world identity between any objects (including spacetime points) at the expense of adopting Lewisian counterpart theory for evaluating *de re* modality claims, and (ii) a modification of the notion of determinism.

The first of these commitments renders impossible the idea that $\langle M, h^*g\rangle$ represents any possible situation, if $\langle M, g\rangle$ already does (for they would share the same spacetime points). The second commitment avoids rendering GR trivially indeterministic.

In this paper, I will defend a version of counterpart substantivalism which differs from Butterfield's in two respects. First, unlike Butterfield ([1989], p. 24), I do not take distinctness of base manifolds as a sufficient condition for distinctness of represented

23 Once one hole diffeomorphism h is constructed (notably, Muller ([1995]) has provided an explicit construction of a hole diffeomorphism applicable to the FLRW model), one could easily make more by diffeomorphically permuting the points within a small enough neighbourhood within the hole and then composing the permutation with h. Generically, there should be uncountably many such permutations available.

24 See Earman and Norton ([1987], pp. 180–1). Also, here is where a 'Lorentzian manifold substantivalist' such as Hoefer ([1996]) would object. They would note that on their view both models can represent the same possibility: since they reject (S-Man), the fact that they share the base manifold is insufficient to imply their representing different possible situations.

25 For forceful criticism of metric essentialism, see Butterfield ([1989], sec. 5), Pooley ([2021], pp. 151–3), and Teitel ([2019]).

spacetimes (that is, the converse of the implication asserted by (S-Man)). Hence, I do not commit myself to a multiplicity of qualitatively identical possible situations.[26] This will be discussed in section 4. Second, I do not fully endorse his discussion of determinism, as I will explain in section 6.

As regards defending this version of counterpart substantivalism: in section 5, I will argue against the claims of Weatherall ([2018]) and others that the hole argument is based on a confusion. But first, in the next section, I shall set the hole argument in a broader setting.

4 The Hole in the Metaphysics of Spacetime

To appreciate the broader impact of the hole argument, it is useful to see how it relates to a general issue concerning the ways in which the models of GR represent possible situations. Consider the class $\mathscr{C}$ of Lorentzian manifolds $\langle M_1, g_1\rangle, \langle M_2, g_2\rangle, \langle M_3, g_3\rangle, \ldots$ which are dynamically possible models attempting to represent general-relativistic possible situations. $\mathscr{C}$ can be partitioned into equivalence classes $C_1, C_2, C_3, \ldots$ under the isometries induced by diffeomorphisms.[27] Moreover, we can partition each C_i into equivalence classes under the relation 'has the same smooth manifold', thereby obtaining $C_{i,M_1}, C_{i,M_2}, C_{i,M_3}, \ldots$.[28]

We may now distinguish three views that connect the classes of models with possible worlds which these models may represent. (To foreshadow a bit: I will endorse the first view):

> (One World). There is only one possible world W_i which, for any representational convention, can be represented by some member of C_i, and if $i \neq j$, then $W_i \neq W_j$.[29]
>
> (One World per Base Manifold). There is only one possible world $W_{i,n}$ which, for any representational convention, can be represented by some member of C_{i,M_n}, and

26 Note that the distinction between metaphysically possible worlds and possible situations from section 2 is relevant here, since this view is compatible with one situation being 'extendible' to two distinct qualitatively identical worlds by settling some non-qualitative propositions that are not a part of GR's subject matter in a different way.

27 Even though it is standard to take isometry as the standard of isomorphism between Lorentzian manifolds, this choice does not go unchallenged. See, in particular, footnote 41.

28 Here, by two smooth manifolds being 'the same', I mean that they are set-theoretically identical. It's also worth noting that this particular order of quotienting is not essential, although it will make the notation invoked in the next paragraph less cumbersome.

29 The order of quantifiers is important. If the universal quantifier 'for any representational convention...' came first, (One World) (and also (One World per Base Manifold) below) would be consistent with the undesirable idea that the class of possible situations representable by models from C_i (or C_{i,M_n}) can vary across representational conventions.

if $i \neq j$ or $n \neq m$, then $W_{i,n} \neq W_{j,m}$.

(All Worlds). For any representational convention, every model of GR (that is: every member of each C_i) represents a distinct possible world.

Roughly speaking, (One World) is the view that there is one possible world to be represented by isometric models, whereas (One World per Base Manifold) is the view that there as many possible worlds to be represented by isometric models as there are base manifolds among these models.

Earman and Norton's ([1987]) naive substantivalist, however, subscribes to (All Worlds). Thus their argument is that since the laws of GR cannot distinguish between different possible situations represented by Lorentzian manifolds related by an isometry induced by the hole diffeomorphism, such a substantivalist must conclude that GR is radically indeterministic. So, (All Worlds) seems not to be a good choice. Indeed, from this perspective, the gist of the early papers defending substantivalism against the hole argument, including Butterfield ([1989]) and Maudlin ([1988]), was to argue that the substantivalist is not committed to (All Worlds).

Butterfield ([1989]) himself must subscribe to (One World per Base Manifold), because, as I mentioned at the end of section 3, he takes the distinctness of base manifolds as a sufficient condition for distinctness of represented spacetimes (that is, the converse of the implication asserted by (S-Man)). In local spacetime theories such as GR, it is a working assumption that all observable properties and relations are reducible to quantities which are invariant under isomorphism.[30] Since isometry is generally taken to be the standard of isomorphism for the models of GR (again: setting aside the matter fields, as I did in section 3), the situations represented by isometric models are qualitatively identical, and might differ, if at all, merely haecceitistically (that is, with respect to non-qualitative propositions that they settle). So, Butterfield ([1989]) is committed to the existence of multiple qualitatively identical possible worlds.[31]

In my view, however, the converse of (S-Man) is not an integral part of a counterpart-theoretic response to the hole argument. One might adopt counterpart theory to deal with *de re* modal propositions in the context of spacetime theories and yet prefer, as I do, to endorse (One World). My reason for this is that I'm uncomfortable with the idea of multiple qualitatively identical possible situations,[32] not only due to their unparsimonious nature,

30 See Norton et al. ([2023], sec. 3).

31 Admittedly, their importance for the debate on determinism is bound to be nullified after choosing an appropriate definition of determinism. See section 6 below.

32 I shall not make any judgements regarding metaphysically possible worlds. When it comes to them, Lewis ([1986], p. 224) himself advised 'that we remain agnostic'.

but mainly because I prefer to set aside the worry of underdetermination that they bring about.[33] Let me stress that this is not to say that there are no non-qualitative propositions that are settled by the possible situations described by GR. For one, propositions about spatiotemporal locations and their properties (such as the value of scalar curvature at a given spacetime point) are by all means non-qualitative and within the subject matter of the theory. But no two situations differ only over the truth-values of such propositions. This makes my response anti-haecceitistic, insofar as possible situations described by GR are concerned.

Now, let's turn to counterpart theory. On my preferred approach, the counterpart relation is a relation defined on manifold points of a pair of models of a given theory, and within a single representational convention. It is supposed to track the relation of physical similarity, as described by the given theory, between spacetime points represented within the models in question.[34] Generally, given any representational convention and any two distinct GR-models $\langle M, g\rangle$, $\langle N, h\rangle$, where both of these models are taken to represent a possible situation in that convention, any possibly partial and possibly multi-valued map between their base manifolds gives us a 'candidate counterpart relation'. The issue is that such a candidate relation might not always be most physically apt to be used to evaluate *de re* modal propositions relativised to GR. Finding a method for specifying such physically apt maps mathematically, even for a relatively small class of models, would indeed make a contribution to our understanding of a given spacetime theory.[35] But for present purposes, such a general method is not needed, because the hole argument is concerned with isomorphic models, so all physically relevant properties specified by the matter and radiation fields are invariant under the isomorphism by default, and the underlying differomorphism gives us a matching between the manifold points that exactly preserves the field values representing these physically relevant properties. So, in this case, the counterpart relation is given by the diffeomorphism. Moreover, since I deny the existence of multiple qualitatively identical possible situations described by GR, such a counterpart relation between manifold points of isometric models will relate points that

33 For a discussion on whether underdetermination of this kind would be epistemically worrisome, see Maudlin ([1993]) and Dasgupta ([2015]). The issue of underdetermination will briefly return in the next section.

