What Price Spacetime Substantivalism?  
The Hole Story

by JOHN EARMAN and JOHN NORTON

1 Introduction
2 Local Spacetime Theories
3 What is Spacetime Substantivalism? Denial of Leibniz Equivalence
4 The Verificationist Dilemma
5 The Indeterminism Dilemma

Spacetime substantivalism leads to a radical form of indeterminism within a very broad class of spacetime theories which include our best spacetime theory, general relativity. Extending an argument from Einstein, we show that spacetime substantivalists are committed to very many more distinct physical states than these theories' equations can determine, even with the most extensive boundary conditions.

1 INTRODUCTION

Since the time of Newton, those who hold a substantivalist view of space and time have had to address the following dilemma. They must either

(a) allow that there are distinct states of affairs which no possible observation could distinguish or

(b) give up their substantivalism.

Thus Leibniz asked Clarke how the world would differ if God had placed the bodies of our world in space in some other way, only changing for example East into West. There would be no discernible difference. Our belief that there was a difference would be based on the 'chimerical supposition of the reality of space itself' (Alexander [1956], p. 26). In the modern context, an analogous dilemma arises for spacetime substantivalists. But with the demise of the verifiability criterion of meaning, it is no longer unfashionable for them to escape the dilemma by simply allowing (a).

Substantivalists were led to this dilemma through their insistence that unobservable spatial and temporal properties of matter (e.g. 'is at position x') are not reducible to observable relational properties of matter (e.g. coincidence, betweenness). Relationists seize upon what they regard as a superfluous inflation of their ontology and force substantivalists to commit

Received December 1986
themselves to the distinctness of observationally indistinguishable states of affairs.

In the context of modern spacetime theories, this overcommitment leads substantivalists to a new dilemma. Either they must reject substantivalism or they must accept a very radical form of indeterminism. The examination of how this dilemma arises is the subject of this paper.

The class of spacetime theories concerned is a very wide and important one. In brief the theories posit a differential spacetime manifold upon which fields are defined. The behaviour of the fields is determined exclusively by partial differential field equations. The class includes Newtonian spacetime theories with all, one or none of gravitation and electrodynamics; and special and general relativity, with and without electrodynamics. What is most significant is that all versions of our best theory of space and time, general relativity, belong to the class. Thus substantivalists must face the indeterminism dilemma if they believe our best theory of space and time.

In developing the dilemma, we shall see that the equations of these theories are simply not sufficiently strong to determine uniquely all the spatio-temporal properties to which the substantivalist is committed. The type of indeterminism involved will be a very radical one indeed. Given some neighbourhood of spacetime we shall see that these theories cannot uniquely determine the fields within the neighbourhood from even the most exhaustive prescription of the fields outside of it. This is true no matter how small the neighbourhood. We have christened this behaviour 'radical local indeterminism'. We believe that this radical form of indeterminism is a very heavy price to pay for a doctrine that adds no new predictive power to our spacetime theories.

The indeterminism dilemma arises from a very general form of gauge freedom in the spacetime theories discussed. This gauge freedom manifests itself in the general covariance of the theories' equations. General covariance can be understood in the usual passive sense as the form invariance of these equations under arbitrary spacetime coordinate transformation. Viewed passively, the choice of a gauge is merely a restriction on the spacetime coordinate systems which can be used. This obscures the connection between determinism and the gauge freedom. However the dual active interpretation of general covariance makes the connection much clearer. It is expressed as a gauge freedom in the theory's models.¹

That this freedom could lead to radical local indeterminism was discovered by Einstein late in 1913 in the form of the so-called 'hole argument'.² He did not see how to deal with the resulting dilemma until

¹ See Stachel [1985] for a treatment of general covariance in this active sense. Stachel also maintains a distinction between absolute and dynamic objects and focuses on the concerns of Einstein's 'hole argument'.

