ON SOME RECENT ATTEMPTED NON-METAPHYSICAL DISSOLUTIONS OF THE HOLE DILEMMA

CHRISTOPHER ARLEDGE AND ROBERT RYNASIEWICZ

1. Introduction

The hole argument of Earman and Norton (1987) is designed to establish a dilemma: either give up the possibility of determinism in a wide class of spacetime theories, which they believe substantivalists[1] must do, or else embrace a principal called 'Leibniz equivalence' that identifies isomorphic models as representing one and the same physical situation, a principle certainly necessary for relationalism but also thought by many to be sufficient for relationalism. It would seem, then, that any resolution of the hole dilemma must engage in major metaphysical maneuvering, as is typical of the enormous literature on the hole argument. Recently, however, a number of authors have claimed metaphysical excursions can be avoided and the dilemma dissolved if sufficient attention is paid to how the mathematics used in the hole argument, and spacetime theories in general, applies to the world[2] (Weatherall 2018, Curiel 2018, Fletcher 2019). We think these attempts are unsuccessful and intend to explain why, but not because we believe resolution of the hole dilemma calls for metaphysical tinkering. For at the end we will suggest a resolution that is non-metaphysical, at least to the extent that basic facts about reference are not metaphysical.

2. Details of the Hole Argument

For those steeped in the literature, nothing is more painfully boring than yet another rehearsal of the hole argument. Nonetheless, for readers unfamiliar with it and for the sake of completeness, we will outline the argument. Veterans can skip ahead, although we will register a few new points that have been by and large overlooked.

2.1. **Einstein's Hole Argument.**

The argument of Earman and Norton (1987) descends from an argument Einstein had hit upon around 1912 in the course of discovering his field equations, an argument he dubbed the *Lochbetrachtung*. Consider a solution to the field equations which is such that the stress-energy tensor vanishes in a[n open] neighborhood [with compact closure]. (This neighborhood is the so-called "hole".) By leaving everything

[1]To be more precise, *manifold substantivalists*, i.e., those who take spacetime to be [represented by] the bare manifold in models/solutions of general relativity (and any other spacetime theory that involves only dynamical fields and no fixed background).
[2]For some earlier attempts at a non-metaphysical dissolution see (Leeds 1995) and (Mundy 1992). But also see (Rynasiewicz 1996) for a critique of these attempts.

outside the hole fixed but by smoothly distorting the manifold inside the hole, one arrives, by general covariance, at second solution, which, though distinct from the first, is isomorphic to it. The conclusion Einstein drew is that general covariance is at odds with the "law of causality" [*Kausalgesetz*], a Machian principle maintaining that the spacetime metric (and thus inertial structure) is determined by specification of the mass-energy distribution on the spacetime manifold. Einstein initially reacted by rejecting the demand of general covariance. Only after a couple of years did he find a way to reinstate the requirement of general covariance while retaining his Kausalgesetz. The key was to identify isomorphic models as corresponding to one and the same physical situation, i.e., to adopt what Earman and Norton call 'Leibniz equivalence'.

Earman and Norton generalize Einstein's argument in such a way that (i) the "hole" need no longer be empty of sources[3] and that (ii) just the possibility of determinism is used in place of Einstein's Kausalgesetz. But before giving the argument, we need the notion of a *pushforward* or *carry along* of a mathematical object, as well as the reciprocal notion of a *pullback*.

2.2. **Pushforwards and Pullbacks.** [4] Let M, N be equipollent sets and $\pi : M \to N$ a bijection. Furthermore, let F_M and F_N be the sets of real valued functions on M and N, respectively. There is a natural extension of π to $\pi^* : F_M \to F_N$ given by $\pi^*(f) = f \circ \pi^{-1}$ for all $f \in F_M$. Dropping the parentheses, $\pi^* f$ is referred to as the *pushforward* or *carry along* of f. In the other direction, there is an extension of the inverse of π to $\pi_* : F_N \to F_M$ such that $\pi_*(g) = g \circ \pi$ for all $g \in F_N$. Again, dropping the parentheses, $\pi_* g$ is called the *pullback* of g.

If M and N are differential manifolds (perhaps of different dimensions), then F_M and F_N are the sets of scalar fields on M and N, respectively. Moreover, the mapping π also induces a mapping[5] $\pi^* : T_p \to T_{\pi(p)}$, where T_p is the tangent space at point $p \in M$, according to the recipe $(\pi^* v)(f) = v(f \circ \pi)$. Let ${T_p}^*$ be the dual of T_p.[6] Let $v \in T_p$ and $w \in {T_{\pi(p)}}^*$. Then π induces a further mapping $\pi_* : {T_{\pi(p)}}^* \to {T_p}^*$ according to the formula $(\pi_* w)(v) = w(\pi^* v)$. Both π^* and π_* can be extended straightforwardly to arbitrary contravariant and covariant tensors, respectively. If π is furthermore a diffeomorphism, then π^*, as well as π_*, can be extended to tensors of arbitrary types (m, n) by

$$(\pi^* A)(w_1, \ldots, w_m; v_1, \ldots, v_n) = A(\pi_* w_1, \ldots, \pi_* w_m; [\pi^{-1}]^* v_1, \ldots, [\pi^{-1}]^* v_n),$$

where A is a tensor of type (m, n), the w_i's covectors and the v_i's contravectors. We can similarly define $\pi_* A$.

[3] I.e., The stress-energy tensor need not vanish

[4] We follow the notation of mathematicians, who systematically write ϕ^* for a pullback and ϕ_* for a pushforward. It is noteworthy that Wald (1984) reverses the two. Whether there is a tradition outside of philosophy physics following this practice, we do not know.

[5] Technically, we should give this mapping a different name to distinguish it from the pushforward of scalar fields. But since there are all sorts of geometric objects on a manifold, we would run out of notation if we tried to give each pushforward its own name.

[6] ${T_p}^*$ is the set of all linear functions from T_p to $\mathbb{R}$.