34 See also footnote 11.

35 A big step in this direction has been recently made by Gomes and Butterfield ([2023b]), who provide a technical notion of counterparthood between spacetimes using sections of an infinite-dimensional fiber bundle of models of GR *in vacuo*, where the fibers are generated by the action of the diffeomorphism group on the base manifold of the model which is fixed throughout the construction. Then, they use the idea of 'threading' between across models from different fibers to get the notion of counterparthood between manifold points of the models.

represent the same physical spacetime point.

Thus, if one sets up a representational convention such that $\langle M, g\rangle$ (say) represents a possible world, then for any for any manifold M' that is diffeomorphic to M but distinct from it, one is free to choose any diffeomorphism $d : M \to M'$ to pushforward the metric g to $d^*g =: g'$, and the resulting isometric model $\langle M', g'\rangle$ will represent the same possible situation, and the spacetime point o represented by $p \in M$ in the former model will be represented by $d(p) \in M'$ in the latter. And any proposition about the properties of o (such as its scalar curvature value) that is warranted by $\langle M, g\rangle$ in virtue of the (mathematical) properties of p there, will be warranted by $\langle M', g'\rangle$ in virtue of the (mathematical) properties of $d(p)$. In a nutshell, such a diffeomorphism tells us which pairs of points of distinct isometric manifolds represent within our representational convention the same spacetime point.

On this picture, however, in accordance with (S-Man), no other model with a base manifold M (or with the diffeomorphic M') will represent any possible world within that same representational convention. And these 'disenfranchised' (that is: non-representing) models include the models which are not isometric to $\langle M, g\rangle$ and $\langle M', g'\rangle$. So, if one wishes to represent some distinct possible world, but remain within that representational convention, one would have to choose a model with yet another base manifold M'', say $\langle M'', g''\rangle$, which is not isometric to the previous two models.[36]

Now, let me turn to the threats against counterpart substantivalism. First, I will address Weatherall's claim that the hole argument is, in general, based on a confusion. In terms of the labels I have adopted above, this argument amounts to the claim that (All Worlds) and (One World per Base Manifold) are ruled out not by philosophical argument of the kind I am pursuing, but by the sheer practice of mathematical physics, so that Earman and Norton's ([1987]) naive substantivalist who accepts (All Worlds) is a strawman. I will disagree with this claim.

36 One may sensibly ask whether this procedure is formalisable. I believe that it can be done insofar as 'formalisability' means a well-defined mathematical operation on the space of models which would pick, for any possible world and without violating (S-Man) and (One World), a model representing that world. Gomes ([2021], ch. 1.2.1.) defines the formal procedure for choosing a representational convention along these lines among equivalence classes (that is, orbits of the group of isometries) of Lorentzian manifolds under isometry. One can simply adopt his approach while insisting that the models from two distinct orbits must contain different base manifolds. So, Gomes' procedure enriched with the operation of choosing a unique base manifold for each orbit would give us the desired result. This operation can be specified, in an obvious way, by a bijection between $\{C_1, C_2, C_3, \ldots\}$ and $\{C_{i,M_1}, C_{i,M_2}, C_{i,M_3}, \ldots\}$ (where i is fixed).

5 The Hole Confusion?

For Weatherall ([2018]), the hole argument 'is based on a misleading use of the mathematical formalism of general relativity' (p. 330). This conclusion, and Weatherall's argument for it, has been endorsed, among others, by Weatherall ([2020]), Fletcher ([2020]), Bradley and Weatherall ([2022]), and Halvorson and Manchak ([forthcoming]). So, whereas the proponents of this argument might endorse various combinations of metaphysical positions described in this paper ((S-Man), (One World), etc.), they ultimately believe that such metaphysical considerations are irrelevant to the hole argument. In this section, I explain why I disagree with these claims.

Following Pooley and Read ([forthcoming]), I distinguish two arguments offered by Weatherall ([2018]) for the above conclusion (these arguments are not completely independent, but also do not stand or fall together).

First, Weatherall ([2018], p. 336) argues that the hole argument hinges on two facts:

(i) two putatively possible worlds represented by models related by isometry (say $\langle M, g\rangle$ and $\langle M, d^{*}g\rangle$) are empirically indistinguishable, and

(ii) $\langle M, g\rangle$ and $\langle M, d^{*}g\rangle$ are distinct objects (at least insofar as the mathematics of GR is concerned).

Now, empirical indistinguishability of putatively possible worlds represented by $\langle M, g\rangle$ and $\langle M, d^{*}g\rangle$ can be established only if we compare these two Lorentzian manifolds using the isometry induced by a given diffeomorphism d. Distinctness of $\langle M, g\rangle$ and $\langle M, d^{*}g\rangle$, however, can be established only if we compare these two Lorentzian manifolds using the identity map on the base manifold M. This, however, poses a problem, because we need to invoke two different standards of comparison and 'one cannot have it both ways' ([2018], p. 338). Weatherall's first argument, then, can be summarized thus:

(a) one cannot use both standards of comparison (that is, isometry and the identity map), and this is presupposed by the hole argument.

Weatherall's second argument is that:

(b) isometry is the 'relevant standard of sameness' for Lorenztian manifolds warranted by 'contemporary mathematics' ([2018], p. 331).

This should also block the hole argument, because we would have no legitimate means of establishing point (ii) above.

I am persuaded by the critique of claim (a) by Pooley and Read ([forthcoming], sec. 4). The crux of their argument is that the formulation of the hole argument which results in the threat of indeterminism (rather than underdetermination) does not require any mention of empirical indistingushability between putatively possible situations, and thus one does not need to 'compare' models using isometry. Indeed, my own presentation in section 3 is an example of this. So, claim (i) is set aside and it is enough to establish point (ii), so as to get the hole argument going.

I am also sympathetic to the critique of (b) by Roberts ([2020]) and Pooley and Read ([forthcoming], sec. 5). But I still think that some comments are in order, for two reasons. First, (b), if true, would block the hole argument independently of whether (a) is true. So (b) deserves further scrutiny. Second, more importantly, there is an independent motivation for (b) which is not extensively discussed in the literature, perhaps because it emerged only recently.

For Weatherall, what makes the relevant standard of sameness truly relevant, is 'the way in which mathematics is used in physics' ([2018], 330, fn. 4), that is, the practice of mathematical physics. This entails the following principle:

> [I]f a particular mathematical model may be used to represent a given physical situation, then any isomorphic model may be used to represent that situation equally well. ([2018], p. 332)

This principle can be interpreted in two ways. On one reading, it is equivalent to what Roberts ([2020], p. 252) calls 'Weak Leibniz Equivalence', (WLE), which states that each of the isomorphic models can be chosen to represent a particular physical situation equally well. On the other reading, Weatherall's principle is equivalent to what Roberts calls 'Strong Leibniz Equivalence', (SLE), which states that any two isomorphic models may be used to represent a particular physical situation equally well at once (or, as I would say, within a single representational convention).[37]

In the case of the hole argument, where the relevant diffeomorphism is the hole diffeomorphism h, (WLE) states that I am equally correct in stipulating that $\langle M, g\rangle$ represents a possible world W, and in stipulating (independently of the previous stipulation) that $\langle M, h^{*}g\rangle$ represents W. (SLE), however, states that having stipulated that $\langle M, g\rangle$ represents W, I am correct in claiming that $\langle M, h^{*}g\rangle$ also represents W. It is clear that (WLE) can be accepted by all parties in this dispute (see footnote 13), including even those who embrace (All Worlds), so it cannot be used to vindicate Weatherall's conclusion in any

37 Pooley ([2021]) discusses a principle called (Models), which is equivalent to the restriction of (SLE) to GR.

way.