² See for example Einstein and Grossmann [1913], pp. 260–1, and a clearer version in Einstein [1914], pp. 1066–7. Stachel was the first to see clearly that Einstein's active reading of general covariance made the 'hole argument' non-trivial. Stachel [1980].
late in 1915. Our purpose here is not to present an historically faithful version of Einstein’s argument, which has been discussed elsewhere (see Norton [1987]). We intend our argument to stand by itself, although we wish to make its ancestry known.

2 LOCAL SPACETIME THEORIES

We begin by describing the general form of spacetime theories in which we shall derive the indeterminism dilemma. These theories posit differentiable manifolds on which geometric objects are defined at every point. A model of one of these theories will always be an $n+1$ tuple $\langle M, O_1, \ldots, O_n \rangle$. $M$ is a differentiable manifold with all the usual intrinsic structure and $O_1, \ldots, O_n$ are $n$ geometric objects, defined everywhere on $M$, for some positive integer $n$.

Each model will satisfy a set of field equations, which are just the vanishing of a subset of the objects defined. That is for some positive integer $k$ less than or equal to $n$, the field equations are

$$O_k = 0, O_{k+1} = 0, \ldots, O_n = 0$$

We require that each of the objects in the field equations be tensors.

Since we allow that some of the objects can be constructed from others already defined, this prescription is sufficiently general to include versions of just about every classical field theory of interest to us. For example special relativistic electromagnetics has models of the form

$$\langle M, g_{ab}, D_a, F_{ab}, j^a, D_b g_{bc}, R^a_{bcd}, D_{[a} F_{bc]}^a, D_a F^{ac} - j^a \rangle$$

$g_{ab}$ is a metric tensor of Lorentz signature, $D_a$ a derivative operator, $F_{ab}$ the Maxwell field tensor, $j^a$ the charge flux and $R^a_{bcd}$ the curvature tensor of the metric $g_{ab}$. For this version of the theory, the field equations are the vanishing of $O_5$ to $O_8$. The vanishing of $O_5$ adapts the derivative operator to the metric and the vanishing of $O_6$ forces $g_{ab}$ to be flat. The final two equations are Maxwell’s equations.

We shall call a spacetime theory a ‘local spacetime theory’ if it has the above form and satisfies the completeness condition:

**Completeness condition** If a spacetime theory has models of the form

$$\langle M, O_1, \ldots, O_n \rangle$$

which satisfy field equations

$$O_k = 0, O_{k+1} = 0, \ldots, O_n = 0$$

then every $n+1$ tuple of this form which satisfies the field equations is a model of the theory.

We consider only local spacetime theories

The dilemmas developed below arise in local spacetime theories. The premier instance of such a theory is our current best theory of space and
time, general relativity. All known formulations of general relativity are local spacetime theories or formulations which reduce to one.\footnote{A variational formulation of general relativity does not have tensor field equations, as required by local spacetime theories. However such field equations are readily derived from its basic action principle.} Thus a spacetime substantivalist who believes general relativity cannot avoid the dilemmas.

Virtually every other classical spacetime field theory can be formulated as a local spacetime theory. We prefer wherever possible to formulate them as such. So we take special relativity to have models $\langle M, g_{ab}, R^*_{bcd} \rangle$ where $g_{ab}$ can be any of many possible Minkowski metrics definable on $M$, which satisfy the field equations $R^*_{bcd} = 0$. Thus the completeness condition is satisfied.

This is by no means a universal practice, especially in older work. Alternatively, one could insist that special relativity deals with just one Minkowski spacetime, which is a pair $\langle N, n_{ab} \rangle$, where $n_{ab}$ is a particular Minkowski metric defined on $N$, an $R^4$ manifold. What is worrisome about this alternate portrayal of special relativity is that it starts out by making unnecessary global assumptions. We must stipulate in the laws of the theory itself what the global manifold topology is to be and incorporate in these laws one of the many Minkowski metrics definable on the manifold.

