Indeterminism and the Hole Argument

Metaphysical Problems of Physics

8th June 2021

Alex Mathie

Three caveats:

1. My teaching style will be different to Erik's---I have some slides rather than a set of notes. These will be uploaded to dropbox/the course website along with the speaker notes and the recording of the session.

2. The content of today's lecture might seem slightly out of keeping with the content we've covered so far. The reason for this is that when Erik asked me to give a lecture and let me decide what the topic would be, I was operating only off the title of the course. Since the hole argument is (to my mind) probably the most prominent example of contemporary spacetime physics bearing directly on a traditional metaphysical question (namely, causation and determinism), that's what I went with. So today's lecture is slightly more applied and less metametaphysics.

3. The lecture was put together with the course prerequisites in mind. That is, it assumes little to no familiarity with the technical details of general relativity. I hope that means it's more accessible to people without a background in physics. However, I'm conscious that the people currently in the Zoom meeting with me *are* people with physics backgrounds, in some cases considerably more physics background than I have. So for those people, this might seem a bit pedestrian. Nevertheless, I hope it's interesting.

Plan

1. **The Hole Argument & Consequences for ‘Metaphysics’**

    1.1 General Relativistic Spacetime Models
    1.2 General Covariance
    1.3 Leibniz Equivalence
    1.4 The Hole Argument

2. **Intermission**

3. **Four Responses to the Hole Argument**

    3.1 Biting the Bullet
    3.2 Sophisticated Substantivalism
    3.3 Metrical Essentialism
    3.4 “Stop Abusing my Math!”

The Hole Argument & Consequences for ‘Metaphysics’

- The first half of this lecture will be on the hole argument and its putative consequence for metaphysics

- I say ‘metaphysics’ in scare quotes not because I’m necessarily scared or sceptical of metaphysics but rather to indicate that these are consequences for what we might think of as being typically *metaphysical* questions. Questions of determinism, causality, the nature of spacetime, possible worlds, and so on.

- So while I echo Erik’s suggestion from lecture 2 that we should try to avoid bringing to this material preconceptions of what metaphysics *is*, there is a sense in which the hole argument bears directly on what are (rightly or wrongly) considered metaphysical questions.

The Hole Argument & Consequences for ‘Metaphysics’

The hole argument (broadly) says something like this:

> *“On a substantivalist reading, our best physical theory of spacetime, general relativity, is indeterministic.”*

In order to understand what this means, we need to clarify:

- What general relativity says
- How the substantivalist must interpret this
- Why this commits us to indeterminism (and what that means)

To give a sense of where this is all going, here’s the broad brush summary of the hole argument: on a substantivalist reading, our best physical theory of spacetime, general relativity, is indeterministic

Can see straight away this is significant news for metaphysics---determinism seems to fail when we think of spacetime as a substance.

In order to unpack this, we need to cover a few different subtopics:

1. First, what does general relativity actually say, i.e. what are the relevant features of general relativity that lead to the hole argument
2. Second, what does it mean to be a substantivalist about spacetime, particularly general relativistic spacetime? I.e. what is the substantivalist committed to in the context of GR?
3. Third, why do these two things sum to the failure of determinism? And what sort of determinism is this?

General Relativistic Spacetime Models

First off, what does general relativity actually say about spacetimes? What are spacetimes, according to general relativity?

The Hole Argument & Consequences for ‘Metaphysics’
General Relativistic Spacetime Models

A *spacetime model* is an n-tuple $\mathcal{M} = \langle M, g_{ab}, \phi_1, \phi_2 \ldots \rangle$, where:

- *M is* a Lorentzian manifold
- g_{ab} is a metric tensor defined on that manifold
- $\phi_{1,\,2,\,\ldots}$ are fields representing the material content of that spacetime
- (This last bit will sometimes be written as the stress-energy tensor, T_{ab})

A ‘spacetime model’ in general relativity is an n-tuple consisting of the following objects:

1. A Lorentzian manifold, M, which is a topological space which is locally homeomorphic to R4
2. A metric tensor, g_ab which is defined on M and encodes the geometric properties of the spacetime and allow us to talk about things like the distance between two points in a tangent space at a the manifold.
3. Some fields, which represent the material content of the spacetime: mass, electric fields and so on. Sometimes rather than writing a number of fields we’ll just specify the stress-energy tensor T_{ab} instead.

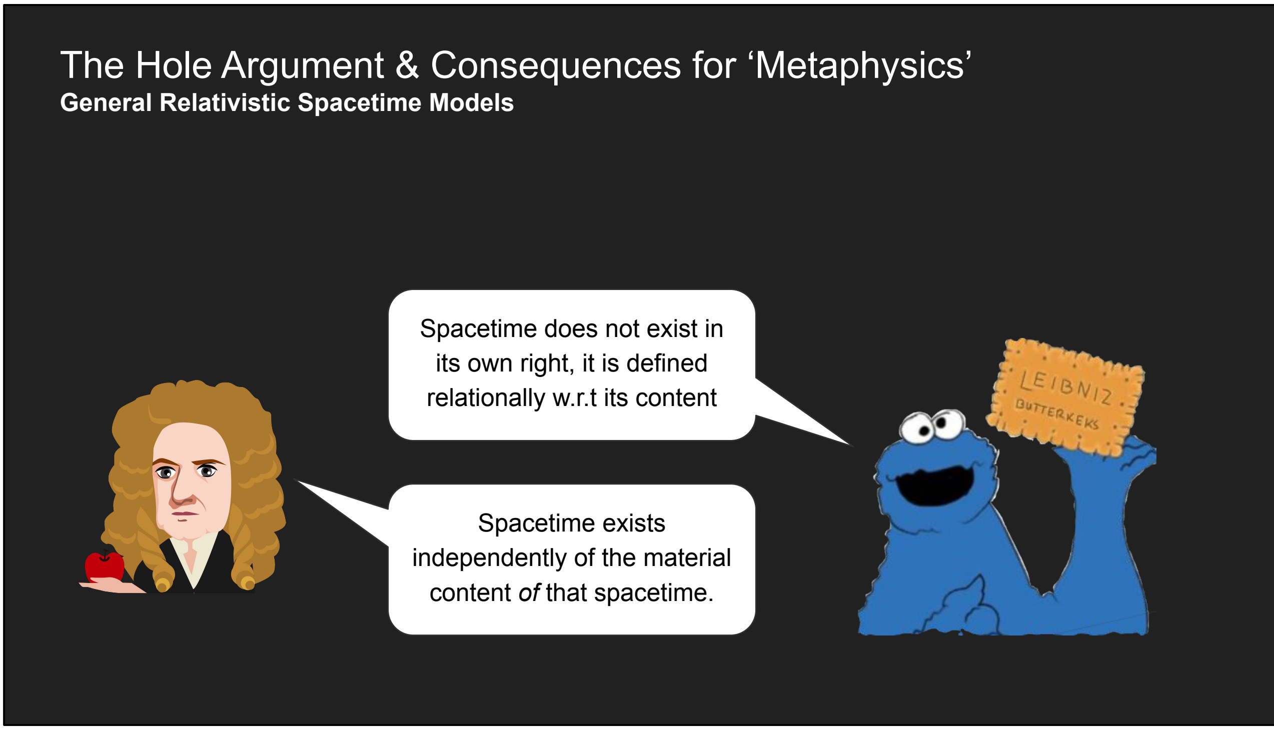
The Hole Argument & Consequences for ‘Metaphysics’