I would like, however, to explore a different way of motivating Weatherall's claim (b): one which is not explicitly focused on the issue of representation. Rather, it is concerned with the correct conception of expressive resources available to a competent interpreter (or user) of a given highly-mathematicised physical theory.

In order to motivate Weatherall's claim (b), Bradley and Weatherall ([2022]) argue that the hole argument presupposes expressive resources that are external to GR, where GR is understood as a theory whose models' structure is 'fully characterized by the structure of [...] pseudo-Riemannian manifold[s]' ([2022], p. 1231). Such external expressive resources are said to be a part of 'semantic metatheory' ([2022], p. 1227), and should not be used by any reasonable interpreter of GR unless sufficient justification is given. Usually, however, no such justification is given, according to Bradley and Weatherall.[38] (Very similar remarks are offered by Halvorson and Manchak ([forthcoming], pp. 24–6).)[39]

Notions such as 'expressive resources' (or 'expressive power') and 'semantic metatheory' are technical concepts used by philosophical logicians and philosophers of language in their analyses of (semi-)formal languages and frameworks. Yet, neither Bradley and Weatherall ([2022]) nor Halvorson and Manchak ([forthcoming]) explicate what exactly they mean by those terms. To make matter precise, I will take 'expressive resources' and 'expressive power' of a theory cast in a given language to be exhausted by the propositions expressible by sentences of that language under an intended interpretation, whereas 'semantic metatheory', in the present context, is best understood as the collection of propositions expressible by sentences in the metalanguage that are not expressible by sentences in the object language (both under intended interpretations). I will now argue, however, that once these notions are made precise, the threat of unjustified semantic ascent proposed by these authors is either untrue or leads to an overly restrictive conception of the language in which GR might be cast.

General relativity, in its standard formalism, is not a formal theory in the sense studied by logicians. The models of GR, as I presented them in section 3, are not models in the model-theoretic sense. They are sets with certain properties, constructed in a background set theory, whose existence is guaranteed by any model (now in the model-theoretic sense) of that background set theory. So, the object language of GR in its standard formalism

38 See Bradley and Weatherall ([2022], pp. 1229, 1231).

39 See their discussion of the difference between the 'theory GR [...] [and the theory] GR+ZF$_m$, where the latter is ZF set theory read in the "material mode" as a theory about concrete possibilia' ([forthcoming], p. 24)). Halvorson and Manchak do not explicitly state what they take 'the theory GR' to be, but their assertion that 'the statement "p can have different metric properties in different models" cannot even be formulated [in the theory GR]' ([forthcoming], p. 25), suggests that they have something like a Mundy-style axiomatisation of Lorentzian geometry in mind (see also footnote 41).

is the language of set theory.[40] Now, it's possible to construct a true sentence of set theory which says that two given isometric Lorentzian manifolds are non-identical, so the language of set theory definitely has expressive resources to distinguish between isometric Lorentzian manifolds. So, it's definitely not true that 'the language of GR' understood as 'the language in which the standard formalism of GR is usually cast', has no expressive resources to carry out the hole argument.

Admittedly, I take Bradley and Weatherall ([2022]) and Halvorson and Manchak ([forthcoming]) to agree with this, but to disagree about what 'the language of GR' should be once the structure of GR-models is taken into account. In particular, they would hold that it should be a language unable to distinguish between objects playing the roles of isometric Lorentzian manifolds on one hand, and yet capable of recovering a working formalism for general relativity (as it is practiced) on the other. I reply, however, that no language would be able to meet both of these desiderata, simply because physicists need an object language where the difference between isometric Lorentzian manifolds is expressible, even for purely formal purposes.[41] Otherwise, one would trivialise such fundamental tools as the application of the Lie derivative to the metric, where a non-vanishing Lie derivative of the metric is presupposed by, among many other things, a standard treatment of variational symmetries using Noether's Theorem and the relativistic definition of a non-rigid continuous body.[42]

Moreover, in some cases, distinguishing between non-isometric models might be rel-

40 Let's assume that the background set theory is pure, that is, it admits no Urelemente and it is not, *pace* Halvorson and Manchak ([forthcoming], p. 24), a theory about 'concrete possibilia'. Thus, the sentences of this theory express propositions that are purely about mathematical objects.

41 Set theory obviously meets the second desideratum, but fails at the first. There are some candidates that meet the first desideratum. For example, in Homotopy Type Theory (HoTT), under the assumption that Lorentzian manifolds are constructed as the so-called 'dependent pairs', isometric Lorentzian manifolds are 'internally' identical (roughly, there's an object in HoTT that expresses the proposition that $\langle \mathtt{m}, \mathtt{g} \rangle$ and $\langle \mathtt{m}, \mathtt{h}^{*}\mathtt{g} \rangle$ are identical), even though they're also 'externally' non-identical ($\langle \mathtt{m}, \mathtt{g} \rangle$ and $\langle \mathtt{m}, \mathtt{h}^{*}\mathtt{g} \rangle$ are not co-substitutable in all contexts, for example $\langle \mathtt{m}, \mathtt{h}^{*}\mathtt{g} \rangle$ cannot be substituted into the definition of $\langle \mathtt{m}, \mathtt{g} \rangle$). For a pedagogic exposition of the HoTT-perspective on the hole argument, see Ladyman and Presnell ([2020]). Alternatively, one might consider an intrinsic axiomatisation of Lorentzian geometry in a higher-order logic as in (Mundy [1992]), where the objects representing spacetime would be models of this formal theory ('models' in the model-theoretic sense, of course). By a suitable representation theorem, such models would be translatable into Lorentzian manifolds *qua* set-theoretic constructions, and model-theoretic isomorphisms would be translatable into set-theoretic isometries. Since isomorphic models are elementarily equivalent, they make exactly the same sentences of the theory true, so there's no way to distinguish between them from within the theory. Notice, however, that on this approach there's also no way to express any comparative claims about any models whatsoever.

