Spacetime Structure

Thomas William Barrett[*][†]

Abstract

This paper makes an observation about the "amount of structure" that different classical and relativistic spacetimes posit. The observation substantiates a suggestion made by Earman (1989) and yields a cautionary remark concerning the scope and applicability of structural parsimony principles.

1 Introduction

There is a story that is often told about the progression of classical spacetime theories.[1] We began long ago with Aristotelian spacetime. Aristotelian spacetime singles out a preferred worldline as the *center of the universe*. Then we moved to Newtonian spacetime and did away with this structure. Newtonian spacetime does not single out a preferred worldline, but it does single out a preferred inertial frame as the *rest frame*. Finally, we moved to Galilean spacetime and again did away with structure. Galilean spacetime does not even single out a preferred inertial frame.

This story provides a sense in which each of these classical spacetimes has "less structure" than its predecessors. It is natural to ask whether this progression towards less structure continues in the transition between classical and relativistic spacetimes. The purpose of this paper is to answer this question by investigating the structural relationships that hold between Galilean, Newtonian, and Minkowski spacetime. There is a precise sense in which Newtonian spacetime has more structure than both Galilean spacetime and Minkowski spacetime. But in this same sense, Galilean and Minkowski spacetime have *incomparable* amounts of structure; neither spacetime has less structure than the other. The progression towards a less structured spacetime therefore does not continue into the relativistic setting.

This discussion of spacetime structure will yield two modest philosophical payoffs. First, it will substantiate a remark made by Earman (1989). Earman

[*]thomaswb@princeton.edu

[†]I am especially grateful to Jim Weatherall for comments on multiple drafts and for all of his help and encouragement. Thanks also to Jeff Barrett, Shamik Dasgupta, Hans Halvorson, David Malament, JB Manchak, and the audience at IPP 2014 for helpful discussion.

[1]See for example Geroch (1978), Earman (1989), and Maudlin (2012).

has suggested, somewhat paradoxically, that Newtonian spacetime is a more natural stepping-stone to relativistic spacetimes than Galilean spacetime is. This discussion will provide one way of making Earman's suggestion perfectly precise. Second, this discussion will also yield a cautionary remark concerning the scope and applicability of the following methodological principle.

Structural parsimony: All other things equal, we should prefer theories that posit less structure.

The structural parsimony principle has been explicitly endorsed by North (2009). This paper will provide an example of two physical theories that posit incomparable amounts of structure. In such cases, a structural parsimony principle is not applicable.

2 Structure preliminaries

We begin by explicating the idea of the "amount of structure" that a mathematical object has. We would like a clear and principled way to say when some mathematical object X has more or less structure than another mathematical object Y. One particularly natural way to compare amounts of structure appeals to the automorphisms, or symmetries, of a mathematical object.

An automorphism of a mathematical object X is an invertible function from X to itself that preserves all of the structure of X. The automorphisms of an object bear a close relationship to the structure of the object. This relationship suggests the following kind of criterion for comparing amounts of structure.

SYM: A mathematical object X has more structure than a mathematical object Y if the automorphism group $\text{Aut}(X)$ is "smaller than" the automorphism group $\text{Aut}(Y)$.

The basic idea behind SYM is clear. If a mathematical object has more automorphisms, then it intuitively has less structure that these automorphisms are required to preserve. Conversely, if a mathematical object has fewer automorphisms, then it must be that the object has more structure that the automorphisms are required to preserve. The amount of structure that a mathematical object has is, in some sense, inversely proportional to the size of the object's automorphism group.[2]

The criterion SYM is intuitive, but it is imprecise. One way to make SYM precise is as follows.

SYM*: A mathematical object X has more structure than a mathematical object Y if $\text{Aut}(X) \subsetneq \text{Aut}(Y)$.

[2]A criterion like SYM is used by North (2009) to compare the structure of Hamiltonian and Lagrangian mechanics and by Earman (1989) to compare various classical spacetime theories. SYM is mentioned explicitly by Halvorson (2011), Swanson and Halvorson (2012), and Barrett (2014).

The condition $\text{Aut}(X) \subsetneq \text{Aut}(Y)$ is one way to make precise the idea that $\text{Aut}(X)$ is "smaller than" $\text{Aut}(Y)$. SYM* makes intuitive verdicts in many simple cases of structural comparison. For example, one can verify that in general SYM* makes the following verdicts:

- A set X has less structure than a group $(X, \cdot)$.
- A set X has less structure than a topological space (X, τ).
- A vector space V has less structure than an inner product space (V, g).

The criterion SYM* is one particularly natural way to explicate the idea that an object X has "more structure" than another object Y. In what follows, we will use SYM* to compare the structure of Galilean, Newtonian, and Minkowski spacetime.[3] Earman has implicitly used SYM* to compare the structure of various classical spacetimes (Earman, 1989, Ch. 2). And indeed, SYM* makes the intuitive verdicts in all of these cases. Galilean spacetime has less structure than Newtonian spacetime, which in turn has less structure than Aristotelian spacetime. This paper simply extends Earman's discussion into the relativistic setting.

3 Spacetime preliminaries

Before applying SYM* to these spacetimes we need some preliminaries.[4] We first present the standard mathematical descriptions of Galilean, Newtonian, and Minkowski spacetime, and then discuss their automorphisms.

3.1 Spacetimes

Spacetime theories begin by specifying a smooth, connected, four-dimensional manifold M. Each point $p \in M$ represents the location of an "event" in spacetime. Galilean, Newtonian, and Minkowski spacetime all have the underlying manifold $M = \mathbb{R}^4$. They then endow $\mathbb{R}^4$ with different geometric structures.