The notion of the pushforward (pullback) of an object is sufficiently general that it can be defined for objects of almost any type, including non-tensorial geometric objects such as the covariant derivative. This is not peculiar to differential geometry, but applies to other categories of mathematical objects, as well, even models of arbitrary type. E.g., let $\mathcal{M} := \langle M, O_1, \ldots, O_n \rangle$ Then $\pi^*\mathcal{M} := \langle M, \pi^*O_1, \ldots, \pi^*O_n \rangle$.

2.3. **Active vs. Passive Views of a Leibniz Shift.** We specialize to the case of the metric tensor of a Lorentzian spacetime. The treatment generalizes to tensors of arbitrary type.

The goal here in relating the active and passive viewpoints is as follows. Let $d : M \to M$ be a diffeomorphism, $\{x^\mu\}$ a coordinate chart around $p \in M$, and $(d_*g)_{\mu\nu}$ the components in $\{x^\mu\}$ of the pushforward of g by d. We're looking for a coordinate chart $\{x'^\mu\}$ such that, at the point p, the equation

$$(d_*g_{\mu\nu})(p) = g'_{\mu\nu}(p) \tag{1}$$

holds, where $g'_{\mu\nu}(p)$ are the components of g in the chart $\{x'^\mu\}$. Then an active transformation, which uses d and its pushforwards, can be be faithfully mirrored by a passive transformation, according to which we view the same object (in this instance g) in different coordinates. We then compute the pushforward of $g(p)$ at the very same point p, thus "moving the metric" but not the base space M.

The key to the solution is that we want the coordinates of $d(p)$ in $\{x'^\mu\}$ to be the same as the coordinates of p in $\{x^\mu\}$, i.e.,

$$x'^\mu(d(p)) = x^\mu(p).$$

This is to say that the mappings x^μ, x'^μ, and d stand in the relation

$$x^\mu = x'^\mu \circ d.$$

Composing each side with d^{-1},

$$x'^\mu = x^\mu \circ d^{-1}.$$

This yields the primed coordinate chart for which equation (1) holds.

Note that if $d^{-1}(p)$ lies in the domain of the chart $\{x^\mu\}$, we then have

$$(d_*g_{\mu\nu})(p) = g_{\mu\nu}(d^{-1}(p)).$$

This conforms with the expectation that, in terms of intrinsic quantities, $d_*g(p) = g(d^{-1}(p))$.

Formulated in terms of fiber bundles, the difference between the active and the passive is that, in the first instance, d induces a mapping of the bundle space that takes the fiber over p of the base manifold to the fiber over its image to its image $d(p)$. In the second instance, each fiber is non-trivially mapped to itself. The result, in both cases, is to map one section to another, the mapping being the same in both cases. Thus, although conceptually different, the outcome of the two is the same and Wald is correct in claiming that in this sense they are equivalent (1984, 439).

2.4. **The Earman-Norton Argument.** Rather than to deal with arbitrary spacetime theories, as do Earman and Norton, let us take the particular case of general relativity from which we can generalize. The solutions to Einstein's field equations, a.k.a. the models of general relativity, have the form $\mathcal{M} := \langle M, g, T \rangle$ where M is a 4-dimensional differential manifold, g a Lorentzian metric on M, and T the stress-energy tensor.[7] Now consider a diffeomorphism $d : M \to M$ that is not an automorphism of $\mathcal{M}$. Call such a mapping a 'Leibniz shift'.

One key fact behind the hole argument is that if $\mathcal{M}$ is a model of general relativity, then any Leibniz shifted model $d^*\mathcal{M}$ is a model of the theory as well. Earman and Norton go to some length to establish this by requiring that the models be fleshed out with additional tensors, defined from g and T, so as to arrive at tensors that are everywhere 0.[8] They then take coordinate representations of these tensors in a chart and examine the coordinate values under the Leibniz shift of that chart. However, all this is unnecessary. That the Leibniz shift of a model is also a model is not a deep fact. For d is an isomorphism of $\mathcal{M}$ onto $d^*\mathcal{M}$, and the space of solutions to a set of equations is closed under isomorphism just as the class of models of any theory in an abstract logic is closed under isomorphism.

Another key fact is that d can be selected such that it is the identity everywhere except on an arbitrarily selected open subset $H \subset M$ with compact closure, and smoothly connects with the identity map on the boundary of H. This suffices to show that the values of the field quantities everywhere outside of H fail to determine those in H.

Now suppose it is held that distinct models represent distinct physical situations, a principle elsewhere know as *model literalism* (Rynasiewicz 1994). Earman and Norton presume that substantivalism is committed to model literalism based on an "acid test" Leibniz proposed to Samuel Clarke in their famous correspondence. If absolute space is something real, then God could have created a different world in which all bodies are displaced a fixed distance in a fixed direction. Generalizing, the acid test is then that substantivalism is wedded to the idea that a Leibniz shifted model represents a state of affairs distinct from that which the original model represents. This leads to a dilemma for the substantivalist. Take any model $\mathcal{M}$ of general relativity that can be foliated into a one-parameter set of global time slices. Select one of these time slices S as the present, and select a temporal orientation in order to distinguish past from future. Place the hole H in the future of S. Then $\mathcal{M}$ and $d^*\mathcal{M}$ agree on all facts up to and including S but fail to agree thereafter. Hence, if they represent distinct situations, as the substantivalist allegedly must hold, then general relativity is a wildly indeterministic theory. Earman and Norton brand this special form of indeterminism as 'radical local indeterminism'. The resulting dilemma for substantivalists is that they must either (a) embrace radical local indeterminism for general relativity and any other spacetime theory that has no fixed background geometry,

[7]To be more perspicuous, we should write, using Wald's (1984) abstract index notation, $\langle M, g_{ab}, T^{ab} \rangle$ as the generic form of a solution to the Einstein field equations. However, the abstract indices are merely heuristic, and to keep them throughout would prove tedious.

[8]In this case form the Einstein tensor G as well as the tensor $G - T$, which, in light of the field equations, is everywhere zero.

or else (b) give up substantivalism by embracing a principle Earman and Norton call 'Leibniz equivalence': isomorphic models represent one and the same physical situation.