42 This point is raised by Landsman ([2023]) and Gomes and Butterfield ([2023a], pp. 22–3) who provide further examples and discuss it in much greater detail. Moreover, if one decides to adopt a Mundy-style axiomatisation as the language of GR, as Halvorson and Manchak ([forthcoming]) seem to lean toward (see footnote 39), one would not be able to express any comparative claims about the models within the object language. Since these models are supposed to represent possible spacetimes, one would not be able, for example, to express results about the topological stability of spacetime's properties.

evant for more than purely formal considerations. As pointed out by Belot ([2018], pp. 966–70), within the class of models that involve given asymptotic boundary conditions (for example that the spacetime metric approaches Minkowski metric, or de Sitter metric, at spatial, or null, infinity), maps between models that might cut finer than isometries are used to search for conserved quantities and to study physically relevant phenomena. Whereas Belot focuses on solutions that are asymptotically flat (that is Minkowski) at spatial infinity, one might also look at asymptotic symmetries for solutions that approach flatness at null infinity. There, for example, the supertranslations from the BMS group, which extends the Poincaré group, might be used to reproduce the gravitational wave memory effect.[43]

To summarise, putting a cap on the expressive resources that a physics-suitable mathematical language might have is, quite clearly, a very radical manoeuvre that is unlikely to accommodate the flexibility with which mathematical physicists use mathematics. So, it is no surprise that, despite various assertions about ‘expressive resources’ and the ascent to ‘semantic metatheory’ being unjustified, Bradley and Weatherall ([2022]) sometimes read Weatherall ([2018]) as pursuing a different project. Namely, one that ‘[...] is not to solve an interpretational problem by adopting some novel formal apparatus. Rather, it is to argue that the (formal) problem allegedly raised by the hole argument is illusory.’ ([2022], p. 1224). This alternative approach is backed up by the claim that ‘as a principle of mathematical practice, mathematical objects are defined only up to isomorphism’,[44] and even though the formalism should not be changed, its correct interpretation (or application) is supposed to be illuminated by the fact that the formalisations of Lorentzian geometry either in higher-order logic, or Homotopy Type Theory, do not contain expressive resources that would distinguish between isometric Lorentzian manifolds.

Yet it is difficult to square these remarks with the claims about expressive resources

43 See, for example, Strominger and Zhiboedov ([2016]). I thank an anonymous referee for bringing this example to my attention. There is, of course, an additional problem posed by Belot’s considerations, namely how the existence of isometric, yet arguably physically inequivalent, models bears on (One World) of section 4. Even though the discussed solutions are used to model isolated subsystems of the universe, Belot ([2018], p. 970) notes that, in the presence of a positive cosmological constant, there are asymptotically de Sitter solutions able to model the whole universe. This does not on its own commit us to (All Worlds), but perhaps it shows that the decision to partition the class $\mathscr{C}$ of models using isometry was too hasty, and yet some more fine-grained notion of isomorphism (such as isometry and appropriate agreement at spatial (or null) infinity) can do the job in such a way that we can retain (One World) in a slightly modified version. These, I admit, are just speculations. But I want to stress that the threat to (One World) posed by Belot ([2018]), unlike my response to the claims of Bradley and Weatherall ([2022]) and Halvorson and Manchak ([forthcoming]), does depend on whether the physicists are right to decouple isometry and physical equivalence in the cosmological sector, and I shall not pursue this matter any further here. For a criticism of Belot’s examples from a different angle, see Luc ([2022]).

44 This is what Bradley and Weatherall ([2022], p. 1226) ascribe to Weatherall ([2018]). See also Weatherall ([2018], p. 331).

and the appeal to semantic metatheory being unjustified. For expressive resources depend on the language (that is 'the formal apparatus'), so insofar as one is working in background set theory (as is commonly assumed in mathematical textbooks), it is Bradley and Weatherall's claim that isomorphic mathematical objects should be treated as identical that counts as a part of semantic metatheory! And, as discussed in previous paragraphs (and footnote 41), the alternatives to set theory that cannot internally distinguish between isometric Lorentzian manifolds are arguably not apt for physics. So I conclude that the considerations about (in)expressibility offered by Bradley and Weatherall ([2022]) and Halvorson and Manchak ([forthcoming]) do not provide a convincing motivation for Weatherall's claim (b). Consequently, they do not vindicate the dissolution of the hole argument on purely mathematical grounds.

On a more irenic note, let me add two comments. First, recasting physical theories (or parts of them) into axiomatised formal theories is generally a noble and illuminating enterprise, even though the interpretational significance of any such formalisation must always be tested against our knowledge of the bounds of a given physical theory as well as its use in practice, both theoretical and experimental. So, in principle, there is nothing wrong with using formal means to reach philosophical conclusions via some form of explication, but any such attempt must be prefaced by a guarantee that the chosen formal apparatus is appropriate for the question at hand. This point will in fact reappear in sections 6.2 and 6.3.

Second, I admit that there is something not quite right with modal propositions relativised to a particular theory T that can be formally explicated only by distinguishing between isomorphic T-models. I do not think, however, that 'inexpressibility' is the most adequate concept for capturing this sentiment, especially when applied to purely mathematical distinctions. In order to find a middle road that implements this disquiet with what is expressible, and yet avoids the error of an unjustified limitation of mathematical apparatus, one would have to construct new semantics for modal propositions relativised to a theory T. Unfortunately, the authors I've criticised offer no such account.

6 Counterpart Substantivalism and Determinism

In this section, I argue that my preferred version of counterpart substantivalism is perfectly compatible with an attractive definition of determinism which avoids the worries raised against counterpart-theoretic notions of determinism by Melia ([1999]) and Belot ([1995], [2018]). This attractive definition is cast in terms of possible worlds, but it also has a few equivalent formal explications for GR formulated in terms of GR-models.

I begin, in section 6.1, with the discussion of the canonical model-based, counterpart-substantivalist definition of determinism for local spacetime theories, namely Butterfield's ([1987], [1989]) (Dm2). I argue that Butterfield's hole-argument-based motivation for (Dm2) is not fully justified, and that the definition can be strengthened in the context of the hole argument without any loss of its counterpart-theoretic spirit. Then, I state how any of these model-based definitions can explicate my preferred worlds-based definition of determinism which I shall call (Dm2-Worlds+). It will differ from the standard gloss of (Dm2) in terms of worlds (called (Dm2-Worlds)).

Then, in section 6.2, I discuss the examples of intuitively indeterministic toy theories, which are allegedly classified as deterministic by counterpart-friendly definitions of determinism. I agree that these examples pose a problem to (Dm2-Worlds), but I also note that they are correctly classified as indeterministic by (Dm2-Worlds+). Then, I consider the worry that admitting theories explicitly describing possible situations that differ only as to where or when things happens should give us reason to do the same in the case of GR. I resist this conclusion: all these toy theories, as they are usually described, are radically different from GR in the way in which they present their modal content. And if they are described as theories that present their modal content in a way GR does, then they are not indeterministic. I will illustrate this claim with an exercise in formalisation in section 6.3. It follows that they should not be treated as a guide to the space of possible situations described by GR.

6.1 Definitions of determinism

(Dm2) was originally defined for all local spacetime theories, and the notion of determinism was relativised to a certain kind of spacetime region. Here is its harmless restriction to the sector of GR subject to the conditions given in footnote 22 (an initial segment is understood as a causal past of some Cauchy surface):[45]

> (Dm2). A theory with models $\langle M, g\rangle$ is deterministic iff for any two models $\langle M, g\rangle$ and $\langle M', g'\rangle$ containing initial segments S and S', and any diffeomorphism $\alpha : S \to S'$:

45 See Butterfield ([1987], p. 29, [1989], p. 9) for the original definition, and Doboszewski ([2019], p. 11) for a useful generalization that makes room for imposing various auxiliary conditions on models and/or regions and/or functions between regions (as I, in fact, did in footnote 22 by restricting my attention to a particular class of GR models). Also, Butterfield ([1987]) explicitly says that S/S' in GR are Cauchy surfaces ('slices' or 'sandwiches'), rather than causal pasts of Cauchy surfaces ('segments'), but this difference is unimportant for our purposes. Anyway, the latter choice meshes better with the idea of 'initial temporal segments' of a world (due to Lewis ([1983], p. 359)), which Butterfield ([1989], pp. 25–6) wishes to capture through (Dm2).

> if $\alpha^*(g) = g'$ on $\alpha(S) = S'$,
> then there exists an isometry $\beta : M \to M'$ such that $\beta^*(g) = g'$ throughout M' and $\beta(S) = S'$.