Galilean spacetime is the tuple $(\mathbb{R}^4, t_{ab}, h^{ab}, \nabla)$. The smooth tensor fields t_{ab} and h^{ab} and the derivative operator ∇ are defined as follows,

$$t_{ab} = (d_a x^1)(d_b x^1)$$
$$h^{ab} = \left(\frac{\partial}{\partial x^2}\right)^a \left(\frac{\partial}{\partial x^2}\right)^b + \left(\frac{\partial}{\partial x^3}\right)^a \left(\frac{\partial}{\partial x^3}\right)^b + \left(\frac{\partial}{\partial x^4}\right)^a \left(\frac{\partial}{\partial x^4}\right)^b$$
$$\nabla \text{ is the coordinate derivative operator on } \mathbb{R}^4,$$

[3] SYM* has some shortcomings. In particular, it is overly sensitive to the set that underlies a mathematical object. There is a sense in which a topological space (X, τ) has more structure than a set Y even when the sets X and Y are distinct. It is not the case, however, that $\text{Aut}(X, \tau) \subset \text{Aut}(Y)$ (since functions from X to itself are different from functions from Y to itself), so SYM* is incapable of capturing this sense. But this will not be problematic for our purposes; all of the spacetimes that we consider have the same underlying set $\mathbb{R}^4$.

[4] The reader is encouraged to consult Malament (2012) for details.

where $d_a x^i$ is the differential of the standard coordinate function $x^i : \mathbb{R}^4 \to \mathbb{R}$ and $\left(\frac{\partial}{\partial x^i}\right)^a$ is the standard i^{th} coordinate vector field on $\mathbb{R}^4$. The coordinate derivative operator ∇ on $\mathbb{R}^4$ is defined to be the unique derivative operator that satisfies $\nabla_a \left(\frac{\partial}{\partial x^i}\right)^b = \mathbf{0}$ for each $i = 1, \ldots, 4$.[5] Importantly, we note that ∇ is flat, in the sense that its curvature field $R^a{}_{bcd} = \mathbf{0}$ everywhere on $\mathbb{R}^4$.

One interprets these geometric structures on Galilean spacetime as follows (Malament, 2012, Ch. 4.1). The field t_{ab} is a "temporal metric." It assigns a temporal length to vectors, and defines a preferred partitioning of Galilean spacetime into "simultaneity slices." The field h^{ab} is a "spatial metric." Given a vector ξ^a, one can use h^{ab} to (indirectly) assign a spatial length to it. Finally, the derivative operator ∇ endows $\mathbb{R}^4$ with a "standard of constancy." It specifies which trajectories through Galilean spacetime are geodesics.

Newtonian spacetime is obtained by adding a preferred notion of "rest" to Galilean spacetime. Specifically, it is the tuple $(\mathbb{R}^4, t_{ab}, h^{ab}, \nabla, \lambda^a)$, where t_{ab}, h^{ab} and ∇ are defined exactly as in Galilean spacetime, and

$$\lambda^a = \left(\frac{\partial}{\partial x^1}\right)^a.$$

The structures t_{ab}, h^{ab}, and ∇ are interpreted as above. The field λ^a singles out a preferred rest frame. It allows one to classify trajectories through Newtonian spacetime as "at rest" or "not at rest." A geodesic $\gamma : I \to \mathbb{R}^4$ with tangent field ξ^a is **at rest** if $\xi^a = c\lambda^a$ for some constant $c \in \mathbb{R}$.

It only remains to define Minkowski spacetime. **Minkowski spacetime** is the pair $(\mathbb{R}^4, \eta_{ab})$, with the Minkowski metric η_{ab} defined by

$$\eta_{ab} = (d_a x^1)(d_b x^1) - (d_a x^2)(d_b x^2) - (d_a x^3)(d_b x^3) - (d_a x^4)(d_b x^4).$$[6]

The metric η_{ab} endows Minkowski spacetime with "lightcone structure." It allows one to classify a vector ξ^a at $p \in \mathbb{R}^4$ as **timelike** (if $\eta_{ab}\xi^a\xi^b > 0$) or **lightlike** (if $\eta_{ab}\xi^a\xi^b = 0$) or **spacelike** (if $\eta_{ab}\xi^a\xi^b < 0$). Timelike vectors at a point $p \in \mathbb{R}^4$ lie on the interior of the lightcone, lightlike vectors lie on the boundary of the lightcone, and spacelike vectors lie outside the lightcone.

3.2 Spacetime automorphisms

In order to use the criterion SYM* to compare the structure of Galilean, Newtonian, and Minkowski spacetime, we need to say what the automorphisms of these three spacetimes are.

An **automorphism of Galilean spacetime** is a diffeomorphism $f : \mathbb{R}^4 \to \mathbb{R}^4$ that satisfies the following three conditions.

[5] For proof that the coordinate derivative operator is unique see (Malament, 2012, Proposition 1.7.11). One can easily verify that Galilean spacetime, so defined, is a classical spacetime in the sense of (Malament, 2012, p. 249).

[6] Note that $\nabla_a \eta_{bc} = \mathbf{0}$, where ∇ is the derivative operator defined above. So the coordinate derivative operator ∇ on $\mathbb{R}^4$ is the unique derivative operator compatible with the Minkowski metric.

(i) $f^*(t_{ab}) = t_{ab}$,

(ii) $f^*(h^{ab}) = h^{ab}$,

(iii) a smooth curve $\gamma : I \to \mathbb{R}^4$ is a geodesic with respect to ∇ if and only if $f \circ \gamma : I \to \mathbb{R}^4$ is a geodesic with respect to ∇.

The first two conditions require that f preserves the temporal metric t_{ab} and the spatial metric h^{ab}. Since a derivative operator is completely characterized by its class of geodesics, the third condition requires that f preserves the derivative operator ∇.[7] An **automorphism of Newtonian spacetime** is an automorphism of Galilean spacetime $f : \mathbb{R}^4 \to \mathbb{R}^4$ that in addition satisfies $f^*(\lambda^a) = \lambda^a$. This additional condition requires that automorphisms of Newtonian spacetime preserve the standard of rest λ^a. Lastly, an **automorphism of Minkowski spacetime** is a diffeomorphism $f : \mathbb{R}^4 \to \mathbb{R}^4$ that preserves the Minkowski metric η_{ab} in the sense that $f^*(\eta_{ab}) = \eta_{ab}$.