General Relativistic Spacetime Models

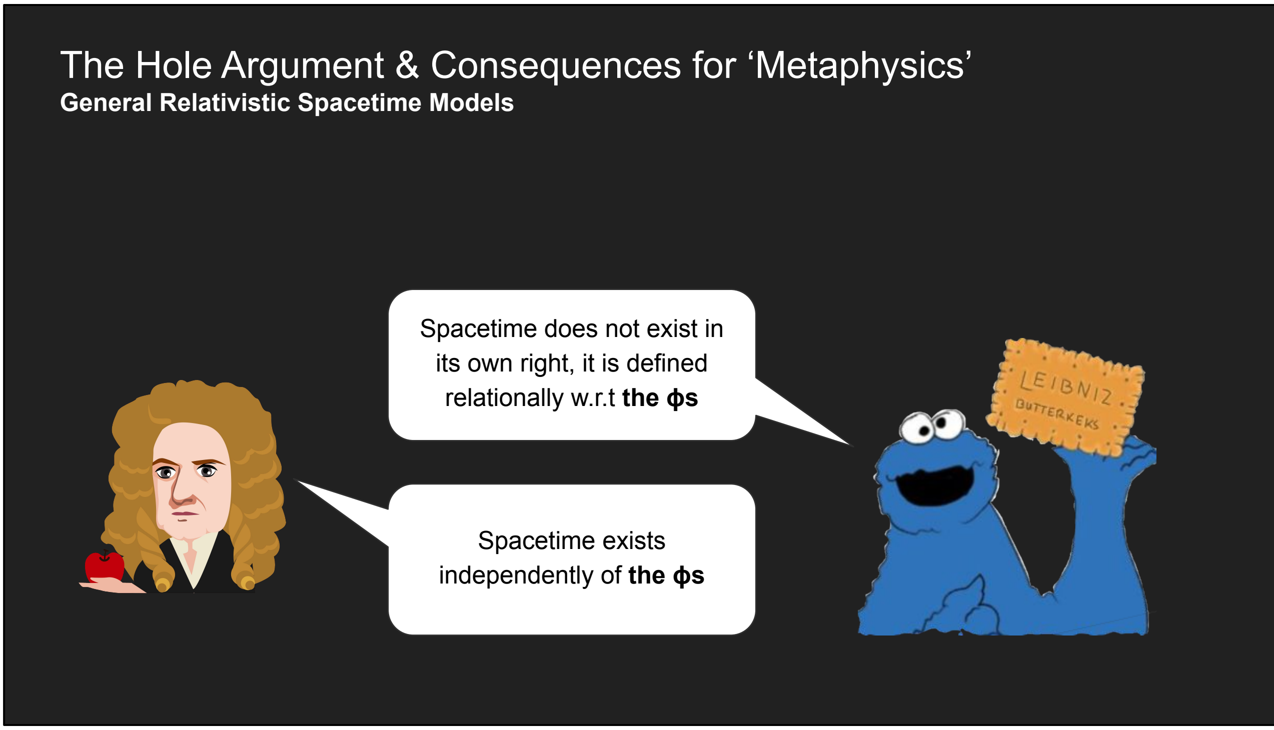
Which part(s) of this spacetime model (if any) represent ‘the spacetime’ itself?

Since the spacetime model typically contains elements representing the material content of that spacetime, there is some latitude in choosing which of the elements represent the spacetime.

In other words, the substantivalist has to say which of element or elements they take to constitute the spacetime substance.



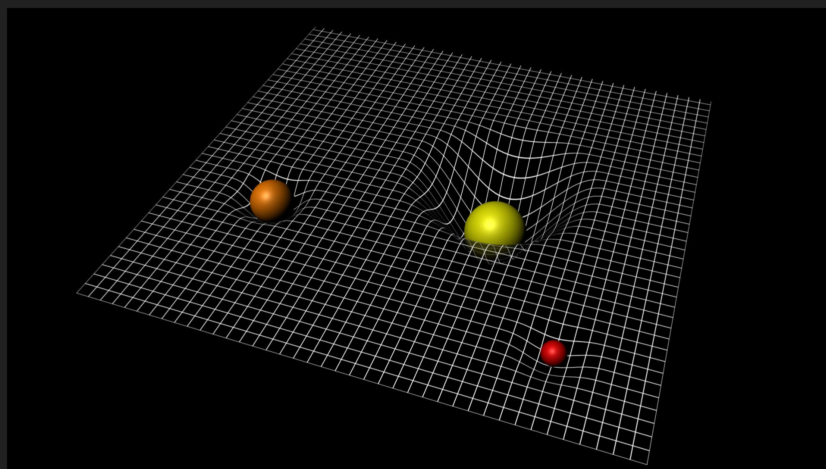
This debate over the nature of spacetime goes back many hundreds of years, and is normally associated with Newton espousing a picture of absolute space, and Leibniz asserting that space was purely relational.



Anachronistically, because neither Leibniz nor Newton/Clarke knew anything about the mathematical framework of general relativity, we might attribute to the Newtonian absolutist the claim that <M, g_ab> represents spacetime independently of any material content, and attribute to the Leibnizian relationalist that spacetime doesn’t exist in its own right, it’s derivative upon matter.

To be a substantivalist, then, one at least has to say which part of a spacetime model is supposed to be the real physical entity ‘spacetime’

The Hole Argument & Consequences for ‘Metaphysics’
General Relativistic Spacetime Models



General relativity makes it *prima facie* tempting to be a substantivalist about spacetime.

After all, in GR, spacetime is a dynamical physical entity that acts on matter and in turn is acted upon by matter.

If we think of relationalism as being roughly committed to the idea that spacetime exists only as a relational object whose existence is abstracted from the material content of the universe then it might be difficult, at least superficially, to see how this tallies with the depiction of spacetime given in general relativity, where spacetime seems to be causally efficacious in the same way that matter is.

The Hole Argument & Consequences for 'Metaphysics'
General Relativistic Spacetime Models

Can't be the whole thing...

It can't be the whole MgT, since that would seem not to respect the distinction between spacetime and the stuff in it (a distinction that the substantivalist is committed to).

So that leaves two options: either the spacetime is the manifold M with the metric g, or it is just the manifold itself

(It makes no mathematical sense to talk about the metric without the manifold, since the metric is defined on a manifold, so taking spacetime to be 'just g' is not an option).

The Hole Argument & Consequences for ‘Metaphysics’
General Relativistic Spacetime Models

Is the manifold spacetime?

“The metric tensor now incorporates the gravitational field and thus, like other physical fields, carries energy and momentum, whose density is represented by the gravitational field stress-energy pseudo-tensor. [...] energy and momentum are carried by the metric in a way that forces its classification as part of the contents of spacetime.”

(Earman & Norton 1987, p.519)

Earman & Norton suggest that spacetime is just the manifold.

Since, they argue, the metric carries energy and momentum just like a physical field, it must be classified as part of the material content of spacetime, not the spacetime itself. This is quite a subtle issue in general relativity that some may object to, but for the moment let’s accept Earman and Norton’s argument.

The manifold, therefore, i.e. a collection of individual spacetime points with certain topological properties just *is* the spacetime. The other elements of the spacetime model, the metric and stress-energy tensors, are supposed to represent the material content within that spacetime.

The Hole Argument & Consequences for 'Metaphysics'

General Relativistic Spacetime Models

For the substantivalist, then, the manifold is ontologically on a par with matter: it represents a real physical entity that exists in its own right.

Why is this a problem?