The success of general relativity has promoted the formulation of spacetime theories as local spacetime theories. Such formulations make comparison between general relativity and these other theories much easier.\footnote{For formulations as local spacetime theories of many versions of Newtonian and special relativistic spacetime theories, see Friedman [1983].}

We also believe that there are good but not compelling reasons to formulate spacetime theories as local spacetime theories. Cosmology has always been a far riskier enterprise than local physics. Since the time of Aristotle, we have found that the weakest part of a physical theory is the global cosmological assumptions it makes. We have learned to our cost that it is better to do local physics first and build one's cosmology from it, rather than the other way round. In rendering theories as a local spacetime theory, we abide by this heuristic. We determine all the fields on the manifold by local field equations, not global stipulation, and we allow the possibility of global topologies other than the usual standby of $R^n$.

**What represents spacetime?**

What structure in spacetime theories represents spacetime? That is, of what does the spacetime substantivalist hold a substantivalist view? We view the manifolds $M$ of the models as representing spacetime.

This view follows naturally from the local formulation of spacetime theories. We take all the geometric structure, such as the metric and derivative operator, as fields determined by partial differential equations. Thus we look upon the bare manifold—the 'container' of these fields—as space-
time. A repeated problem in the literature on spacetime substantivalism is a failure to specify clearly the structure to which substantival properties are ascribed. A welcome exception is Friedman [1983], chapter VI, where the manifold is identified as spacetime and it is argued that we should hold a realist view of it.

The advent of general relativity has made most compelling the identification of the bare manifold with spacetime. For in that theory geometric structures, such as the metric tensor, are clearly physical fields in 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. The pseudo-tensorial nature of this quantity has made its status problematic. But it can still be seen that energy and momentum are carried by the metric in a way that forces its classification as part of the contents of spacetime. Consider, for example, a gravitational wave propagating through space. In principle its energy could be collected and converted into other types of energy, such as heat or light energy or even massive particles. If we do not classify such energy bearing structures as the wave as contained within spacetime, then we do not see how we can consistently divide between container and contained. We might consider dividing the metric into an unperturbed background and a perturbing wave in the hope that the latter alone can be classified as contained in spacetime. This move fails since there is no non-arbitrary way of effecting this division of the metric. Finally, classifying the metric as part of the container spacetime leads to trivialisation of the substantivalist view in unified field theories of the type developed by Einstein, in which all matter is represented by a generalised metric tensor. For there would no longer be anything contained in spacetime, so that the substantivalist view would in essence just assert the independent existence of the entire universe.

In an alternate view usually associated with Newtonian or special relativistic theories, one represents spacetime by the manifold with some additional geometric structure, which we shall call its absolute structure. This view arises most naturally in the older non-local formulations of spacetime theories, in which case the absolute structure is typically posited globally ab initio rather than being defined locally through field equations. Thus if one gives the above global formulation of special relativity, one would probably call spacetime the pair \( \langle N, n_{ab} \rangle \). Or if one had to identify a structure in a Newtonian spacetime which corresponded to the thing about which Newton held his substantivalist view, then that would be the tuple \( \langle N, h^{\kappa}, D_{a}, dt_{a} \rangle \), where \( h^{\kappa} \) is the degenerate metric, \( D_{a} \) the derivative operator and \( dt_{a} \) the absolute time one form (all defined in the usual manner).

---

1 One of us has argued that a primary outcome for Einstein of the principle of equivalence was the recognition that the Minkowski metric \( g_{ab} \) of special relativity was a physical field defined in spacetime, rather than a part of the background of spacetime. Norton [1985].
Our present argument does not address this representation of spacetime since it is commonly associated with non-local spacetime theories already beyond our compass. We note in passing that the hybrid view—using this representation of spacetime within local spacetime theories—still leads to dilemmas of the type discussed below, but they are harder to set up. (See note 2 on p. 522.)