Butterfield ([1989], p. 9) emphasises that he does not wish to strengthen the consequent by requiring β to extend α, because such a definition would be violated by two hole-diffeomorphic models $\langle M, g\rangle$, $\langle M, h^*g\rangle$. His reasoning is this: Let α be a global (that is, $M \to M$) identity map. It's a diffeomorphism. And, for $S = S' := M/O$ (where O is the hole), $\alpha^*(g) = h^*g$. So, the antecedent 'if $\alpha^*(g) = g'$ on ...' is satisfied. Yet, if β were to extend α, it would have to be the case that $\alpha = \beta$, because β is a global map. And it's not the case that $\beta^*(g) = \alpha^*(g) = g = h^*g$ throughout M. So, the consequent of (Dm2) is false.

This reasoning, however, is flawed. In (Dm2), the diffeomorphism α is defined as a map from S to S', not as a map from M to M'. If one says, 'let α be a global identity', one thereby sets $S = M$, and $S' = M$. And if one does that, then one's antecedent is: 'if $\alpha^*(g) = g'$ on $\alpha(M) = M$'. This antecedent is false for $g' = h^*g$, so that hole-diffeomorphic models would not give a counterexample to a stronger version of (Dm2).

But what if we changed (Dm2), so that α would now be defined as a map from M to M', even while the antecedent remains restricted to S and S'? Then, we could say that the antecedent holds if the restriction of α to S drags along the metric appropriately. In that case, however, the gloss on (Dm2) could no longer be 'if there is a local isomorphism, then there is a global isomorphism which need not extend the local one', because what counts as the 'local' isomorphism is the restriction of α to S, not α itself.

Should we strengthen the consequent of (Dm2) then, and require β to extend α (and call it (Dm2+))? Or perhaps we should strengthen it even further, and require β to uniquely extend α (and call it (Dm2++))? In my view, it doesn't matter, because within the sector of GR we're interested in (recall footnote 22), there is no pair of relevant models that would satisfy one of these versions, but not the others. This is a consequence of two mathematical facts:[46]

> (Fact 1). For any two models $\langle M, g\rangle$ and $\langle M', g'\rangle$ containing initial segments S and S':
>
> if $\phi : \langle M, g\rangle|_S \to \langle M', g'\rangle|_{S'}$ is an isometry,
> then there is an isometry $\psi : \langle M, g\rangle \to \langle M', g'\rangle$ such that $\psi|_S = \phi$.

46 (Fact 1) is a direct consequence of the Choquet-Bruhat–Geroch Theorem (it's also sensitive to the assumptions about GR models from footnote 22). For a sketch of the proof and references, see Choquet-Bruhat ([2009], p. 400). For a proof of (Fact 2), see Giulini ([2007], p. 165, fn. 6) or Halvorson and Manchak ([forthcoming], p. 18) (their (Theorem 1)).

> (Fact 2). For any two models $\langle M, g\rangle$ and $\langle M', g'\rangle$ containing initial segments S and S', and any two isometries ϕ and ψ between them:
>
> if $\phi|_S = \psi|_S$, then $\phi = \psi$.

(Fact 1) shows that no pair of models which satisfies (Dm2) would violate (Dm2+). (Fact 2) shows that no pair of models which satisfies (Dm2+) would violate (Dm2++). The converse claims hold just because of the varying logical strengths of these definitions. And I see no intuitive reason to prefer one of them over the others.

I submit that determinism, as a feature of an interpreted theory, is fundamentally not a feature of the theory's models, but rather of the possible situations it describes (even if the relationship between models and situations is quite tight). So, one might wonder whether there is a common gloss in terms of possible worlds that we can give to the family of (Dm2)-style definitions of determinism. As I discussed in section 4, in local spacetime theories such as GR, observable properties and relations are reducible to quantities invariant under isomorphism. For this reason, the original (Dm2) has often been cashed out in the following way:

> (Dm2-Worlds). A theory T is deterministic iff for any possible worlds W and W' described by T: if there is a qualitative agreement between them up to some time, then there is a total qualitative agreement between them.[47]

I do not endorse (Dm2-Worlds) as it stands. I prefer the following variant, which changes the consequent to '... then $W = W'$':

> (Dm2-Worlds+). A theory T is deterministic iff for any possible worlds W and W' described by T: if there is a qualitative agreement between them up to some time, then $W = W'$.

There are two reasons for this modification. First, I endorse (One World) (recall section 4), so that I can't make much sense of two distinct qualitatively identical possible situations described by GR anyway. Second, I agree that there are indeterministic theories (discussed in section 6.2 below) that admit distinct possible situations which differ only about where or when things happen (in general, or after a certain time)—I just don't think that GR is one of them.[48]

47 For this gloss on (Dm2), see Belot ([1995], p. 190), Melia ([1999], p. 656), and Pooley ([2021], p. 155).

48 Also, note that this modified gloss in terms of worlds will apply to each variant of (Dm2), because any global isometry between models representing possible situations establishes, according to (One World), that they represent the same situation (regardless of whether or not it extends the local isometry).

Now let's see how the above discussion relates to the problem often raised against Butterfield's ([1989]) version of counterpart substantivalism, namely that its counterpart-theoretic definition of determinism classifies some intuitively indeterministic theories as deterministic.

6.2 What to make of toy theories

Melia ([1999], p. 661) and Belot ([1995], pp. 190–4, [2018], p. 949) presented intuitively indeterministic toy theories that are nevertheless classified as deterministic by (Dm2-Worlds). I will focus only on Melia's toy theory. Belot's ([1995]) first theory, borrowed from Wilson ([1993], pp. 215–6), has a very similar structure to Melia's and my discussion will apply to it just as well. Belot's ([2018], p. 949) 'swerve theory' is only slightly different (as Belot ([2018], p. 973, fn. 11) himself admits), and nothing important hinges on this difference. The other two theories offered by Belot ([1995]) rely on the idea of same-worldly counterparts and possibilities generated by those: an idea which I find rather exotic in the fairly sanitised context of modality within physical theories, since the most common motivation for this idea turns on some contentious issues regarding consciousness and the self.[49]

I do not wish to dispute the claims that such toy theories show (Dm2-Worlds) to be faulty. But, as we shall soon see, these theories cannot be misclassified as deterministic by (Dm2-Worlds+), precisely because they describe distinct possible worlds which differ, to the future of the relevant time, merely haecceitistically, that is, regarding only facts that are non-qualitative. In that case, the antecedent of (Dm2-Worlds+) will be true, but the consequent false.

Still, one might raise the following worry: if we admit as intelligible some theories which specify possible worlds that differ merely haecceitistically, why shouldn't we interpret GR in the same way? Indeed, coming back to the taxonomy of section 4, why should we choose (One World) over (One World per Base Manifold) or even (All Worlds)? It might seem arbitrary that in the case of some theories we admit that such theories describe distinct, but qualitatively identical possible situations, and in other cases we don't.

I resist this conclusion. For there are two ways in which we can have access to the space of possibilities according to a theory: a direct way, and an indirect way. We can imagine theories that give us a direct access to their possibilities (setting aside any epi-

[49] See Belot ([1995], pp. 192–3) for details. The notion of same-worldly counterparts is taken from Lewis ([1986], pp. 230–2), as is the intuition pump motivating them. The force of my remark, of course, would be substantially diminished have I not distinguished possible situations (described by a theory T) from metaphysically possible worlds where the laws of T hold in section 2.

stemological concerns): they explicitly specify that the world could have been so-and-so if the theory were true, even if it is in fact such-and-such. A different kind of theory would only give us direct access to its space of models, which might not coincide one-to-one with its space of possible physical situations. We can specify how the world could have been if such a theory were true only indirectly, through some conceptual or (meta)physical exegesis of the theory and of its models. GR is indeed an example of the latter kind of theory. And my point in what follows is this: if we treat these toy theories as being of the former ('direct') kind, then they are indeed indeterministic, but also very different from GR. But if we treat them as being of the latter ('indirect') kind, we should first arrive at appropriate formalisations of them, and then it might turn out that we no longer have any reasons to believe that they are indeterministic (unless we enrich them with such gadgets as modal operators, but then they will—again—differ radically from GR). Now let me illustrate these claims with an example.