Later, we'll see some arguments against taking spacetime to be the bare manifold and instead taking it to be the manifold plus the metric. But for now, let's say the substantivalist identifies spacetime with the manifold.

If so, the manifold has equivalent ontological status to the physical content of the spacetime model.

So far this doesn't seem like a huge problem. However, when one takes into account some of the more subtle mathematical properties of GR, things start to look less straightforward.

General Covariance

A second important component of ‘what general relativity says’ concerns a specific mathematical property it exhibits, called general covariance.

The Hole Argument & Consequences for ‘Metaphysics’

General Covariance

The mathematics of general relativity has a property called **general covariance.**

Roughly: the field equations are invariant under arbitrary differentiable coordinate transformations (“diffeomorphisms”)

- The mathematical formalism of general relativity is ‘generally covariant’.
- Roughly what this means is that the field equations in general relativity are invariant not only across inertial reference frames but across any two reference frames related by a differentiable coordinate transformation, which we call a **diffeomorphism**.
- This is prima facie a desirable property for physical theories, since we intuitively want our laws of physics to as free of arbitrariness as we can: we want them to work across different reference frames and plausibly across permutations of spacetime points.

The Hole Argument & Consequences for ‘Metaphysics’
General Covariance

A **diffeomorphism**, *d*, is a bijective map, $d: M \rightarrow M$, with a smooth inverse, that preserves the differentiable structure of *M* (roughly: it takes smooth curves to smooth curves).

Passive diffeomorphisms $d_{passive}: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ (i.e. ‘relabellings’ of spacetime points)

Active diffeomorphisms $d_{active}: M \rightarrow M$ (i.e. ‘rearrangements’ of spacetime points).

- There is an important distinction to be drawn here between active and passive interpretations of diffeomorphism invariance:
    - Passive diffeomorphism invariance is invariance under arbitrary relabelling of spacetime points (e.g. interpreting the action of the diffeomorphism as a move to a different coordinate system)
    - Active diffeomorphism invariance is invariance under a rearrangement of the spacetime points themselves (e.g. moving which points of the manifold instantiate which spatiotemporal relations of the spacetime)
- GR is invariant under active diffeomorphisms (which is not always the case for physical theories).
- In other words, I can take the spacetime points of my manifold, and I can permute them any which way I like (so long as the permutation respects the differential structure of the manifold) and the resulting model (with the appropriately transformed metric and stress-energy tensors) will be empirically equivalent to my original spacetime.

The Hole Argument & Consequences for ‘Metaphysics’

General Covariance

Given any spacetime model $\mathcal{M} = < M, g_{ab}, T_{ab} >$ that is a solution to the field equations, the spacetime model $d^*\mathcal{M} = < M, d^*g_{ab}, d^*T_{ab} >$ (where d^* represents the action of a diffeomorphism) must also be a solution to the field equations.

Given any spacetime model M,g,T that is a solution to the field equations, a related spacetime model d*M, related to M by a diffeomorphism will also be a solution to the field equations.

Since it is a solution to the field equations, it represents a physically possible spacetime according to our best theory of spacetime physics.

The Hole Argument & Consequences for ‘Metaphysics’

General Covariance

So far so good?

GR lends itself to a substantivalist interpretation...

Spacetime is (plausibly) the manifold…

GR is a generally covariant theory…

So far so good.

Nothing I have said so far would be considered particularly worrying for the substantivalist.

We’ve seen that GR plausibly lends itself to a substantivalist interpretation.

We’ve seen that the most obvious candidate for ‘spacetime itself’ is the bare manifold, with the metric and stress-energy tensors being physical fields defined on that spacetime.

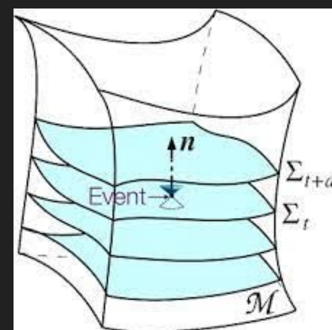
And we’ve seen that GR is a generally covariant theory, which means that it is invariant under a very broad class of transformations called diffeomorphisms.

The Hole Argument & Consequences for 'Metaphysics'

General Covariance

Now consider a special type of diffeomorphism, the "hole diffeomorphism" (following Pooley 2020's construction):

- Suppose that our *d*-dimensional manifold *M* may be foliated into a series of achronal *d-1* dimensional hypersurfaces.
- Let one such hypersurface be called Σ, and let *H* ("hole") be an open subset of *M* that lies entirely to the future of Σ.
- The hole diffeomorphism, d_{hole} , is the identity outside of *H* (i.e. before Σ) and smoothly becomes nontrivial within *H*.



Now let's consider a special type of diffeomorphism which we call the 'hole diffeomorphism'.

- Slice your manifold up into a series of achronal foliations (achronal meaning that every point on a given hypersurface lies neither to the future nor the past of any other point on the hypersurface). This essentially defines a global time function.
- Now choose your favourite slice, and call that one Σ.
- Now define an open subset of the manifold as being entirely to the future of Σ.
- Let the hole diffeomorphism be a diffeomorphism that is the identity outside of this region H, but that varies smoothly to become nontrivial inside H.

The Hole Argument & Consequences for ‘Metaphysics’
General Covariance

$\mathcal{M}$ and $d_{hole}{}^{*}\mathcal{M}$ are both, by general covariance, solutions to the field equations.

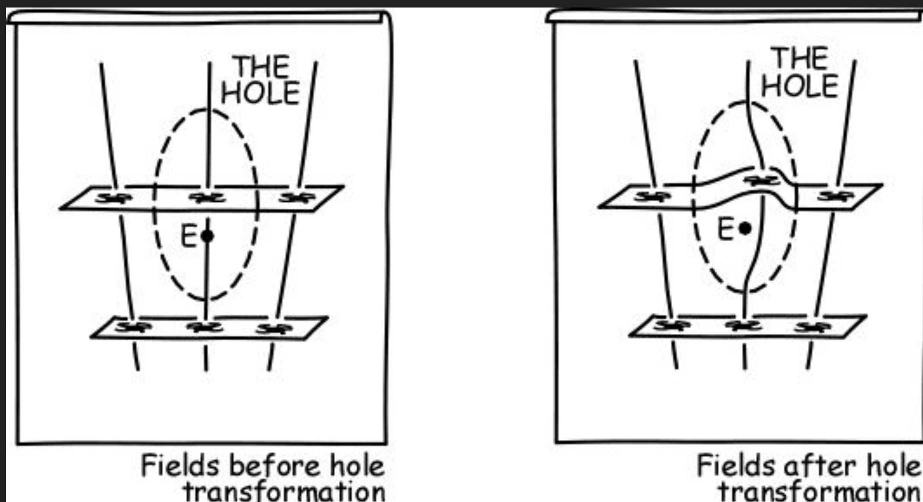
They are at least *mathematically* distinct since $d(p) \neq p$ for some $p \in H$.

But are they *physically* distinct?

By general covariance, any two solutions related by such a hole diffeomorphism are equally valid solutions to the field equations. The hole diffeomorphism is just a type of diffeomorphism, and since GR is invariant under diffeomorphisms, this seems unavoidable.

But the two solutions are at least mathematically distinct in the following sense: within the hole, the diffeomorphism is not the identity, i.e. it maps some spacetime points to others.

Should we think of diffeomorphic solutions as representing *physically* distinct possibilities? By now, you can probably see where this is going: if they *do* represent physically distinct possibilities, the substantivalist is in trouble.