A word of caution. Earman and Norton do not insist that our spacetime theories *must* be deterministic. Rather, at stake is the very possibility of determinism, and, quite sensibly, they suggest that we should not allow metaphysics to commit us to indeterminism (or determinism) *a priori*.

2.5. **Diffeomorphisms Are Not Needed.** As we have seen, it is unnecessary to define vanishing tensors for the hole argument to go through. It is also unnecessary that the Leibniz shifting be done via a diffeomorphism. Any permutation will do as long as we use the permutation to pushforward the atlas of the manifold.

We can think of a smooth manifold M as an ordered pair $\langle S, \mathcal{T}, \mathcal{A}\rangle$, where S is a bare point set, $\mathcal{T}$ the topology, and $\mathcal{A}$ a complete C^∞ atlas on S. In general, a C^∞ atlas $\mathcal{A}$ on M is a collection of mappings φ_α, called *coordinate charts*, such that (a) for each $p \in M$ there exists a chart $\varphi_\alpha \in \mathcal{A}$ with $p \in \mathrm{dom}\ \varphi_\alpha$, i.e., $\{\mathrm{dom}\ \varphi_\alpha\}$ covers S, (b) each φ_α is a bijection with some open $U \subseteq \mathbb{R}^n$, and (c) any two coordinate charts $\varphi_\alpha, \varphi_\beta \in \mathcal{A}$ are C^∞ related in the sense that if $\mathrm{dom}\ \varphi_\alpha \cap \mathrm{dom}\ \varphi_\beta \neq \emptyset$, then $\varphi_\alpha \circ {\varphi_\beta}^{-1}$ and $\varphi_\beta \circ {\varphi_\alpha}^{-}1$ are C^∞ maps. Let $\pi : S \to S$ be any permutation on S, and let $\pi^*\mathcal{T}$ be the pushforward of $\mathcal{T}$, i.e., $\pi^*\mathcal{T} = \{\pi[U] \mid U \in \mathcal{T}\}$. Hence $\langle S, \pi^*\mathcal{T}\rangle$ is homeomorphic to $\langle S, \mathcal{T}\rangle$, π being a homeomorphism. Finally, the pushforward $\pi^*\mathcal{A}$ of the atlas $\mathcal{A}$ is $\{\varphi \circ \pi^{-1} \mid \varphi \in \mathcal{A}\}$. It follows that $\langle S, \pi^*\mathcal{T}, \pi^*\mathcal{A}\rangle$ is diffeomorphic to $\langle S, \mathcal{T}, \mathcal{A}\rangle$, again π being a diffeomorphism. Thus, the hole argument can be run with any permutation of the underlying point set that is the identity map on S outside the hole. That mere permutations of the underlying point set suffice suggests that the hole argument is a special case of a more general consideration. We will return to this in the penultimate section.

2.6. **Metaphysical Responses.** Substantivalists of different stripes have marshalled an array of responses to the hole argument. While the aim of this paper is not to deal with metaphysical replies, we will give a (non-exhaustive) outline of some of the most prominent responses.

One response attempts to add some sort of essentialism to the metric of spacetime. Tim Maudlin (1989, 1990) argues that a physical spacetime *cannot* exist without a set of metrical relations. Since these metrical relations are essential to the spacetime, spacetimes differing only with regards to metrical relations are necessarily non-identical. In effect, the spacetime points have their location in the manifold essentially fixed by the metrical relations. So, for some point $p \in H$, if p is assigned a new metrical value by some diffeomorphism d, then $d * (p) \in H$ is *not* the same point as p and therefore lacks $p's$ essential metrical properties. So any Leibniz shifted model $d * \mathcal{M}$ that does not leave the essential metrical relations intact does not represent a physical state of affairs at all. Hence, the hole argument is obviated, since this is exactly the situation that the hole argument describes.

Another metaphysically motivated response to Earman and Norton is given by Jeremy Butterfield (1989). Using David Lewis' counterpart theory, Butterfield denies the transworld identity of spacetime points. [9] So, if two models contain the same spacetime points, then only one model is a representation of a possible world. This view resolves the hole argument, since a Leibniz shifted model $d * \mathcal{M}$ will contain the same spacetime points as the model $\mathcal{M}$. Considering this within the context of counterpart theory, only one of the models represents a possible physical state of affairs and so no indeterminism. For objections to the modal responses of Maudlin and Butterfield see Earman (1989), Brighouse (1994) and Rynasiewicz (1994).

The previous responses attempt to deny Leibniz equivalence, and yet protect spacetime substantivalism from the hole argument. A set of more sophisticated (and slightly less metaphysical) responses attempt to argue that the substantivalist can accept Leibniz equivalence and yet remain true to substantivalism. Carl Hoefer (1996) argues that the representation of spacetime is given entirely by the metric tensor, and not the manifold + metric (which is Maudlin's view). The role of the manifold is to give continuity and topological properties to the spacetime that is expressed by the metric field, but we should take the metric tensor to be representing the physically salient aspects of spacetime. He calls this view 'metric field substantivalism'. Another variant of this sophisticated substantivalism is advocated by Carolyn Brighouse (1994). She argues that our individuation of spacetime points is done relative to qualitative similarity (in line with the identity of indiscernibles). If this is the standard then isomorphic models are equivalent and this judgment can be rendered independently of a view on the existence of spacetime points. Since this amounts to an admission of Leibniz equivalence, the hole argument is avoided. Other forms of sophisticated substantivalism have been defended by Maidens (1992) and Dasgupta (2011). For some responses to sophisticated substantivalism see Belot (1995) and Brighouse (2018).

3. On Weatherall

Weatherall (2018) contends that excursions into metaphysics are unnecessary in order to come to grips with the hole argument. As he puts it, "if one adopts the standard formalism of relativity theory, on which one represents spacetime with a Lorentzian manifold, the hole argument does not force one to confront a metaphysical dilemma" (p. 344). Instead, he claims, the force of the argument vanishes if one considers carefully how the relevant mathematics applies to the world. His dissolution of the hole argument proceeds by adopting certain theses about applied mathematics.