6.3 An exercise in formalisation

Here is Melia's ([1999], p. 661) theory of four particles:[50]

> There are two duplicate white particles, and two duplicate black particles. Starting from some initial time t_0, there is a fixed time t_1 at which the two black particles travel with a constant speed towards the two different white particles in a straight line.

Melia claims that even though this theory is (Dm2-Worlds)-deterministic, it is nevertheless intuitively indeterministic, because it does not determine which black particle travels to which white particle. Now, I will show that if we were to treat Melia's theory in a way similar to that in which we treat GR (that is, as a theory giving indirect access to its modal content), the intuitive appeal of its indeterminism will disappear. Only if we treat it in a rather special way, can we account for those indeterministic intuitions.

A harmlessly simplified version of Melia's theory postulates only two discrete times: one at which the white and black particles are all spatially separated, and the other at which the white and black particles are adjacent in two colour-mixed pairs. Then, we can attempt to formalise this simplified theory as a dynamical theory. I take a dynamical theory to be a many-sorted first-order formal theory, with a countable and linearly ordered set of sorts $\{\sigma_1, \sigma_2, \ldots\}$ (intuitively: times), whose logical vocabulary contains (among

[50] Melia's original presentation uses names for particles, but I shall refrain from this so as not to make things confusing in my initial formalisation. I return to this issue below.

other standard things) countably many identity relation symbols for objects of the same sort, one for each sort. Non-logical vocabulary is often relativised to sorts, and it may also contain function symbols of the form $\delta_{i,j}$ which denote (possibly partial) functions from objects of sort σ_i to objects of sort σ_j.[51] I wish to focus on dynamical theories in particular, because the structure of their models is analogous (in a way) to the structure of globally hyperbolic models of GR.[52]

Returning to the task of formalising Melia's theory, let's consider a two-sorted dynamical theory, which I call 4P, with non-logical vocabulary $\{W^1, B^1, A^2_{\sigma_2,\sigma_2}, \delta_{1,2}\}$ standing for, respectively, unary predicates 'white' and 'black', a binary predicate 'adjacent' (for ease of exposition, I let unary predicates be unrelativized to sorts), and a function symbol denoting persistence over time.[53] 4P has the following axioms (all of which can be explicitly formalised in our language):

> ($4P_1$) There are exactly four particles of sort σ_1: two are B, two are W, no particle is both W and B.
>
> ($4P_2$) There are exactly four particles of sort σ_2: two are B, two are W, no particle is both W and B, some B-particle stands in relation A with some W-particle, and the other B-particle stands in relation A with the other W-particle.
>
> ($4P_3$) $\delta_{1,2}$ is a bijection such that any argument and its value are either both W or both B.

So much by way of presenting the formalisation 4P of Melia's toy theory.

4P is clearly a categorical theory (that is: all its models are isomorphic), so it must also be (Dm2)-deterministic. But the intuition behind 4P being indeterministic is that the axiom ($4P_2$) does not specify which black particle is adjacent to which white particle at time (sort) σ_2. In terms of models of our formalisation, we may imagine the following two models:

51 For details on many-sorted logics, see Manzano and Aranda ([2022]). For details on using many-sorted logic to formalise dynamical theories in the context of the hole argument, see Halvorson and Manchak ([forthcoming], pp. 20–1).

52 $\delta_{n,m}$ is analogous to a function which, roughly speaking, maps each point on a given spacelike hypersurface Σ_n to the unique point in the neighbouring future spacelike hypersurface Σ_m that lies on the timelike geodesic normal to Σ_n. It's worth noting that whereas a foliation of a globally hyperbolic spacetime contains uncountably many spacelike hypersurfaces, I have taken (as is customary) the set of sorts to be countable. Nothing important hinges on this simplification.

53 Even though I believe that properties are not (collections of) objects (as everyone who has read Frege should!), I assume that predicates refer to collections or tuples of objects, and the models of 4P under consideration are those whose domain consists of particles, and whose extension of (say) W are those particles which are in fact white, etc. I stipulate this in order to have a closer connection between models of 4P and what intuitively count as possible worlds according to Melia's theory.

Model $\mathscr{M}$ has domains: $M_1 = \{a, b, c, d\}$ and $M_2 = \{a', b', c', d'\}$ (where $\delta_{1,2}(x) = x'$), and is such that a, b, a', b' are black, c, d, c', d' are white, and pairs $\langle a', c' \rangle$ and $\langle b', d' \rangle$ are adjacent.

Model $\mathscr{M}^*$ has domains: $M_1^* = \{a, b, c, d\}$ and $M_2^* = \{a', b', c', d'\}$ (where $\delta_{1,2}(x) = x'$), and is such that a, b, a', b' are black, c, d, c', d' are white, and pairs $\langle a', d' \rangle$ and $\langle b', c' \rangle$ are adjacent.

The intuition that 4P is indeterministic is explicated by the fact that $\mathscr{M}$ and $\mathscr{M}^*$ have the same objects, have the same initial segment up to sort σ_1, but different pairs of objects are adjacent when sort σ_2 is included.[54] But, a model is not a possible world, and an obvious comment on behalf of the counterpart substantivalist is to deny that $\mathscr{M}^*$ represents a possible situation if $\mathscr{M}$ already does, at least insofar as the identity of objects across the domains of models matches the identity of objects across putative possible worlds. Let me call this argument 'the metaphysical response'. I believe it is enough to show that 4P fails to be a counterexample to (Dm2-Worlds+).

Alternatively, one might doubt the grounds on which it has been claimed that the object a' (say) from the model $\mathscr{M}$ is the same object as a' from $\mathscr{M}^*$. These kinds of facts are not expressible in 4P. They are part of what the proponents of the 'mathematical dissolution' of the hole argument from section 5 would call 'semantic metatheory' (here expressed in English). As such, these kinds of facts are irrelevant to the interpretation of 4P. Let me call this 'the semantic response'. The difference between the metaphysical response and the semantics response illustrates the contrast between counterpart substantivalism (or metric essentialism for that matter) and the 'mathematical dissolutions' discussed in section 5. And I shall now argue that whereas Melia's argument can be defended against the latter, it is not immune to the former.

The semantic response is sensitive to the way in which 4P was formalised, and a more appropriate formalisation of 4P actually jettisons the considerations on which the response relies. For one might insist that a 4P-theorist should be able to to study the relationships between the models of 4P within the language in which 4P is formalised, analogously to a general-relativist studying the relationships between various GR-models using purely mathematical (that is, fundamentally set-theoretic) tools. In that case, one might further assume that our dynamical theory 4P is not, in fact, a self-standing formalisation used by 4P-theorists, but merely a theory of Urelemente (in our case: proxies representing possible particles) within some standard theory of sets that admits Urelemente, such as ZFU.[55]

54 Equivalently, the difference might be cashed out by two different choices of the $\delta_{1,2}$ function.

55 See Potter ([2004]) for discussion.

Let's call the resulting theory $\mathrm{ZFU_{4P}}$. In that case, 'the models of 4P' would be set-theoretic constructions within $\mathrm{ZFU_{4P}}$ just like GR-models are set-theoretic constructions within some background (pure) set theory. We might further assume that the standard of isomorphism between the models of 4P (*qua* sets in $\mathrm{ZFU_{4P}}$) is given by the maps that formalise the model-theoretic isomorphism between these models (*qua* structures that make 4P true). In that case, '$a' = a'$' is now a sentence in the language of $\mathrm{ZFU_{4P}}$ (and, in fact, a theorem of $\mathrm{ZFU_{4P}}$), so one is fully justified in asserting—within the bounds of the formal apparatus at hand—that the object a' in $\mathscr{M}$ is the same object as a' in $\mathscr{M}^*$ (the same holds for a, b, b', etc.).