(Norton 2019)

Here are two different ways of ‘spreading matter fields over the manifold’, roughly speaking. Since the two different spreadings are diffeomorphic, they are empirically indistinguishable thanks to general covariance.

For a certain spacetime point (i.e. a point on the manifold), ‘E’, we can ask: does the galaxy pass through point E?

The relationalist can justifiably refuse to answer this question, since one cannot speak of spacetime points independent of the material content of spacetime.

Things aren’t so easy for the substantivalist, however: this is exactly the sort of question they should be expected to have an answer to.

Leibniz Equivalence

Getting to the bottom of whether two diffeomorphic models represent physically distinct ways the for the universe to be basically takes us back to Leibniz and Newton/Clarke.

The Hole Argument & Consequences for ‘Metaphysics’

Leibniz Equivalence

The substantivalist maintains that a universe with everything moved three feet East is a fundamentally different physical universe.

Two universes related in this way are related by a ‘Leibniz shift’. They are **Leibniz equivalent** universes.

Rejecting Leibniz equivalence is the ‘acid test of substantivalism’.

There, as we know, the substantivalist wants to maintain that a universe with everything moved three feet East is a fundamentally different universe to the one in which everything stays where it is, *precisely because there is a physical spacetime with respect to which things have moved.*

We say that two such universes are related by a Leibniz shift. They are Leibniz equivalent universes.

The rejection of Leibniz equivalence really constitutes the essential commitment of the substantivalist. If someone accepts Leibniz equivalence, they cannot, by definition, hold a substantivalist view of spacetime.

The Hole Argument & Consequences for ‘Metaphysics’

Leibniz Equivalence

“Whatever reformulation a substantivalist may adopt, they must all agree concerning an acid test of substantivalism, drawn from Leibniz. If everything in the world were reflected East to West (or better, translated 3 feet East), retaining all the relations between bodies, would we have a different world? The substantivalist must answer yes since all the bodies of the world are now in different spatial locations, even though the relations between them are unchanged.”

(Earman & Norton 1987, p.521)

For this reason, Earman & Norton call the rejection of Leibniz equivalence the ‘acid test’ of substantivalism.

The Hole Argument & Consequences for 'Metaphysics'

Leibniz Equivalence

What about relativistic spacetime models?

If diffeomorphisms can be thought as generalisations of "moving things three feet East", then the manifold substantivalist must maintain that two diffeomorphic relativistic spacetime models represent *physically* distinct possibilities.

In other words, they must *deny* the equivalence of diffeomorphic spacetimes.

How does this generalise to the relativistic case?

Well, if diffeomorphisms can be considered the generalisations of things like Leibniz shifts, then it seems as if the substantivalist must deny the equivalence of diffeomorphic spacetimes.

The Hole Argument & Consequences for ‘Metaphysics’

Leibniz Equivalence

“The diffeomorphism is the counterpart of Leibniz' replacement of all bodies in space in such a way that their relative relations are preserved. [...] substantivalists, whatever their precise flavour, will deny:

Leibniz equivalence: Diffeomorphic models represent the same physical situation.”

(Earman & Norton 1987, p.522)

Indeed, this is how Earman and Norton conceive of the substantivalist commitment in GR.

The Hole Argument

Now, all of the pieces are in place:

We know roughly what a spacetime model is.

We know (at least prima facie) which part of it the substantivalist wants to call ‘spacetime’.

Furthermore, we know that general relativity is invariant under active diffeomorphisms.

And we know that we can define a special type of diffeomorphism with a ‘hole’ in it.

Lastly, we know that (at least prima facie) it seems like the substantivalist must regard diffeomorphic spacetime models as representing different physical possibilities, i.e. different ways for the universe to be.

The Hole Argument & Consequences for ‘Metaphysics’
The Hole Argument

Following Pooley (2020):

P1: (Substantivalism)

P1.1 (Plurality)
We can associate to any possible spacetime a plurality of other spacetimes that are (i) instantiations of the very same spatiotemporal properties and involving the very same material content as *P*, but that (ii) differ from *P* only in virtue of which spacetime points instantiate those spatiotemporal properties and locate that material content.

P1.2 (Models)
If the model $\mathcal{M} = \langle M, g_{ab}, T_{ab} \rangle$ can be chosen to represent a spacetime *P* then relative to that choice there is a permissible and natural interpretation of $d^*\mathcal{M}$ according to which it represents a distinct spacetime *P’*.

P2: (Copossibility)
For two models $\mathcal{M}_1$ and $\mathcal{M}_2$ that are solutions to the field equations, then if there is a joint interpretation under which $\mathcal{M}_2$ represents possibility P_1 and $\mathcal{M}_2$ represents possibility P_2, then if P_1 is physically possible so is P_2.

P3: (Determinism)
A theory is indeterministic if it both regards $\mathcal{M}$ and $d^*\mathcal{M}$ as representing distinct possibilities and regards the possibilities as equally possible.

C: Substantivalist GR is indeterministic

So now we can state the hole argument clearly.

This is a dense slide but don’t worry, it’s really just going over things we’ve already seen. It’s useful to state the argument this way because laying out the premises allows us to establish a sort of taxonomy of responses. The responses we’ll consider in the second half of this lecture aim to undermine one or more of the premises of the argument as formulated here.

[Read premises]

Useful to note:

- The distinction between **Plurality** and **Models** should be thought of as the distinction between a *physical* state affairs and the *mathematical representation* of that state of affairs: the idea is that **Plurality** is a denial of *physical* Leibniz equivalence, and **Models** states how we are supposed to identify the natural referent of Leibniz equivalence (and therefore how we are supposed to identify its failure) in the formalism.

- **Copossibility** can be thought of as the demand that one has to be consistent in interpreting the formalism. Supposing the existence of an appropriate joint interpretation, one cannot chop and choose which spacetime models one takes to be ‘physically possible’. (More on this later---it is the essential criticism of metrical essentialism). The idea here is to define (abstractly) the necessity

- of copossiblity and then suppose that M1 and M2 are related by the relevant hole diffeomorphism to force the admission that both are physically possible.

The Hole Argument & Consequences for ‘Metaphysics’

The Hole Argument

Metaphysics!

So now we have quite a precise view of what is at stake for metaphysics.

The Hole Argument & Consequences for ‘Metaphysics’
The Hole Argument

Our best theory of physics seems to point to either:

(a) **The (arbitrarily large) failure of determinism**: complete information up to Σ does not uniquely specify the dynamics after Σ.

(b) **The failure of substantivalism**: spacetime fails to be ontologically on a par with matter (at least in the way characterised above).

In particular, our best theory of physics seems to indicate either:

a. Determinism fails to an arbitrarily large extent, since we can make the hole as large as we like, so that the physical state of most of the universe is rampantly underdetermined, or we can make the hole as small as we like, so that we already have most of the information about the universe, but it’s still not enough to uniquely determine what’s happening in the tiny hole region.

Or:

b. Spacetime resists substantivalism and must be thought of as being ontologically dissimilar to matter.

The Hole Argument & Consequences for ‘Metaphysics’

The Hole Argument

- **Bite the bullet** … Relationalists; indeterminists
- **Deny P1** … Sophisticated substantivalists; Math purists
- **Deny P2** … Metrical essentialists
- **Deny P3** … Denying that this is *really* indeterminism

Now that we’ve set up the argument in this way, we can taxonomise the possible responses like this.