**The Gauge Theorem**

The indeterminism dilemma depends on the following theorem:

**Gauge Theorem** (General covariance): 1 If \( \langle M, O_1, \ldots, O_n \rangle \) is a model of a local spacetime theory and \( h \) is a diffeomorphism from \( M \) onto \( M \), then the carried along tuple \( \langle M, h \cdot O_1, \ldots, h \cdot O_n \rangle \) is also a model of the theory.

**Proof** We need to establish that the vanishing of the field equations

\[
O_k = 0, \quad O_{k+1} = 0, \ldots, O_n = 0
\]

is preserved under diffeomorphism. This follows immediately from the description of the action of the carry along \( h \cdot \) in coordinate terms. For any object \( O_i \) with components \( (O_i)^m \) in some coordinate system \( \{x^m\} \) we have

\[
(O_i)^m = (h \cdot O_i)^{m'}
\]

where the superscript \( m' \) indicates components in the carried along coordinate system \( \{x^{m'}\} = \{h \cdot x^m\} \). Recall that \( O_i \) is a tensor. Therefore \( (O_i)^m = 0 \) and thus \( (h \cdot O_i)^{m'} = 0 \) as well. Therefore \( h \cdot O_i \) vanishes. This argument holds for \( i = k, k+1, \ldots, n \), which establishes that the field equations hold for the carried along tuple.

Notice that the proof depends on the objects being tensors, which have the property of vanishing just in case their components vanish in any coordinate system. This is why we restricted the field equations of local spacetime theories to tensorial equations.

We shall say that the original model and the carried along model are diffeomorphic. Note that the relation of being diffeomorphic divides the set of tuples into equivalence classes.

To see the connection between this gauge theorem and general covariance in its usual passive reading, recall that there is a natural one-one correspondence between diffeomorphisms on \( M \) and coordinate transformations of a particular coordinate system \( \{x^m\} \) of \( M \). Let the diffeomorphism \( h \) map the point \( p \) of \( M \) to \( h(p) \). Then the corresponding coordinate

---

1 This result is not new, although it is commonly known through its passive form. Wald writes 'the diffeomorphisms comprise the gauge freedom of any theory formulated in terms of tensor fields on a spacetime manifold'. Wald [1984], p. 438.
transformation assigns the new coordinates \( \{ x^m \} \) to \( p \), where the values of \( \{ x^m \} \) at \( p \) are equal to the coordinates of \( hp \) in the original coordinate system \( \{ x^m \} \).

Using this correspondence one can translate theorems from the active to the passive language— that is from theorems dealing with diffeomorphisms to theorem dealing with coordinate transformations—and vice versa. The gauge theorem follows immediately from the vanishing of the carry along under arbitrary diffeomorphism of vanishing tensors. This result corresponds to the passive result that the components of these zero tensors remain zero under arbitrary coordinate transformation, which is just the generally covariant transformation law for the components of a zero tensor.

3 WHAT IS SPACETIME SUBSTANTIVALISM?:
DENIAL OF LEIBNIZ EQUIVALENCE

In broad outline, the spacetime substantivalist holds that spacetime can exist independently of any of the things in it. In this form, the thesis is disastrous, because it is automatically denied by every spacetime theory with which we deal. They all postulate that there are always fields at every point in spacetime. That is, they agree that there cannot be unoccupied spacetime events, contrary to the standard position taken by substantivalists against relationists.

We can imagine many less problematic ways of reformulating the substantivalist thesis. We might consider the thesis that spacetime is not reducible to other structures; or the thesis that we must unavoidably quantify over spacetime events in our spacetime theories. Perhaps we might consider a strict realist reading of the models of spacetime theories. Each model no longer represents a physically possible world. Rather each model is a physically possible world, one of them being our world. That is the M of one model of a true spacetime theory is the spacetime of our world.

Fortunately we do not need to settle this reformulation problem. 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.