(1) The equivalence of models of general relativity (or any other theory) is a purely mathematical issue.
(2) Isomorphism is the mathematical standard of equivalence.
(3) Equivalent mathematical models have the same representational capacities.
(4) Isomorphic models can be compared only *relative* to the choice of a particular isomorphism.

The last of these is absolutely essential for Weatherall's purposes.

[9] Lewis' counterpart theory asserts that no object can occupy more than one possible world. Qualitative duplicates across possible worlds are non-identical counterparts. For more cf. Lewis (1968, 1973).

3.1. **Application to the Hole Argument.** The hole argument then falters as follows. Let $\langle M, g\rangle$ be a Lorentzian spacetime[10] and d a Leibniz shift. Thus, $\langle M, g\rangle \neq \langle M, d^* g\rangle$ since $d^* g \neq g$. However, d is an isometry from $\langle M, g\rangle$ to $\langle M, d^* g\rangle$. Isometry in this instance is isomorphism, and thus the two models are equivalent and have the same representational capacities.

To conclude immediately from this that they represent the same physical situation is clearly a nonsequitur.[11] And although Weatherall sometimes appears to draw this inference, (4) above is needed for the argument for his thesis to go through. Let $p \in M$ be such that $g(p) \neq d^* g(p)$. (In the hole argument, the point p lies in the "hole".) It thus appears that the two models represent different physical states of affairs. But according to (4), this comparison is implicitly relative to the identity map 1_M on M. Yet 1_M is not an isomorphism from $\langle M, g\rangle$ to $\langle M, d^* g\rangle$, and therefore, by (4), is not a candidate for comparing the two models. Instead, one should use the isomorphism d. Relative to d, the point corresponding to p is $d(p)$, and it is trivially the case that $g(p) = d^* g(d(p))$ Since this holds for all points (inside and outside the hole), it follows that $\langle M, d^* g\rangle$ represents the same physical situation as $\langle M, g\rangle$. So, radical local indeterminism is avoided.

3.2. **Illustration with the Additive Group of Integers.** In order to make the template of the argument clearer, Weatherall asks us, as a warm-up exercise, to consider the additive group of integers $\langle \mathbb{Z}, +, 0\rangle$, where 0 is selected as a distinguished element since it serves as the additive identity element. Now, because $\langle \mathbb{Z}, +, 0\rangle$ has no automorphisms other than the identity map on $\mathbb{Z}$, any non-trivial permutation generates a distinct but isomorphic model. E.g., let $f : \mathbb{Z} \to \mathbb{Z}$ be given by the rule $f(n) = n + 1$. This induces the structure $\langle \mathbb{Z}, f_*+, f_*0\rangle$, where $f_*0 = 1$ and $(f_*+)(n, m) = n + m - 1$. For ease of notation, write 0_* for f^*0 and $+_*$ for $f^*(+)$, The map f is a group isomorphism since $f(0) = \overline{0}$ and $f(n + m) = f(n)\overline{+}f(m)$. Thus, according to Weatherall's (2), the two structures are equivalent. In virtue of that, according to Weatherall, there is a sense in which both are just the group of integers under addition. But then, asks Weatherall, just what is the identity element of this group? According to the one model, it is the element 0; according to the other, it is $\overline{0} = 1$. So, it is alleged, there appears to be an ambiguity. Weatherall has two responses, one syntactic, the other model-theoretic.

The syntactic response is that we have a case of a trivial notational variation "where we assign the symbols '$\overline{0}$' or '1' to the identity in $\langle \mathbb{Z}, \overline{+}, \overline{0}\rangle$ and the symbol '0' to the identity in $\langle \mathbb{Z}, +, 0\rangle$" (p. 333).

The model-theoretic response brings to bear assumption (4) above, viz., that isomorphic models can be compared only *relative* to the choice of a particular isomorphism. The ambiguity arises from choosing different mappings. Given the same domain of discourse $\mathbb{Z}$

[10]That is to say, g is a Lorentzian metric on M. Weatherall drops the stress-energy tensor T from the models of general relativity on the grounds that T is definable from g using the field equations. We prefer taking the space of all triples $\langle M, g, T\rangle$ and then taking the models of general relativity to be those that satisfy the field equations. However, nothing of substance in evaluating Weatherall's position hinges on the difference.

[11]Of the form: possibly P, therefore P.

in the two models, one is tempted to use the identity map $1_{\mathbb{Z}}$. This might be appropriate for $\mathbb{Z}$ *qua* set, says Weatherall. But *qua* group, the two models should be compared by the group isomorphism f.[12] Relative to f, the integer corresponding to the identity element 0 is $f(0) = 1$. This, according to Weatherall, resolves the ambiguity.

3.3. **Critique of Weatherall.** What are we to make of this? Let's take the warm-up example first.

3.3.1. *Critique Regarding the Additive Group of Integers.* It's pretty clear that the philosophy of math, or of applied math, espoused is heavily influenced by category theory and mathematical structuralism. This is not the occasion to delve into a full scale discussion of these. Suffice it to say that Weatherall's assumption (4) is not a standard part of either, but a tenet of his philosophy of applied math suggested to him by these and rendered evident from the example of the additive group of integers. But what Weatherall took to be uncontroversial is in fact almost the same consideration one of us has used in class as a test case for structuralism.