We’ll consider the first three bullet points in the second half of this lecture (we won’t consider the third, but it is still an option---I include it here for completeness).

Intermission

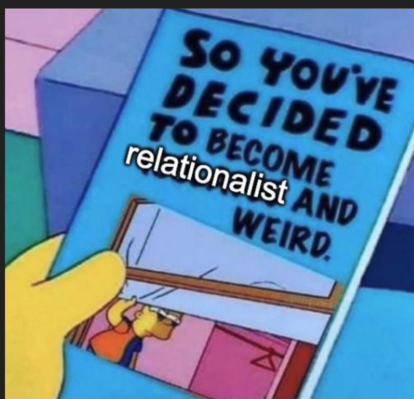
10 minute break

Four Responses to the Hole Argument

The easiest and most boring response to the hole argument is simply to bite the bullet

Four Responses

Biting the Bullet



- The easiest ‘response’ is no response at all
- We could just be relationalists
- Maybe there are reasons not to be…
    - Fields?
    - Metaphysical dogma?

There are really two bullets to bite. Accepting the hole argument means choosing to save substantivalism by sacrificing determinism, or it means choosing to save determinism by sacrificing substantivalism.

So we could just sacrifice substantivalism, and be relationalists. We simply stop maintaining that the bare manifold has ontological parity with the material content of spacetime, and relegate it to derivative existence or eliminate it completely.

But maybe there are reasons not to be.

- We might think that the metric and matter fields presuppose the existence of spacetime points. After all, we can think of fields as *assignments* of certain properties or states to points in space. This argument is made (very aptly, but in a different context) by Hartry Field.

- We might also heed Earman & Norton’s caution that just as determinism should *fail* for reasons of physics, it should also *succeed* for reasons of physics. In the absence of a deciding factor from physics, we should remain agnostic.

Four Responses

Biting the Bullet

- Alternatively, we could just accept indeterminism.
- Is indeterminism really so bad?
- Lots of our physical theories are indeterministic, even classical ones.
- And the sort of indeterminism here might seem particularly benign.
- “[It] is a kind that cannot, in principle, have any empirical consequences.” (Hoefer 1996, p.10)

Alternatively, we could choose to sacrifice determinism in order to save substantivalism.

We might think that such a sacrifice is acceptable, especially since many of our best theories of physics seem to harbour indeterminism of some sort.

But here, the indeterminism isn’t baked into the theory. Rather it comes about as a consequence of a specific metaphysical commitment---substantivalism---and we shouldn’t accept indeterminism purely for metaphysical reasons!

Instead, we might think that the type of indeterminism we have here might be particularly inoffensive since it’s only mathematical indeterminism: diffeomorphic models are empirically identical.

Four Responses
Biting the Bullet

“In one sense the indeterminism is pernicious in that, for every possible spacetime, no matter how small a region one considers, the laws and the rest of spacetime fail to fix the state of that region. In another sense, however, the indeterminism is anodyne. Any two possibilities represented by models that differ by a hole diffeomorphism instantiate the very same global pattern of properties and relations. They are therefore *qualitatively* perfectly alike.”

(Pooley 2020, p.10)

Maybe we can do better, though...

2. Sophisticated Substantivalism

This observation leads naturally to our first proper response, sophisticated substantivalism.

Four Responses

Sophisticated Substantivalism

Why does this kind of indeterminism seem so tolerable?

The ‘difference’ that leads to the indeterminism might not be a difference at all.

Two spacetimes related by a hole diffeomorphism differ only in virtue of which spacetime points instantiate which spatiotemporal relations.

Is this a real difference?

So why does this determinism seem so tolerable?

Because the difference that seems to underpin the determinism isn’t really a *substantive* difference. It has no empirical consequences because it must (by definition) preserve certain properties of the spacetime.

Diffeomorphic spacetimes only differ in terms of which spacetime points are instantiating which spatiotemporal relations.

It seems natural to ask whether this is a real difference?

Four Responses
Sophisticated Substantivalism

If it is, claims like “**that** spacetime point is now **this** spacetime point” have to make some sense.

This presupposes that we may individuate spacetime points in isolation from the specific spatiotemporal relations that they instantiate.

For this to be the case, each spacetime point would need to have an *intrinsic ‘thisness’*, or *haecceity*, that distinguishes it.

If it is to be a substantive difference, then we would have to make sense of the idea that what distinguishes diffeomorphic spacetimes is the fact that “**that** spacetime point is now **this** spacetime point”.

But this presupposes that spacetime points can be individuated from one another without recourse to the specific spatiotemporal properties they instantiate.

In other words, each spacetime point must have some sort of primitive identity, irrespective of what spatiotemporal relations it stands in. That is, it must have haecceitistic properties.

There is a sense in which this argument cedes some ground to the relationalist: spacetime points have purely relational identity, but they still *exist* in a way that the relationalist typically wants to deny.

Four Responses

Sophisticated Substantivalism

But haecceitism is problematic.

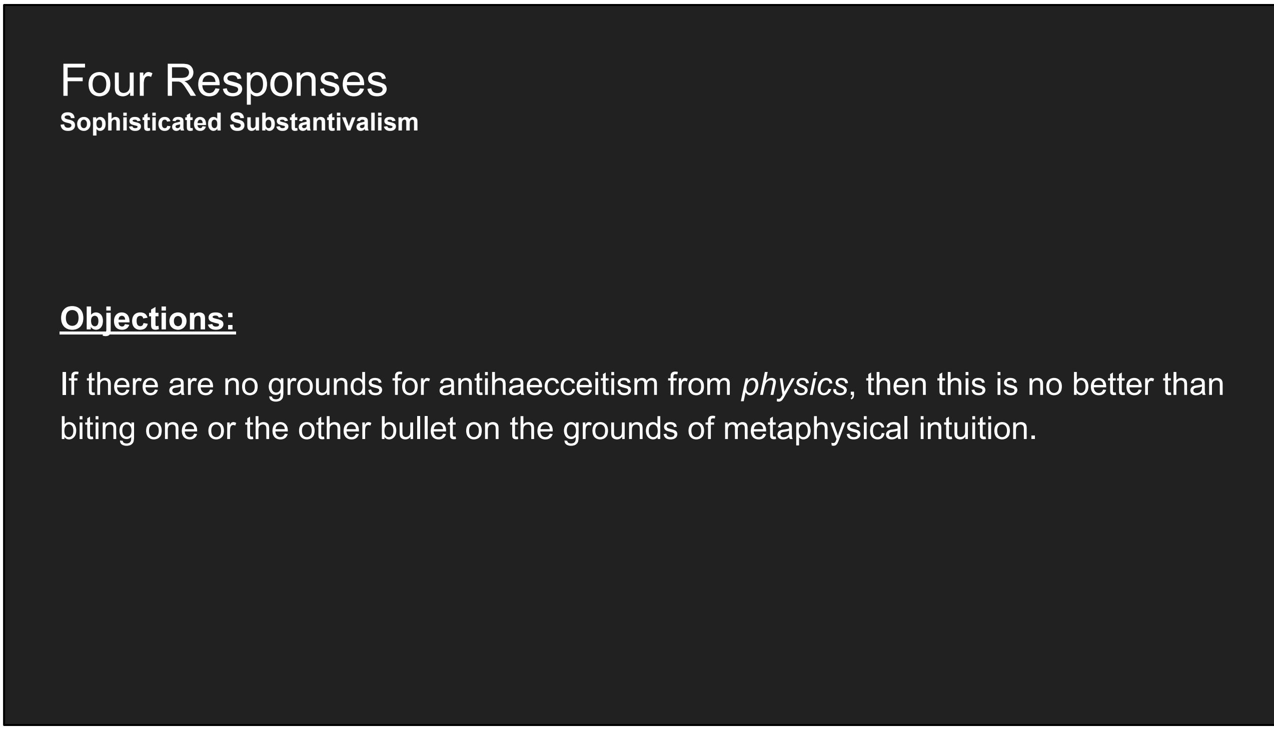
Philosophers have endorsed antihaecceitism on independent grounds, not only on as an ad hoc response to the hole argument.