The necessary agreement of substantivalists on this test is all we shall need to arrive at the dilemmas below. But first we must translate the test into the context of local spacetime theories. The diffeomorphism is the counterpart of Leibniz' replacement of all bodies in space in such a way that their relative relations are preserved. For example, represent two bodies in a local spacetime theory by two spatially small regions of high energy
density in the obvious way. Then all their relative properties, such as the spacetime interval separating them and their relative velocities upon collision, remain unchanged under arbitrary diffeomorphism.

In sum, substantivalists, whatever their precise flavour, will deny:

*Leibniz equivalence*  Diffeomorphic models represent the same physical situation.

This denial already places substantivalists at odds with standard modern texts in general relativity, in which this equivalence is accepted unquestioningly in the specific case of manifolds with metrics.\(^1\) We are now in a position to establish the two dilemmas for spacetime substantivalists.\(^2\)

4 THE VERIFICATIONIST DILEMMA

This dilemma amounts to little more than a restatement of the substantivalists' denial of Leibniz equivalence. To complete the dilemma we need only note that spatio-temporal positions by themselves are not observable. Observables are a subset of the relations between the structures defined on the spacetime manifold. Thus we cannot observe that body b is centred at position x. What we do observe are such things as the coincidence of body b with the x mark on a ruler, which is itself another physical system. Thus observables are unchanged under diffeomorphism. Therefore diffeomorphic models are observationally indistinguishable.

Substantivalists must either deny Leibniz equivalence or deny their substantivalism. That is, they must either

(a) accept that there are distinct states of affairs which are observationally indistinguishable, or

(b) deny their substantivalism.

5 THE INDETERMINISM DILEMMA

To arrive at this dilemma, we need a simple corollary of the gauge theorem:

*Hole corollary*  Let T be a model of some local spacetime theory with

\(^1\) Hawking and Ellis [1973], p. 56; Sachs and Wu [1977], p. 27. This acceptance enables modern treatments of local spacetime theories to avoid radical local indeterminism. Older treatments of classical mechanics and special relativity were not formulated as local spacetime theories. This type of indeterminism was not a problem since they dealt with a single manifold plus absolute structure as the fixed spacetime canvas in which the gauge freedom cannot arise.

\(^2\) Manifold-plus-absolute-structure substantivalists will typically face dilemmas of similar origin. M-p-a-s substantivalists are subject to the Leibniz acid test just in case their absolute structure has symmetries, which is overwhelmingly the case. They must deny that two models represent the same physical situation if they are diffeomorphic under a symmetry transformation. This naturally generalises to the denial of Leibniz equivalence (and the dilemmas below). The generalisation is difficult to avoid. Translational symmetries, for example, can be composed out of hole diffeomorphisms. Thus the affected m-p-a-s substantivalists must deny Leibniz equivalence at least for hole diffeomorphisms, which already is sufficient to yield the dilemmas through the hole corollary of Section 5.
What Price Spacetime Substantivalism? 523

manifold M and H (for hole) any neighbourhood of M. Then there exist
arbitrarily many distinct models of the theory on M which differ from
one another only within H.

Proof Let h be a ‘hole diffeomorphism’, i.e., one which differs from the
identity diffeomorphism within H, but smoothly becomes the identity
on the boundary and outside H. Then from the gauge theorem, the carry
along of T under h satisfies the requirement. Since there are arbitrarily
many hole diffeomorphisms for H, there are arbitrarily many such carry
along models satisfying the requirement.

The name of this corollary stems from Einstein’s original discovery of it in
a specialised form. He considered a matter free hole in a source mass
distribution and showed that the gauge freedom of any generally covariant
gravitational field equation in general relativity allowed multiple metric
fields within the hole.

It now follows immediately that the substantivalists’ denial of Leibniz
equivalence leads to a very radical form of indeterminism in all local space-
time theories, since for a substantivalist the diffeomorphic models of the
hole corollary must represent different physical situations.