There is a particularly natural way to specify automorphisms of a spacetime. One lays down an appropriate smooth vector field ξ^a on the spacetime and then defines an automorphism by "flowing" along the field ξ^a. More precisely, the method is as follows. A smooth vector field ξ^a on a manifold M determines a one-parameter group of diffeomorphisms $\Gamma_t : U \to \Gamma_t[U]$, where $U \subset M$ is some open set. The maps Γ_t are called **flow maps**. Intuitively, the flow map Γ_t takes a point $p \in M$ and moves it t units of parameter distance along the vector field ξ^a.

The flow maps associated with ξ^a are diffeomorphisms, but we also need to say when they preserve the various geometric structures that the manifold M might have. The following two facts do exactly that. The first fact states the conditions under which flow maps preserve various tensor fields on M (Malament, 2012, Proposition 1.6.6).

Fact 1. *Let ξ^a and $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ be smooth fields on M. Then the following conditions are equivalent.*

(1) $\mathcal{L}_\xi \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s} = \mathbf{0}$ (everywhere on M).

(2) For all local one-parameter groups of diffeomorphisms $\{\Gamma_t : U \to \Gamma_t[U]\}_{t \in I}$ generated by ξ^a and all $t \in I$, $(\Gamma_t)^*(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}) = \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$.

The basic idea behind Fact 1 is the following. The Lie derivative $\mathcal{L}_\xi \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ measures how much the tensor field $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ changes along the vector field ξ^a. So it should be that the smooth tensor field $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ is "constant" along ξ^a (in the sense that $\mathcal{L}_\xi \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s} = \mathbf{0}$) if and only if "flowing" along the vector field ξ^a does not change $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$. This is exactly what Fact 1 guarantees.

[7]See (Malament, 2012, Proposition 1.7.8) for proof that a derivative operator is completely characterized by its class of geodesics. Weatherall (2014) also defines automorphisms of a classical spacetime in this way.

The second fact states the conditions under which flow maps preserve derivative operators on M. A proof of Fact 2 is contained in an appendix.[8]

Fact 2. *Let ξ^a be a smooth vector field on M and let ∇ be a derivative operator on M with associated curvature field $R^a{}_{bcd}$. Then the following conditions are equivalent.*

(1) $\nabla_a \nabla_b \xi^m = R^m{}_{bna} \xi^n$.

(2) For all local one-parameter groups of diffeomorphisms $\{\Gamma_t : U \to \Gamma_t[U]\}_{t \in I}$ generated by ξ^a and all $t \in I$, a smooth curve $\gamma : \mathbb{R} \to M$ is a geodesic with respect to ∇ if and only if $\Gamma_t \circ \gamma : \mathbb{R} \to M$ is a geodesic with respect to ∇.

Fact 2 is useful because is typically easier to check whether (1) holds than whether (2) holds. Facts 1 and 2 allow one to say which vector fields on $\mathbb{R}^4$ generate automorphisms of Galilean, Newtonian, or Minkowski spacetime.

4 Spacetime structure

We now have the tools to compare the structure of these three spacetimes. The four simple propositions in this section demonstrate that Galilean, Newtonian, and Minkowski spacetime stand in the following relationships, where the symbol "<" means "has less structure than (according to SYM*)."

$$\begin{array}{ccc} & \not< & \\ \textbf{Galilean spacetime} & & \textbf{Minkowski spacetime} \\ & \not> & \\ \nwarrow & & \nearrow \\ & \textbf{Newtonian spacetime} & \end{array}$$

In what follows, we denote the automorphism group of Galilean spacetime by Aut(Galilean), and similarly for the automorphism groups of Newtonian and Minkowski spacetime.

The first two propositions demonstrate that according to SYM*, Galilean spacetime and Minkowski spacetime have incomparable amounts of structure.

Proposition 1. *It is not the case that Aut(Galilean) $\subset$ Aut(Minkowski).*

Proof. We exhibit an automorphism of Galilean spacetime that is not an automorphism of Minkowski spacetime. Let $p \in \mathbb{R}^4$ and consider the smooth field

$$\xi^a = \left(x^1 - x^1(p)\right)\left(\frac{\partial}{\partial x^2}\right)^a.$$

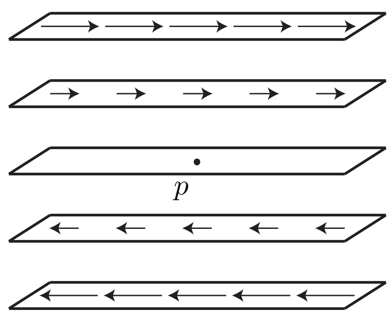
[8] Neither Malament (2012) nor Geroch (1972) contain a proof. One can consult Hall (2004) for another statement of Fact 2 or Poor (1981) for a proof using more advanced tools from the theory of vector bundles.