The sophisticated substantivalist will simply deny that spacetime points have haecceities.

So if this all comes down to haecceity, then maybe the substantivalist can just surrender haecceity.

After all, it’s fairly common to reject haecceitistic differences as being real differences on independent metaphysical grounds not just as an ad hoc response to the hole argument.

So what the sophisticated substantivalist will try to do is simply deny that haecceitistic differences are the sort of differences they are committed to, as substantivalists. The idea is then that they can accept a weakened form of Leibniz equivalence and thereby block the hole argument, because diffeomorphic spacetimes are not *actually* physically distinct from one another.



One objection to sophisticated substantivalism might proceed along roughly naturalistic grounds. We deliberately abstained from biting one of the two available bullets for the reason that it didn't seem right to adjudicate between determinism or substantivalism on the basis of metaphysical intuition at all.

Yet the sophisticated substantivalist, if they don't have an *independent* reason for denying primitive identity to spacetime points (or, stronger, if they don't have an *independent reason from physics* for doing so), then this move is arguably no more legitimate.

Hoefer argues that antihaecceitism is just good metaphysics.

Earman has argued that quantum gravity is unlikely to be compatible with even a sophisticated substantivalism.

Maybe sophistication isn't enough to save the substantivalist.

3. Metrical Essentialism

Let’s look at a different strategy.

Four Responses

Metrical Essentialism

<u>Consider the following intuition:</u>

What it is to *be* a certain spacetime point is to stand in the right sort of metrical relations with other spacetime points (distance *x* from this spacetime point, distance *y* from that one).

Two models related by a hole diffeomorphism are not physically distinct…

- Not just because *haecceities don’t count* (cf. sophisticated substantivalism)
- But because spacetime points bear their metrical properties essentially.

Since these properties come from the metric, spacetime is $< M, g_{ab} >$, not $< M >$.

Recall that Earman and Norton argued that spacetime should be identified with the bare manifold, M, for the reason that the metric seemed to be more like a physical field carrying energy and momentum than it did a spacetime.

Maybe this was wrong.

One of the reasons we might think this can be seen with the following intuition. Maybe what it is to be a certain spacetime point is just to instantiate certain metrical properties, and these come from the metric tensor, g_ab.

This sounds similar to sophisticated substantivalism, but it is importantly distinct: for the SS, the reason spacetime points have certain geometrical properties is because non-geometric haecceitistic differences are rejected. For the metrical essentialist, the reason is that the metrical properties of a given spacetime point are *essential* to it.

And since these properties come from the metric, spacetime is M and g, not just M.

Four Responses
Metrical Essentialism

"[...] we must ask: what are the essential properties of space-time points regarded as substances? Which of their properties might possibly be stripped from them and which not? Perhaps upon answering this question we may find that the hole diffeomorphism generates [...] a description of a distinct but impossible state of affairs"

(Maudlin 1988, p.86)

This line of response is endorsed by Tim Maudlin in his 1988 paper.

[Quote]

Maudlin has some colourful examples: just because I can say "Richard Nixon could have been a ham sandwich" or "Tim Maudlin could have been the Eiffel tower" does not mean that such counterfactuals are metaphysically possible. Being a ham sandwich or the Eiffel tower might go against some of the essential properties of being Richard Nixon or Tim Maudlin.

Four Responses

Metrical Essentialism

For the metrical essentialist, permutations of spacetime points represent **metaphysically impossible states of affairs**.

Pooley takes the metrical essentialist to accept a modified version of **Plurality**

> **P1.1 (Plurality*)**
> We can associate to any possible spacetime a plurality of **<u>other ways for spacetime to be</u>** that are (i) instantiations of the very same spatiotemporal properties and involving the very same material content as *P*, but that (ii) differ from *P* only in virtue of which spacetime points instantiate those spatiotemporal properties and locate that material content.

On this view, while the mathematics of GR may be invariant under permutations of spacetime points, what these permutations represent are **metaphysically impossible states of affairs**.

So the Metrical essentialist cashes out the plurality premise in terms of ‘ways for a spacetime model to be’. GR gives a number of diffeomorphic representations of ways for the world to be.

Four Responses

Metrical Essentialism

But to **reject** a modified version of **Copossibility**:

P2: (Copossibility*)

For two models $\mathcal{M}_1$ and $\mathcal{M}_2$ that are solutions to the field equations, then if there is a joint interpretation under which $\mathcal{M}_2$ represents a certain **'way for the world to be'** P_1 and $\mathcal{M}_2$ represents a certain **'way for the world to be'** P_2, then if P_1 is physically possible so is P_2.

But they deny that these ways for a spacetime model to be represent equivalent physically plausible ways for the universe to be.

Basically, Maudlin attempts to block the inference from a plurality of allowed models in our physics to a plurality of physically possible configurations of the universe.

He attempts to do so by invoking essentialism, with the idea being that since diffeomorphic models change which spacetime points bear which metrical properties, and since these metrical properties are plausibly essential properties of a spacetime point, then d*M seems to change the essences of things, and we should therefore reject it.

Four Responses

Metrical Essentialism

The metrical essentialist claims that:

- The mathematics of GR may be invariant under certain operations
- But since these operations correspond to a change in ‘the way that the world is’ that does violence to the *essential properties* of that world, we should not take them to represent a genuine physical possibility.

Put another way, the metrical essentialist accepts that:

- The mathematics of GR may be transformed in certain ways that preserve physical properties

But they deny that:

- These transformations correspond to metaphysically possible transformations, since they don’t preserve essential properties.

This allows the metrical essentialist to block the hole argument: the mathematics may be indeterministic, but reasons of metaphysics and essence prevent ‘reality’ from inheriting this indeterminism.

Four Responses
Metrical Essentialism

Objections

- Inconsistent interpretation of the mathematical formalism
  - GR is a guide to ontology but not a guide to the essential properties of that ontology?
- Difficulty making sense of even simple counterfactuals
  - The metrical essentialist cannot make sense of (e.g.) “the mass at a spacetime point *could* have been greater than it is”
  - Dynamically allowed spacetimes not isomorphic to ours represent “different possible spacetimes, not different possible states of *this* spacetime” (Maudlin 1989, p.90)
- Lack of substantive positive argument
  - Agreement with an obscure passage of Newton’s *De Gravitatione*
  - Even support from this (Pooley claims) is questionable