To show that the situation is not as clear as he thinks, let us move to a more minimal characterization of groups. It's well known that a first-order axiomatization of group theory requires but one extralogical symbol, viz., a two-place function symbol for group composition. The models thus have the simpler form $\langle G, \circ \rangle$. The axioms assert the existence of an element with the properties of the identity element, and it is then provable that there is a unique element having those properties. That is to say that the identity element is definable in terms of group composition from the axioms. So, it is unnecessary to introduce an individual constant symbol for it. This is not to say that the (pseudo)worry over what element is really the additive identity can no longer be formulated, albeit much less directly. But moving to the minimal setting allows us to focus on the addition function and its pushforward under f. (You'll recall that in the example $f(p) = p + 1$.) We have the models $\langle \mathbb{Z}, + \rangle$ and, using notation initially introduced, $\langle \mathbb{Z}, f_*+ \rangle$. We can ask whether $+$ and f_*+ are one and the same operation on $\mathbb{Z}$. The answer has to be flat out 'no' in light of the fact that f is not an automorphism of $\langle \mathbb{Z}, + \rangle$. Thus, it is inappropriate to say that f_*+ is simply a notational variant on $+$. If that were the case, the two models would be one and the same, which they are not. The bottom line is that the operation $n + m$ on $\mathbb{Z}$ is distinct from the operation $n + m - 1$. There is, of course, the skeptical worry whether '$+$' really refers to $+$ and not to f_*+.[13] But whatever '$+$' refers to, 'f_*+' refers to something else.

[12]For some reason Weatherall insists on a distinction between the map $f : \mathbb{Z} \to \mathbb{Z}$ *qua* set bijection and f *qua* group isomorphism, which he writes as $\overline{f} : \langle \mathbb{Z}, +, 0 \rangle \to \langle \mathbb{Z}, \overline{+}, \overline{0} \rangle$. I suspect that Weatherall has in mind that, in the category of sets, the arrow from the object $\mathbb{Z}$ to itself corresponding to f is distinct from the arrow in the category of groups that goes from $\langle \mathbb{Z}, +, 0 \rangle$ to $\langle \mathbb{Z}, \overline{+}, \overline{0} \rangle$. But in the current context, the language is not that of category theory, but of some standard set theory (say ZF or ZFC) in which we have direct access to the interior of the objects related by arrows. And this leads to oddities in putting the tilde over f. For example, what is the domain or codomain of $\overline{f}$?

[13]This is not a worry *we* have.

3.3.2. *Critique of the Dissolution of the Hole Argument.* Let us first comment on Weatherall's first three theses above before turning to the crucial fourth. Claim (2), that isomorphism is the mathematical standard of equivalence, seems plausible enough, although one needs to be wary. E.g., there are two equivalent ways of formulating the theory of Boolean algebras. One characterizes them relationally as special sorts of partial orderings meeting additional conditions (orthocomplemented lattices). The models have the form $\langle B, \leq \rangle$. The other presents them algebraically so that the models have the form $\langle B, +, \cdot, ^{-}, 0, 1 \rangle$. There can be no question of isomorphism between models of the first type and models of the second even though for each model of one type there is an equivalent model of the other type in the sense that the order relation is definable from the algebraic operations and *vice versa*. This suggests that being equivalent is a bit broader than being isomorphic.

Weatherall's thesis (3), that equivalent mathematical models have the same representational capacities, sounds safe enough, especially *if* we substitute 'isomorphic' for 'equivalent'. But, as we have just seen, equivalence is plausibly broader. A worry is that sameness of representational capacity might be the standard for mathematical equivalence, assuming we had an independent handle on 'representational'. Or, it may be that the handle on sameness of representational capacity derives from mathematical equivalence, in either case rendering (3) uninformative.

In any event, the motivation for (1), that the equivalence of models of general relativity (or any other theory) is a purely mathematical issue, is unclear since there are a number of *prima facie* distinct senses of equivalence used in connection with physical theory, e.g., physical equivalence, empirical or observational equivalence, etc., that are not necessarily a matter of model isomorphism. So, here we need to be especially clear on the sense of 'equivalence of models' prior to claiming that mathematics provides the standard in the form of isomorphism. It may be that Weatherall has in mind the tight standard of *mathematical* equivalence for he writes, "isometry provides the only sense in which ...spacetimes are mathematically equivalent' (p. 337). But then he goes on to claim that the existence of an isometry "provides the *only* sense in which ... spacetimes are *empirically* equivalent" (*Ibid*, emphasis ours), which is dubious. General relativity has models that are arguably observationally equivalent yet differ in global topology. (See (Glymour 1977), (Malament 1977) and (Manchak 2009).[14])

Nevertheless, we should wonder how a criterion dictated by mathematics alone is capable of blocking the hole argument, especially in light of what's said in footnote 9. There we find:

> I take mathematical models to be devices that we use for representational purposes, not objects that stand in some context- or use-independent relation to physical situations that they represent. (p. 332)

We couldn't agree more. But this presents a puzzle. If the way that physicists use mathematical models is context dependent, then we might expect that the physical equivalence of models is also context dependent. If so, how could the equivalence of models be a purely mathematical issue?

[14]This same point is made in an unpublished paper by Roberts (2014).

The crux of Weatherall's argument though is assumption (4), that isomorphic models can be compared only *relative* to the choice of a particular isomorphism. By 'comparison', Weatherall has in mind (at least) point by point comparison. The choice of an isomorphism $d : M \to M$ is supposed to tell us which points to compare with which: corresponding to p is $d(p)$. These are to be treated as if they were the same point. In fact, well after setting up the hole argument by Leibniz shifting a model by d and arguing that the hole argument is blocked by relativizing comparisons of isomorphic models to a particular isomorphism, Weatherall announces he takes his argument to involve a *passive* transformation and not an active map. Now there are a number of points to be made. in response to the preceding, some concerning the active-passive distinction, some concerning the cogency of the need to relativize any comparison to a choice of isomorphism. Let's take these in order.

According to Weatherall,

> "The distinction between passive and active diffeomorphisms . . . is supposed to arise in the context of the action of an induced pushforward map on some field. The passive interpretation of such a map is supposed to correspond to a (mere) re-labelling of of points, effectively leaving the fields unchanged." (p.339)

As we saw in §2.3, points indeed are "relabeled" but we do not have a *mere* relabeling. The fields are *not* left unchanged. That would be just a matter of covariance. Rather the introduction of a new chart facilitates the definition of new, boosted fields. Weatherall goes on to write, "The active interpretation . . . involves keeping the points unchanged (i.e., by somehow fixing the labels of the points, perhaps with a local coordinate system) and assigning fields differently to each point" (*Ibid.*) This a bit misleading. For one, no "fixing of labels" in the form of the introduction of coordinate charts was necessary. For another, the pushforward of a geometric object at point p is defined, not at p, but at $d(p)$. In this sense, it's untrue that the points remain "unchanged".