Let $\Gamma_t : \mathbb{R}^4 \to \mathbb{R}^4$ be a flow map associated with ξ^a. The map Γ_t is an automorphism of Galilean spacetime, but not of Minkowski spacetime. The proof is a simple computation. We first compute that

$$\begin{aligned}\mathcal{L}_\xi t_{ab} &= \xi^n \nabla_n t_{ab} + t_{nb} \nabla_a \xi^n + t_{an} \nabla_b \xi^n \\ &= t_{nb}(d_a x^1)\big(\frac{\partial}{\partial x^2}\big)^n + t_{an}(d_b x^1)\big(\frac{\partial}{\partial x^2}\big)^n \\ &= \mathbf{0}.\end{aligned}$$

The first equality follows from (Malament, 2012, Proposition 1.7.4) and the second and third follow from the definitions of t_{ab}, ∇, and ξ^a. Fact 1 then implies that $(\Gamma_t)^*(t_{ab}) = t_{ab}$. In a similar manner one proves that $(\Gamma_t)^*(h^{ab}) = h^{ab}$. Lastly, one computes that $\nabla_a \nabla_b \xi^c = \mathbf{0}$. Since ∇ is flat (i.e. $R^a{}_{bcd} = \mathbf{0}$), Fact 2 implies that Γ_t preserves ∇. So Γ_t is an automorphism of Galilean spacetime. But one computes in the same manner as above that $\mathcal{L}_\xi \eta_{ab} \neq \mathbf{0}$. Fact 1 then implies that Γ_t is not an automorphism of Minkowski spacetime. □

The vector field ξ^a defined above is sometimes referred to as a "Galilean velocity boost." It can be pictured as follows.[9]



Proposition 1 shows that Galilean velocity boosts are not automorphisms of Minkowski spacetime. A Galilean velocity boost "breaks" the lightcone structure of Minkowski spacetime. And so according to SYM*, it is not the case that Galilean spacetime has more structure than Minkowski spacetime.

Proposition 2 demonstrates the converse. According to SYM*, it is not the case that Minkowski spacetime has more structure than Galilean spacetime. Velocity boosts in Minkowski spacetime — sometimes called "Lorentz boosts" — are not automorphisms of Galilean spacetime. They "break" the preferred simultaneity slice structure of Galilean spacetime.

Proposition 2. *It is not the case that Aut(Minkowski) $\subset$ Aut(Galilean).*

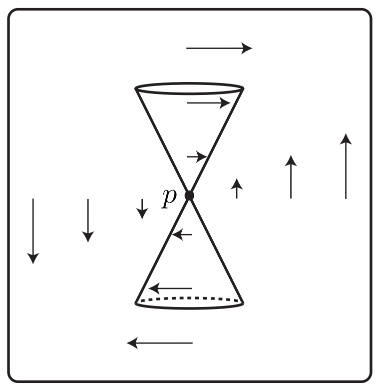
[9]Geroch (1978) sometimes refers to Galilean velocity boosts as "beveling the deck". The thought is that one can picture Galilean spacetime as a deck of cards, where each card represents a simultaneity slice. A Galilean velocity boost "bevels" this deck.

Proof. We exhibit an automorphism of Minkowski spacetime that is not an automorphism of Galilean spacetime. Let $p \in \mathbb{R}^4$ and consider the smooth field

$$\kappa^a = \big(x^2 - x^2(p)\big)\left(\frac{\partial}{\partial x^1}\right)^a + \big(x^1 - x^1(p)\big)\left(\frac{\partial}{\partial x^2}\right)^a.$$

Let $\Gamma_t : \mathbb{R}^4 \to \mathbb{R}^4$ be a flow map associated with κ^a. The map Γ_t is an automorphism of Minkowski spacetime, but not of Galilean spacetime. The proof is again a simple computation. One first computes that $\mathcal{L}_\kappa \eta_{ab} = \mathbf{0}$, so Fact 1 implies that Γ_t is an automorphism of Minkowski spacetime. But one can verify that $\mathcal{L}_\kappa t_{ab} \neq \mathbf{0}$, so Γ_t is not an automorphism of Galilean spacetime. □

The vector field κ^a can be pictured as follows.



Proposition 2 shows that Lorentz velocity boosts like κ^a are not automorphisms of Galilean spacetime. They preserve the lightcone structure of Minkowski spacetime, but they do not preserve the preferred simultaneity slice structure of Galilean spacetime. So according to SYM*, it is not the case that Minkowski spacetime has more structure than Galilean spacetime.

Propositions 1 and 2 provide a precise sense in which Minkowski spacetime and Galilean spacetime have incomparable amounts of structure. Each spacetime has some structure that the other lacks. Minkowski spacetime has lightcone structure that Galilean spacetime lacks. And conversely, Galilean spacetime has preferred simultaneity slice structure that Minkowski spacetime lacks.

The final two propositions show that the structure of Newtonian spacetime *is* comparable with the structure of Minkowski and Galilean spacetime. In fact, Newtonian spacetime has more structure than both. We begin with the simpler of the two. According to SYM*, Newtonian spacetime has more structure than Galilean spacetime.

Proposition 3. *Aut(Newtonian)* $\subsetneq$ *Aut(Galilean).*

Proof. It follows by definition that Aut(Newtonian) $\subset$ Aut(Galilean). Now consider the smooth field ξ^a and the map Γ_t from the proof of Proposition 1.

We already know that Γ_t is an automorphism of Galilean spacetime. But one can easily compute that $\mathcal{L}_\xi \lambda^a \neq \mathbf{0}$, so Γ_t is not an automorphism of Newtonian spacetime. □

Proposition 3 demonstrates that according to SYM*, Newtonian spacetime has more structure than Galilean spacetime. This is perfectly intuitive. After all, Newtonian spacetime is obtained by "adding structure" to Galilean spacetime in the form the vector field λ^a.