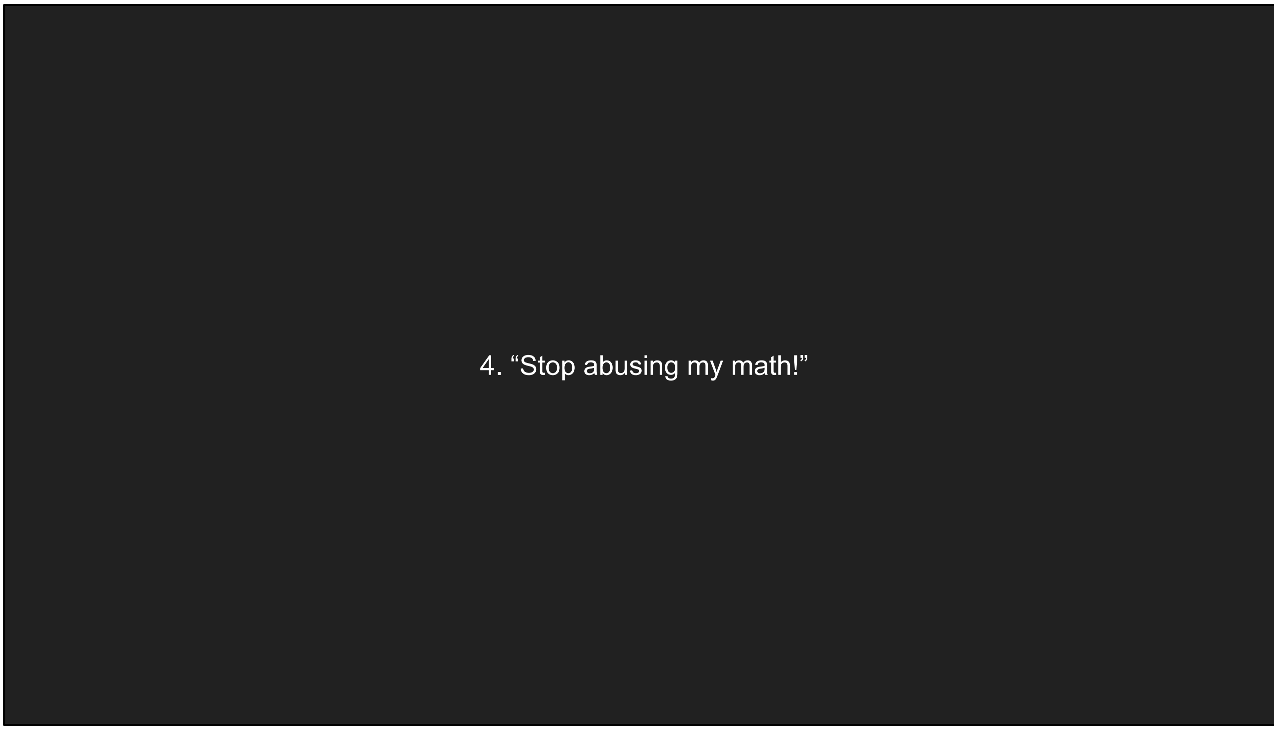
There are several ways we can push back against the metrical essentialist:

1. **First, they seem to be cherry picking which aspects of the mathematical formalism they want to take seriously.**
   a. They want to claim that manifolds of spacetime points exist because they appear in the formalism of general relativity, but they want to deny that the same formalism can determine what counts as an essential property of those spacetime points.

2. **Second, metrical essentialism makes it incredibly difficult to evaluate counterfactuals.**
   a. It just doesn’t seem like a certain spacetime point instantiating a certain set of properties is essential in the same way that Maudlin’s non-Eiffel tower properties might be. “This spacetime point could have had different metrical properties” does seem (prima facie) more plausible than “Tim Maudlin could have been the Eiffel tower” is.
   b. This commitment to essentialism commits the metrical essentialist to the idea that dynamically allowed spacetimes that are not completely isomorphic to our spacetime represent fundamentally different spacetimes, not ways that our spacetime could have been or could later be. This seems highly counterintuitive.

3. **Lastly, there isn’t really a good positive argument in favour of metrical essentialism.**

a. Maybe it works as a way of evading the hole argument. But is it just an ad hoc strategy for saving substantivalism and determinism?
b. Maudlin suggests that his view is in agreement with an obscure passage from Newton’s *De Gravitatione*, but is that really such a win?
c. And Pooley suggests that it is debatable whether this is even the case.



The last type of response we’ll look at is what I’ll playfully call the “stop abusing my math” response.

Four Responses

“Stop abusing my math!”

Does the hole argument respect the mathematics of general relativity?

Curiel (2018) and Weatherall (2018) claim that the hole argument hinges on a misleading or erroneous interpretation of the relevant mathematics.

This response is more technically subtle and involves scrutinising the way in which the hole argument deploys the mathematical formalism of GR.

Four Responses

“Stop abusing my math!”

This type of response rejects **P1.2 (Models)** on the basis that it abuses the mathematical formalism:

> **P1.2 (Models)**
> If the model $\mathcal{M}$ = < *M*, g_{ab}, T_{ab} > can be chosen to represent a spacetime *P* then relative to that choice there is a permissible and natural interpretation of $d^*\mathcal{M}$ according to which it represents a distinct spacetime *P'*.

Namely, these responses reject that there is a “*natural*” joint interpretation of $\mathcal{M}$ and $d^*\mathcal{M}$ according to which they represent physically distinct possibilities.

Very roughly, we can think of these responses as rejecting premise 1.2, Models,

Four Responses

“Stop abusing my math!”

Curiel (2018) argues that:

“Just because the mathematical apparatus of a theory appears to admit particular mathematical manipulations does not *eo ipso* mean that those manipulations admit of physically significant interpretation.” **(p.6)**

“Nothing I can see militates in favour of taking the hole argument as bearing on the ontic status of spacetime points, just because the hole argument by itself provides no independent, clear, and precise criterion for what ‘existence independent of metrical structure’ comes to. That idea has no substantive content on its own.” **(p.9)**

There are multiple plausible candidates for such a criterion, and they pull in opposing ways, Curiel claims.

Four Responses

"Stop abusing my math!"

For Curiel, the argument is simply not well-posed:

- Mathematical arbitrariness is a ubiquitous feature of the formalism of empirically successful scientific theories.
- Diffeomorphic models need not represent anything interesting.
- Moreover, diffeomorphism invariance does not need philosophical explanation since it does not impede investigation of "what is of true physical significance in systems that general relativity models" (Curiel 2018, p.7)
- Be a substantivalist, or don't. Be a relationalist, or don't. It makes no substantive difference to the physical content of GR.

He takes this to indicate that the debate is just ill-posed.

Mathematical arbitrariness features in all of our physical theories. It doesn't necessarily mean that reality is arbitrary or indeterministic, too. To quote Erik from I think lecture 3, "we tell math what to do!"

Diffeomorphic models might just not represent anything of physical import.

And, as we've seen, since diffeomorphisms preserve the empirically significant structures in a spacetime model, diffeomorphism invariance doesn't seem to make a difference to matters empirical. So, he claims, why think that it needs philosophical explanation?

This is really in the spirit of Maxwell: do metaphysics only insofar as it is necessary for the physics, and no further.

In other words, be a substantivalist, or don't be. Be a relationalist, or don't be. It makes absolutely no substantive difference to the physical content of GR, and therefore a 'metaphysical' evaluation of it would be in direct tension with Maxwell's methodology as described in earlier lectures.

Four Responses

"Stop abusing my math!"

But…

- Is there any room left for philosophy, or only for 'careful physics'?
- How do we come to know what is of physical significance and what is not?
- Could there still be non-empirical, non-metaphysical reasons to settle questions like this one? For example: clarity, fruitfulness, beauty, etc.

Even if we are all for metaphysical naturalism, and philosophy respecting the primacy of physics, or whatever, one might worry that Curiel's argument completely obliterates philosophical questions. If we prohibit questions that do not pertain to the way that physical theories arbitrate on empirical matters, then we might think that's what is left is no longer philosophy, but rather 'careful physics'.