Now, concerning the cogency of the need to relativize any comparison to a choice of isomorphism, there's no doubt that if we have a pair of isomorphic models $\langle M, g, T\rangle$ and $\langle M', g', T'\rangle$ such that M and M' are disjoint, then comparisons can be made only relative to some bijection $f : M \to M'$, and if f is an isomorphism, then they have the same representational capacity. Indeed, f can be thought of as providing a criterion of trans-model identity. This is reminiscent of doing counterpart theory on David Lewis's version of possible world metaphysics, which holds that an individual can exist in no more than one possible world. Weatherall's insistence that, when $M = M'$, we still need a mapping relative to which comparisons are made is tantamount to insisting that we must do counterpart theory with mathematical models having the same point set.[15] But why is this necessary when we already have a perfectly good criterion of trans-model identity, to wit, numerical identity? Our judgment that, under the diffeomorphism $d : M \to M$, $d_*\mathcal{M} \neq \mathcal{M}$ presupposes the use of the numerical identity of the manifold points. For if we say, for all $p \in M$, that $d(p)$ is the same as p, then we end up with numerically the same model. Another way of putting this is that the very distinction between an automorphism and a

[15]Only without the restriction that a point can exists in no more than one spacetime model.

Leibniz shift presupposes that we use numerical identity in comparing the original with its pushforward. If we used the diffeomorphism as a criterion for point identity across models, we could never produce a Leibniz shift to begin with. Now, Weatherall might take that as another way of blocking the hole argument. But it would make a complete mess of mathematics and force us to conclude that every permutation of the domain of discourse of a model is a symmetry of the model.

Moreover, whatever plausibility (4) might have had in the context of the hole argument depends crucially on the premise that undressed spacetime points cannot be directly detected. Weatherall's philosophy of applied mathematics, if cogent, should apply to cases in which elements of the domain of discourse represent readily observable items. Take the case of a relational structure $\mathcal{M} = \langle M, R \rangle$, where $R \subseteq M \times M$. Let $M = \{1, 2\}$ and let R be $<$. Suppose furthermore that $\mathcal{M}$ represents the situation in which Thing 1 is smaller than Thing 2, where Thing 1 and Thing 2 are as in the Dr. Seuss story—readily observable. Now let $\pi : M \to M$ interchange 1 and 2. Then, we should all agree the Leibniz shifted model represents the situation in which Thing 2 is smaller than Thing 1. Yet, according to Weatherall's (4), the comparison of the two structures must be done *relative to* the isomorphism π, yielding the verdict that they represent the same situation.

The bottom line is that assumption (4), depending on how it is implemented, leads either to internal inconsistency, or else to triviality. Weatherall tells us,

> "When we say that (M, g) and (M, d_*g) are isometric spacetimes, and thus that they have all the same invariant, observable structure, we are comparing them relative to d. Indeed, we must because, as in the previous example, there is no sense in which 1_M either is or gives rise to an isometry." (pp. 336).[16]

We take it that what Weatherall means is that when 1_M is not an isometry of (M, g) onto (M, d_*g). He goes on, "In other words, relative to 1_M, (M, g) and (M, d_*g) are not equivalent, physically or otherwise." (Ibid).

In using isomorphisms as standards of comparison,if we use a single isomorphism for the entire equivalence class of Leibniz-shifted models, then the use of (4) becomes internally inconsistent. Suppose that $d \neq d \circ d$. Then, Weatherall will tell us that because d is an isomorphism from (M, g) to (M, d_*g), that (M, g) is physically equivalent to (M, d_*g). Furthermore, since d is an isomorphism from (M, d_*g) on (M, d_*d_*g) that (M, d_*g) is physically equivalent to (M, d_*d_*g). But, since $d \neq d \circ d$, d is not an isomorphism of (M, g) onto (M, d_*d_*g), violating the transitivity of equivalence. Thus, using a single isomorphism leads to internal inconsistency.

If, however, Weatherall maintains that we should use different isomorphisms as standards of comparison for different pairs of isomorphic models, the application of (4) becomes trivial in that it says nothing more than that isomorphic models represent the same physical situation, in other words, that we should use Leibniz equivalence. But, this just begs the question, since Weatherall assumes Leibniz equivalence to argue for Leibniz equivalence.

[16]We have altered the notation to conform with ours.

4. Other Actors

At least one author consciously follows Weatherall (Fletcher 2019). Another says he is quite sympathetic to Weatherall's approach (Curiel 2018). and offers up a non-metaphysical treatment of the hole argument. We'd best take a look at these, one at a time.

4.1. **Fletcher.** In his introduction, Fletcher (2019) writes, "I agree with Weatherall's response to the hole argument" (§1). For sure he accepts Weatherall's thesis (3) that mathematically equivalent models have the same representational capacities. But we are a bit puzzled, because when he turns to discuss the hole argument he follows a different course. He does not take it that isomorphic mathematical models must represent one and the same physical situation. "Yet", he says, "if this represented state of affairs, is not unique, then the problem of indeterminism seems to rise again, forcing one to confront a metaphysical dilemma anew" (§5.2). He goes on to say "the error implicit in this concern is the assumption that Lorentzian manifolds, as mathematical models, represent *all* properties of a physical relativistic spacetime. . . " (*Ibid*). As he argues earlier, "some properties are abstracted away – such as the representation of units or additional structure, e.g., additional fields . . . " (*Ibid*). But, how one does this to avoid radical local indeterminism is a bit of a mystery to us. A hint of how this might work, though, might be gleaned from another example in which isomorphic models represent distinct physical situations.