Proposition 4 may not be quite so obvious. According to SYM*, Newtonian spacetime *also* has more structure than Minkowski spacetime. The tuples $(\mathbb{R}^4, t_{ab}, h^{ab}, \nabla, \lambda^a)$ and $(\mathbb{R}^4, \eta_{ab})$ might not appear to bear the same structural relationship to one another as do the tuples $(\mathbb{R}^4, t_{ab}, h^{ab}, \nabla, \lambda^a)$ and $(\mathbb{R}^4, t_{ab}, h^{ab}, \nabla)$. But this is misleading. There is a precise sense in which Newtonian spacetime has more structure than Minkowski spacetime. The structure of Newtonian spacetime allows one to explicitly define the lightcone structure of Minkowski spacetime.[10]

Proposition 4. *Aut(Newtonian)* $\subsetneq$ *Aut(Minkowski).*

Proof. We first show that Aut(Newtonian) $\subset$ Aut(Minkowski). One can easily see that

$$\eta^{ab} = \lambda^a \lambda^b - h^{ab}.$$

Automorphisms of Newtonian spacetime preserve λ^a and h^{ab}, so they also preserve η_{ab} and are automorphisms of Minkowski spacetime. Now suppose for contradiction that Aut(Minkowski) $\subset$ Aut(Newtonian). Proposition 3 then implies that Aut(Minkowski) $\subset$ Aut(Galilean), contradicting Proposition 2. □

5 Conclusion

This discussion takes a step towards explicating some of the relationships that hold between different spacetime theories. We have seen a precise sense in which Newtonian spacetime has more structure than both Galilean spacetime and Minkowski spacetime. But in this same precise sense, Minkowski spacetime and Galilean spacetime have *incomparable* amounts of structure. Neither spacetime has more structure than the other; each has some structure that the other lacks.

The discussion yields two philosophical payoffs. First, it substantiates a somewhat paradoxical suggestion made by Earman (Earman, 1989, p. 34). Earman has suggested that Newtonian spacetime is a more natural stepping-stone to Minkowski spacetime, and to relativistic spacetimes in general, than Galilean spacetime is. One can make Earman's suggestion perfectly precise by considering the comparative structure of these spacetimes. If one begins with Newtonian spacetime, one can arrive at Minkowski spacetime simply by judiciously defining new fields and excising the right structure. But the same cannot be said of

[10] Earman (1989) suggests the idea behind the following proof.

Galilean spacetime. No matter what structure one excises from Galilean spacetime, one will not end up with Minkowski spacetime. Galilean spacetime does not allow one to define the lightcone structure of Minkowski spacetime.

Second, the discussion yields a remark concerning the scope and applicability of the following methodological principle.

Structural parsimony: All other things equal, we should prefer theories that posit less structure.

North (2009) endorses this kind of principle and uses it to argue that we should prefer the Hamiltonian formulation of classical mechanics over the Lagrangian formulation.[11] She explains the principle as follows.

> This is a principle informed by Ockham's razor; though it is not just that, other things being equal, it is best to go with the ontologically minimal theory. It is not that, other things being equal, we should go with the fewest entities, but that we should go with the least structure. (North, 2009, p. 64)

The structural parsimony principle still stands in need of clarification. In order to apply it, one would first need to make precise exactly what the "other things" are and what it would mean for them to be "equal". And furthermore, it is not immediately clear what it might mean to "prefer" one theory over another, nor why a theory's structural parsimony might license this preference.[12]

Before clarifying the structural parsimony principle, however, one can make a cautionary remark about its scope and applicability. There are situations where structural parsimony does not help one to adjudicate between theories. Our discussion has provided a precise sense in which Galilean spacetime and Minkowski spacetime have *incomparable* amounts of structure. It is not the case that Galilean spacetime has less structure than Minkowski spacetime, nor is it the case that Minkowski spacetime has less structure than Galilean spacetime. In cases of incomparable structure, structural parsimony principles simply are not applicable. Physical theories do not always progress in the direction of less structure.[⋆]

References

Barrett, T. W. (2014). On the structure of classical mechanics. *The British Journal for the Philosophy of Science (forthcoming).*

Curiel, E. (2014). Classical mechanics is Lagrangian; it is not Hamiltonian. *The British Journal for the Philosophy of Science.*

[11]See Halvorson (2011), Swanson and Halvorson (2012), Curiel (2014) and Barrett (2014) for other discussions of the relationship between Hamiltonian and Lagrangian mechanics.

[12]Related issues are discussed by Dasgupta (2014).

[⋆]This material is based upon work supported by the National Science Foundation under Grant No. DGE 1148900.

Dasgupta, S. (2014). Symmetry as an epistemic notion (twice over). *The British Journal for the Philosophy of Science (forthcoming).*

Earman, J. (1989). *World Enough and Spacetime: Absolute versus Relational Theories of Space and Time.* MIT.

Geroch, R. (1972). Differential geometry. *Lecture notes.*

Geroch, R. (1978). *General Relativity from A to B.*

Hall, G. S. (2004). *Symmetries and Curvature Structure in General Relativity.*

Halvorson, H. (2011). Natural structures on state space. *Manuscript.*

Malament, D. B. (2012). *Topics in the Foundations of General Relativity and Newtonian Gravitation Theory.*

Maudlin, T. (2012). *Philosophy of Physics: Space and Time.*

North, J. (2009). The 'structure' of physics: A case study. *Journal of Philosophy.*

Poor, W. A. (1981). *Differential Geometric Structures.*

Swanson, N. and Halvorson, H. (2012). On North's 'The structure of physics'. *Manuscript.*

Weatherall, J. O. (2014). Are Newtonian gravitation and geometrized Newtonian gravitation theoretically equivalent? *Manuscript.*

Appendix

This appendix contains a proof of Fact 2, which we restate here for convenience.

Fact 2. *Let ξ^a be a smooth vector field on M and let ∇ be a derivative operator on M with associated curvature field $R^a{}_{bcd}$. Then the following conditions are equivalent.*

(1) $\nabla_a \nabla_b \xi^m = R^m{}_{bna} \xi^n$.

(2) For all local one-parameter groups of diffeomorphisms $\{\Gamma_t : U \to \Gamma_t[U]\}_{t \in I}$ generated by ξ^a and all $t \in I$, a smooth curve $\gamma : \mathbb{R} \to M$ is a geodesic (with respect to ∇) if and only if $\Gamma_t \circ \gamma : \mathbb{R} \to M$ is a geodesic (with respect to ∇).

We will prove this fact using four lemmas. Lemmas 2–4 establish a chain of equivalences that immediately yield the desired result. Lemma 1 will be useful in the proofs of Lemmas 2–4. For ease of exposition we will assume that the vector field ξ^a is complete so that the flow maps $\Gamma_t : M \to M$ are defined on all of M.