So, Melia's argument can be defended against the semantic response by an appropriate re-formalisation of 4P. It is not clear, however, whether a similar trick can defend it against the metaphysical response, to which I now turn.

One might insist that the metaphysical response also hinges on a particular formalisation of Melia's theory, and that a different formulation (not necessarily a formal one), which took seriously the possibility of a different black particle being adjacent to a white particle than the actual one, would be clearly indeterministic. There are two ways in which we can interpret this talk of 'taking the indeterminism seriously'. One way is to say that Melia's theory gives us what I called a direct access to its space of possibilites, and this space includes the possible situation where a goes to c and b goes to d, and a distinct possible situation where (the counterpart of) a goes to (the counterpart of) d, etc. Fair enough: but local spacetime theories, including GR, are not like that at all. So we should not consider theories that give such direct access to be informative in the context of our endeavour to interpret GR.

The other way is to 'embed' indeterminism into the formalism, and arrive at it indirectly, through the space of models. That is certainly possible: but imagine what such a formalisation would need to include. Names are certainly not enough, because they are non-logical vocabulary, and so up for reinterpretation. What could potentially do justice to these intuitions is either:

(A) adding names and modal operators, or

(B) quantifying into modal contexts.

The axiom $(4\mathrm{P}_2)$ would then be replaced by $(4\mathrm{P}_2^{\mathrm{A}})$ or $(4\mathrm{P}_2^{\mathrm{B}})$ respectively (according to whether one adopts strategy (A) or strategy (B):

$(4\mathrm{P}_2^{\mathrm{A}})$ [Let a, b, c, d be names for objects of sort σ_2]. Every particle is either a, or b, or c, or d. a and b are B and distinct, and c and d are W and distinct, and a stands in

A to c, and b stands in A to d, and it's possible that: (a stands in A to d, and b stands in A to c).

($4\mathrm{P}_2^{\mathrm{B}}$) There are exactly four particles of sort σ_2: two are B, two are W, no particle is both W and B, one W-particle stands in A to one B-particle but could have stood in A to the other one, and the other W-particle stands in A to the other B-particle but could have stood in A to the first one.

Certainly, there are ways to formalise Melia's theory in many-sorted quantified modal logic along these lines. But I submit that this again bears little relevance to the issue in question for two reasons.

Firstly, GR and other local spacetime theories have no such syntactic resources, and it is difficult to imagine how and why that would ever change.[56] Secondly, the most straightforward semantics for modal operators in the context of specifying possible situations would presumably be something like the Kripke-style possible-worlds approach. And then a model of any such theory would *ipso facto* represent not a single possible world, but the whole space of possible worlds. More generally, such a model may come with a 'built-in' counterpart relation between objects expressed by, say, the same names referring to counterparts across different worlds. And the key idea underpinning the rejection of (Dm2-Worlds) is that this relation might not depend on qualitative properties and relations of these objects. That being said, not only do these theories still conform to my preferred (Dm2-Worlds+), they are also rather peculiar. Putting aside the issues of relating one Kripke-model to another, they presuppose a 'God's point of view' on the space of possibilties and on the counterpart relations between them. Even if this doesn't make them incoherent, it certainly should make us wary of any purported implications they claim to have about our understanding of theories like GR.

Thus, I conclude that Melia's counterexample does not pose a serious threat to the (Dm2)-family of counterpart-theoretic definitions of determinism for local spacetime theories including GR. And as I said at the start, the same can be said of Belot's ([1995], [2018]) toy theories, which are similar in all relevant respects.

7 Conclusion

I have argued that counterpart substantivalism is an attractive position which protects substantivalism against the hole argument. My preferred version of counterpart substantivalism endorses the following principles:

[56] In this respect, I agree with Halvorson and Manchak ([forthcoming], p. 27) who make similar remarks.

1. (S-Man), which secures an uncontroversially substantivalist position (sections 2 and 3),
2. (One World), which secures parsimony and prevents underdetermination (section 4),
3. (Dm2-Worlds+), which can be explicated by any model-based definition from the (Dm2)-family, and which secures determinism of GR in the globally hyperbolic sector (section 6).

This form of counterpart substantivalism, I have argued, is a genuine response to a genuine philosophical problem posed by the hole argument (see section 5), and it is also immune to challenges raised by Belot and Melia (see section 6).

Acknowledgements

For comments on various drafts of this paper and discussions related to its topic, I want to thank Jeremy Butterfield, Neil Dewar, Henrique Gomes, Klaas Landsman, Tushar Menon, Ruward Mulder, James Read, Bryan W. Roberts, Misha Schmalian, James Owen Weatherall, two anonymous referees, and the participants of the Cambridge-LSE Philosophy of Physics Bootcamp. I am especially grateful to Jeremy Butterfield for his invaluable intellectual and personal support, and to Henrique Gomes for sharing his technical and conceptual expertise in GR. I also want to thank Trinity College, Cambridge for generous financial support during most of the time I spent working on this paper.

Faculty of Philosophy
University of Oxford
Oxford, England
franciszek.cudek@philosophy.ox.ac.uk

References

Arntzenius, F. [2008]: 'Gunk, Topology, and Measure', in D. W. Zimmerman (*ed.*), *Oxford Studies in Metaphysics. Volume 4*, Oxford: Oxford University Press, pp. 225–47.

Belot, G. [1995]: 'New Work for Counterpart Theorists: Determinism', *The British Journal for the Philosophy of Science*, **46**, pp. 185–95.

Belot, G. [2018]: 'Fifty Million Elvis Fans Can't be Wrong', *Noûs*, **52**, pp. 946–81.

Bradley, C. and J. O. Weatherall [2022]: 'Mathematical Responses to the Hole Argument: Then and Now', *Philosophy of Science*, **89**, pp. 1223–32.

Butterfield, J. [1987]: 'Substantivalism and Determinism', *International Studies in the Philosophy of Science*, **2**, pp. 10–32.

Butterfield, J. [1989]: 'The Hole Truth', *The British Journal for the Philosophy of Science*, **40**, pp. 1–28.

Choquet-Bruhat, Y. [2009]: *General Relativity and the Einstein's Equations*, Oxford: Oxford University Press.

Dasgupta, S. [2011]: 'Bare Necessities', *Philosophical Perspectives*, **25**, pp. 115–60.

Dasgupta, S. [2015]: 'Inexpressible Ignorance', *The Philosophical Review*, **124**, pp. 441–80.

Dewar, N. [2022]: *Structure and Equivalence*, Cambridge: Cambridge University Press.

Doboszewski, J. [2017]: 'Non-uniquely Extendible Maximal Globally Hyperbolic Spacetimes in Classical General Relativity: A Philosophical Survey', G. Hofer-Szabó and L. Wroński (*eds.*), *Making it Formally Explicit*, Cham: Springer International Publishing, pp. 193–212.

Doboszewski, J. [2019]: 'Relativistic Spacetimes and Definitions of Determinism', *European Journal for Philosophy of Science*, **9**, available at: <https://doi.org/10.1007/s13194-019-0248–6>.