Consider the case of a single point mass in Minkowski spacetime, that follows an inertial trajectory up until some time t, when the particle swerves and begins to accelerate. The pertinent question to ask is whether this is a genuine case of indeterminism, in other words could the swerve have taken place at a different time or in a different direction. Fletcher proceeds to deal with this *via* the following steps. First is to describe the particle, which can be described by the triple $(\mathbb{R}^4, \eta, \gamma)$ where η is the Minkowski metric and $\gamma : \mathbb{R} \to \mathbb{R}^4$ is the particle worldline. Second, let T be a non-trivial time translation in Minkwoski spacetime by a fixed time.[17] Third, one must select some isomorphism $d : \mathbb{R}^4 \to \mathbb{R}^4$ as a standard of comparison for evaluating relations between models. Fletcher asks us to consider the model $(M, \eta, (T \circ d)[\gamma])$ and asks whether if by pulling back by d^{-1} we arrive at $(\mathbb{R}^4, \eta, \gamma)$. He concludes that, unless $(d^{-1} \circ T \circ d) = T$, then $(\mathbb{R}^4, \eta, (T \circ d)[\gamma])$ "represents" a time translation relative to d.[18] In contrast, if $(T \circ d)$ is the standard of comparison, then "it would not represent any difference at all." (Fletcher 2019, §5.2). He claims that this leads to the conclusion that "the swerve models do not represent exactly one state of affairs..." (§5.2), that is that the swerve theory is indeterministic. This is not just relative to d, but categorically so.[19]

One way to arrive at this conclusion is with the suppressed premise that if there is a difference relative to some isomorphism d, then $(\mathbb{R}^4, \eta, \gamma)$ and $(\mathbb{R}^4, \eta, T[\gamma])$ represent distinct

[17] Fletcher uses both a time translation T and a rotation R of Minkowski space, but for simplicity we will restrict our attention to the time translation.

[18] By "represents" we take it that he means "constitutes" and not represents in the sense in which a model represents a physical situation.

[19] We are not entirely sure what it would mean to say that a theory is indeterministic relative to a choice of isomorphism.

physical situations. But in the context of the hole argument such a standard of comparison is always available. Take for example the identity map. One then wonders why it is not the case in the hole argument that one doesn't similarly conclude that Leibniz-shifted models represent distinct situations. If this is not the missing premise, then we are at a loss as to why we should take time translations relative to some isomorphism other than the Leibniz shift itself.

4.2. **Curiel.** Curiel's (2018) attempted dissolution of the hole argument rests on a distinction between applying a diffeomorphism to the manifold cum metric and applying it to the metric field alone. Both he takes it are mathematically well-defined,. But, as Curiel points out, not every mathematical operation admits a physically significant interpretation. The example he gives is using a real vector space to represent Newtonian space. In this space, 3-vectors represent spatial points. It makes perfect mathematical sense to add the vectors, but the addition of spatial points is not physically significant. This might not be entirely convincing, since the usual treatment of space using vectors is with $\mathbb{R}^4$ as an affine space, not a vector space. In affine space, it makes physical sense to add vectors by placing the tail of a vector at the head of another vector. Nonetheless, the point is well taken. For example, in classical electrodynamics a transformation $A^{'} = A + \nabla\lambda$ of the vector potential A by a scalar function λ is certainly mathematically well-defined, but it doesn't result in a physically significant operation.

His rehearsal of the hole argument casts the pushforward of the metric as applying a diffeomorphism to the metric field *sine* manifold. Curiel takes this to be a mathematically well-defined operation, but not a physically meaningful one. Furthermore, if one interprets the pushforward as a diffeomorphism of the metric *cum* manifold, then one is just giving a diffeomorphic presentation of the same abstract mathematical object.

So, we have two complaints to register. The first concerns diffeomorphic presentations and mappings of the metric *cum* manifold. Curiel has an unorthodox take on the nature of mathematical entities. In general, isomorphic structures are merely different presentations of a single underlying abstract entity. For example, he takes $\mathbb{S}^2$ and $\mathbb{R}^2$ with a point added at infinity as different presentations of one and the same entity (p.453, fn.5) Similarly, a group is an abstract entity with different concrete presentations. And a differential manifold is an abstract object that has different "diffeomorphic presentations". We are not quite sure what to make of this. One reading is that these abstract objects are like Platonic forms standing in a relation of one over the many. Alternatively, the abstract objects might be construed simply as equivalence classes of concrete isomorphic presentations. In a footnote, he mentions that if one takes a manifold to be an equivalence class of its diffeomorphic presentations, then "...the operation underlying the hole argument does not make even mathematical sense." (p.453, fn.5). But it's not very comforting to have a dissolution of the hole argument depend on such unorthodoxy. Perhaps worse is that Curiel simply assumes that isomorphic presentations are physically equivalent., in that they represent the same intrinsic physical structure (p. 453).

Our second complaint cuts deeper. It turns out that what Curiel means by a mapping of the manifold *cum* metric in terms of fiber bundles takes a fiber over a point p to a fiber over point $d(p)$ (Personal correspondence). As we argued in §2.3, this corresponds to the active picture of Leibniz shifts. Alternatively, a mapping of the metric *sine* manifold in this context is given by the assignment $(p, g(p)) \mapsto (p, g(d^{-1}p))$. If the bundle trivializes, this corresponds to the passive picture of Leibniz shifts. Again, as we argued in §2.3, the two points of view yield the same mapping of a section of the bundle to a section of the bundle. So, it turns out that Curiel takes active transformations, but not passive, to be physically meaningful, and moreover, models related by an active transformation are simply assumed to be physically equivalent. As with Weatherall, Curiel simply assumes Leibniz equivalence with no underlying brief for it.

5. Where We Stand

Does all this mean that we think that the only way to cope with the hole argument is to bring some metaphysical doctrine to bear on it?

No. We think that the hole argument is an instance of a general permutation argument used for various purposes by Quine (1969), then Davidson (1979), and then Putnam (1981), but all focused on the problem of reference.[20] Although we do not necessarily endorse their separate diagnoses as to what to conclude, we agree that reference lies at the heart of the matter. Limitations of space do not permit us to rehearse their permutation arguments or to provide a general account for spacetime theories, much can be gleaned from the two most notorious cases: the hole argument and the Leibniz shift argument originally given by Leibniz in his correspondence with Clarke.