Consider first various forms of Laplacian determinism. Suppose that the
spacetime models in question admit global time slices.1 In the Newtonian
setting such a slice is a hyperplane of absolute simultaneity, while in the
relativistic setting it is a spacelike hypersurface without edges. The Lapla-
cian would then like to prove that the laws of physics guarantee that the
state on a time slice S uniquely fixes the state to the future of S; or failing
that, the state on a finite sandwich lying between two slices S and S’ fixes
the state to the future of the sandwich; or failing that, the state on S and to
the past of S fixes the state to the future of S.

If spacetime is substantival, no such proof can be forthcoming within
local spacetime theories. For by the hole corollary with The Hole placed
in the future of S, if \langle M, O_1, O_2, \ldots \rangle is a model of our theory, then there
is another model \langle M, O_1', O_2', \ldots \rangle which is identical with the first up to
and including the instant corresponding to S (i.e., for any p in M which
lies to the past of S, O_1(p) = O_1'(p)) but which diverges from the first to the
future of S.

It is worth noting that, contrary to the common wisdom, Laplacian
determinism typically does not obtain a clean form in Newtonian theories.
See Earman [1986]. Intuitively, Laplacian determinism breaks down
because there is no upper bound on the velocity of causal propagation,
with the result that influences can ‘sneak in’ from spatial infinity without
announcing themselves on the chosen slice S.

In the face of such space invaders one might hope to achieve a non-trivial
form of determinism by shifting from a pure initial value problem to a

1 Otherwise, the global version of Laplacian determinism does not apply.
boundary-initial value problem. That is, the state is specified on S itself and also the walls of a tube which cuts through all the time slices in the future of S. The hope is that these boundary conditions will determine a unique interior for the tube amongst the models of the theory. But assuming substantivalism, the hole corollary dashes these hopes. Just place The Hole within the tube.

By now the reader has no doubt seen that the hole corollary forces substantivalists to conclude that no non-trivial form of determinism can obtain in local spacetime theories. The state within any neighbourhood of the manifold can never be determined by the state exterior to it, no matter how small the neighbourhood and how extensive the exterior specification.

Of course this radical local indeterminism can be escaped easily by just accepting Leibniz equivalence. Then the diffeomorphic models of the hole corollary represent the same physical situation and the indeterminism discussed becomes an underdetermination of mathematical description with no corresponding underdetermination of the physical situation. But accepting Leibniz equivalence entails denying substantivalism.

We emphasise that our argument does not stem from a conviction that determinism is or ought to be true. There are many ways in which determinism can and may in fact fail: space invaders in the Newtonian setting; the non-existence of a Cauchy surface in the general relativistic setting; the existence of irreducibly stochastic elements in the quantum domain, etc. Rather our point is this. If a metaphysics, which forces all our theories to be deterministic, is unacceptable, then equally a metaphysics, which automatically decides in favour of indeterminism, is also unacceptable. Determinism may fail, but if it fails it should fail for a reason of physics, not because of a commitment to substantival properties which can be eradicated without affecting the empirical consequences of the theory.

In sum we have shown that substantivalists must either deny Leibniz equivalence or deny their substantivalism. That is, they must either

(a) accept radical local indeterminism in local spacetime theories or
(b) deny their substantivalism.

Perhaps it is acceptable to save substantivalism in the verificationist dilemma by accepting option (a). But we feel that the price one has to pay in accepting the option (a) in the indeterminism dilemma is far too heavy a price to pay for saving a doctrine that adds nothing empirically to spacetime theories.

Dept. of History and Philisophy of Science, University of Pittsburgh

1 See Hawking and Ellis [1973] for a definition of this concept.
2 We have not concluded here that spacetime is relational, since the literature contains so many conflicting usages of the term 'relationism'. Of course the conclusion is established if relationism is just the negation of substantivalism. But this presupposes far too crude a dichotomy—substantivalism versus relationism—from which discussion of these matters has suffered too long. Relationism is not established if it implies that all motion is the relative motion of bodies, as Leibniz apparently held.
REFERENCES