Duerr, P. M. [2019]: 'Fantastic Beasts and Where (not) to Find Them: Local Gravitational Energy and Energy Conservation in General Relativity', *Studies in History and Philosophy of Modern Physics*, **65**, pp. 1–14.

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

Earman, J. and J. D. Norton [1987]: 'What Price Spacetime Substantivalism? The Hole Story', *The British Journal for the Philosophy of Science*, **38**, pp. 515–25.

Fletcher, S. C. [2020]: 'On Representational Capacities, with an Application to General Relativity', *Foundations of Physics*, **50**, pp. 228–49.

Giulini, D. [2007]: 'Mapping-Class Groups of 3-Manifolds', in B. Fauser, J. Tolksdorf and E. Zeidler (*eds.*), *Quantum Gravity. Mathematical Models and Experimental Bounds*, Basel: Birkhäuser Verlag, pp. 161–202.

Gomes, H. [2021]: 'Why Gauge? Conceptual Aspects of Gauge Theories', PhD Thesis. University of Cambridge.

Gomes, H. and J. Butterfield [2023a]: 'The Hole Argument and Beyond: Part I: The Story so Far', *Journal of Physics: Conference Series*, **2533**, available at: <https://doi.org/10.1088/1742–6596/2533/1/012002>.

Gomes, H. and J. Butterfield [2023b]: 'The Hole Argument and Beyond: Part II: Treating Non-isomorphic Spacetimes', *Journal of Physics: Conference Series*, **2533**, available at: <https://doi.org/10.1088/1742–6596/2533/1/012003>.

Gomes, H. and C. Rovelli [unpublished]: 'On the analogies between gravitational and electromagnetic radiative energy', available at: <https://arxiv.org/abs/2303.14064>.

Halvorson, H. and J. Manchak [forthcoming]: 'Closing the Hole Argument', *The British Journal for the Philosophy of Science*, available at: <https://doi.org/10.1086/719193>.

Hamblin, C. [1958]: 'Questions', *Australasian Journal of Philosophy*, **36**, pp. 159–68.

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

Hoefer, C. [1996]: 'The Metaphysics of Space-Time Substantivalism', *The Journal of Philosophy*, **93**, pp. 5–27.

Ladyman, J. and S. Presnell [2020]: 'The Hole Argument in Homotopy Type Theory', *Foundations of Physics*, **50**, pp. 319–29.

Lando, T. and D. Scott [2019]: 'A Calculus of Regions Respecting Both Measure and Topology', *Journal of Philosophical Logic*, **48**, pp. 525–50.

Landsman, K. [2021]: *Foundations of General Relativity: From Einstein to Black Holes*, Nijmegen: Radbound University Press.

Landsman, K. [2023]: 'Reopening the Hole Argument', *Philosophy of Physics*, **1**, available at: <https://doi.org/10.31389/pop.12>.

Lewis, D. K. [1983]: 'New Work for a Theory of Universals', *Australian Journal of Philosophy*, **61**, pp. 343–77.

Lewis, D. K. [1986]: *On the Plurality of Worlds*, Oxford: Basil Blackwell.

Lewis, D. K. [1991]: *Parts of Classes*, Oxford: Blackwell.

Lewis, D. K. [1998]: 'Relevant Implication', *Theoria*, **54**, pp. 161–74.

Luc, J. [2022]: 'Arguments from Scientific Practice in the Debate About the Physical Equivalence of Symmetry-related Models', *Synthese*, **200**, pp. 1–29.

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

Manzano, M. and V. Aranda [2022]: 'Many-Sorted Logic', in E. N. Zalta and U. Nodelman (*eds.*), *The Stanford Encyclopedia of Philosophy*, Winter 2022, Metaphysics Research Lab, Stanford University, available at: <https://plato.stanford.edu/archives/win2022/entries/logic-many–sorted/>.

Maudlin, T. [1988]: 'The Essence of Space-Time', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1988*, pp. 82–91.

Maudlin, T. [1993]: 'Buckets of Water and Waves of Space: Why Spacetime is Probably a Substance', *Philosophy of Science*, **60**, pp. 183–203.

Melia, J. [1999]: 'Holes, Haecceitism and Two Conceptions of Determinism', *The British Journal for the Philosophy of Science*, **50**, pp. 639–64.

Muller, F. [1995]: 'Fixing a Hole', *Foundations of Physics Letters*, **8**, pp. 549–62.

Mundy, B. [1992]: 'Space-Time and Isomorphism', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1992*, pp. 515–27.

North, J. [2021]: *Physics, Structure, and Reality*, Oxford: Oxford University Press.

Norton, J. D., O. Pooley and J. Read [2023]: 'The Hole Argument', in E. N. Zalta and U. Nodelman (*eds.*), *The Stanford Encyclopedia of Philosophy*, Summer 2023, Metaphysics Research Lab, Stanford University, available at: <https://plato.stanford.edu/archives/sum2023/entries/spacetime–holearg/>.

Pooley, O. [2021]: 'The Hole Argument', in E. Knox and A. Wilson (*eds.*), *The Routledge Companion to Philosophy of Physics*, pp. 145–59.

Pooley, O. and J. Read [forthcoming]: 'On the Mathematics and Metaphysics of the Hole Argument', *The British Journal for the Philosophy of Science*, available at: <https://doi.org/10.1086/718274>.

Potter, M. [2004]: *Set Theory and Its Philosophy*, Oxford: Oxford University Press.

Read, J. [2020]: 'Functional Gravitational Energy', *The British Journal for the Philosophy of Science*, **71**, pp. 205–32.

Roberts, B. W. [2020]: 'Regarding 'Leibniz Equivalence'', *Foundations of Physics*, pp. 250–69.

Roeper, P. [1997]: 'Region-Based Topology', *Journal of Philosophical Logic*, **26**, pp. 251–309.

Russell, J. S. [2015]: 'Possible Worlds and the Objective World', *Philosophy and Phenomenological Research*, **90**, pp. 389–422.

Skow, B. [2008]: 'Haecceitism, Anti-Haecceitism and Possible Worlds', *The Philosophical Quarterly*, **58**, pp. 98–107.

Stachel, J. [2014]: 'The Hole Argument and Some Physical and Philosophical Implications', *Living Reviews in Relativity*, available at: <https://doi.org/10.12942/lrr-2014–1>.

Strominger, A. and A. Zhiboedov [2016]: 'Gravitational memory, BMS supertranslations and soft theorems', *Journal of High Energy Physics*, **86**, available at: <https://doi.org/10.1007/JHEP01(2016)086>.

Teitel, T. [2019]: 'Holes in Spacetime: Some Neglected Essentials', *The Journal of Philosophy*, **116**, pp. 353–89.

Teitel, T. [2022]: 'How to Be a Spacetime Substantivalist', *The Journal of Philosophy*, **119**, pp. 233–78.

Wald, R. M. [1984]: *General Relativity*, Chicago, IL: The University of Chicago Press.

Weatherall, J. O. [2018]: 'Regarding the 'Hole Argument'', *The British Journal for the Philosophy of Science*, **69**, pp. 329–50.

Weatherall, J. O. [2020]: 'Some Philosophical Prehistory of the (Earman-Norton) Hole Argument', *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, **70**, pp. 79–87.

Williamson, T. [2016]: 'Modal Science', *Canadian Journal of Philosophy*, **46**, pp. 453–92.

Wilson, M. [1993]: 'There's a Hole and a Bucket, Dear Leibniz', *Midwest Studies in Philosophy*, **18**, pp. 202–41.