Possibly, Curiel would endorse this idea: possibly, the idea is that in an ideal world, every physicist would be careful and philosophically sophisticated enough to be able to do this kind of careful technical work themselves, but if that is not possible, then philosophers of physics could do it. Possibly the idea is more pragmatic: it is a simple division of labour.

The question remains: is the division of labour accidental or essential?

There are also other questions we might have:

- What account can be given of what is physically significant? How does one determine when to use the theory's formalism as a guide for interpretation and when to constrain interpretation on grounds of physical significance?

- Could there be non-empirical but non-metaphysical ways to arbitrate these disputes? Can we settle the hole argument on grounds of (say) clarity, fecundity, fruitfulness or beauty?

Four Responses

“Stop abusing my math!”

Weatherall (2018) argues that:

- “Familiar mathematical objects—groups, manifolds, graphs—are only defined up to isomorphism” **(p.331)**
- We should apply the same principle to standards to those mathematical objects when and where they apply in physics.

A similar line of argument is given by Weatherall, who also considers the hole argument to be ill-posed.

In particular, Weatherall claims that since the relevant mathematical machinery in the hole argument only deals in identity *up to isomorphism*, so too should interpretations of that mathematical machinery.

In other words, we must apply the same principle to applications of those mathematical objects.

Four Responses
"Stop abusing my math!"

"Once one asserts that spacetime is represent by a Lorentzian manifold, one is committed to taking isometric spacetimes to have the capacity to represent the same physical situations, since isometry is the standard of isomorphism given in the mathematical theory of Lorentzian manifolds. To deny this would be, in effect, to insist that it is some other structure—one that is not preserved by isometries—that represents spacetime in relativity theory."

(Weatherall 2018, p.343)

And this isn't just purism, Weatherall maintains. It's necessary for consistency.

Four Responses
“Stop abusing my math!”

- We *cannot* deny Leibniz equivalence without abusing the relevant mathematics.
- **An argument against substantivalism, but for stronger reasons:**
  Substantivalism comes not just at the cost of determinism, but at the cost of fundamentally abusing the mathematics of GR.
- Weatherall’s argument (like Curiel’s) is that the hole argument is not well-posed.
- The substantivalist must deny that diffeomorphic spacetimes are equivalent which commits them to spacetimes not being general relativistic models, but something else entirely!
- Substantivalists must specify what the desired additional spacetime structure must be.

The idea here is this:

We cannot deny Leibniz equivalence without abusing the mathematical formalism.

This yields an argument against substantivalism on stronger grounds than what we’ve seen so far. Substantivalism not only results in indeterminism, but it also presupposes a fundamental misuse of the mathematics of general relativity.

This makes the hole argument fundamentally ill-posed, since the substantivalist seems to simultaneously maintain that spacetimes are relativistic spacetime models while trying to deny them one of the central features of such models: identity only up to isomorphism.

Possibly, the substantivalist is actually denying that spacetimes are relativistic models at all. But if that is the case, they have to accept a much greater burden of proof: they have to say what the alternative structure of spacetime actually is. They have not done so and it is far from clear that doing so is even possible.

Four Responses

“Stop abusing my math!”

Do Curiel and Weatherall pull in different directions?

- Is the formalism supposed to adjudicate these decisions?
- Or is fruitful contact with empirical reality (“physical significance”) supposed to constrain interpretation of the formalism?

Though Curiel and Weatherall seem to arrive at similar conclusions, they appear to have different understandings of the relationship between physics and mathematics.

Curiel’s argument turned upon the conviction that an interpretation of a physical theory’s mathematical formalism required us to pay attention to which features of that formalism were “physically significant”. While it is difficult to make such a notion precise, the seemed to be roughly that the features of the mathematics that allowed the theory to make fruitful contact with empirical reality is what we should take to be “physically significant”. This ends up undermining the hole argument, since diffeomorphism invariance is simply considered not to have consequences for GR’s empirical content.

The idea here is that the hole argument comes from paying *too much attention* to the mathematical formalism.

Weatherall’s argument, on the other hand, seems to require us to import the mathematical features of the formalism wholesale. On Weatherall’s view, the hole argument comes about from insufficient attention to the formalism. Since isomorphism is the relevant standard of sameness for the mathematics, isomorphic spacetime models must be considered physically identical to one another on pain of abusing the mathematics of the theory.

The idea here seems to be that the hole argument comes from paying *too little attention* to the mathematical formalism.

Perhaps there is a way to reconcile these approaches: both claim that once attention is paid to the way in which physicists and mathematics actually *use* GR, features of theory that undermine the hole argument reveal themselves. For Curiel, these features have to do with the physical significance afforded to different aspects of the formalism. For Weatherall, they have to do with standards of sameness and difference encoded in that formalism. Both accounts hinge upon the conviction that the hole argument may only proceed from a faulty understanding of the interplay between mathematics and physics.

Summary

Summary

- One of our most powerful and well empirically supported theories of physics has important and far-reaching implications for some traditionally ‘metaphysical’ questions.
- Either spacetime is ontologically on a par with matter, but the universe fails to be deterministic...
- ...or the universe is deterministic but spacetime is relegated to an impoverished type of existence (or perhaps no existence at all).

Summary

- As good 'naturalistic' metaphysicians, we should be wary of choosing a bullet to bite without sufficient reason.
- To paraphrase Earman & Norton (1987), determinism should fail because physics demands it, not because we can't bear to relinquish substantivalism. But nor should determinism stay on the basis of metaphysical dogma.
- Nevertheless, there may be room to salvage some type of substantivalism *without* resorting to dogmatic attachment to metaphysical intuition.
- We have seen three (non-bullet biting) strategies for doing so...

Summary

1. **Sophisticated substantivalism** argues that the hole argument does not militate against substantivalism *simpliciter*, only a haecceitistic substantivalism that assigns primitive identity to spacetime points.

   Surrendering haecceitism about spacetime points blocks the hole argument.

2. **Metrical essentialism** argues that it is metaphysically impossible for spacetime points to instantiate different geometrical properties.

   Maintaining that spacetime points bear their geometrical properties essentially blocks the hole argument.

3. **"Stop abusing my math"** responses claim that the hole argument cannot proceed without doing violence to the mathematical formalism in terms of which it is expressed.

   Respecting the relationship between mathematics and physics blocks the hole argument.

The End

- Next week Prof. Neil Dewar on Lewisian counterfactuals
- Discussion group Friday 13:00 - 14:00

References

Curiel, E. (2018). On the Existence of Spacetime Structure. *The British Journal for the Philosophy of Science*, *69*(2), 447-483.

Earman, J., & Norton, J. (1987). What Price Spacetime Substantivalism? The Hole Story. *The British Journal for the Philosophy of Science*, *38*(4), 515-525.

Hoefer, C. (1996). The Metaphysics of Space-Time Substantivalism. *The Journal of Philosophy*, *93*(1), 5-27.

Maudlin, T. (1988, January). The Essence of Space-Time. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* (Vol. 1988, No. 2, pp. 82-91). Philosophy of Science Association.

Norton, John D., "The Hole Argument", *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/sum2019/entries/spacetime-holearg

Pooley, O. (2020). The Hole Argument. *arXiv preprint arXiv:2009.09982*.

Weatherall, J. O. (2018). Regarding the 'Hole Argument'. *The British Journal for the Philosophy of Science*, *69*(2), 329-350.