It will be useful to first set up some notation. Let M be a manifold, and let $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t)$ be a smooth field on M that is also smooth in t. We define a smooth field $\frac{d}{dt}(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t))\big|_{t=0}$ on M. The field $\frac{d}{dt}(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t))\big|_{t=0}$ takes the value

$$\lim_{t \to 0} \frac{1}{t}\big(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t)\big|_p - \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(0)\big|_p\big)$$

at a point $p \in M$ (Geroch, 1972, §21).

Let ξ^a be a smooth field on M with $\{\Gamma_t : M \to M\}_{t \in \mathbb{R}}$ an associated one-parameter group of diffeomorphisms. Let $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ be a smooth field on M. In the following lemma, for ease of notation, we will write $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t) := (\Gamma_t)^*(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s})$.

Lemma 1. *Let ξ^a be a smooth field on M with $\{\Gamma_t : M \to M\}_{t \in \mathbb{R}}$ an associated one-parameter group of diffeomorphisms. Let $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ be a smooth field on M. Then*

$$\frac{d}{dt}\Big(\nabla_n\big(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t)\big)\Big)\Big|_{t=0} = \nabla_n\Big(\frac{d}{dt}\big(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}(t)\big)\Big|_{t=0}\Big).$$

Proof. The proof is essentially the same regardless of the index structure of $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$, so we work with a representative case λ^a_b. Let $p \in M$ with (U, φ) a neighborhood of p. Let $\overline{\nabla}_n$ be the coordinate derivative operator on (U, φ) (Malament, 2012, Prop. 1.7.11). There is a smooth symmetric field C^a_{bc} on M such that

$$\nabla_n \lambda^a_b(t) = \overline{\nabla}_n \lambda^a_b(t) + \lambda^a_m(t) C^m_{nb} - \lambda^m_b(t) C^a_{mn} \tag{1}$$

(Malament, 2012, Prop. 1.7.3). Differentiating both sides of (1) with respect to t yields

$$\begin{aligned} \frac{d}{dt}(\nabla_n \lambda^a_b(t))\big|_{t=0} = \frac{d}{dt}(\overline{\nabla}_n \lambda^a_b(t))\big|_{t=0} + C^m_{nb}\big(\frac{d}{dt}\lambda^a_m(t)\big|_{t=0}\big) \\ - C^a_{mn}\big(\frac{d}{dt}\lambda^m_b(t)\big|_{t=0}\big) \end{aligned} \tag{2}$$

Now writing $\lambda^a_b(t) = \sum_{i=1}^n \sum_{j=1}^n \overset{ij}{\lambda}(t)\big(\frac{\partial}{\partial u^i}\big)^a (d_b u^j)$, we can compute the first term on the right-side of equation (2) using the definition of the coordinate derivative operator $\overline{\nabla}_n$.

$$\begin{aligned} \frac{d}{dt}(\overline{\nabla}_n \lambda^a_b(t))\big|_{t=0} &= \frac{d}{dt}\Big(\sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \big(\frac{\partial \overset{ij}{\lambda}(t)}{\partial u^k}\big)\big(\frac{\partial}{\partial u^i}\big)^a (d_b u^j)(d_n u^k)\Big)\Big|_{t=0} \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \Big(\frac{d}{dt}\big(\frac{\partial \overset{ij}{\lambda}(t)}{\partial u^k}\big)\big|_{t=0}\Big)\big(\frac{\partial}{\partial u^i}\big)^a (d_b u^j)(d_n u^k) \\ &= \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \Big(\frac{\partial}{\partial u^k}\big(\frac{d}{dt}\overset{ij}{\lambda}(t)\big|_{t=0}\big)\Big)\big(\frac{\partial}{\partial u^i}\big)^a (d_b u^j)(d_n u^k) \\ &= \overline{\nabla}_n\big(\frac{d}{dt}\lambda^a_b(t)\big|_{t=0}\big) \end{aligned}$$

The first equality follows from the definition of the coordinate derivative operator $\overline{\nabla}_n$ (Malament, 2012, p. 65), the second from the linearity of $\frac{d}{dt}$, the third from the fact that the coordinates in (U, φ) do not depend on t, and the fourth again from the definition of the coordinate derivative operator.

Plugging this into equation (2) we have that

$$\begin{aligned}\frac{d}{dt}\big(\nabla_n \lambda^a_b(t)\big)\big|_{t=0} &= \frac{d}{dt}(\overline{\nabla}_n \lambda^a_b(t))\big|_{t=0} + C^m_{nb}\big(\frac{d}{dt}\lambda^a_m(t)\big|_{t=0}\big)\\ &\quad - C^a_{mn}\big(\frac{d}{dt}\lambda^m_b(t)\big|_{t=0}\big)\\ &= \overline{\nabla}_n\big(\frac{d}{dt}\lambda^a_b(t)\big|_{t=0}\big) + C^m_{nb}\big(\frac{d}{dt}\lambda^a_m(t)\big|_{t=0}\big)\\ &\quad - C^a_{mn}\big(\frac{d}{dt}\lambda^m_b(t)\big|_{t=0}\big)\\ &= \nabla_n\big(\frac{d}{dt}\lambda^a_b(t)\big|_{t=0}\big)\end{aligned}$$

The first equality is simply equation (2), the second follows from the above computation, and the third follows from (Malament, 2012, Prop. 1.7.3) and the definition of C^a_{bc}. □

Lemma 2. *Let ξ^a be a smooth field on M and ∇ a derivative operator on M with associated curvature field $R^a{}_{bcd}$. Then ∇ commutes with $\mathcal{L}_\xi$ (in their action on any tensor field) iff $\nabla_a \nabla_b \xi^m = R^m{}_{bna}\xi^n$.*

Proof. (Malament, 2012, p. 70). □

Lemma 3. *Let ξ^a be a smooth field on M and ∇ a derivative operator on M. Then the following conditions are equivalent:*

(1) $\mathcal{L}_\xi$ and ∇ commute (in their action on any tensor field).