Take the case of the hole argument. Assume for the sake of argument that we have established a correlation between actual spacetime events and manifold points up through the time-slice S on which $\mathcal{M}$ and $d_*\mathcal{M}$ agree. Ask then, will we ever be able to come to know which model (if either) correctly captures the values of the object fields in the hole H? Prerequisite to this is to be able to identify the individual points in H. This might seem child's play, since at each point there's a coordinate chart in the atlas including that point in its domain so one just has to pick a chart to use. But that overlooks the problem of how to establish a chart physically. One way to do this is with Gaussian normal (or synchronous) coordinates. Assume that somehow a spatial chart has been fixed for a three dimensional neighborhood U of p in the causal past of H on the present hypersurface S Suppose furthermore that the metric tensor is known at each point of U and that we have test particles so that geodesics with tangents normal to S can be traced into the future.[21] Then we can take the spatial components along a geodesic γ to be the same as at the intersection with S and the time coordinate to be the lapsed proper time. Let φ be the resulting chart. Unfortunately, precisely the same physical recipe is to be followed if we want to trace out geodesics $d_*\gamma = d \circ \gamma$ to establish the chart $d_*\varphi$. We are left with mathematical differences where there is no physical difference.

Now take the case of Leibniz's use of a Leibniz shift. If God had placed the entire contents of the universe some fixed distance, say, to the east, would we have a different, or the

[20]That permutations alone suffice was established in §2.4.

[21]For the sake of argument suppose they do not encounter singularities or intersect before or in H.

same physical situation? Weatherall's apparatus of comparisons relative to isomorphisms dictates 'same', *even if* one is a proponent of absolute space. On our view, there is no problem of having secured reference to spacetime points in question. Hence, the adherent of absolute space has no problem answering 'different'.

6. Conclusion

We laud the efforts of Weatherall, Fletcher and Curiel in seeking ways of blocking the hole argument without entering into metaphysical considerations concerning modality and the nature of spacetime. For we, too, believe that the hole argument can be so blocked. However, we have found each of these attempts wanting in one way or another. We conclude, not that metaphysical consideration are inevitable in order to avoid the dilemma posed by the hole argument. Rather, we believe that the problem can be resolved based on considerations of how it is that we are able to refer to spacetime points in general and how in certain crucial cases, such as in the hole argument, we are unable to do so. Hence, Leibniz equivalence is the proper posture to adopt with respect to the Cauchy problem in general relativity, but not with respect to the original acid test posed by Leibniz.

References

[1] Belot, G. [1995]: 'New work for counterpart theorists.' *British Journal for Philosophy of Science* 46(2), 185–195.

[2] Brighouse, C. [1994]: 'Spacetime and Holes', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1994, pp. 117–25.

[3] Brighouse, C. [2018]: 'Confessions of a (Cheap) Sophisticated Substantivalist.' *Foundations of Physics*, https://doi.org/10.1007/s10701-018-0228-2.

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

[5] Curiel, E. [2018]: 'On the Existence of Spacetime Structure', *British Journal for Philosophy of Science*, 69, 447-483.

[6] Dasgupta, S. [2011]: 'The bare necessities'. *Philosophical Perspectives* 25(1), 115–160

[7] Earman, J. [1989] *World Enough and Spacetime: Absolute vs. Relational Theories of Space and Time.* Cambridge: MIT Press.

[8] Earman, J. and Norton, J. [1987]: 'What Price Substantivalism? The Hole Story', *British Journal for Philosophy of Science*, 38, 515-525.

[9] Fletcher, S. [2018]: 'On Representational Capacities with an Application to General Relativity', https://doi.org/10.1007/s10701-018-0208-6.

[10] Glymour, C. [1977]: 'Indistinguishable Spacetimes and the Fundamental Group', in *Foundations of Space Time Theories*, Eds. J. Earman, C. Glymour, J. Statchel, Minneapolis: University of Minnesota Press.

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

[12] Leeds, S. [1995]: 'Holes and Determinism: Another Look', *Philosophy of Science* 62:425-437.

[13] Lewis, D. [1968]: 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy*, 65, pp. 113-26.

[14] Lewis, D. [1973]: *Counterfactuals.* Oxford: Blackwell.

[15] Maidens, A. [1992]: 'EARMAN JOHN S. 1989: World Enough and Space-Time: Absolute versus Relational Theories of Space and Time Cambridge, MA: MIT Press, xiv + 233 pp. (including bibliography and index). *British Journal for Philosophy of Science*, 43, 129-136.

[16] Malament, D. [1977]: 'Observationally Indistinguishable Spacetimes: Comments on Glymour's Paper', in *Foundations of Space Time Theories*, Eds. J. Earman, C. Glymour, J. Statchel, Minneapolis: University of Minnesota Press.

[17] Manchak, J.B. [2009]: 'Can we know the global structure of spacetime?', *Studies in the History and Philosophy of Science Part B: Studies in the History and Philosophy of Modern Physics*, 40:53-56.

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

[19] Maudlin, T. [1990], 'Substances and Spacetimes: What Aristotle Would have Said to Einstein,' *Studies in the History and Philosophy of Science*, 21: 531–61.

[20] Mundy, B. [1992]: 'Space-Time and Isomorphism', *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 1992:515-527.

[21] Roberts, B. [2014]: 'Disregarding the Hole Argument', unpublished manuscript.

[22] Rynasiewicz, R. [1994]: 'The Lessons of the Hole Argument', *The British Journal for the Philosophy of Science*, 45, pp. 407–36.

[23] Rynasiewicz, R. [1996]: 'Is There a Syntactic Solution to the Hole Problem?', *Philosophy of Science*, 63:S55-S62.

[24] Teitel, T. [2019], 'Holes in Spacetime: Some Neglected Essentials,' *Journal of Philosophy* forthcoming.

[25] Wald, R. [1984] *General Relativity*. Chicago: University of Chicago Press.

[26] Weatherall, J. [2018]: 'Regarding the Hole Argument', *British Journal for Philosophy of Science*, 69, 329-350.