(2) For every (local) 1-parameter group of flows $\{\Gamma_t : U \to \Gamma_t[U]\}_{t\in I}$ generated by ξ^a and all $t \in I$, $(\Gamma_t)^*(\nabla_n \lambda^{a_1\dots a_r}_{b_1\dots b_s}) = \nabla_n\big((\Gamma_t)^*(\lambda^{a_1\dots a_r}_{b_1\dots b_s})\big)$.

Proof. We first show that (2) implies (1). Let ξ^a, M and ∇ be as in the statement of the proposition, and let $\lambda^{a_1\dots a_r}_{b_1\dots b_s}$ be an arbitrary smooth tensor field on M. We compute that

$$\begin{aligned}\mathcal{L}_\xi(\nabla_n \lambda^{a_1\dots a_r}_{b_1\dots b_s})_{|p} &= \lim_{t\to 0}\frac{1}{t}\left((\Gamma_t)^*(\nabla_n \lambda^{a_1\dots a_r}_{b_1\dots b_s})_{|p} - (\nabla_n \lambda^{a_1\dots a_r}_{b_1\dots b_s})_{|p}\right)\\ &= \lim_{t\to 0}\frac{1}{t}\left(\nabla_n\big((\Gamma_t)^*(\lambda^{a_1\dots a_r}_{b_1\dots b_s})\big)_{|p} - (\nabla_n \lambda^{a_1\dots a_r}_{b_1\dots b_s})_{|p}\right)\\ &= \left(\frac{d}{dt}\nabla_n\big((\Gamma_t)^*(\lambda^{a_1\dots a_r}_{b_1\dots b_s})\big)\big|_{t=0}\right)_{|p}\\ &= \nabla_n\left(\frac{d}{dt}\big((\Gamma_t)^*(\lambda^{a_1\dots a_r}_{b_1\dots b_s})\big)\big|_{t=0}\right)_{|p}\end{aligned}$$

$$= \nabla_n \left(\lim_{t \to 0} \frac{1}{t} \big((\Gamma_t)^*(\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}) - (\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}) \big) \right)_{|p}$$
$$= \nabla_n (\mathcal{L}_\xi \lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s})_{|p}$$

The first and sixth equalities follow from the definition of the Lie derivative, the second equality from the assumption that (2) holds, the third and fifth from the definition of $\frac{d}{dt}$, and the fourth from Lemma 1. Since $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$ was arbitrary, this means that $\mathcal{L}_\xi$ and ∇ commute.

We now show that (1) implies (2). This direction requires a bit more work. Assume that $\mathcal{L}_\xi$ and ∇_n commute. The proof is essentially the same regardless of the index structure of $\lambda^{a_1 \ldots a_r}_{b_1 \ldots b_s}$, so we work with a representative case λ^a_b. Let ν_a, μ^b, and β^n be vectors at a point $p \in U$ and define $f : I \to \mathbb{R}$ by

$$f(t) = (\Gamma_{-t})^* \big(\nabla_n (\Gamma_t)^* (\lambda^a_b) \big)_{|p} \nu_a \mu^b \beta^n.$$

Note that $f(0) = (\nabla_n \lambda^a_b)_{|p} \nu_a \mu^b \beta^n$. We show that $f'(t) = 0$ for every $t \in I$. This will imply that f is constant, and therefore that for every $t \in I$,

$$f(t) = (\nabla_n \lambda^a_b)_{|p} \nu_a \mu^b \beta^n = (\Gamma_{-t})^* \big(\nabla_n (\Gamma_t)^* (\lambda^a_b) \big)_{|p} \nu_a \mu^b \beta^n.$$

Since ν_a, μ^b, β^n, and p are arbitrary, this will imply that $(\Gamma_t)^*(\nabla_n \lambda^a_b) = \nabla_n (\Gamma_t)^* (\lambda^a_b)$, as desired.

So we compute.

$$\begin{aligned}
f'(t) &= \lim_{s \to 0} \frac{1}{s} \Big((\Gamma_{-t-s})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} \nu_a \mu^b \beta^n \\
&\qquad - (\Gamma_{-t})^* \big(\nabla_n (\Gamma_t)^* (\lambda^a_b) \big)_{|p} \nu_a \mu^b \beta^n \Big) \\
&= \lim_{s \to 0} \frac{1}{s} \Big((\Gamma_{-t-s})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \\
&\qquad - (\Gamma_{-t})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \\
&\qquad + (\Gamma_{-t})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \\
&\qquad - (\Gamma_{-t})^* \big(\nabla_n (\Gamma_t)^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \Big) \\
&= \lim_{s \to 0} \frac{1}{s} \Big((\Gamma_{-t-s})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \\
&\qquad - (\Gamma_{-t})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \Big) && \text{(A)} \\
&\quad + \lim_{s \to 0} \frac{1}{s} \Big((\Gamma_{-t})^* \big(\nabla_n (\Gamma_{t+s})^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \\
&\qquad - (\Gamma_{-t})^* \big(\nabla_n (\Gamma_t)^* (\lambda^a_b) \big)_{|p} (\nu_a \mu^b \beta^n) \Big) && \text{(B)}
\end{aligned}$$

The first equality follows from the definition of the derivative, the second follows by adding and subtracting the same real number, and the third follows from the linearity of limits.

We now compute the terms labelled (A) and (B) above.

$$\begin{aligned}
(\mathrm{A}) &= \lim_{s\to 0}\frac{1}{s}\bigg((\Gamma_{-t-s})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n) \\
&\qquad - (\Gamma_{-t})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n)\bigg) \\
&= \lim_{s\to 0}\frac{1}{s}\bigg((\Gamma_{-t})^*\circ(\Gamma_{-s})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n) \\
&\qquad - (\Gamma_{-t})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n)\bigg) \\
&= \lim_{s\to 0}\frac{1}{s}\bigg((\Gamma_{-s})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}} \\
&\qquad - \nabla_n(\Gamma_{t+s})^*(\lambda^a_b)_{|_{\Gamma_{-t}(p)}}\bigg)\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= \mathcal{L}_{-\xi}\big(\nabla_n(\Gamma_t)^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= -\mathcal{L}_{\xi}\big(\nabla_n(\Gamma_t)^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n)
\end{aligned}$$

The first equality follows from the definition of (A) above, the second from properties of flow maps, the third from the definition of the pullback and the linearity of the limit, the fourth from the definition of the Lie derivative, and the fifth from properties of the Lie derivative.

And likewise, we compute that

$$\begin{aligned}
(\mathrm{B}) &= \lim_{s\to 0}\frac{1}{s}\bigg((\Gamma_{-t})^*\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n) \\
&\qquad - (\Gamma_{-t})^*\big(\nabla_n(\Gamma_{t})^*(\lambda^a_b)\big)_{|_p}(\nu_a\mu^b\beta^n)\bigg) \\
&= \lim_{s\to 0}\frac{1}{s}\bigg(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)_{|_{\Gamma_{-t}(p)}} - \nabla_n(\Gamma_t)^*(\lambda^a_b)_{|_{\Gamma_{-t}(p)}}\bigg)\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= \bigg(\frac{d}{ds}\big(\nabla_n(\Gamma_{t+s})^*(\lambda^a_b)\big)\big|_{s=0}\bigg)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= \nabla_n\big(\frac{d}{ds}(\Gamma_{t+s})^*(\lambda^a_b)\big|_{s=0}\big)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= \nabla_n\bigg(\lim_{s\to 0}\frac{1}{s}\big((\Gamma_{t+s})^*(\lambda^a_b) - (\Gamma_t)^*(\lambda^a_b)\big)\bigg)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\
&= \nabla_n\big(\mathcal{L}_\xi(\Gamma_t)^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}}\cdot(\Gamma_{-t})_*(\nu_a\mu^b\beta^n)
\end{aligned}$$

The first equality follows from the definition of (B), the second from the definition of the pullback and the linearity of the limit, the third from the definition of $\frac{d}{ds}$, the fourth from Lemma 1, the fifth again from the definition of $\frac{d}{ds}$, and the sixth from the definition of the Lie derivative.

Finally putting all of these computations together we have that

$$\begin{aligned} f'(t) &= \text{(A)} + \text{(B)} \\ &= -\mathcal{L}_\xi\big(\nabla_n(\Gamma_t)^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}} \cdot (\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\ &\qquad + \nabla_n\big(\mathcal{L}_\xi(\Gamma_t)^*(\lambda^a_b)\big)_{|_{\Gamma_{-t}(p)}} \cdot (\Gamma_{-t})_*(\nu_a\mu^b\beta^n) \\ &= 0. \end{aligned}$$

The third equality follows from the assumption that ∇_n and $\mathcal{L}_\xi$ commute. □

Lemma 4. *Let $f : M \to M$ be a diffeomorphism and ∇ a derivative operator on M. Then the following are equivalent.*

(1) γ is a geodesic iff $f \circ \gamma$ is a geodesic (with respect to ∇).

(2) For all tensor fields $\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}$, $f^*(\nabla_n\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}) = \nabla_n f^*(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s})$.

Proof. We first remark that (2) holds if and only if

$$f_*(\nabla_n\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}) = \nabla_n f_*(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}) \tag{3}$$

for every smooth field $\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}$. For if (2) holds, then in particular it must be that $f^*(\nabla_n f_*(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s})) = \nabla_n f^*(f_*(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}))$. And this is equivalent to equation (3) since $f^* \circ f_*$ is the identity map. The converse implication is established in precisely the same manner.

We now show that (2) implies (1). Assume that (2) holds and let γ be a geodesic with tangent field ξ^a. We know that $\xi^n\nabla_n\xi^a = \mathbf{0}$, and therefore that $f_*(\xi^n\nabla_n\xi^a) = \mathbf{0}$. By the preceding remark concerning the equivalence of (2) and equation (3), this implies that $f_*(\xi^n)\nabla_n f_*(\xi^a) = \mathbf{0}$, so $f \circ \gamma$ is a geodesic. If $f \circ \gamma$ is a geodesic, a parallel argument shows that γ is a geodesic.

Now assume that (1) holds. Define a new derivative operator ∇' on M by

$$\nabla'_n(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}) = f^*\nabla_n f_*(\lambda^{a_1\ldots a_r}_{b_1\ldots b_s}).$$

We will show that $\nabla' = \nabla$. This will complete the proof again by the remark concerning the equivalence of (2) and equation (3). It will suffice to show that ∇ and ∇' admit exactly the same class of geodesics (Malament, 2012, Prop. 1.7.8).

Let γ be a geodesic with respect to ∇ with tangent field ξ^a. We see that

$$\xi^n\nabla'_n\xi^a = \xi^n(f^*\nabla_n f_*(\xi^a)) = f_*(\xi^n)\nabla_n f_*(\xi^a) = \mathbf{0}.$$

The second equality follows by the definition of the pullback, and the last equality follows since $f \circ \gamma$ is a geodesic by (1). Let γ' be a geodesic with respect to ∇' with tangent field ξ'^a. We see that

$$\mathbf{0} = \xi'^n\nabla'_n\xi'^a = \xi'^n(f^*\nabla_n f_*(\xi'^a)) = f_*(\xi'^n)\nabla_n f_*(\xi'^a).$$

So $f \circ \gamma'$ is a geodesic with respect to ∇, so γ' is too by (1). Therefore ∇ and ∇' admit the same class of geodesics. □

Lemmas 2, 3 and 4 immediately imply Fact 2.